NBER WORKING PAPER SERIES

IMPROVING HOUSEHOLD DEBT MANAGEMENT WITH ROBO-ADVICE

Ida Chak
Karen Croxson
Francesco D'Acunto
Jonathan Reuter
Alberto G. Rossi
Jonathan M. Shaw

Working Paper 30616
http://www.nber.org/papers/w30616

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2022

Ida Chak and Karen Croxson are at the Financial Conduct Authority of the United Kingdom (FCA). Francesco D'Acunto and Alberto Rossi are at Georgetown University. Jonathan Reuter is at Boston College and National Bureau of Economic Research. Jonathan Shaw is at the FCA and Institute for Fiscal Studies. Our experimental design was preregistered in the AEA RCT registry (AEARCTR-0006447) and implemented by the FCA. We thank the FCA for funding this research. We thank Chris Burke and Maura Feddersen for help designing and implementing the experiment. The views expressed in this paper are those of the authors and should not be attributed to the Financial Conduit Authority. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Improving Household Debt Management with Robo-Advice
Ida Chak, Karen Croxson, Francesco D'Acunto, Jonathan Reuter, Alberto G. Rossi, and Jonathan M. Shaw
NBER Working Paper No. 30616
November 2022
JEL No. D14,D91,G51,G53

ABSTRACT

Poor debt-management skills lower financial security and wealth accumulation. Because optimal solutions to credit repayment problems depend on neither risk preferences nor beliefs, loan repayment is a prime application for robo-advising. Vulnerable households, though, tend to distrust new technologies and override suggestions that do not align with ingrained heuristics, such as matching the minimum payment on a credit card balance. Lower adoption rates by these groups might increase rather than reduce wealth inequalities. To assess these trade-offs, we design and implement an RCT in which robo-advice for borrower repayment decisions is offered to a set of representative UK consumers. The availability of free robo-advice significantly improves average loan repayment choices. When their willingness to pay is elicited, many subjects report values larger than the monetary benefits of the tool, perhaps due to lower cognitive and psychological costs decision-makers face when making assisted choices. Non-adopters and overriders report lower trust in algorithms at the end of the experiment. We find no evidence of learning from robo-advice, which barely improves subsequent unassisted choices, even when paired with explicit tips. In fact, robo-advice usage crowds out learning-by-doing, which is highest for those who make all choices unassisted.

Ida Chak
Financial Conduct Authority
12 Endeavour Square
London E20 1JN
United Kingdom
Ida-chak@hotmail.com

Karen Croxson
Financial Conduct Authority
12 Endeavour Square
London E20 1JN
United Kingdom
karen.croxson@fca.org.uk

Francesco D'Acunto
McDonough School of Business
Georgetown University
Washington, DC 20057
francesco.dacunto@georgetown.edu

Jonathan Reuter
Carroll School of Management
Boston College
224B Fulton Hall
140 Commonwealth Avenue
Chestnut Hill, MA 02467
and NBER
reuterj@bc.edu

Alberto G. Rossi
McDonough School of Business
Georgetown University
Washington, DC 20057
agr60@georgetown.edu

Jonathan M. Shaw
Financial Conduct Authority
12 Endeavour Square
London E20 1JN
United Kingdom
and Institute For Fiscal Studies
Jonathan.Shaw@fca.org.uk

A data appendix is available at
http://www.nber.org/data-appendix/w30616
A randomized controlled trials registry entry is available at
https://www.socialscienceregistry.org/trials/6447/

The origins and consequences of growing wealth inequality are among the most debated issues in today's society (Keister (2000); Zucman (2019)). A large literature links the lack of wealth accumulation by low-income households (Campbell, Ramadorai, and Ranish (2019)) to low saving rates and financial returns due to financial illiteracy and costly access to advice (Lusardi and Mitchell (2014), Chalmers and Reuter (2020)). Differences in wealth accumulation also depend on how much debt households accumulate and how (in)efficiently they manage it, which is especially relevant for vulnerable households, who tend to display low financial literacy (Lusardi, Schneider, and Tufano (2011); Drexler, Fischer, and Schoar (2014)) and misunderstand the implications of annual percentage rates (APRs), minimum payments, and late fees (Agarwal, Chomsisengphet, and Liu (2010); Agarwal and Mazumder (2013)). Heuristics and biases in loan repayment can lead to debt spirals and further reduce wealth (Agarwal et al. (2008); Argyle, Nadauld, and Palmer (2020); Gathergood et al. (2019); Lusardi and Tufano (2015); Ponce, Seira, and Zamarripa (2017)).

Although we understand many of these heuristics and biases (Zinman (2015); Agarwal and Zhang (2015); Stango and Zinman (2021)), the question of how to effectively and efficiently minimize their impact remains open. For instance, while many countries have imposed disclosure requirements on lenders, it is unclear that borrowers with low levels of financial literacy are able to process the technical information embedded in disclosures (Adams et al. (2021); Bertrand and Morse (2011); Navarro-Martinez et al. (2011)).[2] Financial literacy programs can improve outcomes (Bu et al. (2020)), but the ability to scale them up is limited by the financial and cognitive costs faced by those who deliver classes and those who attend them.

In this paper, we implement a randomized control trial (RCT) in the United Kingdom to assess the take-up and effectiveness of an alternative means to improve households' loan repayment choices: robo-advice (D'Acunto, Prabhala, and Rossi (2019); D'Acunto and Rossi (2021b)).[3] In principle, loan repayment is a natural setting for robo-advice because, unlike asset allocation decisions, the optimal choice depends on neither risk preferences nor beliefs about expected returns and correlations. Moreover, in contrast to human advice and debt counseling,

[2] Agarwal et al. (2015) show that regulation can improve loan repayment by limiting credit card fees.
[3] The trial was preregistered in the AEA RCT registry (AEARCTR-0006447) and implemented by the Financial Conduct Authority (FCA) of the United Kingdom.

robo-advising is an easy-to-scale technology that can be delivered cheaply through personal devices, making it a viable resource for low-income and vulnerable households (D'Acunto and Rossi (2022, 2021a)). At the same time, algorithmic aversion, which is especially common among the least sophisticated consumers (Niszczota and Kaszás (2020)), might limit demand for robo-advice tools among the demographic groups that would most benefit the most, thereby potentially increasing wealth inequalities rather than reducing them.

In our RCT, subjects allocate funds across loans in nine real-life loan repayment scenarios. Each scenario includes loans that vary with respect to interest rates, balances, and levels of difficulty. The experiment includes three phases, each consisting of three randomly chosen scenarios. In the *pre-intervention* phase, we assess the mistakes subjects make when managing multiple loans independently. This phase confirms earlier evidence that loan repayment mistakes are common and have sizable economic consequences on debtors' (experimental) wealth: the average percentage of savings forgone relative to the optimal choices is 21.9%, which translates into 29 basis points of interest and fees per month and compounds to 3.55% higher payment per year. For the average US family with $6,270 in credit card debt around the time of our RCT,[4] these mistakes would result in an overpayment of $222.42 per year. The magnitudes increase to 37 basis points per month and $249.37 when we focus on the hard versions of the problems. Consistent with intuition and earlier research (Lusardi (2012)), higher financial literacy, numeracy, and patience reduce the extent of mistakes.

In the *intervention* phase, we introduce robo-advice for loan repayment. Subjects work on three new allocation scenarios after being randomly assigned to one of five experimental arms: a control group for whom the task is identical to the pre-intervention phase; groups for whom robo-advice is available for free with or without explicit tips about the financial principles behind optimal strategies ("robo-advice with education"); and groups for whom it can be obtained for a fee with or without education. In all four treatment arms, subjects are told the maximum interest and fees that they could save by using the robo-advisor. In the treatments with paid robo-advice, we elicit subjects' willingness to pay (WTP) for the robo-advice tool in an incentive-compatible

[4] Source: Federal Reserve Survey of Consumer Finances (Link)

way.[5] The *post-intervention* phase mirrors the pre-intervention phase, in which subjects solve three loan management problems without access to robo-advice. Subjects receive no feedback about optimal solutions after their choices, which are used to determine their incentive payments.

The five experimental arms allow us to determine whether (and by how much) exposure to robo-advice improves loan repayment, as well as a series of related questions: (i) which forms of robo-advice are in the greatest demand; (ii) how much subjects are willing to pay to obtain robo-advice; and (iii) which subjects override the robo-advisor's recommendations. The post-intervention phase reveals whether exposure to optimal allocations by the robo-advisor, with or without education, helps borrowers learn the optimal rules for loan repayment or, instead, whether robo-advice needs to be continually provided to improve repayment decisions.

We find that subjects exposed to free robo-advice improve their repayment strategies significantly relative to the control group. Because loan repayment problems have unambiguous ex-ante solutions, any subject who implements the recommended repayment strategy will minimize interest and late fees. As a result, the Treatment on the Treated (TOT) effect—the effect of robo-advising on the subjects who seek it—is large. The average pre-intervention losses (21.9%) decline by 19.6 percentage points. Forgone savings do not completely drop to 0% because 5.7% of treated subjects choose to override the robo-advisor despite actively seeking it. In the cross-section, the TOT effects are heterogeneous and disproportionally benefit subjects with low financial literacy and numeracy, suggesting that robo-advising can level the playing field in household debt management.

Next, we assess the Intention to Treat (ITT) effect—the difference between the loan repayment performance of subjects exposed to robo-advising and the control group, irrespective of whether the tool is used. The ITT effect accounts for the possibility that borrowers who are offered robo-advice might decide to not use it, perhaps because they do not trust algorithms. Despite remaining economically and statistically significant, the ITT estimate is -14.6 percentage points—about 25% less than the TOT estimates. Relatedly, subjects decline free robo-advice in about 25% of the problems for which they are offered it.

[5] As we describe in Section *1.1.*, the fee was drawn randomly from a uniform distribution after subjects provided their WTP to elicit WTP in an incentive-compatible fashion. This method differs from eliciting the demand for robo-advising after announcing a pre-determined fee and implies that many robo-advising seekers did not receive robo-advising in the paid treatment arms.

Consumers will benefit less from paid robo-advice because the best-case outcome switches from no savings forgone to the positive fee paid. Adopting robo-advising is thus only beneficial in consumption-utility terms if the fee that consumers pay is lower than the savings they would have forgone absent the tool. Interestingly, we find that subjects' WTP is, on average, *higher* than the monetary benefits they obtain from it, which may reflect subjects' pessimistic assessment of their performance when unassisted or a willingness to pay to more to avoid the cognitive and psychological costs of making choices on their own.

From a policy perspective, it would be optimal if demand for robo-advice were greatest among the less financially and numerically skilled, who make costlier mistakes in the pre-intervention phase. And, indeed, demand for robo-advice is inversely related to financial and numerical literacy. Everything else equal, it is also inversely related to confidence in one's skills and positively related to trust in robo-advice.[6] Financially literate subjects have a lower WTP, while men and more trustful subjects are willing to pay more. (Low) trust in algorithms is strongly correlated with overriding robo-advice, which is never optimal in our setting, and the desire to interact with a human advisor.

Finally, we ask whether robo-advice helps subjects to learn about optimal loan repayment or, instead, limits learning-by-doing by reducing their experience solving loan management problems. While we find better post-intervention performance across all of the treatment arms, the largest difference is for the control group, who had to work through more problems without assistance before reaching the post-intervention phase. We detect neither learning by imitation nor from the educational tips bundled with robo-advice.

1 Experimental Design and Procedure

In this section, we describe the design of the experimental tasks and the experimental procedure, including the recruitment of subjects, their characteristics, and the sequence of actions in the experiment.

[6] We elicited trust after the experiment to avoid priming subjects before assessing robo-demand. Being exposed to robo-advising during the experiment might affect trust. We cannot say if subjects' pre-determined trust increases their WTP or if higher WTP and hence more likely exposure during the RCT increases reported trust.

1.1. Experimental Design

To implement our experimental design, we created 27 loan repayment problems, which we report among the experimental materials in the Online Appendix. In each problem, subjects received an amount of money denoted in pounds that they were asked to allocate fully across multiple loans to minimize the sum monthly interest payments or, in problems that include minimum payment amounts, to minimize the sum of monthly interest payments and late fees.

We first designed nine problems that differed based on the number of loans, the mixture of APRs, and loan balances, which we label A to I.[7] We then designed three alternative versions of each problem: easy, medium, and high difficulty. The easy version focuses on loan amounts and APRs; the medium and hard difficulty versions introduce minimum payments, which trigger late fees when left unpaid. The most difficult version mimics real-world scenarios by introducing additional information not needed to calculate optimal repayment amounts. We report the three versions of a sample problem in Figure A.1 in the Online Appendix.

The experiment consisted of three ordered phases, executed sequentially within the same session. We summarize the phases and the sequence of subjects' actions in Figure 1. Over the course of the experiment, loan repayment problems A through I are presented to each subject in random order. Within each of the three phases, subjects are randomly presented with one easy problem, one medium problem, and one hard problem.

In the *pre-intervention phase*, each subject solved three problems without any external inputs or information beyond that required to describe the problem. Subjects did not receive any feedback on whether or how the interest and late fees associated with their repayment strategy differed from the optimal repayment strategy.

The experimental conditions were administered in the *intervention phase*. Subjects were randomly assigned to one of five experimental arms: (i) a control group for whom the task was identical to the pre-intervention phase (group 1); (ii) a group in which robo-advice was available for free (group 2); (iii) a group in which robo-advice was available for free and included an

[7] In the UK, the APR is an effective annual interest rate. According to the FCA: "if you borrowed £100 and the loan APR is 56%, after a year, you would pay back £156 in total." Therefore, to calculate the monthly interest payment for a given loan, we need to calculate $r_{monthly} = (1 + APR)^{(1/12)} - 1$.

explanation of the robo-advisor's proposed allocation ("robo-advice with education"; group 3); (iv) a group in which robo-advice could be obtained for a fee (group 4); and, (v) a group in which robo-advising with education could be obtained for a fee (group 5). In all four treatment arms, subjects are told the maximum interest and fees that they could save by using the robo-advisor (i.e., the maximum possible interest and fees minus the minimum possible interest and fees).

The intervention phase introduced three features absent in the pre-intervention phase: access to robo-advice, education about the strategies proposed by the robo-advisor, and elicitation of subjects' willingness to pay for the robo-advising tool. Robo-advice consisted of an automated proposal for allocating the amount subjects had at their disposal across the multiple loans in the problem. The robo-advisor allocation coincided with the optimal allocation, which minimized the interest and fees that subjects had to pay on their portfolio of loans over the next month. Note that this allocation can be computed without risk and uncertainty. For example, in the easy version of trial E, the subject has 1,000 pounds to allocate between three loans: a 1,040.55 balance with an APR of 22.5%, a 466.74 balance with an APR of 45.9%, and a 879.04 balance with an APR of 49.5%. In the absence of minimum payments, the robo-advisor would allocate 879.04 to the loan charging 49.5%, thus extinguishing the loan subject to the higher APR, and the remaining 120.96 to the loan charging 45.9%. This allocation results in the best-case monthly interest payment of 28.81 pounds. (Allocating everything to the loan with an APR of 22.5% would have resulted in the worst-case monthly interest payment of 45.57 pounds.) Subjects who chose to receive free robo-advice (within treatment groups 2 and 3) and subjects who chose to receive paid robo-advice and whose willingness to pay (WTP) exceeded the cost of the tool (within treatment groups 4 and 5) found the suggested best-case allocation automatically filled in on their problem screen. Subjects were told that the robo-advisor's allocations were optimal, but that they were free to accept or modify this allocation before moving to the next screen.

Education consisted of brief explanations of the strategies proposed by the robo-advisor. We report examples of the explanations among the experimental materials in the Online Appendix. For instance, in the easy version of problem E described above, subjects were exposed to the education steps reported in Figure A.2. Subjects who accepted free robo-advising treatment with education (group 3) and subjects who took robo-advising with education for a fee and whose willingness to pay exceeded the cost of the tool (group 5) were shown the educational text before the allocation proposed by the robo-advising tool appeared on their screen.

In the treatments with paid robo-advice, we elicited WTP in an incentive-compatible fashion. Rather than being asked whether they were willing to pay a pre-specified fee, subjects who sought paid robo-advice were asked to state their WTP for access to the robo-advice tool, using a slider that ranged from 0 pounds to the maximum possible savings in interest and fees for their loan repayment problem. Subjects were told: "You should respond truthfully—after you say how much you are willing to pay, the actual price of the assistant is picked randomly. If what you said you're willing to pay is higher than the random price, you'll buy the automated assistance at that price. If what you said you're willing to pay is lower than the random price, you won't buy the automated assistance and will pay nothing." Subjects were not told how the random price would be chosen. For each subject-problem pair, we drew the price randomly from the uniform distribution ranging from zero to the maximum possible savings in interest and fees. If the subjects' reported willingness to pay was higher than this price, subjects obtained the robo-advising tool. Otherwise, they did not.

Our price setting mechanism is clearly not comparable to the way prices would be set in field applications, which would be a pre-set fee common to any interested user. Our aim, though, is to elicit WTP in an incentived fashion rather than assessing how many consumers would accept robo-advice in the field based on different potential prices. Yet, knowing the truthful WTP of each subject informs us on who would take robo-advice for each potential price that was attached to this service in the field. Overall, only 36.6% of the subjects who sought paid robo-advice actually received it. This is because the average WTP (scaled by the maximum possible savings) was 34.7% while the average random price (scaled by the maximum potential savings) was 48.6%.

Crucially, subjects could see the problem *before* reporting their willingness to pay for the robo-advising tool, which is a realistic feature since, in the field, borrowers can seek advice based on the perceived difficulty of managing their loan portfolio. Moreover, because subjects should have chosen their willingness to pay based on the expected difference between the savings they could have made on their own and the maximum savings they could have made when using the robo-advising tool (which the problem told them explicitly), observing the loan repayment problem and assessing its complexity was important for the subjects to make an informed inference about their willingness to pay.

The *post-intervention phase* followed the same design as the pre-intervention phase. Subjects had to solve three problems of different difficulty levels without access to robo-advice. Before the end of the survey, subjects answered a few questions in a debriefing survey, which we discuss in more detail later.

1.2. Experimental Procedure and Subject Pool

We pre-registered the experimental design and procedure, including details about subject recruitment and target sample sizes, in the AEA RCT Registry (trial ID: AEARCTR-0006447).[8]

Subject recruitment aimed to obtain a UK nationally representative sample of adults aged 18 and above.[9] The target sample size was 4,500 subjects; based on the FCA's budget and timeline for completing the experiment phase, the delivered sample size was 3,423. Recruitment was run by a third-party provider that engaged in quota control on several demographic characteristics to maintain the sample's representativeness.

The experimental data were collected in Summer 2020 in the form of a survey administered through the online platform Qualtrics. In addition to recording subjects' choices in the survey tasks, the recruiter provided us with the demographic characteristics that we control for in the analysis. The randomized assignment of subjects across experimental arms should ensure that observable and unobservable characteristics are uncorrelated with the assignment rule.

After accessing and signing the consent form, subjects read the instructions describing the task of solving a loan repayment problem. The instructions specified that the main scope of the research was to understand how people manage their loans. The instructions briefly described the notion of debt and interest, indicated that subjects would need to work on nine loan repayment problems, and provided subjects with a screenshot of an example problem. Subjects were told

[8]The proposed procedures were approved by the Financial Conduct Authority (FCA) Internal Ethics Review Board on May 6, 2020.

[9]Individuals below 18 years of age were excluded due to the low likelihood of direct access to credit.

that an optimal repayment strategy existed. Subjects assigned to the robo-advising arms and the robo-advising with education arms also read an additional line of instructions reporting that they might be offered help in the form of an automated assistant in some scenarios. Finally, the instructions indicated how the incentive-compatible component of subjects' payment would be computed based on subjects' choices. In particular, subjects were told "For every hypothetical £8 you save in the task, you'll win 10p." The bonus component was capped at £2.

After reading the instructions, subjects proceeded to execute sequentially the three phases described above across nine subsequent screens. We report examples of loan repayment problem screens in Figure A.1 of the Online Appendix. For each debt account, subjects knew the nature of the loan, its current balance and interest rate, as well as the minimum payment and late fee associated with missing the minimum payment, when any existed. Subjects were instructed to allocate a certain amount across the loans described in a table on the screen to minimize interest and fee payments for each problem.

In the post-experiment survey, we elicited individual-level characteristics that earlier research suggests help predict financial decision-making: financial literacy, numeracy, risk tolerance, patience, and generalized trust. Subjects also were asked to describe positive and negative aspects of the robo-advice tool, what types of loans they had used over the past 12 months, whether they had fallen behind on any of these accounts, and whether they had a preference for working with human advisors or "automated assistants," and whether they were willing to pay more, less, or the same to work with a human advisor.

A common concern with experimental studies like ours, which, although incentivized, do not involve high stakes, is that subjects might not take the tasks seriously. This possibility is limited in our setting because subjects are part of a registered pool with Qualtrics, and hence shirking in our study would affect employability (i.e., expected revenues) in all future opportunities. To assess the concern more directly, we measure the time subjects spent on the survey. If shirking, subjects could have finished in less than 3 minutes by always allocating the available amount to any single account and answering the debriefing questions at random. Table A.1 of the Online Appendix reports the total amount of time that subjects spent on the experiment as a function of their performance. We do not detect evidence of systematic shirking.

2 Repayment Without and With Free Robo-Advice

To quantify performance of each subject in each loan repayment problem, we compute the *percentage of savings forgone* as follows:

$$\%\ Savings\ Forgone_{i,p} = \frac{Interest\ Paid_{i,p} - Minimum\ Interest\ Payable_p}{Maximum\ Interest\ Payable_p - Minimum\ Interest\ Payable_p}, \quad (1)$$

whereby $Maximum\ Interest\ Payable_p$ is the (certain and riskless) amount of interest and late fees that subjects would have paid over the next month had they chosen the strategy that resulted in the highest possible interest and late fee payment in problem p; $Minimum\ Interest\ Payable_p$ is the amount of interest and late fees that subjects would have paid had they chosen the optimal loan repayment strategy in problem p; and $Interest\ Paid_{i,p}$ is the actual amount of interest and late fees that subject i paid when solving the loan repayment allocation problem p. The range in possible interest and late fee payments varies across problems A through I; in six of the problems, the range also varies between the low difficulty version and the medium and high difficulty versions.[10]

We focus on $\%\ Savings\ Forgone$ as the main outcome of interest because it does not depend on the specific design of each problem, allowing comparisons across problems. As the subject's repayment strategy improves from worst case to best case, $\%\ Savings\ Forgone_{i,p}$ improves from 100% to 0%. During the pre-intervention phase, we also focus on the fraction of subjects that do not correctly specify the optimal allocation and the fraction of subjects whose performance is worse than the expected performance when allocating money across loans at random. We summarize performance on the loan repayment problems during the pre-intervention phase in Table A.2.

[10] While the medium and high difficulty versions of problems A, F, and I include late fees, there are no feasible allocations of money across the accounts in these three problems that result in subjects incurring any late fees.

2.1. Baseline Variation in Borrower Repayment Performance

We begin by determining the demographic characteristics that correlate with the quality of loan repayment skills. focusing on the control group, who lack access to robo-advice:

$$\%\ Savings\ Forgone_{i,p} = \alpha_p + X_i\beta + \epsilon_{i,p}, \tag{2}$$

where $\%$ *Savings Forgone*$_{i,p}$ reflects the performance on loan repayment problem p by control subject i, X_i is a vector of individual characteristics, and α_p is a trial-specific fixed effect.

In the first column of Table 1, X_i includes financial literacy fixed effects (terciles based on the number of correct answers to five financial literacy questions), educational attainment fixed effects (six categories), and age category fixed effects (seven categories, ranging from "18-24" to "75+").[11] We estimate the equation via OLS; standard errors are clustered on subject. As expected, financial literacy predicts loan repayment skills. For subjects in the middle tercile of financial literacy, the estimated effect is -9.5 percentage points; for those in the top tercile, it is -14.2 percentage points. Relative to the average member of the control group (mean of the dependent variable is 19.7%), those in the middle tercile of financial literacy forgo 48% less savings due to loan repayment mistakes, while those in the top tercile forgo 71% less. Note that the estimated effects of financial literacy are incremental to any effect of general education, for which we also control. Unreported coefficients on the education fixed effects reveal that the most educated outperform the least educated by 5.6 percentage points. Unreported coefficients on the age range fixed effects reveal a non-monotonic relation, with the best performance by subjects aged 45-54 and 55-64.

We expand the set of investor characteristics in the second column to include self-assessed numerical literacy fixed effects (terciles), self-assessed financial confidence fixed effects (terciles), self-assessed patience fixed effects (terciles), self-assessed risk aversion fixed effects (terciles), general knowledge test score fixed effects (terciles), a dummy variable for whether the investor trusts algorithms at least as much as human advisors, a dummy variable for whether the subject

[11] All variables and categories are defined in Table A.3. Including additional controls such as the number of children and adults in the respondent's family, the number of years the respondent has owned a home for, the respondent's occupation, and the region the respondent lives in does not alter the results of any of the findings of the paper either statistically or economically.

identifies as male, and a dummy variable for whether the individual has incurred any debt over the past 12 months.[12]

While the coefficients on the financial literacy fixed effects are slightly attenuated when we include additional proxies for problem solving skills, they remain economically and statistically significant. In addition, the estimate for high numerical literacy suggests that having greater familiarity with performing computations—above and beyond financial literacy and general education—has spillover effects on consumers' performance in loan repayment problems. Greater patience and higher general knowledge are also associated with better performance. However, we find little evidence that recent debt experience or financial confidence are associated with better performance. Trust in algorithms also has little predictive power. This result is important in our setting because it shows that a dimension that might—and does, as we show below—help explain the take-up of robo-advice does not at the same time capture pre-existing differences in subjects' ability to solve loan repayment problems. Finally, we find that gender is unrelated to savings forgone, which suggests that when financial literacy and other demographic characteristics are held constant, men and women behave similarly. This is not surprising in the context of loan repayment where subjects make non-risky choices because systematic differences in the financial decisions of men and women typically relate to choice under risk (e.g., Barber and Odean (2001) and D'Acunto (2019)).

2.2. Robo-Advice and Borrower Repayment Performance: Treatment on the Treated

In this section, we estimate the treatment on the treated (TOT), that is, the average difference between the percent savings forgone of subjects who saw the robo-advised suggestions and subjects who did not. When robo-advice is provided for free, because the tool always recommends the loan repayment allocation that minimizes interest and fees, subjects can only forgo savings if they override the recommended allocation. In contrast, the loan repayment allocations chosen by subjects who have no access to the tool should capture any mistakes borrowers make and biases they exhibit when facing loan repayment problems. Therefore, we

[12] To measure general knowledge, we ask subjects a set of 5 questions—three questions on geography, one about population sizes, and one about when the iPhone first launched.

can interpret the TOT results for free robo-advice as quantifying the effects of individual biases in loan repayment allocations and not just the (mechanical) benefits of robo-advice on individual decisions.

We estimate a multivariate specification of the following form, restricting the sample to trials 4 through 6 for subjects in the control group and both free robo-advice treatment arms:

$$\%\ Savings\ Forgone_{i,p} = \alpha_p + \beta_1 Robo\ Advice_{i,p} + \delta Education\ Treatment_p + X_i\gamma + \epsilon_{i,p}, \quad (3)$$

where $\%\ Savings\ Forgone_{i,p}$ reflects the performance on loan repayment problem p by subject i; $Robo\ Advice_{i,p}$ is a dummy variable that equals one if subject i was offered and accepted robo-advising with or without education when facing problem p, and zero otherwise; $Education\ Treatment_p$ is a dummy variable that equals one if subject i was assigned to a treatment arm that educates the subject on the principles of financial management that the robo-advisor relied upon when making its recommendations; and the problem fixed effects and demographic controls $X_{i,p}$ are defined as in equation (2). As a result of the problem fixed effects, TOT and other estimated coefficients are identified using variation within a given repayment problem and difficulty level (low versus medium or high). We report the results in columns (3)-(4) of Table 1.

The estimated TOT coefficient for free robo-advising is -19.6 percentage points (significant at 1-percent level), which is almost exactly equal to the overall average savings forgone by the control group (19.7% across all nine trials), indicating that the vast majority of subjects who are provided with robo-advising for free follow its recommendations. The concern that participants will override robo-advice recommendations, especially the most vulnerable households, appears to be minimal when robo-advising is provided for free. Regarding the control variables, we continue to find that higher financial and numerical literacies are associated with better performance, but the estimates are attenuated. There are no differences in performance associated with random assignment to robo-advice or robo-advice with education.

Note that while the randomized assignment of subjects across treatment arms should eliminate any concerns that treatment is correlated with observed or unobserved individual-level characteristics, subjects who seek robo-advice within a treatment arm may differ from subjects who do not seek robo-advice along dimensions that are correlated with their performance on loan repayment problems. To address this concern, in Table A.5 of the Online Appendix, we report two-stage least-squares specifications in which $Robo\ Advice_{i,p}$ is instrumented with a dummy variable indicating whether subject i was assigned to the free robo-advice treatment arm. These instrumental variables results are quantitatively similar to the OLS results.

2.3. Robo-Advice and Borrower Repayment Performance: Intention to Treat

Next, we assess the Intention to Treat (ITT) effects, which compare the average savings of subjects offered robo-advising or robo-advising with education, irrespective of whether they sought or received the advice, with the average savings of control subjects. The ITT effects account for the TOT effects and the fraction of subjects who, despite being offered access to free robo-advice, prefer not to see the recommended allocations. This estimation is relevant not only to understand consumers' willingness to accept robo-advice but also for policy purposes. Indeed, if a governmental or private party wanted to assess the effects of proposing robo-advice to the broader population, it should want to capture the extent to which consumers would seek and follow such advice.

Pooling trials 4 through 6, subjects seek access to free robo-advice in 75.3% of the problems.[13] However, demand is consistently higher for free robo-advice without education (80.5%) than for free robo-advice with education (69.9%). We estimate a multivariate OLS regression similar to equation (3) but in which the $Robo\ Advice_{i,p}$ dummy variable equals one for all subjects who were assigned to an experimental arm with robo-advice. The ITT estimate in column (4) is -14.6 percentage points (statistically significant at the 1-percent level), which is approximately three-quarters of the TOT estimate in column (3). The positive estimated effect of the Education Treatment is consistent with lower demand for robo-advice with education among some subjects who would have benefitted from robo-advice.

[13] We summarize demand for and utilization of robo-advice in Figure A.3 and Table A.4.

2.4. Robo-Advice and Borrower Repayment Performance: Heterogeneity

Beyond the average effects of exposure to robo-advice for loan management, we are interested in determining whether its effects vary by demographics, particularly whether more vulnerable decision-makers are made better off or worse off relative to more sophisticated ones by the availability of robo-advice.

This trade-off arises because the treatment effect of robo-advice on borrower repayment performance is likely to differ in the cross-section of subjects based on underlying characteristics for at least two reasons. On the one hand, individual characteristics might predict subjects' performance in loan repayment tasks, and less sophisticated and more vulnerable decision-makers are likely to perform worse when unassisted. Because the implementation of robo-advice is the same for everybody (and, by construction, proposes the unique optimal choice to everybody), the improvement induced by robo-advice should be higher for less sophisticated subjects. On the other hand, less sophisticated subjects might be warier of technologies they do not understand and with which they are unfamiliar, especially if recommendations clash with common rules of thumb, such as paying the minimum payment of a credit card each month.

An important caveat of this analysis is that our pre-registration did not specify a set of characteristics that we would use to test for heterogeneous treatment effects. We use the characteristics we elicited in the debriefing survey. In Figure 2, each panel refers to one of such characteristics. Each bar represents the estimated TOT effect $\hat{\beta}_1$ for a version of equation 3 that interacts the dummy variable for being exposed to free robo-advising with dummy variables for each tercile of a given characteristic. For each characteristic, tercile 1 refers to lower values for each characteristic and tercile 3 to higher values. The control variables mirror column (3) of Table 1.

We estimate monotonic (and statistically significant) heterogeneous treatment effects for three of the eight characteristics: financial literacy, self-assessed numeracy, and risk aversion. In each case, lower values are associated with larger TOT estimates, with the largest differences for financial literacy. While the figure does not report the 95% confidence interval for each estimate to guarantee readability, the difference between estimates for the first and third tercile are statistically significant for all three characteristics. Interestingly, free robo-advising for loan repayment not only levels the playing field across subjects who have different levels of financial literacy but it makes low-literacy subjects better off in levels because a higher share of low-literacy subjects implements the robo-advice fully relative to high-literacy subjects, who override advice more often. This result speaks to the trade-off between the potentially positive and negative relative and absolute effects of access to robo-advising for less sophisticated consumers. The results for patience are also worthy of mention. Even though we do not detect a monotonic heterogeneous pattern, we do detect an economically and statistically significant differential effect between the effects for subjects in terciles 1 and 2, which suggests that the most impatient subjects gain relatively more from robo-advice than others. Note, however, that the confidence intervals around the estimates for tercile 1 and 3 overlap.

For general knowledge, the estimated effects are monotonic but with overlapping confidence intervals. For the remaining characteristics—confidence in one's own ability to manage their finances, recent experience with debt, and trust in algorithms—the estimated effects are non-monotonic, with overlapping confidence intervals.

3 Willingness to Pay for Robo-Advice

Next, we study subjects' WTP for robo-advice. Our incentive-compatible procedure to elicit subjects' WTP (described in section *1.1.*) allows us to measure this dimension, which is usually unobserved in field settings, where any consumer whose WTP is higher than a pre-set fee would make the same choice of taking up robo-advice. WTP is elicited as the maximum amount of potential savings in each trial that subjects are willing to pay to observe robo-advised allocations. The subject states an amount in pounds, which we then convert into a percentage of savings forgone by dividing the problem-specific denominator of equation 1.

In Panel A of Figure 3, we plot the distribution of the WTP of subjects who were willing to bid to obtain robo-advice in trials 4-6 against the distribution of the amounts of savings that the same subjects had forgone in their unassisted choices before robo-advice was available (trials 1-3). If subjects, on average, were willing to bid up to the amounts of savings forgone when unassisted, we should not observe that the WTP distribution lies to the right of the distribution of amounts forgone.

In contrast, Panel A of Figure 3 reveals two different patterns. First, subjects' WTP is quite heterogeneous—so much so that the distribution of WTP is almost uniform. This pattern motivates us to assess the cross-sectional characteristics that correlate with WTP in the next section. Second, subjects' WTP appear to be substantially higher than the savings they forwent when making choices unassisted; that is, subjects appear to be willing to pay significantly more than the monetary benefits they obtain from robo-advice, which by construction cannot be higher than the difference between their unassisted allocations and the best possible allocations. We further assess subjects' valuation of robo-advice in Table A.6 of the Online Appendix. Although we find that lower levels of financial literacy are associated with both higher savings forgone during the pre-intervention trials and higher WTP, we also find that the spread in WTP is substantially higher than the spread in average unassisted performance.

Subjects' high WTP for robo-advice for loan repayment could have several non-mutually exclusive explanations. On the one hand, subjects might overestimate the monetary benefits of robo-advising if they underestimate their performance in loan repayment problems when unassisted. In our experiment, subjects purposefully did not receive feedback on their performance after each trial, which would have been an unrealistic feature given that in the field consumers do not observe the differences between their loan repayment choices and the optimal choices. Subjects might thus believe that they forgo more money than they actually do whenever they make repayment choices on their own. On the other hand, subjects might value robo-advising more than its monetary value if they also value other aspects of using robo-advice, such it allowing them to avoid paying the cognitive costs and potential stress and anxiety of solving quantitative problems on their own that have a direct effect on their wealth without ever learning whether their choices are "right" or "wrong." The latter explanation resonates with the fact that, when asked in the debriefing survey at the end of the experiment, about 71% of subjects liked robo-advice because it was quick and easy to use and 30% because it allowed them to avoid thinking about how to solve the problems. Moreover, about half of those who used the robo-advising tool claimed they did so either because they could not beat a machine or because the scenarios were best suited for an algorithm to solve. About 9% of users, who were concentrated among the least financially literate, stated that they like to delegate their financial choices. And, 29% of users, who were concentrated among the most financially literate, reported other reasons in free text format, the majority of which indicates curiosity in comparing their solutions with those the algorithm would propose.

The motivations of non-adopters are also worth discussing. Among those who decided not to use the tool, about 22% indicated that they found it too expensive. Instead, 54% claimed they wanted to make their own decisions, and 21% claimed that they could make better decisions than the algorithm. The latter result raises a red flag regarding consumers' abilities to understand their limitations when it comes to loan repayment because, by construction, subjects could only have matched the quality of the tool's choices. Even more concerning is the fact that the conviction that one is able to beat the algorithm does not vary substantially by financial literacy—26% of those who chose not to take robo-advice among the least financially literate argued that they could make better choices unassisted versus 31% of the non-adopters among the most financially literate.

Another way to assess the extent to which users value robo-advice is to estimate Equation (3) for subjects in the paid robo-treatment arms (with and without education) and those in the control group after replacing the outcome variable with the % $Net\ Savings\ Forgone_{i,p}$, computed as the sum between % $Savings\ Forgone_{i,p}$ and the willingness to pay of subject i in problem p, which is also measured as a fraction of the potential savings forgone. And, indeed, we find that the TOT estimates for subjects who bid for and received robo-advice (column (5) of Table 1) flip to positive and economically large (30.8 percentage points). This result suggests that, on average, subjects are willing to pay more for robo-advice than the monetary benefits they receive in terms of reduction of savings forgone relative to unassisted subjects, which is captured by the TOT estimate in column (3) of Table 1. Note that these results do not imply that implementing robo-advice for a fee would make consumers lose money, on average, because a real-world application would fix the same price for all robo-advice, which is likely substantially lower than the WTP of most subjects.

In the ITT estimation of column (6) of Table 1, all subjects in the paid robo-advice treatment arms are compared to control subjects, irrespective of whether their WTP was higher than the randomly-generated robo fee. Even in this case, on average, bidding subjects save less than the unassisted subjects in the control group once their WTP is taken into account. Only 22.8% of the subjects end up receiving paid robo-advice recommendations, but the average WTP within this group is 60.1%. This result confirms that, on average, users who were asked to provide a value for the robo-advising tool in our setting valued it more than the monetary benefits that the robo-advisor actually provided.

4 What Explains Take-up and Compliance with Robo-Advice?

Our analysis of consumers' WTP to receive robo-advice for loan repayment begets further inquiry into the characteristics that correlate with a higher valuation for and greater compliance with robo-advisor's recommendations.

When consumers assess their valuation of robo-advice, they should estimate the interest and fees they are unlikely to be able to avoid when unassisted, as well as other potential benefits of delegating their choice to the algorithm, including any potential cognitive and psychological costs (effort in computing, anxiety, doubtfulness, cognitive dissonance) of having to solve loan repayment problems without any feedback

To shed additional light on consumers' demand for robo-advice, we estimate the following cross-sectional specifications at the subject level (one observation per subject):

$$Robo\_Choice_i = \alpha + X_i\beta + \epsilon_i, \tag{4}$$

where $Robo\_Choice_i$ is the decision to demand free robo-advice in column (1), the decision to demand paid robo-advice in column (2), the willingness to pay for robo-advice for those who demand it in column (3), the decision to override the robo-advice recommendation in column (4), and a dummy variable that indicates whether the subject is willing to pay more for a human advisor in column (5). In each column, we focus on subjects' decisions in trial 4—the trial where subjects report their willingness to pay before seeing the savings associated with adopting robo-advice.

Demand for free robo-advice is uncorrelated with measures of financial literacy. This is not surprising because free robo-advice allows more and less sophisticated consumers to costly minimize interest and fees without exerting any effort. In contrast, when robo-advice is offered for a fee, subjects with higher financial literacy are less likely to demand it than others. Low-financial-literacy subjects thus understand that their loan repayment skills are relatively low and value robo-advising more than others. Consistently, in column (3), we find that WTP for robo-advice is substantially lower—economically and statistically—for subjects with higher financial literacy. We also find evidence (significant at the 10-percent level) that subjects with higher self-assessed levels of risk aversion have higher WTP.

Demand for robo-advice, whether free or paid, is higher for subjects who report greater trust in algorithms at the end of the experiment. Moreover, while men are not more likely to demand paid robo-advising (column (2)), they have a systematically higher WTP for robo-advice conditional on bidding for it (column (3)). The estimated coefficients on the Educational Treatment dummy variable are negative but statistically insignificant in columns (1) and (2). However, in Figure A.3, we see that demand for robo-advice with education declines more across the three trials than robo-advice without education, and that the decline is particularly large for free robo-advice. If we expand the sample to include trials 5 and 6, we find that demand for free robo-advice with education is 10.4 percentage points lower than demand for free robo-advice without education. The difference remains 3.0 percentage points when we focus on paid robo-advice, but gains statistical significance within the larger sample.

Approximately 5 percent of the subjects who receive free or paid robo-advice choose to override the recommended allocations. In column (4), we find that subjects are less likely to override robo-advice when the advice is free and trust in algorithms is higher. Neither financial literacy, financial confidence, nor numeracy appear to play a role. In the final column, we ask which of the subjects who received robo-advice in trial 4 state that they would prefer to pay more for a human advisor. Stated willingness to pay more for a human advisor is lower for subjects with high numeracy skills and trust in algorithms. It is also *increasing* in financial literacy. Although we lack direct evidence, a possible explanation is that financially literate subjects derive more value from interacting with an expert whose suggestions they can try to understand and have explained and to whom they can propose own ideas and are, therefore, willing to pay more for this service.

5 Learning About Loan Repayment? Learning-by-doing, Imitation, and Education

An open question when assessing the overall effects of providing robo-advice for loan repayment is whether consumers using robo-advice learn better loan management strategies over time. First, even absent any advice, consumers might learn by doing as they think more and more about the trade-offs of allocating money across different accounts. Second, the fact that robo-advising exposes subjects to optimal allocations might allow them to learn by imitation. Consumers might infer the principles that the robo-advisor uses to allocate money across accounts and apply them in subsequent unassisted choices. Third, combining robo-advice with an explicit description of the principles behind optimal allocations might increase the likelihood that consumers learn how to replicate the robo-advisor's strategies when unassisted.

Based on these three possibilities, we designed our RCT to allow for a direct estimate of the incremental effect of robo-advice with and without education. In the last three trials of the experiment, subjects once again solved three loan repayment problems without assistance. We report the average performance of each group of subjects during the pre-intervention and post-intervention phases in Panel B of Figure 3.

First, we find that subjects learn by doing. All groups, including controls, exhibit better performance in the post-intervention trials than in the pre-intervention trials. The average percent forgone decreases by 2.6 percentage points for the control group (p-value 0.002), 1.0 percentage points for the free and paid robo-advice treatments without education (p-value 0.101), and 2.1 percentage points for the free and paid robo-advice with education (p-value 0.002).[14]

Second, we do not find evidence of incremental learning from exposure to robo-advice. Subjects in the control group—who had to think actively through the problems they faced not

[14] These estimates include neither subject fixed effects nor trial-by-difficulty fixed effects. When we include both sets of fixed effects, the estimated improvements across trials are 2.4 percentage points (p-value 0.001), 1.1 (p-value 0.035), and 1.8 percentage points (p-value 0.002), respectively.

only in the pre-intervention phase but also in the intervention phase—performed at least as well during the post-intervention phase. While these differences are economically small, they call into question the extent to which subjects learn by imitation from the recommendations of robo-advisors or robo-advisors plus education. If anything, providing subjects access to robo-advising might cause them to pay less attention to the loan repayment problems, thereby learning less from them. In this case, robo-advising interventions could successfully improve debt management only if they were offered continuously.

Whether ordinary consumers might learn more from repeated exposures to robo-advising over a long period when managing real-world debt is an open question for future research. Note, however, that due to the low cost of designing robo-advice for borrowers and delivering it to consumers' personal devices, the cognitive costs of learning might dominate the costs of delivering ongoing advice in this context. Subjects' incentives to learn to save on the cost of future advice are minimal.

6 Conclusions

We design and implement an RCT to assess the extent to which robo-advice for loan repayment is adopted, valued, and used by consumers, many of whom may have difficulty minimizing the interest and fee payments in real-world situations. We find that access to free robo-advice significantly improves repayment decisions. The high levels of demand and low levels of overriding among the least sophisticated, who benefit the most from the tool, suggest that robo-advice for loan repayment is likely to decrease wealth inequalities rather than contribute to them.

We also find that the average willingness to pay for robo-advice is higher than its monetary benefits (the interest and fees it lets borrowers save), which suggests that consumers might be excessively pessimistic about their performance in loan repayment tasks when unassisted or might value non-monetary benefits of the application, such as avoiding the cognitive and psychological costs of making choices that are consequential to their wealth. In the cross-section, higher financial literacy decreases the willingness to pay for robo-advice, while higher levels of risk aversion may increase it. Non-adopters and overriders are both more likely to state that they do not trust algorithms at the end of the experiment.

Finally, we find that subjects learn by doing as they improve their loan repayment skills after seeing more problems and thinking through the trade-offs of allocating their budgets across different accounts. If anything, exposure to robo-advised allocations limits this improvement, even when the principles of debt management used by the robo-advisor are spelled out explicitly. To the extent that consumers are unable to infer good loan repayment practices from robo-advisor's recommendations, they will need repeated access to the tool.

Consumers' high valuation of the benefits of robo-advice for debt management and the simplicity of its execution are at odds with the lack of real-world implementations. Market-based solutions are inherently limited because lenders' profits are a non-monotonic function of borrowers' debt. While too much debt may result in default, which is costly for borrowers and lenders, too little debt reduces lenders' revenues; as such, lenders have an incentive to let households accumulate debt and pay interest and fees so long as doing so does not trigger default (Jorring (2018); Ru and Schoar (2016)). Consequently, the supply of robo-advice tools focused on loan repayment might require government intervention or entry by private providers unaffiliated with lenders (D'Acunto, Rossi, and Weber (2019)).

References

Adams, P., S. Hunt, C. Palmer, and R. Zaliauskas. 2021. Testing the effectiveness of consumer financial disclosure: Experimental evidence from savings accounts. *Journal of Financial Economics* .

Agarwal, S., S. Chomsisengphet, and C. Liu. 2010. The importance of adverse selection in the credit card market: Evidence from randomized trials of credit card solicitations. *Journal of Money, Credit and Banking* 42:743–54.

Agarwal, S., S. Chomsisengphet, N. Mahoney, and J. Stroebel. 2015. Regulating consumer financial products: Evidence from credit cards. *The Quarterly Journal of Economics* 130:111–64.

Agarwal, S., J. C. Driscoll, X. Gabaix, and D. Laibson. 2008. Learning in the credit card market. Tech. rep., National Bureau of Economic Research.

Agarwal, S., and B. Mazumder. 2013. Cognitive abilities and household financial decision making. *American Economic Journal: Applied Economics* 5:193–207.

Agarwal, S., and J. Zhang. 2015. A review of credit card literature: perspectives from consumers. *Report, UK Financial Conduct Authority. URL: https://bit. ly/2TWt0v1* .

Argyle, B. S., T. D. Nadauld, and C. J. Palmer. 2020. Monthly payment targeting and the demand for maturity. *The Review of Financial Studies* 33:5416–62.

Barber, B. M., and T. Odean. 2001. Boys will be boys: Gender, overconfidence, and common stock investment. *The quarterly journal of economics* 116:261–92.

Bertrand, M., and A. Morse. 2011. Information disclosure, cognitive biases, and payday borrowing. *The Journal of Finance* 66:1865–93.

Bu, D., T. Hanspal, Y. Liao, and Y. Liu. 2020. Financial literacy and self-control in fintech: Evidence from a field experiment on online consumer borrowing. Tech. rep., SAFE Working Paper.

Campbell, J. Y., T. Ramadorai, and B. Ranish. 2019. Do the rich get richer in the stock market? evidence from india. *American Economic Review: Insights* 1:225–40.

Chalmers, J., and J. Reuter. 2020. Is conflicted investment advice better than no advice? *Journal of Financial Economics* 138:366–87.

D'Acunto, F. 2019. Identity and choice under risk. *Available at SSRN 3466626* .

D'Acunto, F., P. Ghosh, R. Jain, and A. G. Rossi. 2020. How costly are cultural biases? evidence from fintech. *Working Paper* .

D'Acunto, F., N. Prabhala, and A. G. Rossi. 2019. The promises and pitfalls of robo-advising. *The Review of Financial Studies* 32:1983–2020.

D'Acunto, F., and A. G. Rossi. 2022. Fintech and robo-advising: The transformation role of ai in personal finance. *Journal of Economics and Management Strategy (special survey issue)* .

———. 2021a. New frontiers of robo-advising: Consumption, saving, debt management, and taxes. *Machine Learning in Financial Markets: A Guide to Contemporary Practice, Cambridge University Press,* .

———. 2021b. Robo-advising. *Palgrave Macmillan Handbook of Technological Finance* .

D'Acunto, F., A. G. Rossi, and M. Weber. 2019. Crowdsourcing peer information to change spending behavior. Working Paper.

Drexler, A., G. Fischer, and A. Schoar. 2014. Keeping it simple: Financial literacy and rules of thumb. *American Economic Journal: Applied Economics* 6:1–31.

Gathergood, J., N. Mahoney, N. Stewart, and J. Weber. 2019. How do individuals repay their debt? the balance-matching heuristic. *American Economic Review* 109:844–75.

Jorring, A. 2018. The costs of financial mistakes: Evidence from us consumers. *University of Chicago Manuscript* .

Keister, L. A. 2000. *Wealth in america: Trends in wealth inequality.* Cambridge University Press.

Lusardi, A. 2012. Numeracy, financial literacy, and financial decision-making. Tech. rep., National Bureau of Economic Research.

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., D. J. Schneider, and P. Tufano. 2011. Financially fragile households: Evidence and implications. Tech. rep., National Bureau of Economic Research.

Lusardi, A., and P. Tufano. 2015. Debt literacy, financial experiences, and overindebtedness. *Journal of Pension Economics & Finance* 14:332–68.

Navarro-Martinez, D., L. C. Salisbury, K. N. Lemon, N. Stewart, W. J. Matthews, and A. J. Harris. 2011. Minimum required payment and supplemental information disclosure effects on consumer debt repayment decisions. *Journal of Marketing Research* 48:S60–77.

Niszczota, P., and D. Kaszás. 2020. Robo-investment aversion. *Plos one* 15:e0239277–.

Ponce, A., E. Seira, and G. Zamarripa. 2017. Borrowing on the wrong credit card? evidence from mexico. *American Economic Review* 107:1335–61.

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

Ru, H., and A. Schoar. 2016. Do credit card companies screen for behavioral biases? Tech. rep., National Bureau of Economic Research.

Stango, V., and J. Zinman. 2021. We are all behavioral, more or less: A taxonomy of consumer decision making .

Zinman, J. 2015. Household debt: Facts, puzzles, theories, and policies. *Annual Reviews of Financial Economics* 7:251–76.

Zucman, G. 2019. Global wealth inequality. *Annual Review of Economics* 11:109–38.

Figure 1

Experimental Design

Phase	Trials	Group 1	Group 2	Group 3	Group 4	Group 5
Pre-intervention	1-3	← No robo-advice →				
Intervention	4-6	No robo-advice	Free robo-advice	Free robo-advice with education	Paid robo-advice	Paid robo-advice with education
Post-intervention	7-9	← No robo-advice →				

- Subjects are randomly assigned into the five experimental arms.
- Each subject solves nine debt-management games, in random order. Within each phase, each subject solves one easy, one medium, and one hard version, also in random order.
- Trials 1-3 allow us to determine baseline performance on debt-management games and correlation with characteristics.
- Trials 4-6 allow us to estimate treatment-on-treated and intention-to-treat effects separately for free and paid robo-advice tools. They also allow us to measure willingness to pay (WTP).
- Trials 7-9 allow us to test for learning effects.

Figure 1 describes the experimental design of the RCT discussed in the paper.

Figure 2: Treatment on the Treated: Heterogeneous Effects

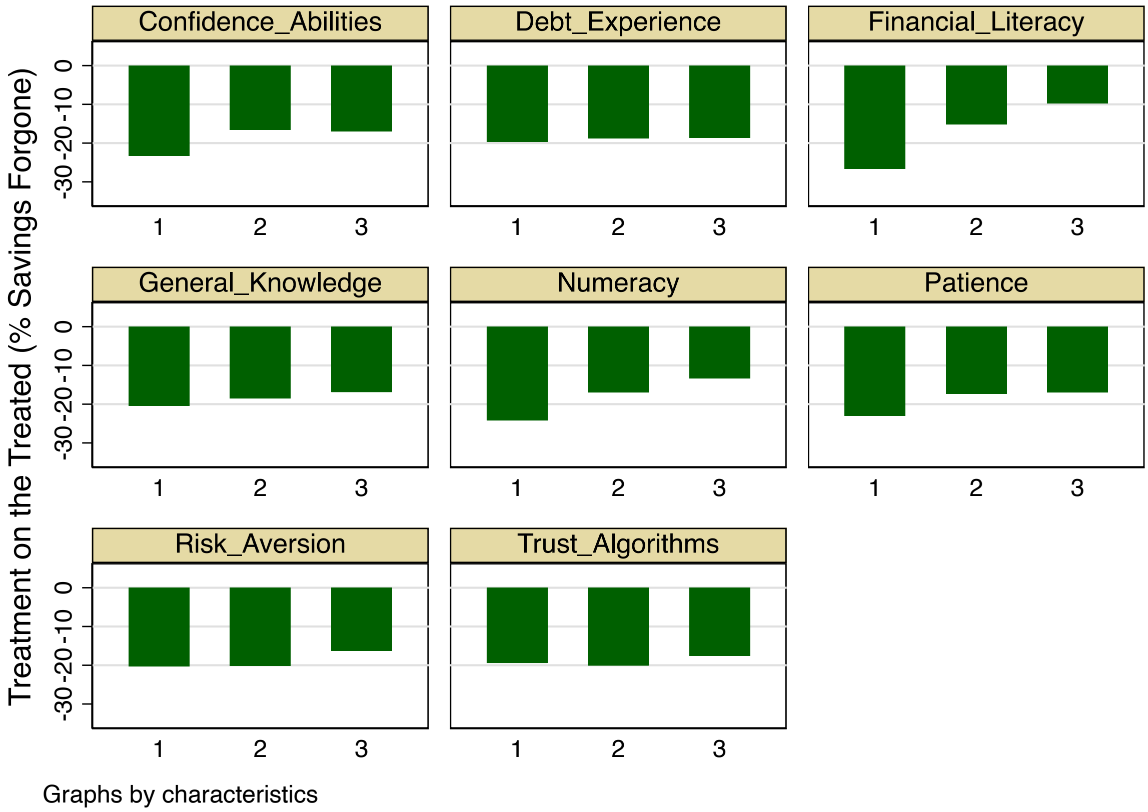
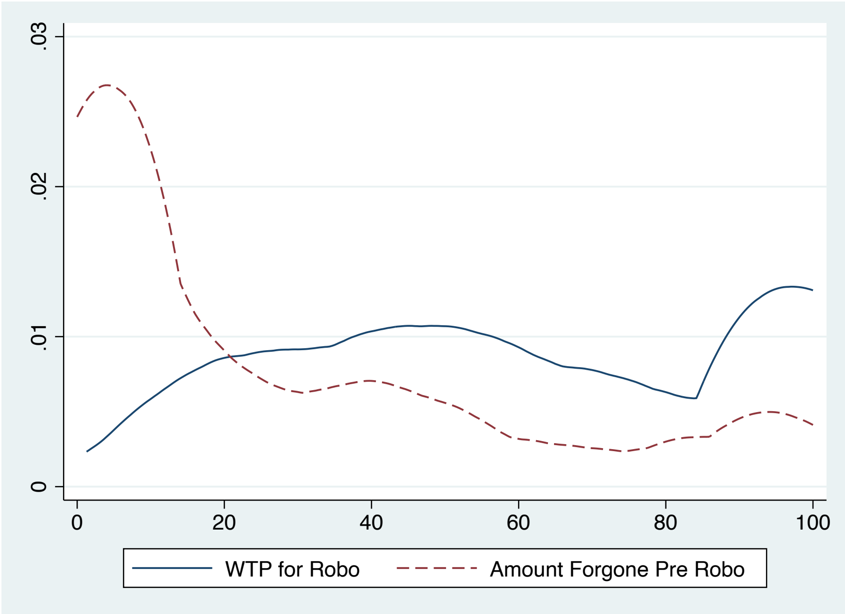


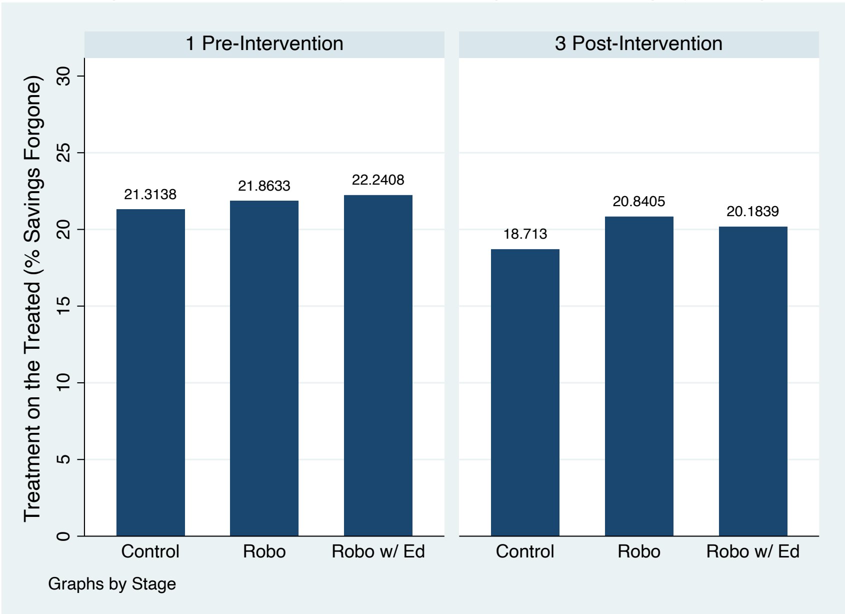
Figure 2 plots the estimated TOT effect of free robo-advice across terciles of subjects sorted based on a set of characteristics we elicit in the post-experiment debriefing survey. We describe the construction of the terciles in Table A.3. The treatment effect is measured in terms of the percentage point reduction of the share of potential savings subjects forgo when they do not make the optimal allocation decision in a loan repayment problem.

Figure 3: Willingness to Pay for Robo and Amounts Forgone Without Robo

Panel A. WTP for Robo and Amounts Forgone Without Robo



Panel B. Learning About Loan Repayment Strategies? Learning-by-doing vs. Education



Panel A of Figure 3 plots the distribution of the willingness to pay (WTP) elicited with an incentive-compatible mechanism from subjects who were proposed robo-advising for pay and received it against the distribution of the savings forewent by the same subjects in the pre-intervention phase, when they implemented allocations unassisted and robo-advising was not available. Panel B of Figure 3 plots the average percent savings forgone during the pre-intervention and post-intervention stages for three groups of subjects: the control group, those in the treatment arms offering free or paid robo-advice without education, and those in the treatment arms offering free or paid robo-advice with education.

Table 1. Effects of Robo-Advising: For Free and For Pay

Dependent variable: *% Savings Forgone*	Unassisted Choices (Control) (1)	(2)	Free Robo-Advising TOT (3)	ITT (4)	Paid Robo-Advising (WTP adjusted) TOT (5)	ITT (6)
Robo-Advice			-19.55***	-14.60***	30.84***	16.53***
			(-28.71)	(-17.54)	(20.36)	(13.99)
Education Treatment			0.82	2.51***	4.79***	-0.58
			(1.29)	(3.74)	(4.43)	(-0.42)
Medium Financial Literacy	-9.51***	-7.53***	-4.66***	-4.22***	-7.13***	-8.21***
	(-8.30)	(-6.19)	(-6.60)	(-5.64)	(-6.08)	(-6.70)
High Financial Literacy	-14.19***	-11.51***	-7.32***	-6.93***	-11.39***	-13.71***
	(-11.33)	(-8.35)	(-8.23)	(-7.61)	(-7.25)	(-8.49)
Medium Numeracy		-0.39	-2.11***	-1.77**	-2.68**	-2.69*
		(-0.28)	(-2.62)	(-2.10)	(-1.98)	(-1.89)
High Numeracy		-4.26***	-2.87***	-2.20**	-4.87***	-5.49***
		(-3.09)	(-3.43)	(-2.56)	(-3.62)	(-3.78)
Medium Financial Confidence		-1.76	-0.54	-0.17	-3.23***	-3.66***
		(-1.43)	(-0.71)	(-0.21)	(-2.71)	(-2.90)
High Financial Confidence		-0.32	-1.25	-0.89	-1.11	-1.21
		(-0.20)	(-1.25)	(-0.86)	(-0.74)	(-0.73)
Medium Patience		-4.28***	-1.44**	-1.48**	-3.06**	-2.92**
		(-3.72)	(-2.03)	(-1.99)	(-2.71)	(-2.47)
High Patience		-3.15**	-0.51	-0.26	-1.45	-1.99
		(-2.04)	(-0.53)	(-0.26)	(-1.00)	(-1.31)
Medium Risk Aversion		1.73	1.13	0.80	1.74	2.57**
		(1.52)	(1.59)	(1.56)	(2.20)	(-3.36)
High Risk Aversion		0.26	-0.42	-0.37	0.24	0.98
		(0.20)	(-0.51)	(-0.43)	(0.18)	(0.70)
Medium General Knowledge		0.05	0.09	-0.828	0.08	0.02
		(0.04)	(0.12)	(-1.11)	(0.07)	(0.02)
High General Knowledge		-2.77**	-0.93	-0.75	-1.26	-1.21
		(-2.27)	(-1.04)	(-0.08)	(-0.86)	(-0.77)
Trust in Algorithms		-0.24	-0.72	-2.77***	0.87	1.76*
		(-0.23)	(-1.10)	(-4.09)	(0.98)	(1.69)
Male		0.49	0.76	0.91	0.84	2.28**
		(0.47)	(1.19)	(1.36)	(0.82)	(2.11)
Has Debt Outside RCT		-2.35	-0.74	-0.74	-3.99***	-4.34***
		(-1.64)	(-0.92)	(-0.86)	(-3.06)	(-3.16)
Problem FE	✓	✓	✓	✓	✓	✓
Education Category FE	✓	✓	✓	✓	✓	✓
Age Category FE	✓	✓	✓	✓	✓	✓
Adj. R^2	0.223	0.237	0.255	0.176	0.231	0.182
N	6,228	5,940	5,841	5,841	5,679	5,679

Table 1 reports the estimated coefficients in the following set of specifications:

$$\%\ Savings\ Forgone_{i,p} = \alpha_p + \beta_1 Treatment_{i,p} + X_i\gamma + \epsilon_{i,p},$$

where $\%\ Savings\ Forgone_{i,p}$ is defined in Section 2.2., X_i includes the set of individual-level characteristics indicated in the table, α_p is a set of fixed effects for each problem by difficulty level, and $Robo\text{-}Advice_{i,p}$ defines the experimental condition subjects face. Columns (1)-(2) are estimated using trials 1-9 for the control group. The specifications in columns (3)-(6) are estimated using data from trials 4-6 for subjects in the free robo-advising experimental arms ((3)-(4)) and subjects in the paid robo-advising experimental arms ((5)-(6)) . In columns (3) and (5), $Robo\text{-}Advice_{i,p}$ is a dummy that equals one if subject i obtained robo-advising in trial p, and zero otherwise (TOT). In columns (4)-(6), it equals one if subject i was assigned to one of the robo-advice treatment arms, irrespective of receiving it, and zero otherwise (ITT). In columns (5)-(6), $\%\ Savings\ Forgone_{i,p}$ also includes subjects' elicited WTP for robo-advice (scaled by maximum potential savings). We report t-statistics below estimated coefficients. Standard errors are clustered at the subject level.

Table 2. Robo-Advising Demand, Willingness to Pay, and Compliance

	Wants Free Robo? (1)	Wants Paid Robo? (2)	Willingness to Pay (3)	Robo Override? (4)	Prefers Human? (5)
Robo-Advice (for Pay)				0.040**	0.003
				(2.40)	(0.09)
Education Treatment	-0.024	-0.030	0.883	0.012	-0.022
	(-1.17)	(-1.45)	(0.49)	(1.00)	(-0.86)
Medium Financial Literacy	-0.030	-0.067***	-7.953***	-0.009	0.068**
	(-1.25)	(-2.84)	(-3.70)	(-0.65)	(2.26)
High Financial Literacy	-0.012	-0.097**	-16.333***	-0.014	0.077*
	(-0.35)	(-2.62)	(-6.07)	(-0.86)	(1.81)
Medium Numerical Literacy	0.003	-0.020	0.984	-0.025	0.027
	(0.11)	(-0.73)	(0.38)	(-1.64)	(0.76)
High Numerical Literacy	-0.054*	-0.036	-2.797	-0.019	-0.071**
	(-1.79)	(-1.23)	(-1.07)	(-1.17)	(-2.00)
Medium Financial Confidence	-0.039	-0.032	-2.989	0.013	0.041
	(-1.62)	(-1.37)	(-1.34)	(0.91)	(1.33)
High Financial Confidence	-0.038	-0.119***	-0.208	0.009	-0.025
	(-1.05)	(-3.29)	(-0.07)	(0.51)	(-0.58)
Medium Patience	0.012	0.021	-0.005	-0.005	0.044
	(0.50)	(0.90)	(-0.00)	(-0.38)	(1.50)
High Patience	-0.015	0.038	-1.272	0.036*	-0.006
	(-0.44)	(1.27)	(-0.48)	(1.67)	(-0.16)
Medium Risk Aversion	0.031	-0.017	2.607	0.028**	0.031
	(1.31)	(-0.69)	(1.25)	(1.99)	(1.02)
High Risk Aversion	-0.002	-0.016	4.461*	-0.009	0.040
	(-0.07)	(-0.57)	(1.76)	(-0.63)	(1.16)
Medium General Knowledge	-0.025	0.028	-0.661	0.013	-0.040
	(-1.02)	(1.25)	(-0.32)	(0.94)	(-1.34)
High General Knowledge	-0.049	0.005	-0.932	-0.017	0.017
	(-0.51)	(0.16)	(-0.33)	(-1.02)	(0.45)
Trust in Algorithms	0.138***	0.064***	3.264*	-0.031**	-0.195***
	(6.20)	(2.99)	(1.79)	(-2.28)	(-7.15)
Male	-0.014	0.006	7.240***	-0.019	-0.020
	(-0.65)	(0.27)	(3.72)	(-1.61)	(-0.78)
Has Debt Outside RCT	0.049*	-0.028	-0.337	0.018	0.000
	(1.88)	(-1.13)	(-0.15)	(1.23)	(0.01)
Problem Type FE	✓	✓	✓	✓	✓
Education Category FE	✓	✓	✓	✓	✓
Age Category FE	✓	✓	✓	✓	✓
Adj. R^2	0.133	0.094	0.089	0.042	0.057
N	1,287	1,233	1,028	1,376	1,366
Mean Dependent Variable	0.815	0.834	29.118	0.052	0.336

Table 2 reports the estimated coefficients in the following set of specifications:

$$Robo\_Choice_i = \alpha + X_i\beta + \epsilon_i,$$

where $Robo\_Choice_i$ is a variable that indicates a different choice about robo-advising in each column. The dependent variable in column (1) is a dummy variable that equals one if subject i sought free robo-advice in trial 4 and zero otherwise (this variable is not defined for subjects who were not offered robo-advice). The dependent variable in column (2) is similar except the focus is on paid robo-advice. The dependent variable in column (3) is the subject's reported willingness to pay (WTP) for robo-advice in trial 4, and the sample is limited to subjects who seek paid robo-advice. WTP is measured as the percentage of maximum potential savings from choosing the optimal allocation that the subject is willing to pay to the robo-advisor; it equals zero for the 15.5% of subjects who demand paid robo-advice but then set WTP to zero. The samples in columns (4) and (5) are limited to subjects that receive free or paid robo-advice. The dependent variable in column (4) is a dummy variable that equals one if subject i received a proposed allocation from the robo-advisor but changed the allocation, and zero otherwise. The dependent variable is column (5) is a dummy that equals one if subject i declared in the debriefing survey that they are willing to pay more to interact with a human advisor (not all subjects answered this question). X_i includes the set of individual-level characteristics indicated in the table. We report t-statistics below estimated coefficients. Standard errors are clustered at the subject level (and hence equivalent to Huber-White standard errors since these specifications only include one observation per subject).