BEHAVIORAL FINANCE: INTERVENTIONS ENCOURAGING RETIREMENT SAVINGS REVIEW PROTOCOL

Highlights

- The objective of this systematic review is to determine the quality of existing causal evidence on behavioral interventions that encourage employees to save more for retirement.
- The review focuses on interventions grounded in behavioral economics that aim to increase retirement savings behaviors. These interventions leverage observations about how employees' choices differ from the predictions of standard economic models in order to affect behavior.
- Research using causal designs is of primary importance to this topic area. Other types of research provide background and context, but are not systematically reviewed.

Introduction

The topic area for this evidence review protocol is interventions designed using insights from behavioral economics for the purpose of increasing employees' savings for retirement. Behavioral economics is the study of how and why a person's choices depart from the predictions of standard economic models. This field integrates insights from psychology and economics to better understand how contextual, cognitive, social, and emotional factors interact with economic decision making. Behavioral economists have identified many potential causes for the divergence of observed choices and predicted behavior, including the following:[1]

- **Status quo bias.** People make certain choices simply because those choices are viewed as the standard, or default, option. For example, a person might be automatically enrolled in a health maintenance organization by his or her employer but have the option to switch to a preferred provider organization (PPO). Even though the employee prefers the PPO, he or she might not switch from the default plan.
- **Overconfidence.** People can have unrealistic expectations of their performance in the future. For example, it is well known that most people estimate that they have above-average driving skills.[2] This could lead to less-than-optimal rates of insurance coverage or people believing they can drive when slightly impaired.
- **Procrastination.** People may continually put off taking certain actions, always believing they will do so eventually. For example, one can always suppose that he or

[1] See www.ideas42.org for details. Thaler, Richard H. (2009). *Nudge: Improving Decisions About Health, Wealth, and Happiness.* New York: Penguin Books. Kahneman, Daniel. (2013). *Thinking, Fast and Slow.* New York: Farrar, Straus and Giroux.

[2] See Svenson, Ole. (1981). "We are all less risky and more skillful than our fellow drivers." *Acta Psychologia, 47*: 143–148.

she will begin exercising tomorrow. The person is not actively choosing to never begin to exercise, but this is the result of the series of small "I will start tomorrow" choices.

- **Choice overload.** Faced with numerous options, people can become overwhelmed and fail to make the optimal choice or not choose any option at all. For example, a commonly cited research paper[3] found that people were 10 times more likely to purchase jam at an upscale supermarket when given a choice among 6 varieties than when given a choice among 24 varieties.
- **Hassle factors.** Seemingly small annoyances related to time or effort can cause people not to take action even when there is a large anticipated gain from a traditional economic perspective. For example, lower-income high school students express a desire to go to college, yet they do not fill out the Free Application for Federal Student Aid (FAFSA), even though tens of thousands of dollar in aid and higher income after graduation from college are on the table. If you help them get over the hassle by helping them fill out the form, they are not only more likely to complete the form, they are more likely to go to college.[4]
- **Use of rules of thumb.** People can use rules of thumb or other simple tools to try to simplify complexity and make decisions. For example, people commonly spend three months salary on an engagement ring, because this is seen as the standard amount; however, this level of spending might be inappropriate for many people.

Many of these behavioral elements can be at play in determining employees' decisions to save for retirement. These forces have spawned a large literature on potential behavioral interventions to increase retirement savings rates. This review addresses the following research questions about the effectiveness of such interventions:

- Which interventions based on insights from behavioral economics are effective at encouraging employees to save more for retirement in the short run? In the long run?
- Do increases in retirement savings resulting from behavioral interventions crowd out other types of savings or lead people to take on higher levels of debt?
- How do behavioral interventions affect the distribution of retirement savings across the population?

Because socioeconomic characteristics imply different savings needs (for example, Social Security is structured so that, when retired, lower-income people will receive payments that represent a higher share of their pre-retirement earnings than will higher-income people), this review also addresses the following:

- Do the impacts of the interventions of interest vary with an employee's income level, gender, race, or marital status?

[3] Iyengar, Sheena S., and. Lepper, Mark R. (2000). "When choice is demotivating: Can one desire too much of a good thing?" *Journal of Personality and Social Psychology, 79*(6), 995.

[4] See Bettinger, Eric P., Long, Bridget Terry, Oreopoulos, Philip, & Sanbonmatsu, Lisa. (2012). "The role of application assistance in college decisions: Results from the H&R Block FAFSA experiment." *Quarterly Journal of Economics, 127*(3), 1205-1242.

The rest of this evidence review protocol sets forth the criteria by which existing research is determined to be eligible for review, rules for using causal evidence guidelines to rate the quality of causal evidence, and review procedures. Appendix A provides further details on the methods used to identify studies for potential inclusion in the review.

Eligibility Criteria

Clearinghouse for Labor Evaluation and Research (CLEAR) staff members identified research studies for potential inclusion in this topic area using a broad literature search (see Appendix A for details). For this topic area, all identified research that meets eligibility criteria undergoes a second-level review (see the CLEAR Policies and Procedures for more information about review types). The eligibility criteria include the following:

1. **Does the research examine an intervention grounded in behavioral economics or developed based on insights from both psychology and economics?** This topic area explores only interventions that aim to use insights from behavioral economics. Interventions of interest include those developed to overcome behavioral biases (for example, making people aware that the default savings option might not be appropriate for them) or exploit such biases to change behavior (for example, changing the default savings option to a higher contribution rate).
2. **Does the research examine an outcome of interest?** To be eligible for review, research must consider retirement savings or a closely related outcome in one of the following domains:
   - Fund participation and fund choice (for example, enrollment in a 401(k) or allocation of assets to equities)
   - Savings (for example, amount saved or amount contributed to 401(k))
   - Hypothetical decision-making (any outcomes that relate to a choice to save or a choice between investment options but do not involve actual savings decisions)
   - Eligible outcomes include both long- and short-run measures
3. **Is it a study of effectiveness?** To be eligible for review, the research must use quantitative methods to assess the effectiveness of a particular intervention.[5] Studies reviewed under this topic area may use a randomized controlled trial or quasi-experimental design (including regression, instrumental variable, and interrupted time series [ITS] analyses).
4. **Does the research examine a population of interest?** The research can examine any group of employed people and/or their families.
5. **Was the research conducted in a relevant place?** All research must have been conducted in English using data from a country classified as a developed economy.[6]

[5] Causal studies in this topic area were reviewed according to CLEAR Causal Evidence Guidelines, Version 2.0. The full set of guidelines is available at http://clear.dol.gov. CLEAR also has guidelines for reviews of descriptive and implementation research; however, those guidelines were not be applied to this topic area.

[6] See http://www.un.org/en/development/desa/policy/wesp/wesp_current/2012country_class.pdf for a list of these countries.

6. **Was the research published in the relevant time frame?** For this topic area, CLEAR reviewed research published in 1996 or after.

Causal Evidence Guidelines

This topic area includes reviews of both experimental and nonexperimental causal research. CLEAR assesses the quality of evidence for randomized controlled trials (RCTs) using an adaptation of the Institute of Education Science's What Works Clearinghouse (WWC) standards.[7] RCTs can receive a High causal evidence rating if there are no obvious confounds to the RCT design and if the level of attrition is low. This topic area will apply the WWC liberal attrition standard. The liberal standard was chosen because the interventions of interest to this topic area involve small changes in conditions, which should not produce large, purposeful changes in the survey population.

If CLEAR determines that an RCT cannot be rated as providing high causal evidence, the research is reviewed using the nonexperimental causal evidence guidelines developed by CLEAR.

Nonexperimental Causal Evidence Guidelines Specific to the Topic Area

In collaboration with a technical work group of experts, Mathematica Policy Research developed a set of evidence guidelines to use in reviewing studies with nonexperimental designs, including but not limited to instrumental variables, difference-in-differences, fixed and random effects, matching comparison group designs, and regression analyses. Nonexperimental designs that meet the applicable evidence guidelines receive a Moderate causal evidence rating; this rating indicates that there is evidence the research establishes a causal relationship between the intervention being examined and the outcomes of interest, but there might be other factors that were not included in the analysis that could affect the outcomes of interest. Nonexperimental designs that do not meet the guidelines receive a Low causal evidence rating, which indicates that we cannot be confident that the estimated effects are attributable to the intervention being examined.

ITS designs are commonly used in the literature of interest to this topic area; another set of causal evidence guidelines covers such designs. ITS designs can receive a High, Moderate, or Low causal evidence rating depending on how many of the specified criteria the study meet.

Causal evidence guidelines for nonexperimental studies are tailored to the topic area of interest. In particular, the topic area protocol sets forth the specific types of control variables that have to be included in nonexperimental regression analyses (other than those using fixed effects) for research to receive a Moderate causal evidence rating. The topic area protocol also describes whether changes in group composition should be a concern for the review and the time period that must be covered by pre-intervention data if the study uses an ITS design.

Control variables. The control variables for the behavioral interventions encouraging retirement savings protocol were developed in consultation with a topic area expert. The employee-level control variables required for all studies include the following:

[7] See http://ies.ed.gov/ncee/wwc/InsidetheWWC.aspx for details.

- Age
- Gender
- Income level or some proxy for income, such as education or occupation

Research including individual-fixed effects will also meet this requirement.

Many interventions within this topic area take place at the firm level. Research using firm-level data must include comparable measures of employees' characteristics. In addition, research that examines employees' behavior using data from multiple firms must include controls for firm size and the structure of the retirement plan offered by the firm.

Research focusing on interventions conducted outside the workplace (for example, during tax preparation) is exempt from this requirement. Firm-level-fixed effects will also be considered sufficient controls for firm size and retirement plan characteristics.

Regression methods that incorporate a matching design, in which statistical methods are used to create a comparison group that is as similar as possible to the group receiving the intervention, must match on the previously listed control variables or, if they do not match on them, must include them as controls in the regression.

Changes in group composition. Changes in group composition resulting from an intervention are potentially a concern for studies with nonexperimental designs in this topic area. For instance, a difference-in-differences analysis comparing the average change in savings rates among employers in treatment and comparison groups could be biased if the savings of participants who changed employers were not included in the post-intervention outcome measure. Thus, studies with nonexperimental designs and analysis at the group level in this topic area must meet Criterion Regression.4 to receive a Moderate causal evidence rating.

Pre-intervention data. To satisfy Criterion ITS.2, an ITS design must use data drawn from a sufficiently long period of time before an intervention's implementation. For the behavioral finance topic area, data must cover at least one year before the implementation of the intervention.

Review Procedures

Each research paper or report identified as eligible for review against causal evidence guidelines is assigned to a reviewer who has been certified by CLEAR to understand and be able to apply its standards with fidelity. The reviewer reads the study in detail, applies the causal evidence guidelines to determine the design's causal evidence rating, and documents all aspects of the review in a standardized review guide. In particular, the review guide contains supporting information for the rating, details of the study sample and intervention, and any other pertinent information.

If the reviewer assigns a rating of High or Moderate causal evidence, a second reviewer then reviews the research to confirm such a rating is warranted.[8] The principal investigator (PI) or another reconciler resolves any discrepancies between the two reviewers' ratings to determine a

[8] For studies with ITS designs, the second review is conducted by the topic area PI or another senior reviewer.

final causal evidence rating. If the first reviewer assigns a rating of Low, the PI examines the review guide and confirms that the rating is appropriate.

When a research paper or report does not contain sufficient information to determine a causal evidence rating, CLEAR may contact the authors to gather this information; whether this step is undertaken depends on the age of the study and the quantity of information that would have to be gathered (so as not to overly burden authors). Authors receive a minimum of four weeks to respond and reasonable requests for extensions are granted. Information provided by the authors is incorporated into the review and factors into the causal evidence rating. If the authors do not provide the relevant information, the design receives the highest rating that can be determined with the information available in the report.

APPENDIX A

LITERATURE SEARCH

Studies in this topic area are identified by conducting a literature search in Scopus, which covers 19,500 peer-reviewed journals, 400 trade publications, 360 book series, and articles in press from more than 3,850 journals.[9] Studies that have not yet been published are identified by searching the Social Science Research Network, which contains abstracts on more than 464,100 scholarly working papers and forthcoming papers.[10] The search parameters for both searches include the following:

- The document contains one of the following phrases in the title or abstract:
  - "Retire*" and "save*" within two words (for example, "save for retirement")
  - "Retire*" and "saving*" within two words (for example, "retirement savings")
  - 401(k)
  - IRA
  - pension
- The document also contains one of the following phrases in the title or abstract:
  - "behavioral economic*"
  - "behavioral finance"
  - "behavioral intervention"
  - "behavioral science"
  - "psych*" and "economic*" within two words (for example, psychology and economics)
  - "neuroeconomic*"
  - "nudge"
  - "bounded" and "ration*" within two words (for example, "rationality is bounded")

In addition, studies are identified by searching the websites of the Behavioral Economics Working Group at the National Bureau of Economic Research, the Russell Sage Behavioral Economics and Consumer Finance Working Group, ideas42, the Employee Benefits Security Administration at the U.S. Department of Labor, the United Kingdom Behavioral Insights Team (Nudge Unit), and the Employee Benefit Research Institute.

Finally, CLEAR identified studies based on citations in reviews of the literature on behavioral interventions and retirement savings, such as Thaler (2009) and Kahneman (2013).[11]

[9] See "Content Overview." Available at http://www.info.sciverse.com/scopus/scopus-in-detail/facts.

[10] See "SSRN's Objective and Commitments to Users." Available at http://www.ssrn.com/.

[11] Thaler, Richard H. (2009). *Nudge: Improving Decisions About Health, Wealth, and Happiness.* New York: Penguin Books. Kahneman, Daniel. (2013). *Thinking, Fast and Slow.* New York: Farrar, Straus and Giroux for details.

APPENDIX B

REFERENCES

Studies with a high evidence rating

Choi, J., Laibson, D., & Madrian, B. (2011). $100 bills on the sidewalk: Suboptimal investment in 401(K) plans. *The Review of Economics and Statistics*, *93*(3), 748-763.

Related reports:
Choi, J.J., Laibson, D., & Madrian, B.C. (2005). $100 bills on the sidewalk: Suboptimal investment in 401(K) plans. National Bureau of Economic Research working paper No. 11554. Cambridge, MA: NBER.

Duflo, E., & Saez, E. (2003). The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment. *Quarterly Journal of Economics*, *118*(3), 815–842.

Goda, G., Manchester, C., & Sojourner, A. (2012). What will my account really be worth? An experiment on exponential growth bias and retirement saving. National Bureau of Economic Research working paper 17927. Cambridge, MA: NBER.

Related reports:
Goda, G.S., Manchester, C.F., & Sojourner, A. (2014). What will my account really be worth? Experimental evidence on how retirement income projections affect saving. *Journal of Public Economics*.

Hershfield, H., Goldstein, D., Sharpe, W., Fox, J., Yeykelis, L., Carstensen, L., & Bailenson, J. (2011). Increasing saving behavior through age-progressed renderings of the future self. *Journal of Marketing Research*, *48*(SPL), S23–S37.

Iyengar, S., & Kamenica, E. (2010). Choice proliferation, simplicity seeking, and asset allocation. *Journal of Public Economics*, *94*, 530-539.

McKenzie, C., & Liersch, M. (2011). Misunderstanding savings growth: Implications for retirement savings behavior. *Journal of Marketing Research*, *68*, S1–S13.

Studies with a moderate evidence rating

Benartzi, S., Peleg, E., & Thaler, R. (2007). Choice architecture and retirement savings plans. Los Angeles, Ca. SSRN working paper.

Morrin, M., Inman, J., Broniarczyk, S., Nenkov, G., & Reuter, J. (2012). Investing for retirement: The moderating effect of fund assortment size on the 1/n heuristic. Fox School of Business Research Paper No. 14-009, 1–38.

Studies with a low evidence rating

Beshears, J., Choi, J., Laibson, D., & Madrian, B. (2010). Simplification and saving. Cambridge, MA: National Bureau of Economic Research.

Related reports:
Beshears, J., Choi, J., Laibson, D., & Madrian, B. (2013). Simplification and Saving. *Journal of Economic Behavior & Organization, 95,* 130-145.

Brown, J., Lang, N., & Weisbenner, S. (2007). Individual account investment options and portfolio choice: Behavioral lessons from 401(K) plans. National Bureau of Economic Research working paper No. 13169. Cambridge, MA: NBER.

Related reports:
Brown, J.R., Lang, N., & Weisbenner, S. (2007). Individual account investment options and portfolio choice: Behavioral lessons from 401(K) plans. *Journal of Public Economics, 91*(10), 1992–2013.

Burman, L., Coe, N., Dworsky, N., & Gale, W. (2008). Effects of public policies on the disposition of pre-retirement lump-sum distributions: Rational and behavioral influences. CentER Discussion Paper Series No. 2008-94.

Carroll, G., Choi, J., Laibson, D., Madrian, B., & Metrick, A. (2009). Optimal defaults and active decisions. *The Quarterly Journal of Economics*, *124*(4).

Choi, J., Laibson, D., & Madrian, B. (2005). Are empowerment and education enough? Underdiversification in 401(k) plans. *Brookings Papers on Economic Activity, 2005*(2), 151-213.

Choi, J., Haisley, E., Kurkoski, J., & Massey, C. (2012). Small cues change savings choices. National Bureau of Economic Research working paper 17843. Cambridge, MA: NBER.

Choi, J., Laibson, D., & Madrian, B. (2009). Mental accounting in portfolio choice: Evidence from a flypaper effect. *American Economic Review*, *99*(5), 2085-2095.

Choi, J., Laibson, D., Madrian, B., & Metrick, A. (2004). For better or for worse default effects and 401(k) savings behavior. National Bureau of Economic Research, 81-126.

Choi, J., Laibson, D., Madrian, B., and Metrick, A. (2006). Saving for retirement on the path of least resistance. In Edward McCaffrey and Joel Slemrod (Eds.), *Behavioral Public Finance*. New York: Russell Sage.

Related reports:
Choi, James, Laibson, David, Madrian, Brigette, and Metrick, Andrew. (2006). Saving for retirement on the path of least resistance. In Edward McCaffrey and Joel Slemrod (Eds.), *Behavioral Public Finance.* New York: Russell Sage.

Iyengar, S., Huberman, G., & Jiang, W. (2003). How much choice is too much? Contributions to 401 (k) retirement plans. Pension Research Council working paper.

Related reports:

Iyengar, S., Jiang, W., and Huberman, G. (2004). How much choice is too much? Contributions to 401(k) retirement plans. In O. Mitchell & S.P. Utkus (Eds.) *Pension Design and Structure: Lessons from Behavioral Finance* (pp. 83–95). Oxford, UK: Oxford University Press.

Laibson, D., Choi, J., & Madrian, B. (2009). Reducing the complexity costs of 401(k) participation through Quick Enrollment. *Developments in the Economics of Aging*, 57-82.

Liang, N., & Weisbenner, S. (2002). Investor behavior and the purchase of company stock in 401(k) plans—the importance of plan design. National Bureau of Economic Research working paper 9131. Cambridge, MA: NBER.

Lusardi, A., Keller, P. A, & Keller, A. M. (2009). New ways to make people save: A social marketing approach. National Bureau of Economic Research working paper 14715. Cambridge, MA: NBER.

Madrian, B., & Shea, D. (2001). The power of suggestion: Inertia in 401(k) participation and savings behavior. *The Quarterly Journal of Economics*, *116*(4), 1149-1187.

Thaler, R., & Benartzi, S. (2004). Save More Tomorrow™: Using behavioral economics to increase employee saving. *Journal of Political Economy*, *112*(S1), S164-S187.

Thrift Savings Plan. (2012). Participant behavior and demographics: Analysis of 2008–2012.

Wenger J., & Weller, C. (2011). Boon or bane?: 401(k) loans and loan provisions. Available at http://ssrn.com/abstract=1941411.

Related reports:

Wenger, J.B., and Weller, C.E. (2014). Boon or Bane: 401(k) loans and employee contributions. *Research on Aging, 36*(5), 527–556.