

Financial Education, Financial Competence, and Consumer Welfare

Executive Summary

Sandro Ambuehl,
Stanford University

B. Douglas Bernheim,
Stanford University

Fulya Ersoy,
Stanford University

Annamaria Lusardi,
GFLEC, The George
Washington University
School of Business,
TIAA Institute Fellow

David Zuckerman,
Stanford University

Employers routinely attempt to address deficiencies in financial decision making among their employees by deploying educational interventions that are brief and laden with motivational rhetoric. The object of the rhetorical elements is to compensate for brevity by making the material engaging, memorable, and actionable. Prior studies of these interventions evaluate their benefits using criteria that suffer from serious conceptual and practical limitations. In contrast, our research employs a novel method for assessing the quality of decision making that respects each consumer's underlying tastes and focuses instead on mistakes arising from misconceptions about opportunities. We deploy this method to evaluate a pair of educational interventions targeting two critical topics in personal finance: compound interest and portfolio allocation. Both interventions appear to be effective based on conventional outcome measures; they significantly increase performance on financial literacy tests and, in the case of compound interest, change behavior in a direction that appears to counteract a known bias. Nevertheless, upon applying our methods, we discover that neither intervention improves the quality of financial decision making. In both cases, some consumers benefit while others are harmed; overall, the effect on the average quality of decision making is neutral. We resolve the apparent tension between these findings by studying one of the interventions in greater depth to determine whether its effects are attributable to the substantive elements of instruction or to its motivational and rhetorical components. We find that substance accounts for the improvement in performance on literacy tests, while rhetoric accounts for the changes in financial choices. Thus, our analysis shows that motivational rhetoric can have the unintended effect of distracting from substance and promoting an indiscriminate one-size-fits-all response. Conventional methods for evaluating these programs—unlike our methods—fail to detect these deficiencies.

Any opinions expressed herein are those of the authors, and do not necessarily represent the views of TIAA, the TIAA Institute or any other organization with which the authors are affiliated.

Introduction

Despite the regularity with which people make financial decisions, financial literacy, defined as the ability to understand basic principles of business and finance, is low throughout the world (Lusardi & Mitchell, 2014). For example, according to one worldwide survey, only 35% of adults are aware that diversification across multiple investments reduces risk (Klapper et al., 2015). Many employers attempt to bolster financial literacy among their employees by offering financial education. A 2013 survey of retirement plan sponsors by Aon Hewitt found that 77% of providers offered on-site seminars or meetings focusing on critical financial topics (Austin & Evens, 2013).

Economists often assess the effectiveness of financial education by asking whether it either improves tested knowledge or alters behavior in ways that appear to countervail presumed biases. The former strategy is appealing only if the analyst is confident that financial knowledge actually translates into better decisions. The latter strategy requires the analyst to take a stand on the relative merits of the consumer's various options. Because the best choice usually depends on a mix of objective facts and subjective judgments, such stands can be difficult to defend. While the analyst may be an expert on financial principles, each consumer is the only expert concerning his or her own preferences.

Many evaluations of financial decision making implicitly attribute discrepancies between the consumer's choices and the analyst's prescriptions to the consumer's relative lack of financial expertise rather than to differences in tastes, even though there is typically little or no evidence to support this critical assumption. The paternalistic judgments that guide these evaluations leave no room for the possibility that one size may not fit all.

As an example, consider the question of whether households save sufficiently. Basic economic principles teach us that prescriptions for rates of saving should depend not only on the intertemporal budget constraint, the appreciation of which requires expertise, but also on subjective and inherently personal judgments concerning the relative importance of consumption in the present

and the future. Determining whether low saving reflects misapprehended constraints or personal tastes is challenging, to say the least. Two households may agree about the long-term material benefits of austerity, yet disagree about whether the reward justifies the sacrifice. As a result, rigorous evidence-based foundations for the commonplace and seemingly reasonable notion that most people save "too little" are surprisingly difficult to identify.

Taking consumers' underlying preferences into account is easier said than done, especially in complex economic settings that they understand poorly. Yet in our research, we have developed methods for evaluating the quality of financial decision making that respect each consumer's underlying tastes and focus instead on mistakes arising from misconceptions about opportunities. We use these methods in controlled laboratory settings to investigate the effects of educational interventions targeting two distinct classes of financial principles. We find that both interventions perform well according to conventional methods of evaluation; they improve financial literacy as assessed through standard tests, and one appears to countervail presumed behavioral biases. However, our methods reveal that neither intervention improves the average quality of decision making. Further investigation pinpoints the precise reasons for these failures. By deploying our methods more broadly, we hope to discover the keys to designing interventions that address the causes of poor decision making more effectively.

Defining quality of decision making

We say that a consumer displays financial competence with respect to targeted financial principles if she makes equivalent choices from equivalent opportunity sets in contexts where the targeted principles govern the equivalence. In practical terms, we compare measures of the consumer's willingness to accept (WTA) for two equivalent claims on future income, where one is a simplified version of the other. (The WTA for a claim is defined as the smallest amount of money the consumer would accept in exchange for the claim.) The simple version of the claim states the future income transparently. The complex version packages the claim as an income-generating asset. We design the asset so that knowledge of the targeted financial principles

is required to infer the associated income, and hence to understand the equivalence between the simple and complex versions. Someone who both possesses and fully operationalizes that knowledge will consistently ascribe the same value to both claims regardless of their preferences and/or other decision biases. Thus, when a consumer's WTAs for equivalent claims differ systematically, the magnitude of the divergence provides an intuitively appealing measure of her competence to make good decisions in contexts involving the pertinent principles.

Apart from its intuitive appeal, our measure of financial competence is interpretable as an index of economic welfare: it indicates the extent to which the consumer's incomplete operational command of the principles that govern the equivalence exposes her to losses.¹ This point is most easily understood through an example. Imagine a consumer's WTA for a given claim on future income is \$20 when framed transparently, but \$15 when framed as a financial instrument. Her choices indicate that if she fully understood the instrument, her WTA would be \$20 rather than \$15. The difference of \$5 represents the cost of the most severe error she could make as a consequence of her misconceptions. Specifically, she would be willing to sell the instrument for any price exceeding \$15, even though its true value to her is \$20.

Contributions

Our research makes two main contributions.

First, we have developed a new method for evaluating the quality of financial decision making—one that offers important advantages over previous approaches. Some existing studies focus on the propensity for consumers to select dominated options in settings wherein certain choices are objectively bad irrespective of preferences (Ernst et al. 2004, among others). While this approach has intuitive appeal, the frequency of dominated outcomes can be a poor indicator of economic welfare. Consider, for

instance, the problem of portfolio allocation. Convincing everyone to invest all their assets in Treasury bills would reduce the frequency of dominated outcomes to zero, inasmuch as other options achieve higher average returns only at the cost of greater risk, but it would make many people worse off. Other studies attempt to draw inferences about the quality of decision making from highly detailed models of consumers' preferences and opportunities (see, for instance, Song 2015)). Unfortunately, the construction of such models necessitates much more restrictive assumptions than many economic analysts are willing to make or accept. Still, other studies examine the consistency of a consumer's choices across non-equivalent decision problems (see, for instance, Choi et al., 2014). A limitation of this method is that it cannot detect bad choices that result from consistent misunderstandings. Our approach overcomes all of the aforementioned limitations.

Our second main contribution is to document the potential pitfalls of the types of brief interventions commonly used for financial education in the workplace, and to demonstrate that conventional methods of evaluation may fail to detect their deficiencies. To compensate for their brevity, these programs generally focus on simple heuristics accompanied by highly motivating messages. The intent is to make the material engaging, memorable, and actionable. Yet compelling rhetoric may also distract from substance and promote a one-size-fits-all response, which may be excessive for some and even directionally inappropriate for others. We conduct experiments involving two well-designed interventions addressing, respectively, compound interest and portfolio allocation, two central topics of enormous practical importance. Both interventions significantly improve test scores, and they change behavior in directions that appear to countervail known or presumed biases. However, according to our measure of financial competence, neither improves the average quality of financial decision making. We account for

¹ Reliable welfare analysis is potentially challenging because consumers may suffer from additional decision-making biases falling outside the scope of analysis. Yet we have proven that our measure of financial competence remains interpretable as an index of economic welfare even when the consumer suffers from other decision biases, known and/or unknown; see Ambuehl, Bernheim and Lusardi (2016) for details. General foundations for our approach to assessing the quality of decision making can be found in Bernheim and Rangel (2009) and Bernheim (2016).

these findings by demonstrating that the observed changes in behavior occur in response to the motivational and rhetorical components of these interventions, rather than to their substance.

Next, we summarize the results obtained to date for two research projects concerning, respectively, compound interest and portfolio allocation. Both projects involve experiments and share a similar design. Subjects made a collection of incentivized decisions after viewing an educational video that differs between the treatment and control groups. Then we administered a test to assess their knowledge of the targeted topics, as well as a survey to elicit attitudes, self-reported decision strategies, and other information. We conducted the two experiments with different subject pools (Amazon Mechanical Turk participants vs college students), in different settings (online versus laboratory), and employing varying degrees of abstraction in decision tasks. Despite these differences, we find similar results.

Project I: Evaluating an educational intervention on compound interest

Our first project examines the effects of a financial education module focusing on compound interest. Sound financial decision making requires an operational understanding of this topic. Yet previous studies have shown that people typically underestimate compounding, a phenomenon known as exponential growth bias (Wagenaar & Sagaria, 1975, Levy & Tasoff, 2016). The consequences of this bias for personal financial security are potentially severe.

The educational module we study is based on a popular book, *The Elements of Investing: Easy Lessons for Every Investor*, by Burton Malkiel and Charles Ellis (2013). It resembles typical workplace financial interventions with respect to its brevity, reliance on motivational rhetoric, and use of simple prescriptions.² In addition to describing the mechanics of compounding, it also

provides a useful rule of thumb for making interest calculations (the “Rule of 72”).³

We recruited subjects for this experiment from an online platform, Amazon Mechanical Turk, and assigned them randomly to one of four treatments groups. Each group viewed a different video. One (the “Full” group) viewed the entire video. A second (the “Substance-Only” group) viewed a modified version that omitted rhetorical material such as exhortations and motivational quotes, but retained all substantive instruction. A third (the “Rhetoric-Only” group) viewed a modified version that omitted all substantive material beyond the basic explanation of compounding, but retained all motivational appeals. A fourth (the “Control” group) viewed an educational video on an unrelated financial topic (index funds). Our purpose in exploring these varied treatments was to pinpoint the mechanisms through which the intervention affects financial decisions.

After viewing the instructional video, subjects were presented with a collection of decision tasks. Unbeknownst to them, each task was presented twice, once simply framed and once complexly framed, in random order. As an example, we elicited the value each subject attached to \$10 invested for 36 days at an interest rate of 2%, compounded daily. Separately, we elicited the value the subject attached to \$20 in 36 days. Depending on their patience and confidence in the experimenter, people discount these types of payments to differing degrees, and many discount them substantially. However, regardless of their preferences and expectations, if they understand compound interest, their valuations for these two claims should be identical. In contrast, if subjects suffer from exponential growth bias, their valuations should be higher with simple framing than with complex framing.

² As an example, it instructs viewers that “Albert Einstein is said to have described compound interest as the most powerful force in the universe.”

³ The Rule of 72 states that, to a close approximation, the doubling period for an investment equals 72 divided by the percentage rate of return.

Results

The Full treatment substantially increases financial knowledge test scores, as well as the frequency with which subjects say they used the Rule of 72 when making decisions in the complex frame, and the frequency with which subjects say they relied on mathematical calculations when making decisions. Furthermore, these effects are mainly driven by the substantive material in the module rather than by the rhetorical material, as one would expect of an effective intervention. Figure 1 displays the results for test scores. Subjects in the Full treatment correctly answered 68% of the questions, compared with only 39% in the Control treatment. Results for the Substance-Only group are similar to those for the Full treatment, while results for the Rhetoric-Only group more closely resemble those for the Control treatment.

Next, we ask whether subjects' valuations differ systematically between simply and complexly framed claims. We refer to the difference between these valuations as the *framing distortion*. Figure 2 shows that the average framing distortion is negative and statistically significant for the Control group, which means that valuations are typically higher with simple framing. This finding is consistent with previous studies in that it reflects underestimation of compound interest (exponential growth bias). Remarkably, Figure 2 also shows that the Full treatment essentially eliminates the average framing distortion.⁴ On its face, this finding suggests that the intervention may indeed improve the quality of decision making. However, other findings in Figure 2 suggest that such inferences are premature. In particular, one can see that the effect is essentially the same when the substantive portions of instruction are removed, and largely disappears when the motivational rhetoric is removed. It follows that the change in the average framing distortion is entirely attributable to the rhetoric rather than to the substance. This finding is not encouraging, because the rhetoric has a one-size-fits-all flavor that is at odds with the observed heterogeneity in tastes and initial biases.

The problem with taking a simple average of the framing distortions is that instances of overvaluation and undervaluation can cancel out, creating the false impression that errors are small and/or infrequent. For instance, if one decision yields a \$5 overvaluation and another yields a \$5 undervaluation, the average framing distortion is zero, even though each decision involves a valuation error that is \$5. To assess the impact of our treatments on the quality of decision making more accurately, we focus next on the average magnitudes of the decision errors by taking absolute values before averaging. According to this measure, the average valuation error in the preceding example is \$5, exactly as it should be. The results in Figure 3 confirm our concerns. Notice in particular that the Full treatment yields only a small and statistically insignificant improvement in the average quality of financial decision making. Further investigation reveals the reason—in keeping with the one-size-fits-all nature of the motivational rhetoric, the effect of the Full treatment is indiscriminate. Ideally, an educational intervention should increase or decrease valuations in the complex frame according to whether the individual undervalues or overvalues those claims. Yet, we find that complex valuations increase for everyone, regardless of their initial biases. The effect is beneficial for some people and harmful for the others.

Further analysis sheds light on the mechanisms behind these patterns. The Substance-Only treatment has an unmistakable effect on effort: subjects take significantly more time to make decisions than in the Control treatment, and their choices change in measurable ways, even though they do not significantly improve. Apparently, these subjects at least attempt to apply the knowledge they acquire, albeit unsuccessfully. However, once the rhetorical elements of instruction are added, subjects revert to making decisions as quickly as in the Control treatment, which indicates that the rhetoric overwhelms the substance. Even though subjects process the substantive material (as indicated by tests of knowledge), they operationalize only a blunt heuristic concerning the extraordinary power of compounding when making consequential decisions.

⁴ The Full treatment eliminates the average framing distortion by changing valuations with complex framing. Valuations with simple framing are unaffected, just as one would expect.

Our analysis has important implications for the types of brief, rhetoric-laden educational interventions commonly used in the workplace. As we have said, motivational rhetoric is used to make the material engaging, memorable, and actionable. We have shown that it can have the unintended effect of distracting from substance and promoting an indiscriminate one-size-fits-all response that benefits some and harms others. Alarming, conventional methods for evaluating these programs—unlike our methods—fail to detect these deficiencies.

Project II: Evaluating an educational intervention on portfolio allocation

Our second project examines the effects of a financial education intervention focusing on risk and portfolio allocation. These crucial topics are difficult for most people to grasp. Most people do not understand even the most basic principles concerning risk (Lusardi, 2015; Klapper et al., 2015; Coppola et al., 2017; Lusardi et al., 2017), and research shows that they bear substantial economic costs as a result. For example, Calvet et al. (2007) find that the annual costs of under-diversification amount to as much as 4.5 percent of disposable income for one in ten Swedish investors.

The educational module we study is based on two popular finance books, *A Random Walk Down Wall Street*, by Burton Malkiel (2013), and *The Elements of Investing: Easy Lessons for Every Investor*, by Burton Malkiel and Charles Ellis (2013). It resembles typical workplace financial interventions with respect to its brevity (approximately 30 minutes), reliance on motivational rhetoric, and use of simple prescriptions (such as “diversify, diversify, diversify”). It conveys important concepts concerning portfolio diversification, provides simple examples, and illustrates important ideas with figures and animation.

We recruited subjects for this experiment from two pools: Amazon Mechanical Turk and Stanford University students. We assigned subjects randomly either to a Treatment group, which viewed our video on risk and portfolio allocation, or to a Control group, which viewed a video on an unrelated financial topic.

After watching the instructional video, subjects were presented with a collection of decision tasks. Unbeknown to them, each task was presented twice (as in the previous experiment), once simply framed and once complexly framed, in random order. As an example, we elicited the value each subject attached to a lottery that pays \$5 if a fair coin comes up head and \$15 if it comes up tails (See Figure 4). Separately, we elicited the value the subject attached to a portfolio containing five units (“shares”) of an asset that pays (per unit) nothing for heads and \$1 for tails, and five units of an asset that pays (per unit) \$1 for heads and \$2 for tails (see Figure 5). Depending on their risk preferences, different people value the same lottery differently. However, regardless of their preferences, if they understand the principles of portfolio allocation, their valuations for these two claims should be identical. In contrast, if subjects are averse to options they fear they may not understand, their valuations should be higher with simple framing than with complex framing.

A possible objection to evaluating educational interventions based on the types of abstract decision tasks shown in Figures 4 and 5 is that the instructional material does not specifically target them. Because risk is such a complex topic, the typical intervention attempts to convey simple rules and heuristics that may be inappropriate in some circumstances, but that tend to perform well in the types of environments people actually encounter (Gigerenzer, Todd, and the ABC research group, 1999). We therefore conduct our experiment in two different settings. The *Abstract* setting involves decision tasks resembling those shown in Figures 4 and 5, designed to explore decision-making performance on tasks that we intentionally endow with particular features—for example, ones in which equally weighted portfolios perform well, and ones in which they perform poorly. In contrast, the *Naturalistic* setting involves decision tasks that mimic realistic portfolio decisions—we determine the distribution of outcomes based on actual historical returns, and present information to subjects formatted so as to resemble materials used by financial services companies. This setting enables us to evaluate the effectiveness of the educational intervention in the types of environments for which the instruction was specifically designed.

Results

The financial education intervention substantially increases scores on tests of knowledge concerning risk and portfolio principles. Figure 6a depicts average scores for the Treatment and Control groups for our Abstract experiment, which used subjects recruited through Amazon Mechanical Turk. Figure 6b shows the same information for our Naturalistic experiment, which used Stanford students. The difference in subject pools accounts for the baseline difference in scores between the Control groups. In Figure 6a, the Treatment improves the average score from 57% to 74%. The magnitude of the improvement in Figure 6b is similar.

Next, we ask whether subjects' valuations differ systematically between simply and complexly framed tasks. Figures 7a and 7b show that the average framing distortion, defined as in the previous experiment, is negative for both Control groups, which means that valuations are typically higher with simple framing. This finding is consistent with the notion that people are averse to features of a claim that render it more complex and difficult to understand. Notably, the Stanford sample is considerably less complexity averse than the Amazon Mechanical Turk sample—a finding that potentially reflects the relative sophistication and/or numeracy of these groups. The educational intervention reduces the average framing distortion in the Abstract setting, but increases it in the Naturalistic setting, though neither effect is statistically significant. The possibility that education may increase complexity aversion, as found in Figure 7b, is particularly intriguing. It suggests that, when tasks are sufficiently complex, a little bit of knowledge may cause consumers to perceive portfolio investments as more risky, thereby compounding an existing bias.

To assess the impact of the intervention on the quality of decision making, we follow the same procedure as in our first experiment, focusing on the average magnitudes of the decision errors by taking absolute values before averaging. Our findings appear in Figure 8. Here we see almost no difference between the treatment and control groups for either the Abstract or Naturalistic settings. Responses to the intervention appear to be largely unsystematic, bearing little or no relation to the

consumer's initial bias. Some consumers benefit while others are harmed; overall, the effect on the average quality of decision making is neutral.

Good financial decision making does not necessarily require strong financial skills if individuals recognize their own limitations and can identify appropriate advisors. Unfortunately, research shows that low skill tends to go hand-in-hand with a poor ability to evaluate skill (Kruger and Dunning, 1999). Even if education fails to improve financial decision making ability, it may improve outcomes if it alerts consumers to their limited understanding and motivates them to seek help. To explore these issues, we assess each subject's willingness to pay to replace the choice they made in a randomly selected complexly framed task with their corresponding choice in the associated simply framed task—an option that simulates “ideal advice.” Our findings appear in Figure 9. In the Abstract setting, we find that the educational intervention significantly increases the willingness to pay for ideal advice, which means it makes consumers more aware of their limitations. Surprisingly, we do not find the same effect in the Naturalistic setting, where we see no indication that education makes consumers more aware of their limitations.

Conclusion

Summary

Employers routinely attempt to address deficiencies in financial decision making among their employees by deploying educational interventions that are brief and laden with motivational rhetoric. Prior studies of these interventions evaluate their benefits using criteria that suffer from serious conceptual and practical limitations. In contrast, our research employs a novel, non-paternalistic, and easy-to-implement method for assessing the quality of decision making. We deploy this method to evaluate a pair of educational interventions targeting two critical topics in personal finance: compound interest and portfolio allocation. Both interventions appear to be effective, based on conventional outcome measures: they significantly increase performance on financial literacy tests and,

in the case of compound interest, change behavior in a direction that appears to counteract a known bias. Nevertheless, upon applying our methods, we discover that neither intervention improves the quality of financial decision making. In both cases, some consumers benefit while others are harmed; overall, the effect on the average quality of decision making is neutral. We resolve the apparent tension between these findings by studying one of the interventions in greater depth to determine whether its effects are attributable to the substantive elements of instruction or to its motivational and rhetorical components. We find that substance accounts for the improvement in performance on literacy tests, while rhetoric accounts for the changes in financial choices. Thus, our analysis shows that that motivational rhetoric can have the unintended effect of distracting from substance and promoting an indiscriminate one-size-fits-all response. Conventional methods for evaluating these programs – unlike our methods – fail to detect these deficiencies.

Policy relevance and implications

Conventional criteria for evaluating financial education interventions can yield misleading conclusions concerning the effects on the quality of decision making. Because our method does not suffer from the same limitations, it provides useful tools for evaluating existing interventions, as well as for guiding the design and development of future interventions. Of particular note, our findings provide an important cautionary note concerning the use of motivational rhetoric and exhortations. In making educational interventions engaging, memorable, and actionable, their designers may inadvertently diminish the potentially beneficial effects of their substantive elements. Striking the proper balance is likely to prove challenging.

Three classes of strategies for improving the quality of financial decision making merit further analysis and discussion. The first is to devise pedagogical methods that more effectively induce consumers to apply “book knowledge” of financial principles to their consequential decisions. One possibility is to emphasize practical exercises and provide feedback. The second strategy is to employ interventions that rely on powerful one-size-

fits-all exhortations while targeting populations known to suffer from biases that the intervention countervails (“targeted de-biasing”). A final strategy is to simplify the framing of naturally occurring decision problems, either by developing and deploying better tools for visualizing opportunities and consequences, or by requiring suppliers of financial products to characterize them in simple terms.

Next steps

We envision many directions for subsequent research, some of which we are already pursuing. First, we are using our methods to evaluate whether peer-to-peer communication improves the quality of financial decision making. Most adults rely on friends and family, rather than trained professionals, for advice and guidance (Bernheim, 1998, Lusardi, 2004). It is therefore important to know whether communication among non-experts helps or hinders decision making, and whether knowledge transfers successfully from one person to another within this domain. Second, we are planning to bring our method into the field to assess a full-scale employer-based financial education program. We hope that this project will yield insights concerning the relative effectiveness of different pedagogical approaches, and thus provide useful guidance for future program design. Third, in collaboration with a leading U.S. asset management firm, we are currently using our methods to assess the financial competence of medium to high net worth individuals (investable assets > 250k). Fourth, we are deploying our methods to investigate the role of motivated reasoning in financial decision making. For example, once a prospective home buyer becomes emotionally invested in a particular house, is he or she more likely to conclude, upon reviewing materials on mortgages, that houses in that price range are affordable? If so, does the use of just-in-time education reduce the quality of decision making? And what can be done to improve decisions in light of these psychological biases? Additional applications of our methods promise to yield important insights concerning these and other issues.

References

- Ambuehl, S., Bernheim, B. D., & Lusardi, A. (2016). The effect of financial education on the quality of decision making. *NBER Working Paper*, 20618.
- Austin, R. & Evens, W. (2013). 2013 trends & experience in defined contribution plans. *Aon Hewitt*.
- Bernheim, B. (1998). Financial illiteracy, education, and retirement saving. In Mitchell, O. S. and Schieber, S. J., editors, *Living with Defined Contribution Pensions - Remaking Responsibility for Retirement*, chapter 3. University of Pennsylvania Press.
- Bernheim, B. (2016). The Good, the Bad, and the Ugly: A Unified Approach to Behavioral Welfare Economics. *Journal of Benefit-Cost Analysis*, 7(1), 12-68.
- Bernheim, B. & Rangel, A. (2009). Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics. *Quarterly Journal of Economics*, 124(1): 51–104
- Calvet, L. E., Campbell, J. Y., & Sodini, P. (2007). Down or out: Assessing the welfare costs of household investment mistakes. *Journal of Political Economy*, 115(5):707–47.
- Choi, S., Kariv, S., Müller, W., & Silverman, D. (2014). Who is (more) rational? *American Economic Review*, 104(6):1518–550.
- Coppola, M, G. Langley, M. Sabatini & R. Wolf (2017), When will the penny drop? Money, financial literacy and risk in the digital age. *International Pension Papers* 1/2017
- Ernst, K., Farris, J., & King, U. (2004). Quantifying the economic cost of predatory payday lending. Technical report, Center for Responsible Lending.
- Gigerenzer, G., Todd, P.M., & the ABC Research Group. (1999). *Simple Heuristics That Make Us Smart*. New York: Oxford University Press.
- Klapper, L., Lusardi, A., & van Oudheusden, P. (2015). Financial literacy around the world: Insights from the Standard & Poor's Rating Services Global Financial Literacy Survey.
- Kruger, J. and Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6):1121–134.
- Levy, M. & Tasoff, J. (2016). Exponential growth bias and lifecycle consumption. *Journal of the European Economic Association*, 14(3):545–83.
- Lusardi, Annamaria (2004). "Savings and the Effectiveness of Financial Education" in Olivia S. Mitchell and Stephen Utkus (eds.), *Pension Design and Structure: New Lessons from Behavioral Finance*, Oxford: Oxford University Press, pp. 157–184.
- Lusardi, A. (2015), Risk literacy. *Italian Economic Journal*, 1(1):5-23.
- Lusardi, A. & Mitchell, O. (2014). The economic importance of financial literacy: Theory and evidence. *Journal of Economic Literature*, 52(1):1–44.
- Lusardi, A., Oggero N., & Yakoboski, P. (2017), The TIAA-Institute-GFLEC Personal Finance Index: A new measure of financial index.
- Malkiel, B. G. (2003). *A random walk down Wall Street*. New York: Norton.

- 
- Malkiel, B. G. & Ellis, C. D. (2013). *The Elements of Investing: Easy Lessons for Every Investor*. Wiley, New Jersey.
- Song, C. (2015). Financial illiteracy and pension contributions: A field experiment on compound interest in China. *Unpublished Manuscript*.
- Wagenaar, W. M. & Sagaria, S. D. (1975). Misperception of exponential growth. *Perception and Psychophysics*, 18(6):416–22.

Appendix

Figure 1: Effect of treatment on test scores, compound interest project

The mean number of test questions answered correctly (1 to 5) is depicted for each treatment group separately. 95% confidence intervals depicted in gray.

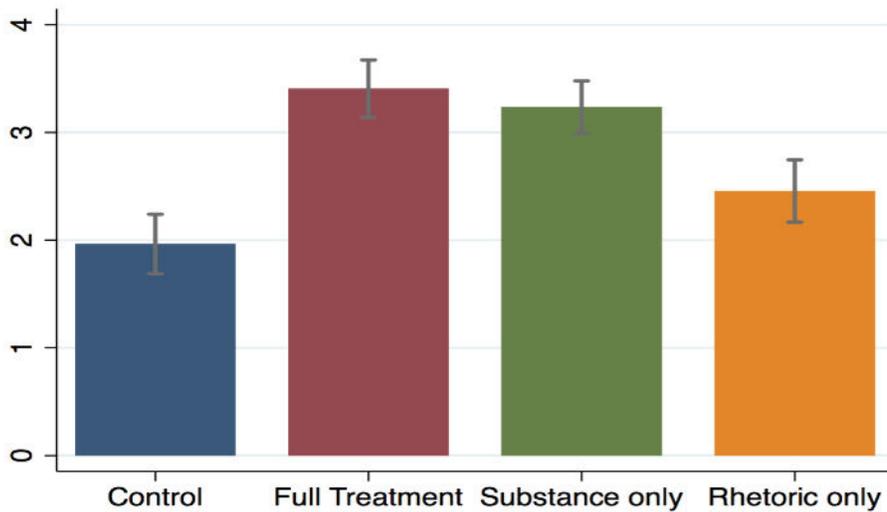


Figure 2: Effect of treatment on framing distortion, compound interest project

A subject's valuation in a simple frame question is subtracted from his/her valuation in the equivalent complex frame question and normalized by the amount of the future reward. It is averaged across subjects and questions to form the framing distortion. A negative number indicates exponential growth bias. 95% confidence intervals depicted in gray.

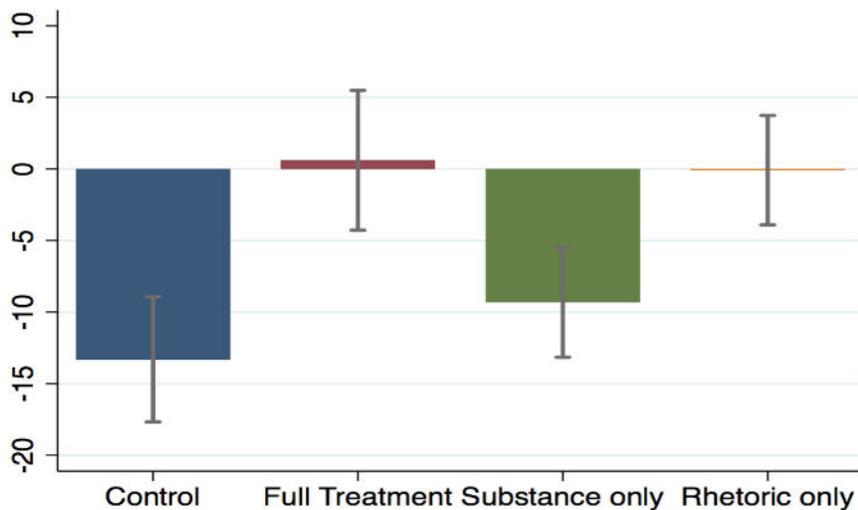


Figure 3: Effect of treatment on quality of decision making, compound interest project

The absolute difference between a subject's valuation in the simple and equivalent complex frame question, normalized by the amount of the future reward. It is averaged across subjects and questions to form the quality of decision making. 0 means perfect quality of decision making. 95% confidence intervals depicted in gray.

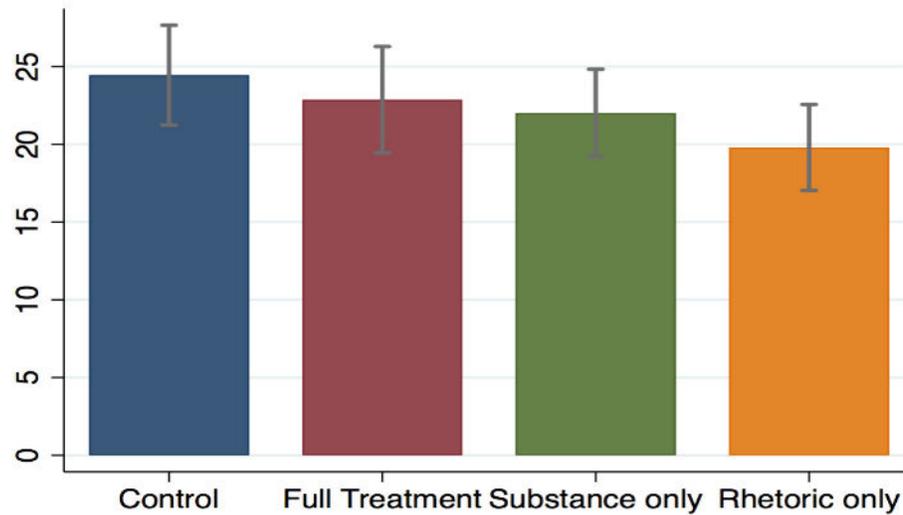


Figure 4: Lottery in simple frame

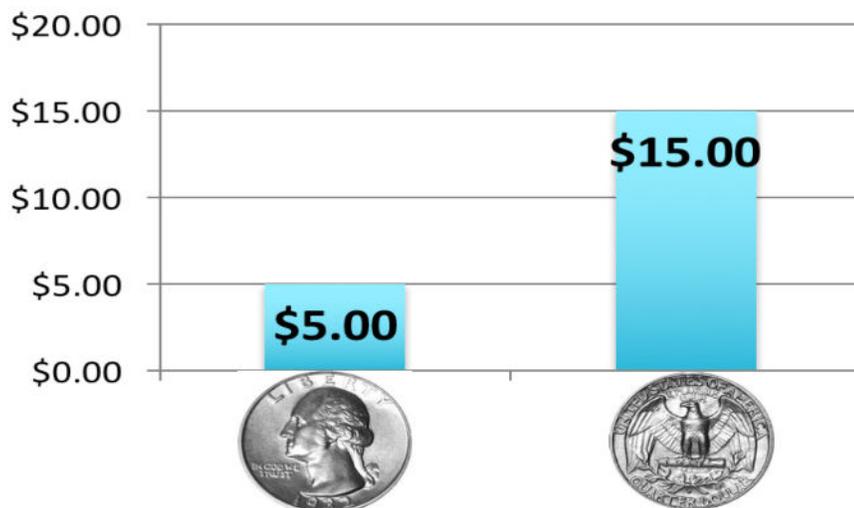


Figure 5a: Each unit of lottery 1



Figure 5b: Each unit of lottery 2

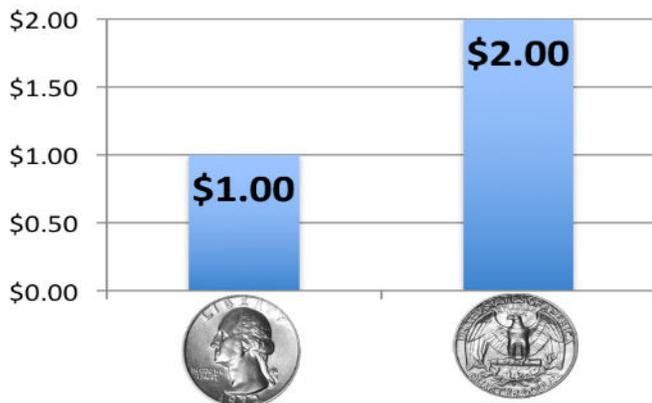


Figure 5c: Lottery in complex frame

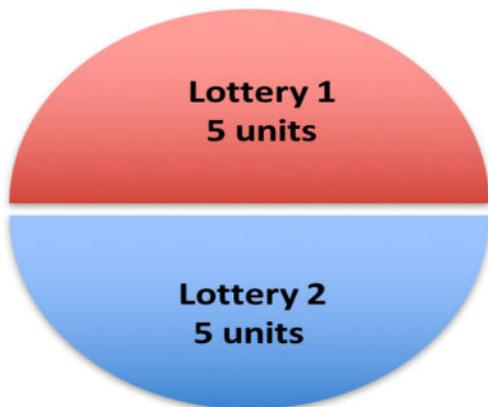


Figure 6a: Effect of treatment on test scores, abstract setting, risk project

The mean number of test questions answered correctly (1 to 10) is depicted for control and treatment group separately. 95% confidence intervals depicted in gray.

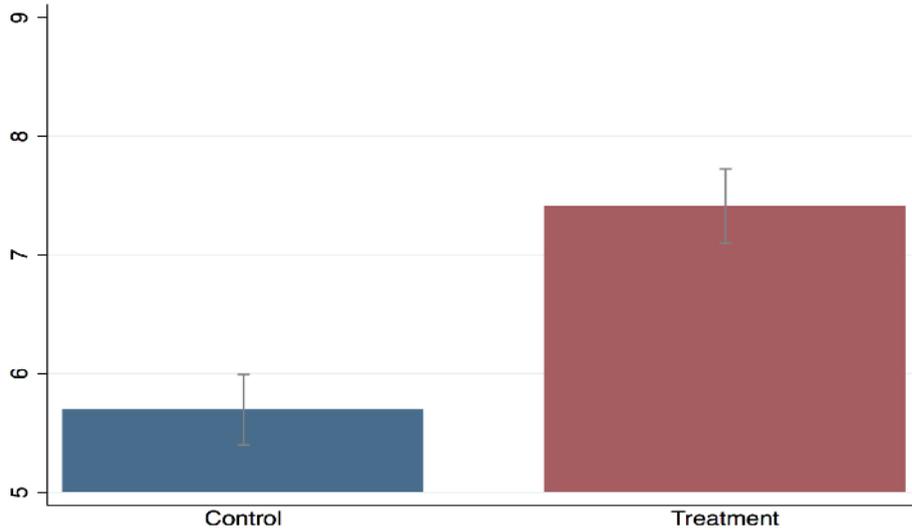


Figure 6b: Effect of treatment on test scores, naturalistic setting, risk project

The mean number of test questions answered correctly (1 to 10) is depicted for control and treatment group separately. 95% confidence intervals depicted in gray.

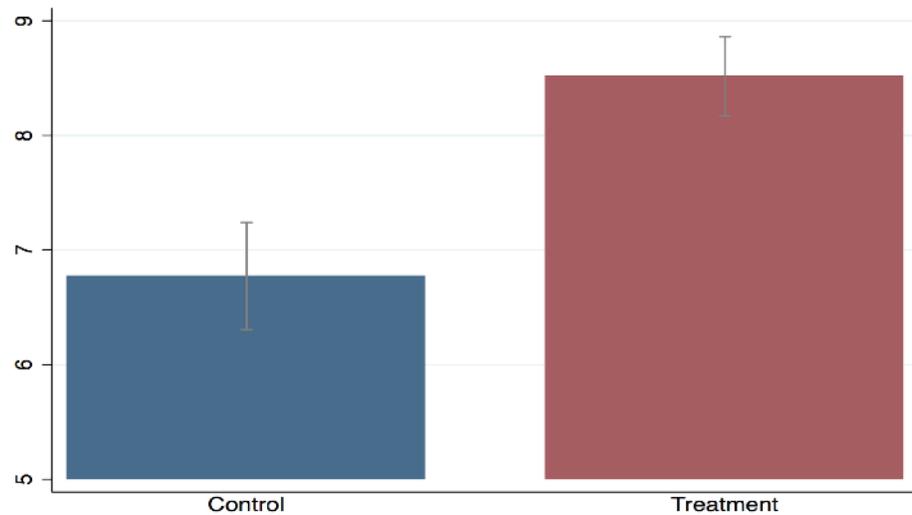


Figure 7a: Effect of treatment on framing distortion, abstract setting, risk project

A subject's valuation in a simple frame question is subtracted from his/her valuation in the equivalent complex frame question. It is averaged across subjects and questions to form the framing distortion. A negative number indicates complexity aversion. 95% confidence intervals depicted in gray.

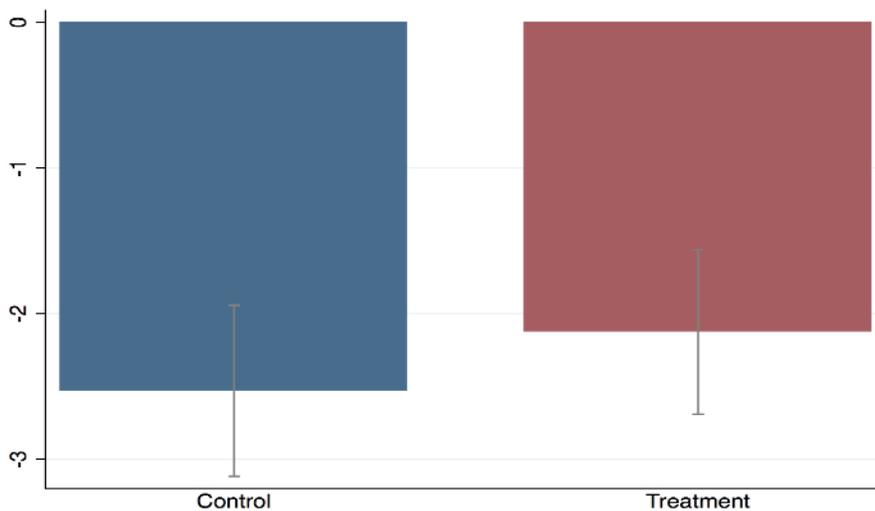


Figure 7b: Effect of treatment on framing distortion, naturalistic setting, risk project

A subject's valuation in a simple frame question is subtracted from his/her valuation in the equivalent complex frame question. It is averaged across subjects and questions to form the framing distortion. A negative number indicates complexity aversion. 95% confidence intervals depicted in gray.

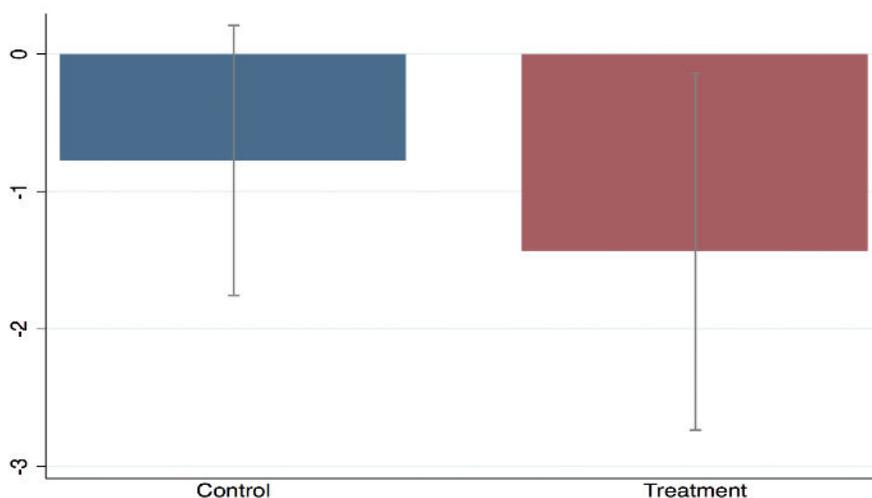


Figure 8a: Effect of treatment on quality of decision making, abstract setting, risk project

The absolute difference between a subject's valuation in the simple and equivalent complex frame question. It is averaged across subjects and questions to form the quality of decision making. 0 means perfect quality of decision making. 95% confidence intervals depicted in gray.

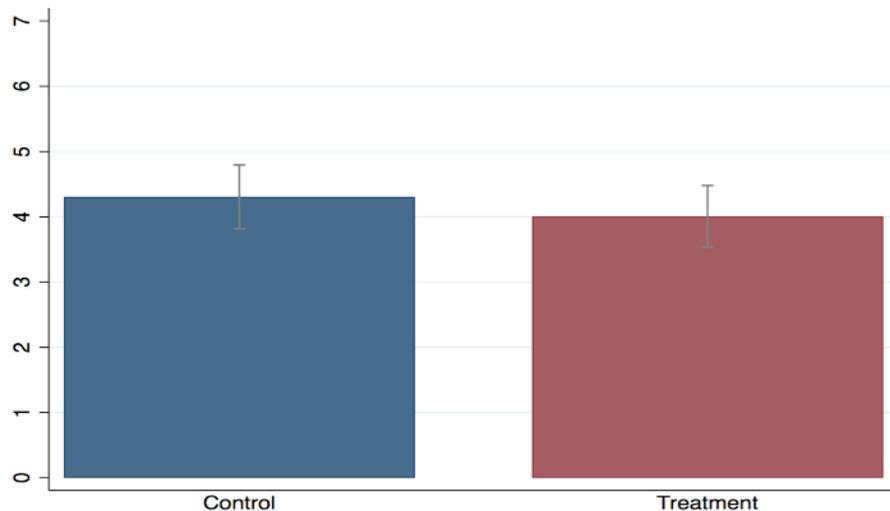


Figure 8b: Effect of treatment on quality of decision making, naturalistic setting, risk project

The absolute difference between a subject's valuation in the simple and equivalent complex frame question. It is averaged across subjects and questions to form the quality of decision making. 0 means perfect quality of decision making. 95% confidence intervals depicted in gray.

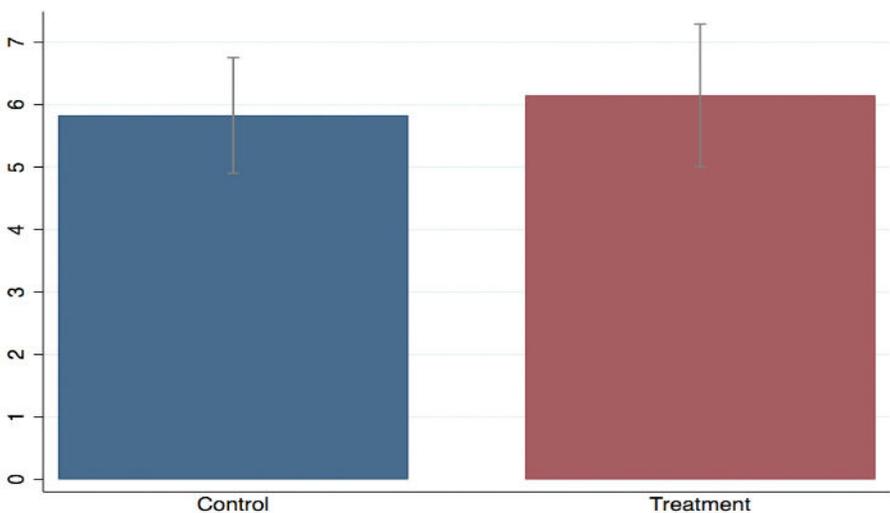


Figure 9a: Effect of treatment on confidence, abstract setting, risk project

The average willingness to pay (\$) to switch from the valuation in the complex frame to the valuation in the equivalent simple frame problem. A higher number signifies lower confidence. 95% confidence intervals depicted in gray.

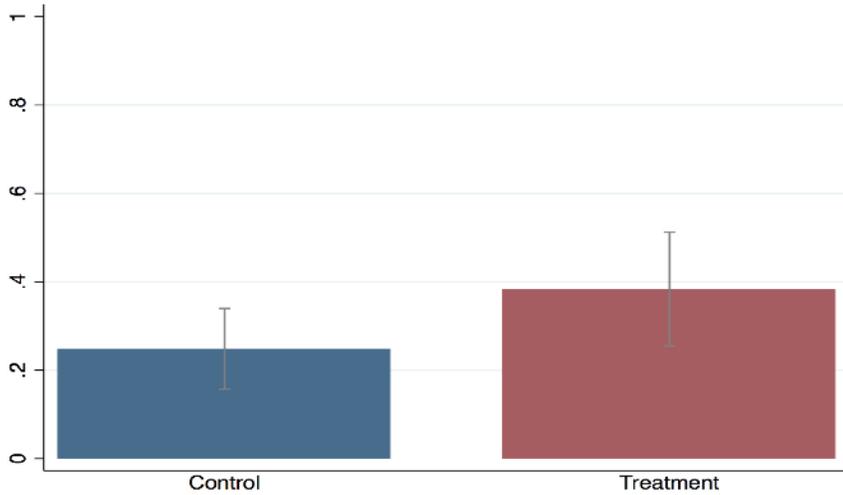
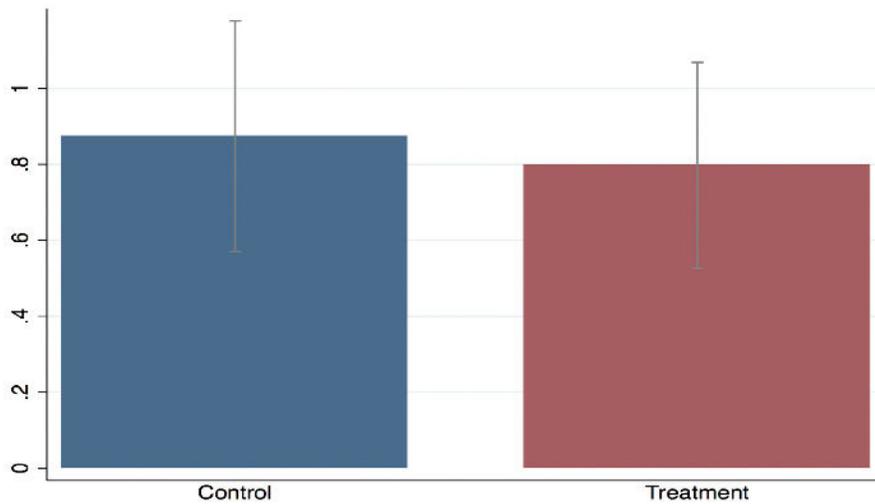


Figure 9b: Effect of treatment on confidence, naturalistic setting, risk project

The average willingness to pay (\$) to switch from the valuation in the complex frame to the valuation in the equivalent simple frame problem. A higher number signifies lower confidence. 95% confidence intervals depicted in gray.



About the Authors

Sandro Ambuehl is an Assistant Professor of Economics at the Department of Management at UTSC and at the Rotman School of Management, University of Toronto. Professor Ambuehl's research specializes in behavioral and experimental economics. One of his areas of research examines what policies help people make good financial decisions. Another line of research concerns designing policies for the exchange of goods about which people have strong ethical intuitions, as is the case, for example, with living organ donation, or when incentive systems place a price on the natural environment. He addresses these questions using a combination of controlled experiments and economic theory. His research has been featured in the popular press, including the Washington Post, the Financial Times, and the Frankfurter Allgemeine Zeitung.

B. Douglas Bernheim is the Edward Ames Edmonds Professor of Economics in the Department of Economics at Stanford University, as well as Department Chair. After completing an A.B. in Economics from Harvard University and a Ph.D. from the Massachusetts Institute of Technology, he joined the Stanford faculty as an Assistant Professor in 1982. He moved to Northwestern University's J. L. Kellogg Graduate School of Management in 1988, and to Princeton University in 1990, before returning to Stanford in 1994. His awards and honors include election as a fellow of the American Academy of Arts and Sciences, election as a fellow of the Econometric Society, a John Simon Guggenheim Memorial Foundation Fellowship, and an Alfred P. Sloan Foundation Research Fellowship. Professor Bernheim's work has spanned a variety of fields, including public economics, behavioral economics, game theory, contract theory, industrial organization, political economy, and financial economics. A notable contribution of his include conducting the earliest economic analyses of financial education. Professor Bernheim is a Research Associate of the National Bureau of Economic Research, a Senior Fellow of the Stanford Institute for Economic Policy Research (SIEPR), and Co-Director of SIEPR's Tax and Budget Policy Program. He has also served as the Director of the Stanford Institute for Theoretical Economics (SITE), and as Co-Editor of the *American Economic Review*. He is currently serving as Co-Editor of the *Handbook of Behavioral Economics*.

Fulya Ersoy is a Ph.D. candidate in Economics at Stanford University. She holds a B.A. in Management from Bogazici University and an M.A. in Economics from Sabanci University. She spent a semester visiting Maastricht University during her undergraduate studies. Her primary research interests are the economics of education and experimental economics. In particular, she utilizes behavioral insights, experimental methods, and empirical methods to understand how students make their choices about their effort, educational attainment, or field of study. Her current work aims to explain how students' beliefs about the effort-rewards relationship play a role in students' effort choices.

Annamaria Lusardi is the Denit Trust Chair of Economics and Accountancy at the George Washington University School of Business. She previously taught at Dartmouth College, Princeton University, the University of Chicago Harris School of Public Policy and Booth School of Business, and Columbia Business School. She also was a visiting scholar at Harvard Business School.

Lusardi has won numerous awards for her economic and financial literacy research, including a research fellowship from the University of Chicago Harris School of Public Policy, a faculty fellowship from the John M. Olin Foundation, and a junior and senior faculty fellowship from Dartmouth College. She earned her B.A. from Bocconi University in Milan and Ph.D. from Princeton University.



David Zuckerman is currently a third year graduate student in the Economics department at Stanford University. He holds a B.A. in Physics and Economics from University of Rochester and an M.A. in Economics from Stanford University. His main research interests lie within the realm of experimental and behavioral economics. He is especially interested in “experienced-based decision making,” in which people’s decisions are primarily guided by past outcomes they have experienced. Another avenue of his research lies is in financial learning, studying the effect of interventions and past experiences on financial behavior. He hopes to continue research in both of these domains, using economic theory to guide his thinking.