2015

Investigating the role of objective and subjective task complexity on retirement investment plan earnings judgments

Natasha Hardy
Michigan Technological University, njhardy@mtu.edu

INVESTIGATING THE ROLE OF OBJECTIVE AND SUBJECTIVE TASK COMPLEXITY ON RETIREMENT INVESTMENT PLAN EARNINGS JUDGMENTS

By

Natasha J. Hardy

A THESIS

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

In Applied Cognitive Science and Human Factors

MICHIGAN TECHNOLOGICAL UNIVERSITY

2015

This thesis has been approved in partial fulfillment of the requirements for the Degree of MASTER OF SCIENCE in Applied Cognitive Science and Human Factors.

Department of Cognitive and Learning Sciences

Thesis Advisor: *Susan Amato-Henderson*

Committee Member: *Kelly Steelman*

Committee Member: *Michele H. Miller*

Department Chair: *Susan Amato-Henderson*

Dedication

To Alexander Hardy, my husband.

Table of Contents

List of Figures

List of Tables

Acknowledgements

Thank you to my committee, my advisors, and mentors for pulling, pushing, and prodding me to the finish line.

A special thanks to my ACSHF compadres Margo, Saima, Joel, and Patrick for your support, commiseration and advice.

Abstract

This research addressed the effect of retirement plan task complexity on retirement plan earnings estimates. Past research has shown that increased task complexity results in more decision-making errors as well increased use of heuristics, or rules of thumb, which can result in non-optimal outcomes such as under-saving or disproportionate equity/income balances (Benartzi & Thaler, 2001, 2007; Maynard & Hakel, 1997). This research used two experiments to test whether individuals would judge a retirement investment plan with high task complexity to be more profitable than a plan with low task complexity - a non-normative and potentially costly bias. Experiment 1 used retirement plans based on theoretical models while Experiment 2 used materials that were ecologically representative. In both studies participants judged a retirement plan with high task complexity to be more likely to return higher earnings than a retirement plan with low task complexity; this finding was unaffected by financial literacy and numeracy, which were expected to have a de-biasing effect. Subjective task complexity was found to be a significant predictor of earnings estimates, independent from estimates of plan risk and stability. These findings have practical and theoretical implications. Individual investors may be susceptible to the high task complexity of retirement investment plans which could lead to paying more fees. Benefits administrators can use this information to design and present retirement investment plan options in a way that potentially can mediate this bias for complexity.

Keywords: *Task complexity, Retirement plan decision-making, Heuristics.*

Introduction

Americans have seen a slow but steady shift from defined benefit retirement plans (e.g., pensions) to defined contribution retirement plans (e.g., 401(k)) in the last fifty years (Broadbent, Palumbo & Woodman, 2006). This translates into an increased need for individuals to gather, comprehend, and utilize complex financial information in order to make their own retirement decisions (Medill, 2000). Reported decreased confidence in ability to retire among retirement-age workers (Yakoboski & Dickemper, 1997) and inadequate retirement savings (Lusardi, 2000, 2001; Munnell & Golub-Sass, 2007) suggest that Americans are ill-prepared to handle these decisions.

One possible reason for Americans' decreased confidence and low savings rates is that the task of retirement investment decision-making is too complex for the average person. Standard economic theories of saving treat individuals as rational decision-makers who will stabilize their style of living throughout their lifetime by first saving and then withdrawing those savings upon retirement in order to maintain their standard of living through their lifetime (Benartzi & Thaler, 2007). However, researchers in decision-making acknowledge that individuals make decisions in the context of bounded rationality; i.e., individuals are constrained by time, the amount of information available, and their own cognitive abilities (Simon, 1987). The more complex the task, the more likely individuals are to use simplified decision strategies (i.e., heuristics) to make decisions (Swait & Adamowicz, 2001), and the more likely they are to make sub-optimal decisions (Benartzi & Thaler, 2007; De Palma, Myers

& Papageorgiou, 1994). This research seeks to understand how the level of task complexity in retirement investment plans affects subjective judgments of earnings; namely, whether consumers judge an investment plan with high task complexity to be capable of higher earnings than an investment plan with low task complexity. Understanding the relationship between task complexity and individual judgments of earnings can potentially add to the research on heuristics and biases as well as suggest better retirement investment plan designs.

Previous research has shown that the use of heuristics affects decisions to initiate retirement investments, how money is invested, and the amount of money invested in retirement investment plans (Benartzi & Thaler, 2007). For example, employees often use simplified strategies to determine how much of their pay to contribute to their retirement savings account; Benartzi and Thaler (2007) found that contributions in multiples of 5 (e.g., 5%, 10%, and 15%) were used more often than expected by chance. Other heuristic contribution strategies included selecting the minimum or maximum amount allowed by the plan design or contributing the same amount as the employer match. While some of these heuristic-based strategies result in adequate saving, many of them contribute to a tendency to under-save.

Another common heuristic that is used when selecting an investment portfolio is the 1/N or naïve diversification rule. 1/N is a simple strategy for diversifying one's portfolio by splitting up contributions equally among the investment options (Benartzi

& Thaler, 2001). Formulaically, DeMiguel, Garlappi, and Uppal (2007) represented this by:

$$W_t^{ew} = \frac{1}{N}$$

where W_t^{ew} is the portfolio weight give the number of assets.

Huberman and Jiang (2006) reported that this strategy was utilized by approximately 50% of investors. For example, TIAA-CREF once offered just two investments options: stocks and bonds. The majority of plan participants allocated their investment dollars equally between these two investments (Samuelson and Zeckhauser, 1988). Famed economist and Nobel laureate, Harry Markowitz, reported that he also divided his investments equally between equities and bonds rather than adopting the more sophisticated investment strategies espoused by his own research (Zweig, 1998). Although research has shown that the 1/*N* strategy can outperform more complex optimal portfolio models such as the mean-variance model posited by Markowitz (1952) and Bayesian diffuse-prior portfolio (Barry, 1974), use of the 1/*N* heuristic is likely to lead to unbalanced portfolios by over-investing in company stocks or through inappropriate equity/income balances (Benartzi & Thaler, 2007).

Another heuristic decision strategy individuals rely on is the status quo or default heuristic. The default heuristic is used when individuals passively accept the default option rather than making an active choice (Mitchell & Utkus, 2003). In a

study of the Swedish Premium Pension Scheme, which consists of national,

mandatory and voluntary pension savings, the authors found that over 55% of respondents opted for the default plan (Hedesström, Svedsäter, & Gärling, 2007). Moreover, additional research has shown that once individuals select a plan, they are unlikely to make any changes to it when given the choice (Fry, Heaney, & McKeown, 2007). This strategy is likely to lead to unbalanced portfolios that are either too risky or, more likely, too conservative since plan defaults tend to err on the side of caution (Choi, Laibson, Madrian, & Metrick, 2005).

Trend chasing, another heuristic behavior, involves buying stocks when they are on the rise and selling when they drop (Andreassen & Kraus, 1990). Trend chasing is typically attributed to use of the representativeness heuristic. This heuristic posits that individuals will attempt to categorize a problem based on how similar it is to a reference (i.e., representative) category (Ayton & Fischer, 2004; Hastie & Dawes, 2010). When using this heuristic, individuals create extensive generalizations regarding an event using very little information about its actual features (Kahneman & Tverskey, 1972). For financial decision-making, individuals may seek to "match" systematic patterns in investments that do not actually exist (Hirshleifer, 2001). These efforts to exert control over investments can result in sub-optimal outcomes when individuals over-react to the cyclical changes that naturally occur with investments over time, such as selling underperforming stocks at a loss or purchasing 'hot' (e.g., trending upward) stocks for a premium.

All of these heuristic decision-making patterns are consistent with individual performance on tasks with high complexity. That is, to a point, increased complexity increases the amount of effort an individual invests in performing a task. However, if complexity increases too much, effort will actually decrease and individuals will revert to heuristic-based strategies resulting in an inverted U-shaped relationship between effort and complexity (Schroder, Driver, & Streufert, 1967). The point where effort will begin to decrease is a function of the individual's level of ability (Wood, 1986) as well as the domain. To gain a better understanding of when heuristics work, Todd and Gigerenzer (2012) proposed a basic taxonomy to determine the conditions under which a heuristic will succeed or fail. This taxonomy included the following environmental cues: uncertainty (e.g., how accurately an outcome can be predicted), correlation between environmental cues, number of observations, and the variability of the weight assigned to each cue. For example, a one-reason heuristic such as the hiatus heuristic, which is often used to predict time between customer purchasing activity, involves the cues of uncertainty and redundancy. Uncertainty refers to not being able to predict future purchases whereas redundancy refers to previous purchase behavior. If this taxonomy is applied to retirement investment plans, there would be a high level of uncertainty (e.g., not being able to predict future income needs), few observations (e.g., low familiarity with the task), and increased variability of cue weights (e.g., how much importance is attached to different elements of a retirement plan). According to Todd and Gigerenzer's taxonomy, this would translate into higher variability in decision-making. This variability is associated with an

increased likelihood of decision failure - a non-optimal outcome. Thus, decision-makers in the domain of retirement investments are not only likely to use heuristics due to high task complexity, but they are also likely to have an increased risk of using heuristics that fail.

In order to clarify how task complexity affects decision outcomes – both normative (e.g., optimal) and non-normative – the following section reviews how task complexity has been manipulated and operationalized in the relevant domains of information processing, personality, and organizational development.

Research in the domain of information processing has manipulated task complexity by increasing overall information load and information diversity in differing choice sets (Campbell, 1988). In order to operationalize this relationship, Campbell (1988) proposed an ordered taxonomy with sixteen degrees of task complexity that vary on four factors: multiple pathways to a desired goal, multiple acceptable end-states, conflicting interdependence, and uncertainty. Complexity was then objectively determined by the presence or absence of attributes, as well as the total number of attributes. Swait and Adamowicz (2001) formalized this typology as:

$$\text{Equation 1: } \boldsymbol{H(X) = (H(\pi_x) = -\sum^{J}_{j=1} \pi(x_j) log \pi(x_j) \geq 0}$$

where J is the number of alternatives and $\pi(x)$ is the probability distribution. Using this equation, they found that the number of alternatives within a choice set directly affects the level of complexity. This finding is empirically corroborated by DeShazo

and Fermo (2002) in the domain of organizational development. The authors used a large data set (N = 3900) to determine the effects of varying levels of task complexity on choice consistency. They varied the number of alternatives (between 2 and 9) and the number of attributes within each alternative (between 4 and 9) for an economic valuation task (i.e., assigning a value to a national park) and found that the number of errors - defined as sub-optimal preference ordering – increased when the number of alternatives and attributes increased. Maynard and Hakel (1997) used a similar formulae with an employee scheduling task to determine how high and low task complexity affected performance, which was operationalized as variance from maximum estimated profit. Seventy-six percent of participants in the simple condition were able to achieve maximum performance while only 20% of participants in the complex condition achieved maximum performance. Further, they found that subjective complexity, defined as participants' self-perceptions of the difficulty of the task, was a unique predictor of performance.

Subjective complexity for a task can stem from several individual factors such as the level of importance placed on the task and how much the outcome of the task depends on the individual's efforts (Hackman & Oldham, 1975). In a study of performance in highly complex scenarios in a nuclear control room, subjective task complexity was measured with a 7-point scale on eight dimensions: root-cause difficulties, spread of information, ambiguous information, coordination, guidance information, attention demand, severity for plant safety, and temporal demand (Braarud, 1998). The study showed that subjective task complexity, as measured by

the NASA-TLX, predicted a portion of task performance. The NASA-TLK consists of six subscales of subjective workload including mental demands, physical demands, own performance, effort, and frustration. While the criterion for this research was plant safety, this study is relevant to the current research as it supplies some confirming evidence that subjective task complexity can negatively affect performance. A study in a more relevant domain, personality and individual differences, showed that subjective task complexity was related to the perceived difficulty of completing a task (Koren, and Zakay, 2001). The decision task for this study was selecting a college major and was operationalized as the number of alternatives taken into account when making the decision and the number of features each alternative contained. A participant's overall complexity index (e.g., subjective task complexity) was calculated by multiplying the number of alternatives by the number of features each participant listed. This measure correlated moderately with perceived difficulty (r = .32, p< 0.01) and highly with the number of alternatives (r = .79, p< 0.01) and the number of features (r = .56, p< 0.01). Subjective task complexity increased based on the number of features participants listed for each alternative, but an increase in the number of alternatives decreased the number of features participants were willing to consider. This is consistent with the inverted U-shaped relationship between task complexity and effort (e.g., individuals will expend effort as task complexity increases but switch to simpler strategies when task complexity increases too much; Keller and Staelin, 1987). The authors suggested that this is because subjective judgments of task complexity stem from an individual's

development of their own mental representation of the problem. Because individuals create their own mental representations, their subjective perceptions of task complexity are expected to be different as well. In addition, Keller and Staelin (1987) found that subjective task complexity moderated the relationship between task complexity and cognitive ability – individuals with lower cognitive ability were more likely to perceive a task to be complex and their performance on the task was decreased. Due to the effects of both subjective complexity and cognitive ability on task performance, both of these constructs are measured in the current research. Rather than using a broad measure of cognitive ability, measures of financial literacy and numeracy are used as they are purported to aid individuals in financial decision-making (Van Rooij et al., 2011).

Financial literacy is the knowledge of financial and economic concepts and numeracy is the capacity to understand and apply probabilities and numerical information (Hung, Parker, & Yoong, 2009). These concepts are thought to help individuals cope with complex financial products and services (Van Rooij, Lusardi, & Alessie, 2007). Individuals with higher levels of financial literacy are more likely to engage in saving behaviors, a primary determinant of retirement wealth (Lusardi & Mitchell, 2005). Alternatively, individuals with low levels of financial literacy are less likely to plan for retirement, less likely to buy stocks (van Rooij, Lusardi, & Alessie, 2007), and less likely to choose mutual funds with low fees (Hastings & Tejeda-Ashton, 2008). Numeracy does not measure the same concepts as financial literacy although the two are related. Numeracy is a more broad and basic measure of

general cognition or of specific numerical skills and metacognitive behaviors (Cokely et al., 2012; Hung, Parker, & Yoong, 2009).

Conclusion of Literature Review

A review of the literature shows that task complexity can be manipulated by increasing the cognitive burden on an individual through the number of alternatives, and number of attributes within each alternative. Cognitive burden is also increased by reading comprehension level (Pollock, Chandler, & Sweller, 2002). Subjective task complexity is a measure of how task complexity is perceived by the individual in a psychological sense and has been found to correlate modestly with task complexity and to have a measurable, negative effect on performance. Task complexity increases individual's reliance on heuristic decision-making strategies such as dividing their money equally among investments or relying on administrator-set defaults. Use of heuristics can result in both normative and non-normative biases (Gigerenzer & Gaissmaier, 2011; Keller & Staelin, 1987; Marewski, Gaissmaier, & Gigerenzer, 2010) and, in the domain of investing, seem to result in more non-normative biases (Benartzi & Thaler, 2001; 2007).

Overview of Current Research

The current research investigates the effects of task complexity on retirement investment decisions by examining subjective judgments of retirement investment plan earnings between low task complexity and high task complexity retirement investment plans. By doing so, the effects of task complexity on retirement

investment planning may be clarified, which could be beneficial for the development of a decision aid for retirement planning.

For example, Benartzi and Thaler (2007) have designed a decision aid for retirement plan benefits administrators called the Save More Tomorrow™ plan which takes into account decision-makers tendency to use heuristics that often result in disproportionate allocations between stocks and bonds. The Save More Tomorrow™ plan uses a combination of automatic adjustments and expert advice to create an optimum plan for each participant. If increased task complexity in retirement plans is found to affect individual judgments of plan earnings, the results could be used to advise administrators to create more balanced portfolio options that do not create a mental trade-off between complexity and earnings.

In order to investigate the effects of task complexity on retirement investment decisions, the current research manipulated task complexity (e.g., high v. low) and collected participants' subjective judgments regarding investment plan profitability (i.e., earnings) and valuation (i.e., how much one is willing to pay for plan administration) while controlling for subjective riskiness and volatility. In addition, financial literacy and numeracy are measured as a proxy for cognitive abilities in order to examine the effect of these abilities on judgments of task complexity and plan profitability.

The following hypotheses delineate the goals of Experiment 1:

*Hypothesis 1***.** Overall, it was hypothesized that the participants would rate the subjective task complexity of the high task complexity plan higher than the low task complexity plan.

*Hypothesis 2***.** Participants would judge the high task complexity plan to be capable of higher earnings than the low task complexity plan. This difference would be attributed to subjective complexity, not other potentially relevant predictors suck as not as risk or plan cost.

*Hypothesis 3***.** An interaction between financial literacy/numeracy and plan complexity was expected such that individuals who score higher on financial literacy and numeracy tests will be less likely to judge the high task complexity plan to be capable of higher earnings than the low task complexity plan.

The following experiment tested these hypotheses by using two financial management plans - one with a high level of task complexity and one with a low level of task complexity – as well as several measures of subjective complexity. Financial literacy and numeracy were measured to examine potential interaction effects of numeric abilities, earnings, and task complexity.

Experiment 1

Experiment 1 examined the effect of task complexity on subjective judgments of retirement investment plan earnings using a single factor design, manipulating order between subjects (i.e., low task complexity first, high task complexity first)

with several dependent variables including: subjective task complexity, retirement plan earnings, volatility, and estimated losses, and willingness-to-pay for plan management. Individual differences in financial literacy and numeracy were measured in order to examine the potential moderating effects of financial and numerical knowledge between subjective task complexity and expected earnings.

Method

Participants

Two hundred and fifty-six individuals were recruited via Amazon Mechanical Turks to complete the online study. Two hundred and forty-two participants (Mean age = 33.21, *SD* = 11.86) completed the study and were included in the analysis. Of these, 50% were men, 48% were women, and 2% chose not to respond. The majority of participants had at least some college or trade school and earned between $20,000 and $70,000 per year. The decision to use Mechanical Turk participants was motivated in part by the subject matter, which was directed more toward older adults rather than the typical undergraduate participant often found in commonly available subject pools. In addition, it has been found that Mechanical Turk participants exhibit the same cognitive biases as conventional participants (Amir, Rand, & Gal, 2012; Goodman et al., 2013; Horton, Rand, & Zeckhauser, 2011) which are relevant to this study.

Materials

The materials included the following: (1) two retirement portfolio management plans, one with a low level of task complexity and one with a high level of task complexity; (2) a set of two questions measured subjective judgments of plan earnings, which was the key dependent variable; (3) eleven questions measured subjective task complexity, riskiness, and willingness-to-pay; (4) 3 questions measured financial literacy; (5) an 8 question surprise memory test of plan components served as a manipulation check for task complexity; (6) 7 questions from two numeracy tests measured participant numeracy and lastly, (7) demographic questions. Each of these is described in detail below.

Portfolio management plans

The high task complexity portfolio management plan was titled the Quantitative Portfolio Management Plan (hereafter referred to as high task complexity or HTC plan). The theoretical foundation for the task complexity of the HTC plan is the Markowitz Portfolio Theory, which mathematically derives the optimal level of return for a minimal amount of risk by carefully selecting the component assets. The low task complexity plan was titled the Regular Portfolio Management Plan (hereafter referred to as the low task complexity or LTC plan) was modeled after the *1/N* rule, also referred to as naïve diversification (Benartzi & Thaler, 2001) which uses a heuristic strategy to select portfolio components. Each plan was equally matched for number of rules and words and can be seen in Figure 1.

Regular Portfolio Management Plan	Quantitative Portfolio Management Plan
Rule # 1: Select twenty different stocks that have about the same amount of risk.	Rule # 1: Select a range of investment instruments that are not perfectly correlated.
Rule # 2: Divide your investment dollars equally among each of the twenty stocks.	Rule # 2: Determine risk to return ratio using the Capital Asset Pricing Model.
Rule # 3: Keep portfolio management fees low by purchasing and holding stocks.	Rule # 3: Determine risk tolerance by calculating the standard deviation of return.
Rule # 4: Check stocks once every year, when you file your taxes.	Rule # 4: Assess portfolio volatility by testing the correlations of component assets.

Figure 1: Display of the Regular Portfolio Management Plan and the Quantitative Portfolio Management Plan.

The complexity of each plan was also manipulated using reading level; the Flesch Reading Ease Score for the HTC plan was 23.3 which corresponded to a rating of very difficult to read. The score for the LTC plan was 64.1 which corresponded to a rating of standard/average difficulty. This approach is consistent with the information processing theories of task complexity: increasing the reading comprehension level should increase the cognitive burden and, subsequently, task complexity (Pollock, Chandler, & Sweller, 2002).

Questions regarding plan earnings

There were two questions that measured earnings judgments: 1) if you invest $1000.00, how much will each plan earn in one year and, 2) which plan would you choose if you wanted to earn $500 quickly?

Questions to elicit subjective judgments

Eleven questions were designed to measure subjective task complexity, riskiness, and willingness-to-pay for the plan. There are few existing measures of subjective task complexity and none were pertinent to the domain of retirement investments, so these items were developed for the purpose of this study. The questions were intended to uncover how the mentally demanding and complex the plans were for each participant. In total, seven of the questions were designed to measure subjective task complexity. Three questions were intended to measure individual judgments regarding the financial stability of each plan (e.g., How risky do you think each plan is; If the stock market were to drop by 50% how much would each plan lose; and How stable is each plan). Lastly, one question measured valuation of each plan (e.g., how much one is willing to pay to have someone administer each plan). Wording for all questions can be seen in Table 1 and a complete listing of all questions and their corresponding scales can be found in Appendix A.

Table 1: Questions regarding subjective plan judgments.

Questions regarding subjective plan judgments	
Earnings	In an average year, if you invest $1000.00 how much do you think each plan will earn?
Subjective Complexity	How complicated do you think each plan is?
Subjective Complexity	How hard to you think it would be to find any errors made by the financial planner for each plan?
Subjective Complexity	How difficult would it be to report each plan on your taxes?
Subjective Complexity	How difficult would it be for you to explain this plan to your partner or spouse?
Subjective Complexity	How much anxiety does the thought of selecting and starting each of these investment plans give you?
Subjective Complexity	How much time do you think it will take to set up each of these plans?
Riskiness	How stable do you think each of these plans will be over time?
Riskiness	If the stock market were to drop by 50% how much would each plan lose?
Riskiness	How risky do you think each plan is?
Earnings	Which plan would you choose if you wanted to earn $500 quickly?
Willing-to-Pay	How much would you be willing to pay per year to have someone administer each of these plans for you?
Subjective Complexity	How hard do you think each plan is to remember?

Financial Literacy Test

A 3-item financial literacy test was developed by Lusardi and Mitchell (2005) for use in the 2005 DNB Household Survey. This tool has been used extensively by Lusardi and Mitchell (2007a, 2007b, 2007c, 2008, 2009, & 2011) and by van Rooij, Lusardi and Alessie (2007). This test was selected because of its abbreviated format

and large body of supporting research. The complete financial literacy test is included in Appendix B.

Memory Test

Previous research in human factors has shown that individuals will display better memory for simple information than complex information (Wichman & Oyasato, 1983) therefore the memory test was used as a manipulation check to examine the level of complexity of the HTC plan components versus the LTC components. The surprise memory test consisted of four questions regarding the HTC plan and four questions regarding the LTC plan - one question for each rule. For each plan there were 2 multiple choice, one True/False and one fill in the blank question, in order to provide a wider-range of difficulty (e.g., recognition is easier than recall). For example, one question was:

"Which of the following is a rule in the Quantitative Portfolio Management Plan?

A) Determine the return on investment using the Capital Asset Pricing Model.

B) Determine the risk to return ratio using the Capital Asset Pricing Model.

C) Assess the risk to return ratio using the Principle Asset Costing Model and

D) Assess the return on investment using the Principle Asset Costing Model."

The memory questions were in a blocked format with all of the questions regarding the HTC plan grouped together and all the questions for the LTC plan

grouped together. Participants saw the HTC plan memory questions first if they saw the HTC plan first and vice versa for the LTC plan. The complete memory test can be found in Appendix D.

Numeracy

The adaptive Berlin Numeracy Test (BNT) was used to assess numeracy, in combination with the three item numeracy test by Schwartz, Woloshin, Black, and Welch (1997) as prescribed in Cokely, Galasic, Schulz, Ghazal and Garcia-Retamaro, (2012). This test has a maximum of four questions and has been shown to be the strongest single predictor of one's understanding of everyday risks (e.g., risk literacy), after controlling for a wide range of other individual differences in personality, values, cognitive styles, and cognitive abilities (Cokely et al., 2012). The Chronbach alpha for the adaptive BNT could not be calculated; however, the authors conducted a principal component analysis which showed that all four items loaded highly on a single factor and explained approximately 45% of the variance. The test-retest reliability for the BNT is $r = .91$. The complete numeracy measures can be seen in Appendix C.

Lastly, participants answered demographic questions regarding age, sex, income, and education level.

Procedure

Experiment 1 was administered via Unipark online experiment platform and took an average of 11.5 minutes to complete. Participants began by reading a short informational paragraph regarding their role as participants in the experiment. Following this they completed the university-required informed consent form and clicked to indicate that they were over 18 years of age and willing to participate in the experiment. The complete informed consent form, as approved by MTU IRB M0650, can be seen in Appendix E.

Participants were then shown the two retirement investment plans side by side. The HTC plan and LTC plan were experimentally counterbalanced so that approximately half the participants saw the HTC plan on the left and half saw the LTC plan on the left. Participants were randomly assigned to each group using a Unipark algorithm.

After viewing the plans, participants were required to check a box for each plan indicating that they had read each one completely. Following this, thirteen questions regarding subjective judgments of each plan were presented in a series. Twelve of the thirteen questions were in the form of Likert-style scale responses either from 1 to 5, 1 to 7, or 1 to 9 while one of the questions was a binary choice.

Following the portfolio management questions, participants completed a three question financial literacy test and a computerized adaptive numeracy test as well as a 3-item numeracy measure.

Lastly, participants completed demographic questions including income, education, age, and sex. Participants were then debriefed, thanked and given an alphanumeric code to enter into Amazon Mechanical Turks to receive their payment for participating, which was approximately $0.50 USD.

Data Coding

Data was withdrawn from the Unipark online experiment platform into Microsoft Excel to be recoded using Visual Basic for Excel and then exported to PASW 18 SPSS for data analysis procedures. The plan judgments were all scale judgments, with the exception of one question, which was binary response. Therefore, the data for the HTC and LTC plan judgments were in an appropriate format for data analysis. Data for the one binary response question was re-coded where 1 = HTC plan and 2 = LTC plan.

The financial literacy scale was first coded right/wrong (e.g., 1 = right and 2 = wrong) for each item and then aggregated into a percent correct response. The same procedure was used for the memory test.

The Berlin Numeracy Test was coded based upon the number of questions answered correctly (out of four) and combined with the score on the Schwartz numeracy items (out of three).

Results

A Pearson r bivariate correlation matrix was calculated between participants' judged memory and their performance on the surprise memory test. There was a significant correlation between performance on the memory test and participants' estimated ability to remember all of the investment plan components $r(238) = .131$, $p < .04$. Individuals who judged their memory for plan components to be low were more likely to perform poorly on the memory test.

A confirmatory factor analysis was conducted using Principle Component Analysis with Varimax to test the assumption that there are two main factors that exist within this pool of questions: complexity factors and financial factors (see Table 2). Because of the way the complexity factors were selected we would expect a juxtaposition of those judgments against judgments regarding earnings. Factor analysis revealed two components: component 1 consisted of how difficult to find errors, how difficult to report on taxes, how difficult to explain to spouse or partner, how much anxiety, how much time to set up, how much one would pay for each plan, how difficult each plan is to remember, and how complicated is each plan. Component 2 consisted of how stable over time, how much will each plan earn, how much each plan will lose, and how risky is each plan.

Table 2: Key factor loadings of component analysis.

Component Matrix		
	Component	
	1	2
How hard to find errors	**.504**	.116
How difficult to report	**.686**	.057
How difficult to explain	**.768**	.209
How much anxiety	**.816**	-.093
How much time to plan	**.712**	.186
How stable over time	**.647**	-.421
How much plan will earn	.141	**.595**
How much plan will lose	.401	**-.690**
How risky is each plan	.699	**-.452**
Pay for administration	.367	**.511**
How hard to remember	**.778**	.172
How complicated	**.849**	.171

The results of this analysis confirm the existence of a component which could be termed "complexity" and a second component which could be termed "financial considerations" (e.g., earnings and losses).

Hypothesis 1 expected that participants would rate the subjective task complexity of the HTC plan higher in comparison to the LTC plan. Paired sample *t*-tests were conducted to examine participant judgments of the seven elements of subjective complexity. For all seven items, participants judged the HTC plan to have significantly higher subjective task complexity than the LTC plan, indicating that the task complexity manipulation was successful. A delineation of the mean rating and

standard deviation for each of the subjective complexity components can be seen in Table 3.

Table 3: Summary of paired sample t-tests with significant results flagged *p<.001.

Comparison of HTC/LTC plans	**Scale**	**HTC *M(SD)***	**LTC *M(SD)***	***t*-value**	***p*-value**
How hard to you think it would be to find any errors made by the financial planner for each plan?	1= not hard at all and 7 = very hard	5.07(1.47)	3.38(1.56)	10.65	<.001*
How difficult would it be to report each plan on your taxes?	1= not difficult at all and 7 = very difficult	4.87(1.62)	2.71(1.48)	15.1	<.001*
How difficult would it be for you to explain this plan to your partner or spouse?	1= very easy and 7 = very hard	5.55(1.42)	2.68(1.62)	20.06	<.001*
How much anxiety does the thought of selecting and starting each of these investment plans give you?	1= no anxiety at all and 7 = a lot of anxiety	5.14(1.58)	3.24(1.63)	13.39	<.001*
How much time do you think it will take to set up each of these plans?	1= less than one hour and 5 = 6 or more hours	3.74(1.06)	2.43(1.04)	14.92	<.001*
How complicated do you think each plan is?	1= not complicated and 7 = very complicated	5.71(1.27)	2.85(1.59)	21.57	<.001*

Independent sample t-tests were used to check for differences between female and male rankings for the twelve questions that included scaled responses. Results indicated that for seven variables there were significant differences (see Table 4).

Table 4. Difference in female and male subjective complexity and riskiness

Variable	*M Female*	*SD Female*	*M Male*	*SD Male*	*t*(237)	*p*
Anxiety Complex	5.39	1.57	4.94	1.54	2.22	0.028
Stability Complex	4.56	1.45	4.03	1.42	2.82	0.005
Losses Simple	5.98	2.52	6.15	2.42	2.85	0.005
Riskiness Complex	5.30	1.40	4.58	1.53	3.81	<.001
Administration Complex	3.55	1.79	2.97	1.64	2.59	0.010
Remember Complex	5.73	1.36	5.33	1.33	2.28	0.024
Complicated Complex	2.93	1.27	2.75	1.24	2.75	0.006

Females overall ranked the complex plan as more likely to cause anxiety, more stable over time, more risky, more difficult to remember, more complicated and were willing to pay more to have someone administer the complex plan compared to male responses. They also rated the simple plan as less likely to lose money if the stock market dropped as compare to male responses.

For Hypothesis 2, the expectation was that subjective task complexity would affect judgments regarding plan earnings such that the HTC plan would be predicted to earn more than the LTC plan, regardless of subjective riskiness or willingness-to-pay. First, paired sample t-tests were used to determine whether participants judged that the HTC plan was capable of higher earnings than the LTC plan and whether participants estimated that they would pay more for the HTC plan than the LTC plan. In addition, a chi square test assessed the binary response question, “Which plan

would you choose if you wanted to earn $500.00 quickly?" Results of these tests showed that participants rated the HTC plan as capable of higher earnings and also estimated that they would be willing to pay more for the HTC plan. There was no significant difference between judgments for which plan would earn money more quickly. Table 5 summarizes the results for these three questions.

Table 5: Participant responses to questions regarding earnings/willingness-to-pay.

Comparison of HTC/LTC plans	**Scale**	**HTC *M(SD)***	**LTC *M(SD)***	***t*-value**	***p*-value**
Which plan would you choose if you wanted to earn $500 quickly?	Binary Choice				0.75
How much would you be willing to pay per year to have someone administer each of these plans for you?	1= up to $25 and 7 = over $300	3.26(1.74)	2.44(1.47)	8.89	<.001*
In an average year, if you invest $1000.00 how much do you think each plan will earn?	1= $0-$24 and 9 = $200 or more	5.53(2.01)	4.57(1.90)	7.22	<.001*

Next, a hierarchical multiple regression analysis was used to determine if participant's earnings estimates were a result of their subjective judgments of plan task complexity rather than volatility (e.g., taking on additional risk is expected to result in higher earnings) was tested. Riskiness, stability and expected losses were the predictor variables for earnings estimates in Model 1, and the amount participants were willing to pay for administration (e.g., plan cost) was added as another predictor in Model 2. Model 3 incorporated plan complexity as a predictor, which for the sake of parsimony was reduced to the question, "How complicated do you think each plan is?" Tests for multicollinearity denote a low level of multicollinearity (*VIF* = 1.607

for expected losses, 2.504 for riskiness, 1.880 for stability, 1.127 for plan cost, 1.375 for complicatedness).

Results for the regression analysis provide a measure of support for Hypothesis 2; earnings judgments are not affected by judgments of riskiness, stability or expected losses. The best fitting model for predicting earnings estimates was model 2 (R^2= .198, F(5, 237) =14.735, p < .001). A summary of the regression models is presented in Table 6.

Table 6: Hierarchical regression analysis for variables predicting earnings estimates (N= 242). *p < .05

		Model 1			Model 2	
Variable	β	***SE β***	***β***	β	***SE β***	***β***
Riskiness	0.205	0.083	0.235*	0.121	0.08	0.138
Stability	-0.086	0.077	-0.095	-0.117	0.071	-0.129
Losses	-0.168	0.063	-0.214*	-0.104	0.058	-0.133
Plan cost				0.598	0.087	0.413*
Complex						
R^2		0.038			0.186	
F value		3.129			47.71	

		Model 3	
Variable	β	***SE β***	***β***
Riskiness	0.1	0.081	0.114
Stability	-0.13	0.072	-0.143
Losses	-0.097	0.058	-0.124
Plan cost	0.582	0.088	0.402*
Complex	0.065	0.07	0.064
R^2		0.185	
F value		0.851	

Because of the significant role willingness-to-pay (i.e. plan cost in Table 6) played in predicting earnings judgments, a hierarchical multiple linear regression was constructed to determine if participant estimates regarding willingness-to-pay varied as a function of volatility (e.g., riskiness, expected losses, stability) and/or task complexity. Measures of volatility – riskiness, expected losses, and stability – were included in Model 1. Plan complicatedness was added to Model 2. The best fitting

model for predicting willingness-to-pay was a linear combination of expected losses and plan complicatedness (R^2= .078, F(1, 237) = 6.098, p< .001). A summary of the models can be seen in Table 7. This provides additional support for Hypothesis 2 which expected that participants would judge the HTC plan to be capable of higher earnings than the LTC plan because of subjective complexity and not because of judged riskiness or willingness-to-pay. While willingness-to-pay estimates accounted for nearly 20% of the variance in earnings judgments, willingness-to-pay was itself partially predicted by subjective task complexity.

Table 7: Hierarchical regression analysis for variables predicting willingness-to-pay (N = 241) *p < .05.

	Model 1			Model 2		
Variable	β	*SE β*	*β*	β	*SE β*	*β*
Riskiness	0.140	0.057	0.233	0.083	0.059	0.138
Stability	0.051	0.053	0.082*	0.018	0.053	0.029
Losses	-0.107	0.043	-0.196*	-0.086	0.043	-0.158*
Complicatedness				0.158	0.051	0.226*
R2		0.056			0.093	
F value (Change in R2)		4.708			9.748	

Finally, Hypothesis 3 examined whether individuals with higher financial literacy and higher numeracy were less likely to show a bias for the HTC plan. In order to examine the relationship between financial literacy, numeracy and judgments of plan complicatedness and plan earnings, a bivariate correlation matrix was first calculated using Pearson *r*. Financial literacy was significantly positively correlated with complicatedness judgments as well as significantly positively correlated with earnings judgments. Numeracy was not significantly correlated with complicatedness judgments or with earnings judgments (see Table 8).

Table 8: Pearson r correlations between financial literacy, numeracy, subjective complicatedness, and subjective earnings complex & LTC plans. N= 238, * p < .05.

	Financial Literacy	BNT & Schwartz	Earnings Score	Complicatedness Score
BNT & Schwartz	.14*			
Earnings Score	.15*	.12		
Complicatedness Score	.20*	.12	.14*	

To determine the nature of any potential interaction between complexity, earnings, and financial literacy and between complexity, earnings, and numeracy, separate moderation analysis for numeracy and financial literacy were performed using simple linear regression where main variables and interaction variables (Numeracy*complexity and Financial literacy * Complexity) were entered. Results indicated no significant interaction effect, neither for numeracy and complexity (R^2 = .034, β = .144, p = .43) nor for financial literacy and complexity (R^2 = .04, β = .26, p = .18). It was expected that individuals with higher levels of numeracy and higher levels of financial literacy would be more likely to rate the LTC plan as more efficacious (e.g., more likely to result in higher earnings). These results do not support this hypothesis.

Experiment 1 Discussion

Results indicated that subjective task complexity may have an effect on retirement investment decisions in that subjective task complexity judgments explain some variance of willingness-to-pay which, in turn, explains a portion of the variance in earnings judgments. The fact that there are significant differences in judgments between earnings capabilities for the LTC and HTC plan has practical implications: the Cohen's effect size value shown in Table 9 (d =0.49) indicate that participants judged the HTC plan to be capable of earning approximately 13% more per year than the LTC plan on an investment of $1000.00.

Table 9: Effect sizes for difference between *c*omplex and LTC plan judgments.

Item title	Cohen's *d*
Hard to find errors	1.09
Difficult to report	1.37
Difficult to explain	1.89
Likely to cause anxiety	1.18
Amount of time	1.24
Stability over time	0.57
Result in earnings	0.49
Result in losses	0.11
Level of risk	0.76
Pay to administer	0.51
Difficult to remember	1.94
How complicated	1.99

We can also see that participants were willing to pay approximately $20.00 per year more to have someone administer the HTC plan than the LTC plan - a difference of nearly 12% $t(241) = 8.887$, $p < .001$. This corroborates consumer

research which has documented a positive association between quality and price (i.e., a higher cost items is thought to be a higher quality item) (Monroe & Chapman, 1987).

Recall that Hypothesis 1 expected to see a bias toward the HTC plan in that participants would rate the HTC plan as able to earn more money than a low task complexity investment plan, given an equal investment of dollars. This hypothesis was supported: judgments of plan earnings were significantly higher for the HTC plan than the LTC plan; earnings estimates for the HTC plan amounted to over $10.00 more than the LTC plan. Assuming a 30 year investment horizon and a modest interest rate of 7%, this would result in an additional $11,411 for the HTC plan versus the LTC plan. However, there was no difference between judgments regarding which plan participants would select in order to earn $500 quickly. It is possible that the use of the word 'quickly' created some ambivalence about investment strategy: over the long term, a complex investment strategy may result in higher earnings but for short term earnings, participants gambled that either plan was likely to be successful. It is also possible that the response format was responsible for these judgments - recall that this was the only question formatted as a binary response. It is possible that using scale values for this question would have increased the sensitivity of this question and given more insight into these judgments.

In addition, there was a significant difference between men and women in subjective task complexity and riskiness judgments. Related research has documented

significant differences between men and women with regards to retirement investment decisions. Women have been shown to be more reticent regarding risk (Sunden & Surette, 1998) and more likely to jeopardize their retirement due to poor planning (Lusardi & Mitchell, 2008). The data from this experiment appear to support these findings.

In general, higher earnings are the compensation for accepting higher risk (Lakonishock, Shleifer, & Vishny, 1994). However, when the independent contribution of complexity, riskiness, stability, and losses to judgments in earnings were tested, only plan complexity and losses contributed to earnings estimates. It was surprising that expected losses accounted for an almost equal amount of the variance as complicatedness (R^2 change = .035) as participants judged both the HTC and LTC plans as equally likely to lose money if the stock market fell by 50%. The inference being that although individuals will judge a HTC plan to be capable of higher earnings than a LTC plan, both a simple and a complex investment strategy are likely to lose money at the same rate.

In addition, a model with willingness-to-pay as the dependent variable showed that elements of volatility did not predict willingness-to-pay but subjective complicatedness did. This allows some tentative conclusions to be drawn. Namely, participants are willing to pay more for a complex investment plan because they judge that a HTC plan will result in higher earnings. Participants were also willing to pay more for the high task complexity plan; data suggested that this preference was

unrelated to the volatility of the plan. Cronley, Posavac, Meyer, Kardes and Kellaris (2005) found that individuals are often influenced by price when there is a high amount of information and individuals have few cues from which to infer the actual quality of their purchase. That is to say, individuals believe that higher price equals higher quality. Within this experiment it may be that although individuals judge the complex plan as more likely to result in higher earnings, they are willing to pay more to have someone administer this plan for them because perceive the complex plan to be of higher quality. This inference is supported by Stewart et al. (2003) who proposed that individuals, because they are constructing their preferences as they go in unfamiliar domains, will be unable to make differentiated decisions regarding magnitude as they have no reference point or context to help them make those judgments. In the context of retirement investment decisions, this finding has some implications. If individuals are willing to pay more for a complex investment strategy that they believe will result in higher earnings, this may result in selection of an inappropriate investment strategy. For example, an investment strategy that seems simple on the face, such as a target-date fund, can have a wide range of associated fees depending on whether it is an actively or passively managed fund. The payment of fees is not inherently associated with increased performance (Delva & Olson, 1998). A heuristic that results in a bias toward a higher-cost plan will have long term financial repercussions.

Financial literacy and numeracy both involve understanding and interpreting numerical information and it was expected that individuals who were more numerate

and more financially literate would be less likely to be confused by financial concepts and therefore less likely to exhibit a bias for the HTC plan. However, results indicated no interaction effects between complexity and financial literacy or numeracy on earnings estimates, indicating that neither financial literacy nor numeracy had a de-biasing effect on participant's earnings judgments. It may be that individuals who lack financial literacy skills and numeracy skills are less able to differentiate between the simple and complex plans (i.e., they see them as both quite complex) while individuals who are even slightly more sophisticated about financial concepts are able to make a finer distinction between the two retirement plans and will trust the more complex investment plan.

The aim of the following experiment was to extend the examination of the role that complexity plays in individual judgments regarding retirement investment decisions and the potential practical implications related to individual investment decisions. While the previous study was based on a control design (i.e., abstract, highly controlled but ecologically inspired materials), Experiment 2 was designed to be more representative of materials individuals might see when meeting with a financial advisor. By reproducing this experiment with ecologically representative materials, it is possible to test the generalizability and robustness of the judgment that individuals will display a bias wherein they judge a complex investment strategy to be more efficacious than a simple one in terms of earnings.

Experiment 2

The following experiment used a modified procedure to rule out potential differences from plan costs as well as to further examine financial literacy among a representative sample of participants. First, a representative panel of subjects were recruited within the age range of individuals who would be making retirement investment decisions for the year 2050. Second, two retirement plans with high and low task complexity were again used but equal plan fees were included in each plan. Thus, differences in judgments of earnings could not be affected by differences in judgments of willingness to pay. Lastly, financial literacy was measured using a more extensive, 13-item test.

Hypothesis 1: Using a representative sample of participants to test the effects of high and low retirement plan task complexity on judgments of plan earnings it was expected that subjective task complexity would be significantly and positively related to judgments of plan earning ability.

Hypothesis 2: It was expected that financial literacy, as measured by a 13-item financial literacy test, would affect earnings judgments after controlling for subjective task complexity, such that individuals who score higher in financial literacy (e.g., scored in the top two quartiles) would be less likely to judge the HTC plan to be capable of higher earnings than the LTC plan.

Lastly, this research sought to extend Experiment 1 by investigating a potential side-effect of high task complexity – the decision not to choose. Tverskey

and Shafir (1992) suggested that conflicted choice aptly describes an individual's reaction to complex decisions – it may be that individuals will respond to high task complexity by deferring their decision-making to a later time (Dhar, 1997). The results of delaying the start of retirement savings can be high; the longer retirement savings is delayed, the less likely an individual will be able to save enough to adequately smooth consumption into retirement (Banks, Blundel, & Tanner, 1998). Tversky and Shafir (1992) proposed that the addition of a third, no-choice option to a choice set of two differing but acceptable options will increase the likelihood that an individual will defer making a decision. Therefore, the following hypothesis was included in Experiment 2.

Hypothesis 3: When presented with a choice to defer (e.g., put off making a decision) alongside both the HTC plan and the LTC plan, participants would choose to defer. If participants must choose only between deferring their choice and either the HTC plan OR the LTC plan, they would choose the retirement plan over deferring.

Method

Participants

Participants were 103 paid Survey Monkey panelists between the ages of 25 and 39; the cost per participant was approximately $6.00. 81 participants completed the study and were included in the analysis. Of these, 26% were male, 65% were female and 9% chose to not respond. 74% of participants had either a Bachelor degree or a graduate degree. Household income ranged from less than $20,000 per year (17% of respondents) to over $150,000 (6%). Most respondents (77%) had a household income between $20,000 and $150,000 per year.

Materials

The experiment included the following materials: (1) two retirement portfolio management plans which served as the basis for the independent variable, one plan had a low level of task complexity and one had a high level of task complexity; dependent variables included (2) four questions regarding subjective task complexity and two questions regarding judgments of potential plan earnings; (3) a 13 question financial literacy test; (4) three questions regarding risk; (5) four questions regarding intent to defer; and (5) demographic questions. Each of these is described in detail below.

Portfolio management plans

The two portfolio management plans were developed through meetings with advisors from Edward Jones and Charles Schwab who had 10+ years of experience. These retirement plans used asset terminology and fund allocations matched for level of risk from real portfolios available through TIAA-CREF, which is the largest financial services group in the United States.

The low task complexity (LTC) plan was a target date fund that consisted of three asset classes and one investment option and was titled the Target Investment Plan (see Table 10). A target date fund is a mix of funds that employs a riskier mix of assets while the individual is younger and automatically shifts toward a more conservative investment strategy as the target retirement date approaches, in order to first grow the funds and then to protect them. In this manner, a target date fund is a simplified strategy for planning for retirement – one need simply calculate their anticipated retirement date and then select a target date fund that matches that year. The fund shown in Table 9 was for the year 2050 and was appropriate for the participants, who were between the ages of 25 and 39.

Table 10. Low task complexity retirement plan.

Target Investment Plan		
Asset Allocation		
Domestic equity	67%	Lifecycle 2050 Fund
International equity	23%	
Fixed income	10%	
Total:	**100%**	
Plan fee: 1.03%		

The high task complexity (HTC) plan consisted of six asset classes (i.e., real estate, large cap stocks) with a total of 10 investment options and was titled the Active Investment Plan (see Table 11). This type of investment strategy not only requires more effort in order to understand the available options, but also more management over time as assets must be manually shifted to more conservative options as retirement approaches.

Table 11. High Task Complexity Retirement Investment Plan

Active Investment Plan		
Asset Allocation		
Real estate	8%	TIAA Real Estate **OR** Morgan Stanley Global Real Estate Portfolio
Fixed income	13%	PIMCO Real Return Fund Institutional Class Shares
Guaranteed	7%	TIAA Traditional
Large-cap stocks	27%	Wells Fargo Advantage Large Cap Growth Fund **OR** T. Rowe Price Equity Index 500 Fund
Mid/Small cap stocks	22%	Columbia Mid-Cap Index Fund Class Z **OR** Small-Cap Blend Index Fund - Institutional Class
International stocks	23%	TIAA-CREF International Equity Index Fund **OR** PIMCO Foreign Bond Fund (Unhedged)
Total:	**100%**	
Plan fee: 1.03%		

High and low task complexity were manipulated using the number of dimensions within each plan and the number of alternatives available within each dimension. The HTC condition contained six dimensions and ten alternatives while the LTC condition contained three dimensions with only one alternative. The higher number of elements in the complex plan increased the amount of information processes needed - such as reading, comparing, and choosing - thus required more cognitive effort (Payne, Bettman & Johnson, 1988). Both plans included an equal percentage-based fee so the key independent variable was task complexity.

Subjective judgments of complexity and earnings

There were four questions regarding the subjective complexity of each plan, These elements included: how complicated each plan seems, how much anxiety is caused by selecting each plan, how difficult it would be to remember plan components, and how difficult it would be to explain to others.

Each of the questions to measure subjective task complexity were rated on a 7-point Likert-type scale.

Participants answered two questions regarding earnings potential for each investment plan. Responses were made on 7-point Likert-type scales.

Q1. In an average year, if you invest $1000.00 how much do you think each plan will earn?

Q2. What percentage rate of return do you think each plan has?

Investment Plan Risk Assessment

The three risk assessment questions were intended to serve as an indicator for how well individuals udnerstood the role that market volatility and inflation can have on rates of return. Therefore the following three questions were adopted from the risk assessment process used by Edward Jones:

1. *All investment plans have the potential for losses. Suppose you have 10 years until you start making withdrawals from your retirement savings and your*

portfolio fell by 20% (i.e., a $1000 initial investment would now be worth $800). How would you react?

2. *Investment plans typically experience different rates of return (i.e., profits). What percentage rate of return do you think each plan has?*
3. *Inflation is the rise of prices over time and can eat into your investment returns. How do you think each of these plans will perform in comparison to the inflation rate?*

Each of the questions to measure risk were rated on a 7-point Likert-type scale.

Financial Literacy Test

A 13 question financial literacy test from Lusardi and Mitchell (2007b) was used to measure financial literacy. Five questions on the test involve basic financial literacy, such as the ability to calculate interest and eight questions cover more sophisticated financial concepts such as stocks and mutual funds. This tool has been used extensively by Lusardi and Mitchell (2007a, 2007b, 2007c, 2008, 2009, & 2011) as a measure of financial literacy and by van Rooij, Lusardi and Alessie for use in the 2005 DNB Household survey. The Cronbach alpha for the 13-item measure is 0.76, indicating that the instrument is internally consistent. The Pearson *r* correlation indicates the test-retest reliability for this instrument ranges between .74 and .80, indicating a high level of reliability. The precise wording of all 13 questions can be found in Appendix B.

Deferring Retirement Plan Selection

There were four within-subjects questions regarding deferring retirement plan selection, which are shown verbatim below. The first question asked participants whether they would prefer to select the HTC or LTC retirement plan in order to determine if one option was more attractive to participants than the rest. The next two questions were regarding deferring when offered either the HTC or LTC plan and were intended to determine if one plan would lead more participants to defer. The fourth question offered participants the option to select either the HTC plan, the LTC plan, or select to defer their decision. Whether the participant saw the active plan first or the target plan first was dependent on which of the two forms of the experiment they were taking.

Question 1: Active v. Target. As you peruse the Active and Target Investment Plans, the benefits coordinator asks which option you would like to choose. Would you prefer to: Select the Target Investment Plan, Select the Active Investment Plan.

Question 2: Target v. Defer. As you peruse the Active and Target Investment Plans, the benefits coordinator tells you that they recently decided to eliminate the Active Investment Plan. Would you prefer to: Select the Target Investment Plan, Wait until you have looked for other retirement plan options.

Question 3: Active v. Defer. As you peruse the Active and Target Investment Plans, the benefits coordinator tells you that they recently decided to eliminate the Target

Investment Plan. Would you prefer to: Select the Active Investment Plan, Wait until you have looked for other retirement plan options.

Question 4: Active v. Target v. Defer. As you peruse the Active and Target retirement plans, the benefits coordinator mentions that they are planning to add more retirement plan options soon. You now have three options available. Would you prefer to: Select the Active Retirement Plan, Select the Target Retirement Plan, Wait until you learn more about the new options.

Demographics

The demographic questions included age, sex, income, and education level. In addition, there were two questions regarding what type of investment instruments participants use, if any, as well as whether they manage their investments themselves or someone else manages them (e.g., a financial planner). The purpose of these questions was to examine the relationship between financial literacy and actual financial choices.

Design and Procedure

A within-subject experimental design was used where participants saw the HTC and LTC plans side-by-side. The within subject design was selected as more ecologically valid than a between subjects design (e.g., having participants view and respond to only one level of task complexity): in real life retirement plan selection an individual is frequently presented with multiple plans and must make a selection among those options.

The HTC plan and the LTC plan were experimentally counterbalanced; however due to high dropout rates in one condition, 65% of participants saw the HTC plan first and only 45% saw the LTC plan first. The possible causes for the high drop-out rate are discussed in the study limitations.

Participants were told that they are making a retirement investment plan decision at their new job and that they have two investment plan options. After viewing both retirement savings plan options, participants answered the four subjective task complexity questions and two plan earnings questions, followed by the 13 question financial literacy test. Then, they completed the three risk assessment questions. The risk assessment questions were asked after the financial literacy test in order to prevent any learning effects, as the risk assessment questions provided some information regarding how stocks and bonds work and the nature of inflation which were covered in the financial literacy test. Participants answered four questions regarding selecting a plan or deferring. Finally, participants answered demographic questions.

Data Coding

Data was first downloaded from Survey Monkey to Excel to be recoded using Visual Basic and then exported to PASW 18 SPSS for data analysis. The subjective plan questions and risk questions were all in Likert-style scale (e.g., interval) format and were in an appropriate format for data analysis and were summed into a total subjective complexity score and risk score. The financial literacy test was scored as a

single weighted average of correct/incorrect responses which were then factor analyzed per the authors' method (Lusardi, 2007b).

Results

In order to validate the task complexity manipulation, the four elements of subjective task complexity were compared between the HTC and LTC group using paired sample t-tests. For all four components, participants judged the HTC plan to have significantly higher subjective task complexity than the LTC plan, indicating that the task complexity manipulation was successful. The mean rating and standard deviation for each of the subjective complexity components are presented in Table 12.

Table 12. Summary of paired sample t-tests with significant results flagged p <0.05*

Subjective judgment questions	**HTC** ***M(SD)***	**LTC** ***M(SD)***	***t*-value**	***p*-value**
How much time do you think it will take to set up each plan?	3.41(1.79)	2.74(1.67	4.27	<.001*
How much anxiety does the thought of selecting and starting each of these investment plans give you?	4.73(1.90)	4.20(1.93)	2.55	0.013*
How complicated do you think each plan is?	4.76(1.77)	3.71(1.81)	4.90	<.001*
How difficult do you think each plan is to remember?	4.78(1.93)	4.24(2.07)	2.88	0.005*

Paired sample t-tests were also used to determine whether there was a difference in judged earnings between the HTC and LTC plans. Results indicated that

participants judged the HTC plan (M = 3.88, SD =1.63) as likely to earn more than the LTC plan (M = 3.50, SD =1.52) on the question, “In an average year, if you invest $1000.00 how much do you think each plan will earn?” $t(73) = 3.12$, $p < .001$ but there was no significant difference in their judgments for the question, “What percentage rate of return do you think each plan has?” $t(72) = 1.14$, $p = .257$. For this experiment, there were not a significant difference in responses for male and female respondents.

A multiple linear regression with the components of subjective task complexity as the predictor variables and dollar earnings as the dependent variable tested whether subjective task complexity is a significant predictor of earnings. First, a bivariate correlation matrix was calculated to determine the strength and direction of the relationship between the components of subjective task complexity and earnings. Results indicated that all elements of subjective complexity were significantly correlated at the $p < .01$ level except time*earnings and difficulty*earnings.

Table 13. Pearson r correlations between subjective task complexity and earnings N=81, **p < .01.

	Anxiety	Time	Complicated	Difficult
Time	0.30**			
Complicated	0.45**	0.53**		
Difficult	0.54**	0.41**	0.69**	
Earnings	0.24**	0.10	0.38**	0.39

The multiple linear regression showed that subjective task complexity significantly predicted earnings. The results indicated that two predictors explained 16.8% of the variance (R^2 = 16.8, F(1,72)=8.36, p =.001. It was found that subjective judgments of plan complicatedness (β=.24, p=.001) and difficulty remembering (β=.-.23, p<.001) significantly predicted earnings. Next, a hierarchical multiple regression analysis assessed whether a participant's earnings estimates were a result of their subjective judgments of plan task complexity rather than risk (e.g., taking on additional risk is expected to result in higher earnings). Risk was entered into the model in step one and subjective complexity in step two. Tests for multicollinearity denote a low level of multicollinearity (*VIF* between 1.02 and 2.5 for all variables). Results of the regression showed that risk judgments accounted for a very small portion of the variance (R^2 = .08, F(3, 64) = 3.02, p = .036) while subjective complexity accounted for almost a quarter of the variance (R^2 = .22, F(7, 60) = 3.73, p = .002), indicating that Hypothesis 1 – that participants would rate the subjective task complexity of the HTC plan higher than the LTC plan - was supported.

A one-way ANCOVA was conducted to test Hypothesis 2 regarding whether financial literacy moderated the relationship between subjective task complexity and

judgments of plan earning ability such that individuals who scored higher in financial literacy were less likely to judge the HTC plan to be capable of higher earnings than the LTC plan. Financial literacy was divided into quartile scores and its effect on earnings judgments was tested while controlling for subjective complexity. Results showed no significant effect of financial literacy on earnings after controlling for subjective complexity $F(3, 68) = 2.586$, $p = .061$, indicating that Hypothesis 2 was not supported.

Recall that Hypothesis 3 posited that, when presented with a choice to defer (e.g., put off making a decision) alongside choosing either the HTC plan or the LTC plan, participants would choose to defer. If participants must choose only between deferring their choice and either the HTC plan OR the LTC plan, they would choose the retirement plan over deferring. An analysis of the frequency of the selection of the active versus target plan, the active plan versus the decision to defer, and the target plan versus the decision to defer indicated that both options were selected with equal frequency (see Table 14).

Table 14. Summary of selection frequency for Hypothesis 3

Component 1	% Selected	Component 2	% Selected
Active	53%	Target	47%
Active	49%	Defer	51%
Target	52%	Defer	48%

A Chi-Square Test was used to test whether participants would choose to defer if offered both retirement plans in addition to an option to defer and results

indicated that each option was selected with equal probability ($p = .44$), indicating that Hypothesis 3 was not supported.

In order to further assess participants' selection patterns for the four deferral questions, crosstab analyses were conducted in SPSS to examine conditional probabilities. Results showed that participants were relatively consistent in their responses: If they selected the Active plan in question 1, they would select the active plan in following questions if it was an option. Likewise for selecting the Target Plan. The pattern showed that if the first option selected was not available in the second question, the choice would be to defer (e.g., In question 1: Active v. Target, the Active plan was selected, in Question 2: Target v. Defer, the Defer option was selected). So, while the results seen were not as expected, there is evidence that participants were consistent in their selection patterns, and would elect to defer if their first option was not available.

Experiment 2 Discussion

The central thesis of this research was that high and low task complexity in retirement plans affects individual judgments of plan earnings capability, and this thesis was supported; participants judged the HTC plan to be capable of higher earnings than the LTC plan. In addition, subjective task complexity was found to be a unique predictor of earnings judgments, even after judged risk was accounted for. The Cohen's *d* effect sizes for subjective task complexity and judged earnings range from .24 to .58 indicating that these results are have practical significance. The judged

difference in earnings between the HTC and LTC plan is approximately $12.50. For a 30 year investment time horizon, which is appropriate for these participants, with a modest rate of return of 7% compounded monthly - and a yearly contribution of $12.50 - this would result in a difference of over $23,000.

Table 15. Cohen's d for subjective task complexity and earnings.

Item Title	Cohen's *d*
Likely to cause anxiety	0.27
Amount of time	0.39
How complicated	0.58
Difficult to explain	0.27
Amount of earnings	0.24

In Experiment 1, a similar effect from subjective complexity was found but there was an additional effect from willingness-to-pay, a proxy for the cost of the plan. Experiment 2 controlled for this effect by assigning an equal plan fee to both the HTC and LTC plan. The concept that individual decisions are affected by task complexity is not novel (Bettman et al., 1993; Dhar, 1997; Keller & Staelin, 1987; Tversky & Shafir, 1992), however, to the best of the author's knowledge this is the first set of experiments that have shown that task complexity can create a supposition of value, which could be labeled a "complexity bias" - an expectation that more complex investment strategies result in higher earnings.

From a theoretical standpoint, this non-normative bias may be a mechanism of the representativeness heuristic wherein individuals seek order in random information

(Tverskey & Kahneman, 1974). Use of this heuristic has been found to increase reliance on past investment performance while simultaneously bypassing important information regarding expected returns and risk. A tendency to seek order in a random situation may lead individuals to believe that complex problems, such as retirement planning, are better addressed by complex solutions.

The relationship between subjective task complexity and earnings were expected to be moderated by financial literacy such that individuals with higher financial literacy would be less susceptible to bias and less likely to judge the HTC plan as capable of higher earnings. However, it was found that there was not a significant relationship between subjective task complexity and financial literacy. This does not align with the findings of Maynard and Hakel (1997) who found that individuals with lower cognitive ability were more likely to judge a task with high complexity to be subjectively more complex. It may be that financial literacy was an inadequate proxy for cognitive ability; a measure such as the Weschler Adult Intelligence Scale which is a well-validated measure of general mental ability, may have shown different results. Little has been written about the relationship between cognitive ability and financial literacy; Cole and Shastry (2009) reported that financial literacy education did not affect financial market participation while cognitive ability does, but their measure of financial literacy was participation in an educational course. Future research on the psychometric properties of financial literacy, particularly construct validity, would be useful and has been previous recommended by van Rooij et al, (2007).

These findings shed some perspective on financial literacy and financial literacy education. Willis (2008) states, "… high financial literacy can be necessary for good financial decision making, but is not sufficient; heuristics, biases, and emotional coping mechanisms that interfere with welfare-enhancing personal finance behaviors are unlikely to be eradicated through education, particularly in a dynamic market." In this research, financial literacy was found to be negatively related to whether participants had a retirement savings plan (-.235, p = .047) and there was no significant relationship between financial literacy and the amount participants had actually saved for retirement. These results support Willis (2008) - financial literacy may be inadequate to prevent individuals from making sub-optimal retirement investment choices because it is an inadequate amount of education for the ever-changing and complex financial decisions consumers must make.

Finally, it was expected that when choosing between one retirement plan option and a no-choice option, individuals would select the retirement plan but when presented with both retirement plans and a no-choice option, individuals would select the no-choice option. This was not the case. It may have been that choosing between the HTC plan and the LTC plan does not create a conflicted choice; Tversky and Shafir (1992) state that, "When one option is better than another in all essential respects, there is no conflict and choice is easy." It is also possible that this result is a limitation of the within-subjects study design; a between-subjects design for this section of the experiment may have decreased interference from each participant answering all four deferred choice questions. The selection pattern that emerged from

analyzing the conditional probabilities indicated that participants strove to maintain consistency in their responses throughout the questions (i.e., if Active plan was selected in first question, Active plan would be selected whenever that option was present). Future research would benefit from adapting the experimental design of Tversky and Shafir (1992) where each participant answers only one question: either from two options (e.g., retirement plan OR defer) or three options (e.g., retirement plan A, retirement plan B, OR defer). In this way, it would be simpler to determine whether adding an additional retirement plan option is more likely to increase conflicted choice and result in deferring retirement savings.

The results of this research have implications for benefit plan administrators, legislators, and consumer protection agencies. By knowing that individuals have a proclivity to choose more complex investment strategies, plan administrators can mitigate the effects of this by setting default options that favor a simplified investment strategy, imposing penalties for opting out or by offering simple heuristic devices to aid individuals who are building their own retirement portfolio (e.g., if you are *x* years old, *xx*% of your contributions should be in equities and *xx*% in income funds).

Study Limitations and Future Research

This study was limited by the data collection procedure. In order to control the length of the survey, which increased the cost of paying for participants, the counterbalancing was divided into two separate studies - one group was solicited to

respond to the HTC first condition and a second group was solicited to participate in the LTC first condition. This created an issue as there was no way to control whether a participant took both versions of the survey. It appeared that participants self-selected out of taking the study twice as there was a high drop-out rate in the version that was posted second. In addition, the IP addresses for all participants were cross-checked and any completed responses with duplicate IP addresses were deleted (N = 3). This decision was disadvantageous to the study in that it limits the conclusions that can be drawn from the data and resulted in fewer usable responses. In that the results of Experiment 2 corroborated Experiment 1, it may be that these limitations were not overwhelming.

The high level of attrition from both forms of the experiment is also cause for concern. It may be that the length was too long, but it is also possible that the subject matter caused certain types of individuals to drop out which would result in a sample that was not adequately representative of the population. In that the majority of the panel was college educated, it is more likely that the results are over-estimating, rather than under-estimating the capabilities of the population.

Lastly, this research assumes that participants are employing a compensatory model (Swait & Adamowicz, 1999). Research has shown that individuals may adopt different decision strategies based on the context and complexity of the task; future research should incorporate different choice models into the experiment. Other models might include elimination-by-aspects, non-compensatory, or weighted

additive. In that a compensatory model is a rational decision-making model, it makes sense to explore other options that may be more representative of how individuals make decisions (Payne, 1976).

Given these limitations, there is more research that can and should be done in this domain. Americans are faced with many decisions that have a high level of task complexity and long-term financial implications such as selecting insurance or choosing a mortgage type. If the complex-is-better bias extends to other domains, particularly those that involve taking on debt, the result could be non-optimal. For example, a homebuyer may choose a more complex hybrid Adjustable Rate Mortgage (ARM) over a simple fixed-rate mortgage but not understand the interest rate implications or the probability of higher future payments with an ARM. Future research is needed to determine if the complex-is-better bias extends to other domains as well as whether its effects are still seen when selecting among more than two retirement investment plans.

References

Amir, O., Rand, D. G., & Gal, Y. K. (2012). Economic games on the Internet: The effect of $1 stakes. *PLoS ONE,* 7(2), e31461. doi:10.1371/journal.pone.0031461

Andreassen, P. B., & Kraus, S. J. (1990). Judgmental extrapolation and the salience of change. *Journal of Forecasting, 9,* 347-372.

Ayton, P., & Fischer, I. (2004). The hot hand fallacy and the gambler's fallacy: Two faces of subjective randomness? *Memory &Cognition, 32,* 1369-1378.

Banks J., Blundell R. & Tanner S. (1998). Is there a retirement-savings puzzle? *American Economic Review,* 88, 769-788.

Barry, C. B. (1974). Portfolio analysis under uncertain means, variances, and covariances. *Journal of Finance, 29,* 515-522.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of personality and social psychology*, *51*(6), 1173.

Benartzi, S., & Thaler, R. S. (2001). Naïve diversification strategies in defined contribution savings plans. *American Economic Association, 91,* 79-98.

Benartzi, S., & Thaler, R. H. (2007). Heuristics and biases in retirement savings behavior. *The Journal of Economic Perspectives*, *21,* 81-104.

Bettman, J. R., Johnson, E. J., Luce, M. F., & Payne, J. W. (1993). Correlation, conflict, and choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*(4), 931.

Braarud, P. Ø. (1998). Complexity factors and prediction of crew performance (HWR-521). *Halden, Norway: OECD Halden Reactor Project.*

Broadbent, J., Palumbo, M., & Woodman, E. (2006). The shift from defined benefit to defined contribution pension plans–implications for asset allocation and risk management. *Reserve Bank of Australia, Board of Governors of the Federal Reserve System and Bank of Canada.*

Campbell, D. J., & Gingrich, K. F. (1986). The interactive effects of task complexity and participation on task performance: A field experiment. *Organizational Behavior and Human Decision Processes, 38,* 162-180.

Campbell, D. J. (1988). Task complexity: A review and analysis. *Academy of management review, 13,* 40-52.

Choi, J. J., Laibson, D., Madrian, B. C., & Metrick, A. (2005). Saving for retirement on the path of least resistance. *Rodney L. White Center for Financial Research Working Papers, 9*.

Christelis, D., Jappelli, T., & Padula, M. (2010). Cognitive abilities and portfolio choice. *European Economic Review, 54*, 18-38.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.)* Hillside, NJ: Erlbaum.

Cokely, E. T., Galesic, M., Schulz, E., Ghazal, S., & Garcia-Retamero, R. (2012). Measuring risk literacy: The Berlin numeracy test. *Judgment and Decision-making, 7,* 25-47.

Cole, S. A., & Shastry, G. K. (2009). *Smart money: The effect of education, cognitive ability, and financial literacy on financial market participation* (pp. 09-071). Boston, MA: Harvard Business School.

Cronley, M. L., Posavac, S. S., Meyer, T., Kardes, F. R., & Kellaris, J. J. (2005). A selective hypothesis testing perspective on price-quality inference and inference-based choice. *Journal of Consumer Psychology, 15,* 159-169.

DeMiguel, V., Garlappi, L., Nogales, F.J., & Uppal, R. (2009). A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms. *Management Science, 55,* 798-812.

DePalma, A., Myers, G. M., & Papageorgiou, Y. Y. (1994). Rational choice under an imperfect ability to choose. *The American Economic Review*, 419-440.

DeShazo, J. R., & Fermo, G. (2002). Designing choice sets for stated preference methods: the effects of complexity on choice consistency. *Journal of Environmental Economics and management*, *44*(1), 123-143.

Dhar, R. (1997). Consumer preference for a no-choice option. *Journal of Consumer Research, 24*, 215-231.

Fry, T., Heaney, R., & McKeown, W. (2007). Will investors change their superannuation fund given the choice?. *Accounting & Finance*, *47*(2), 267-283.

Ghazal, S., Cokely, E. T., & Garcia-Retamero, R. (2014). Predicting biases in very highly educated samples: Numeracy and metacognition. *Judgment and Decision Making*, *9*, 15-34.

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision-making. *Annual Review of Psychology*, *62*, 451-482.

Goodman, J. K., Cryder, C. E., & Cheema, A. (2013). Data collection in a flat world: The strengths and weaknesses of Mechanical Turk samples. *Journal of Behavioral Decision Making*, *26*, 213–224.

Hackman, R., & Oldham, G. (1975) Development of the job diagnostic survey. *Journal of Applied Psychology*, *60*, 159-170.

Hastie, R., & Dawes, R. M. (Eds.). (2010). *Rational choice in an uncertain world: The psychology of judgment and decision-making*. Thousand Oaks, CA: Sage.

Hastings, J. and L. Tejeda-Ashton. 2008. Financial literacy, information, and demand elasticity: Survey and experimental evidence from Mexico.” *NBER Working*

Paper n. 14538.

Hedesström, T. M., Svedsäter, H., &Gärling, T. (2004). Identifying heuristic choice rules in the Swedish premium pension scheme. *The Journal of Behavioral Finance*, *5*, 32-42.

Hirshleifer, D. (2001). Investor psychology and asset pricing. *The Journal of Finance*, *56*, 1533-1597.

Horton, J. J., Rand, D. G., & Zeckhauser, R. J. (2011). The online laboratory: Conducting experiments in a real labor market. *Experimental Economics*, 14, 399–425.

Huberman, G., & Jiang, W. (2006). Offering vs. choice in 401(k) plans: Equity exposure and number of funds. *Journal of Finance, 61,* 763-801.

Hung, A. A., Parker, A. M., & Yoong, J. K. (2009). Defining and measuring financial literacy. *RAND Labor and Population Working Paper Series,* WR-708.

Kahneman, D., & Tverskey, A. (1972). A subjective probability: A judgment of representativeness. *Cognitive Psychology, 3,* 430-454.

Keller, K. L., & Staelin, R. (1987). Effects of quality and quantity of information on decision effectiveness. *Journal of Consumer Research, 14,* 200-213.

Lakonishok, J., Shleifer, A., & Vishny, R. W. (1994). Contrarian investment, extrapolation, and risk. *The Journal of Finance, 49*(5), 1541-1578.

Lipkus, I.M., Samsa, G., & Rimer, B.K. (2001). General performance on a numeracy

scale among highly educated samples. *Medical Decision-making, 21*, 37-44.

Lusardi, A. (2000). Saving for Retirement: The Importance of Planning. *Research Dialogue, 66*, TIAA-CREF Institute.

Lusardi, A. (2001). Explaining why so many people do not save. *Center for Retirement Research Working Paper*, (2001-05).

Lusardi, A., & Mitchell, O.S. (2005). Financial literacy and planning: Implications for retirement wellbeing. *University of Michigan Retirement Research Center Working Paper*.

Lusardi, A., & Mitchell, O.S. (2007a). Baby boomer retirement security: The role of planning, financial literacy, and housing wealth. *Journal of Monetary Economics, 54*, 205-224.

Lusardi, A. & Mitchell, O.S. (2007b). Financial literacy and retirement planning: New evidence from the Rand American Life Panel. *Pension Research Council Working Paper*, WP2007-3.

Lusardi, A., & Mitchell, O.S. (2007c). Financial literacy and retirement preparedness: Evidence and implications for financial education. *Business Economics*, *42*, 35-44.

Lusardi, A., & Mitchell, O.S. (2008). How much do people know about economics and finance? Financial illiteracy and the importance of financial education.

Policy Brief, 5.

Lusardi, A., & Mitchell, O.S. (2009). How ordinary consumers make complex economic decisions: Financial literacy and retirement readiness. *NBER Working Paper,* 15350

Lusardi, A., & Mitchell, O. S. (2011). Financial literacy and planning: Implications for retirement wellbeing. *National Bureau of Economic Research,* WP 2005-108.

Marewski, J. N., Gaissmaier, W., & Gigerenzer, G. (2010). Good judgments do not require complex cognition. *Cognitive processing*, *11*, 103-121.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, *7*, 77-91.

Maynard, D. C., & Hakel, M. D. (1997). Effects of objective and subjective task complexity on performance. *Human Performance, 10,* 303-330.

Medill, C. E. (2000). The individual responsibility model of retirement plans today: Conforming ERISA policy to reality, *Emory LJ*, *49*, 1.

Mitchell, O., & Utkus, S. (2003). Lessons from behavioral finance for retirement plan design. *Pension Research Council Working Paper,* WP 2003-6.

Monroe, K. B., & Chapman, J. D. (1987). Framing effects on buyers' subjective product evaluations. *Advances in Consumer Research, 14(1),* 193-197.

Munnell, A. H., & Golub-Sass, F. (2007). Is there really a retirement savings crisis? *An NRRI Analysis*, 1-9

Payne, J. W. (1976). Task complexity and contingent processing in decision making: an information search and protocol analysis. *Organizational Behavior and Human Performance, 16*, 366-387.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision-making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 534-552.

Peters, E., Västfjäll, D., Slovic, P., Mertz, C.K., Mazzocco, K., & Dickert, S. (2006).

Pollock, E., Chandler, P., & Sweller, J. (2002). Assimilating complex information. *Learning and instruction, 12*(1), 61-86.

Numeracy and decision-making. *Psychological Science, 17*, 407-413.

Samuelson, W., & Zeckhauser, R. (1988). Status quo bias in decision-making. *Journal of Risk and Uncertainty, 1*, 7-59.

Schroder, H., Driver, M., & Streufert, S. (1967).*Human information processing*. New York: Holt, Rinehart and Winston.

Schwab, D., & Cummings, L. (1976). A theoretical analysis of the impact of task scope on employee performance. *Academy of Management Review, 1*, 23-35.

Schwartz, L. M. L., Woloshin, S. S., Black, W. C. W., & Welch, H. G. H. (1997). The role of numeracy in understanding the benefit of screening mammography. *Annals of Internal Medicine, 127*, 966–972.

Shiloh, S., Koren, S., & Zakay, D. (2001). Individual differences in compensatory decision-making style and need for closure as correlates of subjective decision complexity and difficulty. *Personality and Individual Differences, 30*, 699-710.

Simon, H. A. (1987). *Bounded rationality. The New Palgrave: utility and probability*, 15-18.

Slovic, P. & Lichentstein, S. (1971). Reversals of preference between bids and choices in gambling decisions. *Journal of Experimental Psychology, 89*, 46-55.

Swait, J., &Adamowicz, W. (2001). The influence of task complexity on consumer choice: a latent class model of decision strategy switching. *Journal of Consumer Research, 28*, 135-148.

Swait, J., & Adamowicz, W. (2001). Choice environment, market complexity, and consumer behavior: a theoretical and empirical approach for incorporating decision complexity into models of consumer choice. *Organizational Behavior and Human Decision Processes, 86*(2), 141-167.

Todd, P. M., & Gigerenzer, G. (2012). *Ecological rationality: Intelligence in the world*, 3-30.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124-1131.

Tversky, A., & Shafir, E. (1992). Choice under conflict: The dynamics of deferred decision. *Psychological Science, 3*, 358-361.

Van Rooij, M., Lusardi, A., & Alessie, R., (2007). Financial literacy and stock market participation. *CFS Working Paper*, No. 2007/27.

Wichman, H., & Oyasato, A. (1983). Effects of locus of control and task complexity on prospective remembering. *Human Factors, 25*, 583-591.

Willis, L. E. (2008). Against financial literacy education. *Iowa Law Review*, 94, 8-10.

Yakoboski, P., & Dickemper, J. (1997). Increased saving but little planning: Results of the 1997 Retirement Confidence Survey. *EBRI Issue Brief, 191*, 1-21.

Zweig, J. (1998). Five investing lessons from America's top pension fund, *Money, 27,* 115-18.

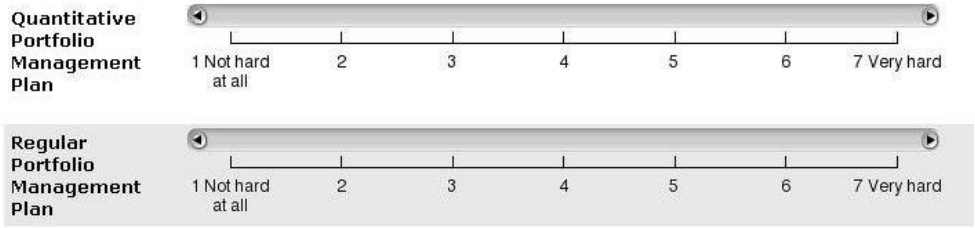
Appendices A through G

Appendix A

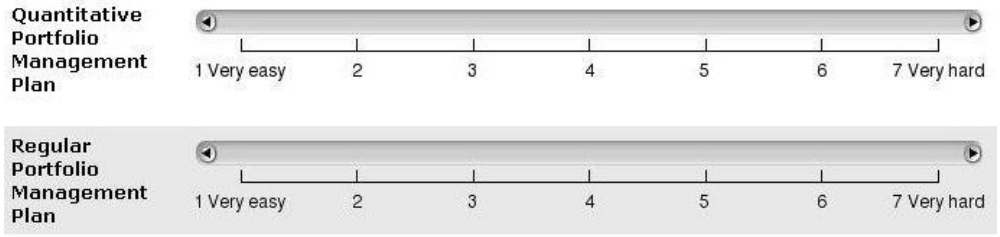
Quantitative Portfolio Management Plan Questions

Regular Portfolio Management Plan Questions

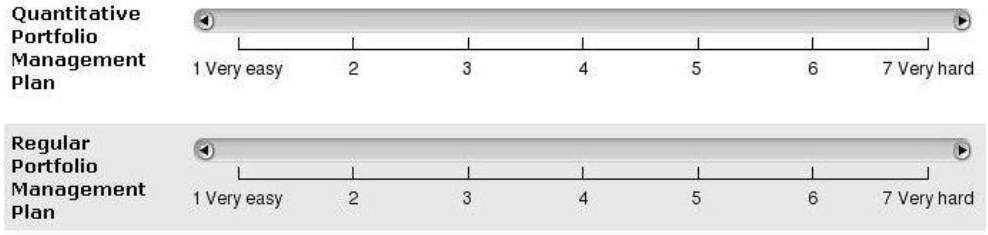
1. How hard do you think it would be to find any errors made by the financial planner for each plan?



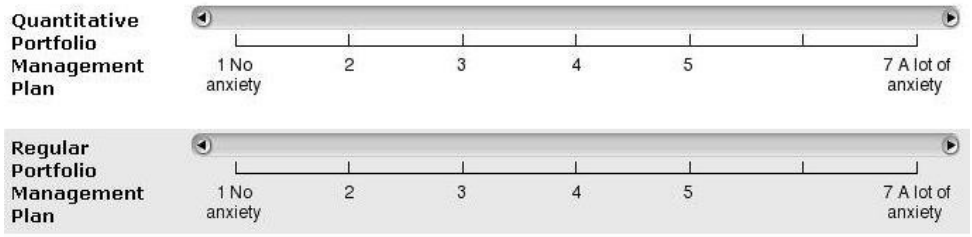
2. How difficult would it be to report each plan on your taxes?



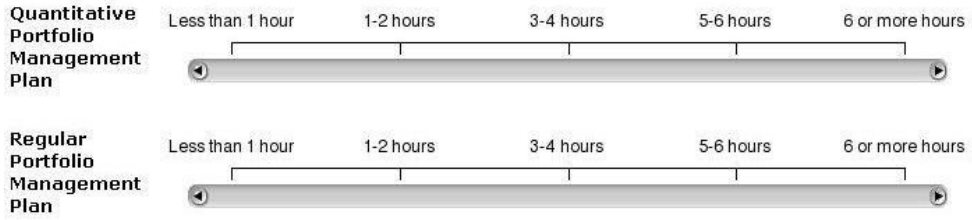
3. How difficult would it be for you to explain this plan to your partner or spouse?



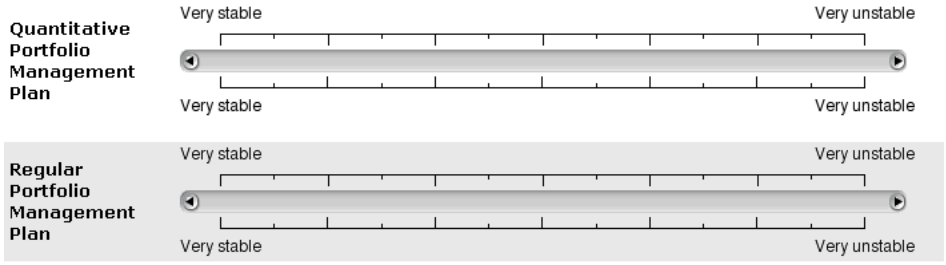
4. How much anxiety does the thought of selecting and starting each of these investment plans give you?



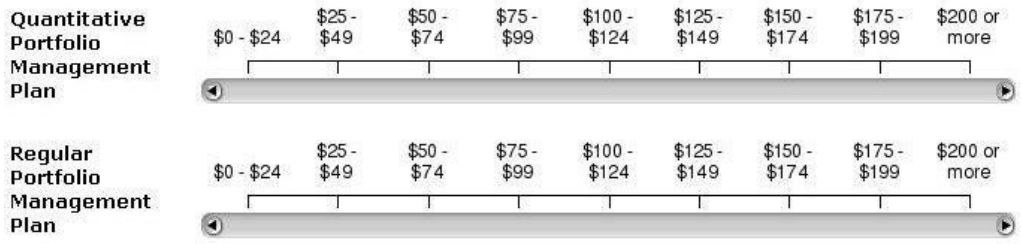
5. How much time do you think it will take to set up each of these plans?



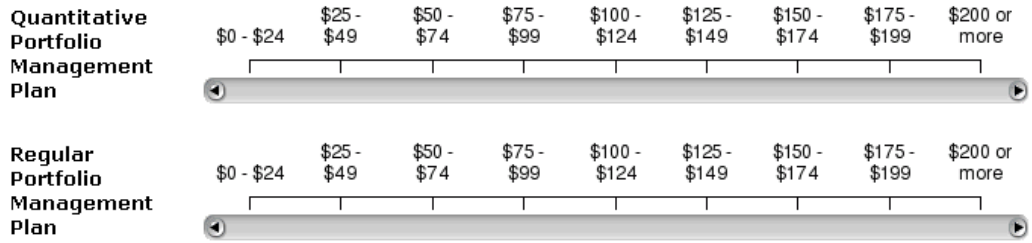
6. How stable do you think each of these plans will be over time?



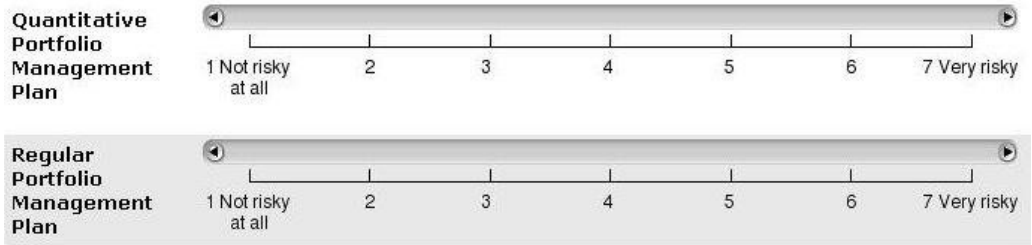
7. In an average year, if you invest $1000.00 how much do you think each plan will earn?



8. If the stock market were to drop by 50% how much would each plan lose?



9. How risky do you think each plan is?

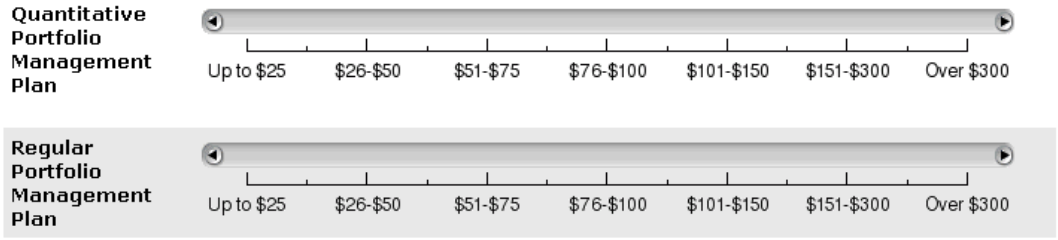


10. Which plan would you choose if you wanted to earn $500 quickly?

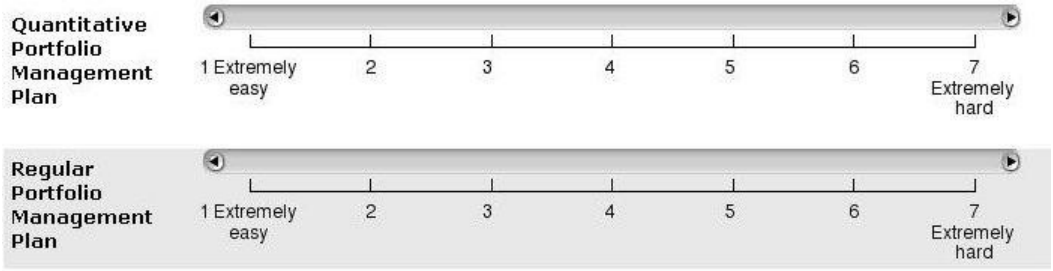
◯ Quantitative Portfolio Management Plan

◯ Regular Portfolio Management Plan

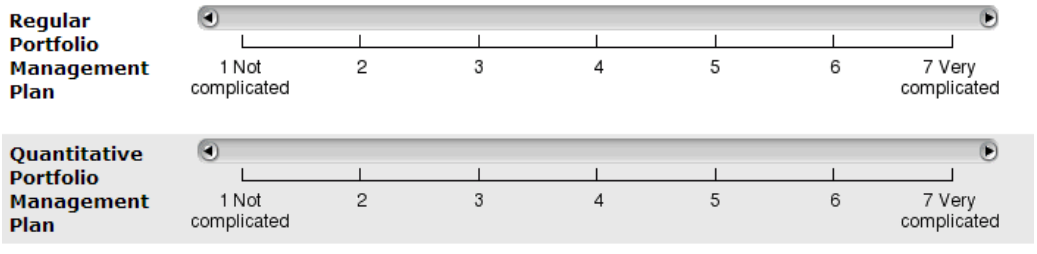
11. How much would you be willing to pay per year to have someone administer each of these plans for you?



12. How hard do you think each plan is to remember?



13. How complicated do you think each plan is?



Appendix B

Financial Literacy Question (3-item measure)

1. Numeracy: Suppose you had $100 in a savings account and the interest rate was 2 percent per year. After years, how much do you think you would have in the account if you left the money to grow? (i) More than $102; (ii) Exactly $102; (iii) Less than $102; (iv) Do not know (DK); (v) Refuse.

2. Compound Interest: Suppose you had $100 in a savings account and the interest rate is 20 percent per year and you never withdraw money or interest payments. After 5 years, how much would you have on this account in total? (i) More than $200; (ii) Exactly $200; (iii) Less than $200; (iv) DK; (v) Refuse.

3. Inflation: Imagine that the interest rate on your savings account was 1 percent per year and inflation was 2 percent per year. After 1 year, how much would you be able to buy with the money in this account? (i) More than today; (ii) Exactly the same; (iii) Less than today; (iv) DK; (v) Refuse.

Appendix C

Numeracy Tests

(Berlin Numeracy Test; see Cokely et al., 2012)

(Lipkus Numeracy Scale; see Lipkus et al., 2001)

Berlin Numeracy Test Items

1) Out of 1,000 people in a small town 500 are members of a choir. Out of these 500 members in the choir 100 are men. Out of the 500 inhabitants that are not in the choir 300 are men. What is the probability that a randomly drawn man is a member of the choir? Please indicate the probability in percent.

   ____%

2) Imagine we are throwing a five-sided die 50 times. On average, out of these 50 throws how many times would this five-sided die show an odd number (1, 3, or 5)?

   _____ out of 50 times.

3) Imagine we are throwing a loaded die (6 sides). The probability that the die shows a 6 is twice as high as the probability of each of the other numbers. On average, out of these 70 throws how many times would the die show the number 6?

   _____ out of 70 throws.

4) In a forest 20% of mushrooms are red, 50% brown and 30% white. A red mushroom is poisonous with a probability of 20%. A mushroom that is not red is poisonous with a probability of 5%. What is the probability that a poisonous mushroom is red? Please indicate the probability in percent.

   ____%

Lipkus Numeracy Scale Items

Imagine that we flip a fair coin 1,000 times. What is your best guess about how many times the coin will come up heads in 1,000 flips?

_____ out of 1,000

In the BIG BUCKS LOTTERY, the chance of winning a $10 prize is 1%. What is your best guess about how many people would win a $10 prize if 1,000 people each buy a single ticket to BIG BUCKS?

_____ out of 1,000

In ACME PUBLISHING SWEEPSTAKES, the chance of winning a car is 1 in 1,000. What percent of tickets to ACME PUBLISHING SWEEPSTAKES win a car?

_____ %

Appendix D

Surprise Memory Test Questions

(Correct answers highlighted)

Quantitative Questions

1) How do you select investment instruments in the Quantitative Portfolio Management Plan?

   a) Select a range of investment instruments that are weakly correlated.

   b) Select a series of investment instruments that are perfectly correlated.

   c) ***Select a range of investment instruments that are not perfectly correlated***.

   d) Select a series of investment instruments that are weakly correlated.

2) Which of the following is a rule in the Quantitative Portfolio Management Plan?

   a) Determine the return on investment using the Capital Asset Pricing Model.

   b) Determine the risk to return ratio using the Capital Asset Pricing Model.

   c) Assess the risk to return ratio using the Principle Asset Costing Model.

   d) Assess the return on investment using the Principle Asset Costing Model.

3) Is this one of the rules of the Quantitative Portfolio Management Plan?
   "Determine risk tolerance by calculating the expected utility of return."

   a) Yes

   b) No

4) Please fill in the information to complete this Quantitative Portfolio Management Plan rule.

   a) "To determine portfolio volatility, test the correlations of: ***Composite Assets*****."**

5) How do you select stocks for the Regular Portfolio Management Plan?

   a) Select thirty different stocks that have about the same amount of risk.

   b) Select twenty different stocks that have about the same amount of risk.

   c) Select thirty different stocks that have about the same amount of return.

   d) Select twenty different stocks that have about the same amount of return.

6) Please fill in the blanks to complete this rule for the Regular Portfolio Management Plan.

   a) "Divide your investment dollars: ***Equally among each of the twenty stocks***."

7) Is the following one of the rules from the Regular Portfolio Management Plan? "Keep portfolio management fees low by buying and trading stocks."

   a) Yes

   b) No

8) Which of the following is a rule from the Regular Portfolio Management Plan?

   a) Check stocks once every year before you file your taxes.

b) Check stocks once every year when you file your taxes.

c) Check stocks at the beginning of the year and when you file your taxes.

d) Check stocks at the end of the year and before you file your taxes.

Appendix E

Experiment 1 Informed Consent and Debriefing Documentation

Michigan Technological University

Office of Compliance,
Integrity, and Safety
302 Lakeshore Center
1400 Townsend Drive
Houghton, MI 49931
906.487.2902

MEMO

DATE: September 24, 2012

TO: Edward Cokely, PhD, CLS

FROM: Joanne Polzien, Executive Director [redacted]

RE: M0650
TITLE: [360755-1] Cognitive Regulation in Judgment and Decisions Under Uncertainty

SUBMISSION TYPE: Continuing Review/Progress Report

STATUS: **RENEWAL APPROVAL**

Thank you for your submission of renewal materials for this research study. **The Institutional Review Board (IRB) has reviewed your request for *renewal, with change(s)*, and has APPROVED your submission.** All research must be conducted in accordance with this renewal approval.

APPROVAL DATE: September 24, 2012
EXPIRATION DATE: September 27, 2013

A pdf of this signed memo and any stamped approved documents, if applicable, have been placed in the review details under "Board Documents" for this project.

Please remember that informed consent is a process beginning with a description of the study and assurance of participant understanding followed by a signed consent form. Informed consent must continue throughout the study via a dialogue between the researcher and research participant. Federal regulations require each participant receive a copy of the consent document.

Please note the following in order to comply with federal regulations and IRB policy:

1. Any modification to previously approved materials or personnel must be approved by this office prior to initiation. Please use the **Update / Modification / Completion** form found in the IRBNet library for modifications and follow the steps in **INSTRUCTIONS: Modification during approval**. This includes, but is not limited to changes in procedures, personnel, study location, participant selection process, etc.
2. All Unanticipated Problems Serious Adverse Events to participants or other parties affected by the research must be reported to this office within two days of the event occurrence. Please use the **Unanticipated Problems or Serious Adverse Events** form found in the IRBNet library and our web site.

   All instances of non-compliance or complaints regarding this study must be reported to this office in a timely manner. There are no specific forms for this report type.

3. All required research records must be securely retained in either paper or electronic format for a minimum of three years following the closure of the approved study. This includes signed consent documents from all participants.
4. This project requires continuing review by our office on an annual basis.

If you have any questions, please contact the Compliance, Integrity, and Safety Office at 906.487.2902 or send your message via email through IRBNet using the Send Project Mail feature.

Informed Consent Form

The purpose of this research is to assess judgments and perceptions, and to understand how we make decisions. There are several potential benefits of this research including understanding of the human condition as well developing custom-tailored decision-support systems. The study will take no more than one hour to complete. You will be asked to complete a computerized decision task, during which eye-tracking, audio, and video data may be collected.

There are no known risks to participating in this study. All data will be given an anonymous code number to identify it and no personally identifying information will be connected with the data. Data will be destroyed after 10 years unless express written consent is given.

Your participation in this study is voluntary and you may stop at any time.

Your participation in this study is voluntary and if you choose not to complete the study at any time please inform the researcher and an alternative task (e.g., read an article) will be given.

If you have any question regarding the research please contact Dr. Edward Cokely, ecokely@mtu.edu. The Michigan Tech Institutional Review Board has reviewed my request to conduct this project. If you have any concerns about your rights in this study, please contact Joanne Polzien of the Michigan Tech-IRB at 906-487-2902 or email jpolzien@mtu.edu .

I have read and understand the above information and am willing to participate in this research study.

Print Name:_____ **Signature:**_____ **Date:**_____

DATE OF IRB APPROVAL: 09-24-12
IRB NUMBER: M0650 360755-1
PROJECT EXPIRATION DATE: 09-27-13

Informed Consent Form for Online/Community Participants

The purpose of this research is to assess judgments and perceptions, and to understand how we make decisions. There are several potential benefits of this research including understanding of the human condition as well developing custom-tailored decision-support systems. The study will take no more than one hour to complete. You will be asked to complete a computerized decision task.

There are no known risks to participating in this study. All data will be given an anonymous code number to identify it and no personally identifying information will be connected with the data. Data will be destroyed after 10 years unless express written consent is given.

Your participation in this study is voluntary and you may stop at any time.

If you have any question regarding the research please contact Dr. Edward Cokely, ecokely@mtu.edu. The Michigan Tech Institutional Review Board has reviewed my request to conduct this project. If you have any concerns about your rights in this study, please contact Joanne Polzien of the Michigan Tech-IRB at 906-487-2902 or email jpolzien@mtu.edu . **I have read and understand the above information and am willing to participate in this research study.**

□ **Agree**

DATE OF IRB APPROVAL: 09-24-12
IRB NUMBER: M0650 360755-1
PROJECT EXPIRATION DATE: 09-27-13

Debriefing

We are ethically required by our university to tell you what this experiment was about. The purpose of this experiment was help us understand more about the individual factors that affect financial decision making. There is a considerable body of research investigating this topic but there are still many issues that are not yet fully understood. Thank you for your participation in this survey. Your input is greatly appreciated and helpful!

Appendix F

Experiment 2 Informed Consent for Online Participants

Michigan Technological University

Office of Compliance,
Integrity, and Safety
302 Lakeshore Center
1400 Townsend Drive
Houghton, MI 49931
906.487.2902

MEMO

DATE: February 27, 2015

TO: Susan Amato-Henderson, PhD, Cognitive and Learning Sciences
FROM: Cheryl A Gherna, CIP, Coordinator Regulatory Review Boards

[redacted]

RE: M1259, [695818-1]
TITLE: INVESTIGATING THE ROLE OF OBJECTIVE AND SUBJECTIVE TASK COMPLEXITY ON RETIREMENT INVESTMENT PLAN DECISIONS

SUBMISSION TYPE: New Project

STATUS: **APPROVED, Exempt Status**

Thank you for your submission of materials for this research study. Your application to use human subjects in research or classroom situations has been reviewed and determined to have an Exempt status.

This approval is based on no greater than minimal risk to research participants. All research must be conducted in accordance with the approved submission; if ANY changes are made in the protocol or conditions set forth in the application, the principal investigator must obtain a separate approval before the changes are implemented, other than to avoid any immediate harm to the subject.

When a research project is determined by the Office of Compliance, Integrity, and Safety (CIS) to meet the criteria for an Exempt status, it means that it is exempt from annual continuing review by the Institutional Review Board (IRB). The investigator, however is required to report to the IRB ANY changes, revisions, or modifications in the research activity including personnel changes. ANY changes made may cause the research status to change from Exempt to Expedited or Full Review and must be determined and documented by the CIS Office.

All Unanticipated Problems / Serious Adverse Events to participants or other parties affected by the research must be reported to this office within two days of the event occurrence. All instances of non-compliance or complaints regarding this study must be reported to this office in a timely manner. Please use the INSTRUCTIONS and FORM: Unanticipated Problem / Serious Adverse Event Form found both on our web site and the IRBNet Library.

This approval applies only for this project, and only under the conditions and procedures described in the application. If/when changes become necessary but are not limited to: changes in protocol, personnel, study location, participant recruitment, etc., as set forth in this approval, you must follow the INSTRUCTIONS and submit the FORM for Change Request during approval found in the IRBNet Library. **You must receive notification of approval PRIOR to implementing the change(s).**

If you have any questions, please contact the Compliance, Integrity, and Safety Office at 906.487.2902 or send your message via email through IRBNet using the Send Project Mail feature.

Informed Consent for Online Participants

The purpose of this research is to assess judgments and perceptions and to understand how we make decisions that affect our lives. There are many potential benefits of this research including understanding human thought processes as well as developing decision aids to support individual decision-making. This study will take no more than 20 minutes to complete and you will be asked to complete a computerized decision task. There are no known risks to participating in this study. All data will be given an anonymous code number to identify it and no personally identifying information will be connected with the data. Data will kept for three years unless express written consent is given. Your participation is voluntary and you may stop at any time. Compensation will be provided for your time through Survey Monkey's Audience portal. If you have any questions regarding this research, please contact Dr. Susan Amato-Henderson, slamato@mtu.edu. The Michigan Technological University Institutional Review Board has reviewed my request to conduct this project. If you have any concerns about your rights in this study, please contact the Office of Compliance, Integrity, and Safety at 906-487-2902 or email IRB@mtu.edu. By clicking you agree that you have read and understand the above information and are willing to participate in this research study.

○ Agree

Next

Appendix G

Target and Active Retirement Plan Questions

Subjective Complexity

How much anxiety does the thought of selecting and starting each of these investment plans give you?

How much time do you think it will take to set up each of these plans?

How complicated do you think each plan is?

How difficult would it be for you to explain this plan to your partner or spouse?

Earnings

In an average year, if you invest $1000.00 how much do you think each plan will earn?

What do you think the annual average rate of return will be for each investment plan?

Deferred Choice

As you peruse the Active and Target Investment Plans, the benefits coordinator asks which option you would like to select. Would you prefer to: Select the Target Investment Plan, Select the Active Investment Plan

As you peruse the Active and Target Investment Plans, the benefits coordinator tells you that they recently decided to eliminate the Active Investment Plan. Would you prefer to: Select the Target Investment Plan, Wait until you have looked for other retirement plan options.

As you peruse the Active and Target Investment Plans, the benefits coordinator tells you that they recently decided to eliminate the Target Investment Plan. Would you prefer to: Select the Active Investment Plan, Wait until you have looked for other retirement plan options.

As you peruse the Active and Target retirement plans, the benefits coordinator mentions that they are planning to add more retirement plan options soon. You now have three options available. Would you prefer to: Select the Active Retirement Plan, Select the Target Retirement Plan, Wait until you learn more about the new options

Risk

All investment plans have the potential for losses. Suppose you have 10 years until you start making withdrawals from your retirement savings and your portfolio fell by 20% (i.e., a $1000 initial investment would now be worth $800). How would you react?

Investment plans typically experience different rates of return (i.e., profits). What percentage rate of return do you think each plan has?

Inflation is the rise of prices over time and can eat into your investment returns. How do you think each of these plans will perform in comparison to the inflation rate?

Appendix H

Financial Literacy Test (13-item Measure)

1. Numeracy: Suppose you had $100 in a savings account and the interest rate was 2 percent per year. After 5 years, how much do you think you would have in the account if you left the money to grow? (i) More than $102; (ii) Exactly $102; (iii) Less than $102; (iv) Do not know (DK); (v) Refuse.

2. Compound Interest: Suppose you had $100 in a savings account and the interest rate is 20 percent per year and you never withdraw money or interest payments. After 5 years, how much would you have on this account in total? (i) More than $200; (ii) Exactly $200; (iii) Less than $200; (iv) DK; (v) Refuse.

3. Inflation: Imagine that the interest rate on your savings account was 1 percent per year and inflation was 2 percent per year. After 1 year, how much would you be able to buy with the money in this account? (i) More than today; (ii) Exactly the same; (iii) Less than today; (iv) DK; (v) Refuse.

4. Time Value of Money: Assume a friend inherits $10,000 today and his sibling inherits $10,000 3 years from now. Who is richer because of the inheritance? (i) My friend; (ii) His sibling; (iii) They are equally rich; (iv) DK; (v) Refuse.

5. Inflation/Money Illusion: Suppose that in the year 2010, your income has doubled and prices of all goods have doubled too. In 2010, how much will you be able to buy with your income? (i) More than today; (ii) The same; (iii) Less than today; (iv) DK; (v) Refuse.

6. Stock Market Functioning: Which of the following statements describes the main function of the stock market? (i) The stock market helps to predict stock earnings; (ii) The stock market results in an increase in the price of stocks; (iii) The stock market brings people who want to buy stocks together with those who want to sell stocks; (iv) None of the above; (v) DK; (vi) Refuse.

7. Knowledge of Mutual Funds: Which of the following statements is correct? (i) Once one invests in a mutual fund, one cannot withdraw the money in the first year; (ii) Mutual funds can invest in several assets, for example invest in both stocks and bonds; (iii) Mutual funds pay a guaranteed rate of return which depends on their past performance; (iv) None of the above; (v) DK; (vi) Refuse.

8. Interest Rate/Bond Prices Link: If the interest rate falls, what should happen to bond prices? (i) Rise; (ii) Fall; (iii) Stay the same; (iv) None of the above; (v) DK; (vi) Refuse.

9. Safer: Company Stock or Mutual Fund: True or false? Buying a company stock usually provides a safer return than a stock mutual fund. (i) True; (ii) False; (iii) DK; (iv) Refuse.

10. Riskier: Stocks or Bonds True or false? Stocks are normally riskier than bonds. (i) True; (ii) False; (iii) DK; (iv) Refuse.

11. Long Period Returns: Considering a long time period (for example 10 or 20 years), which asset normally gives the highest return? (i) Savings accounts; (ii) Bonds; (iii) Stocks; (iv) DK; (vi) Refuse.

12. Highest Fluctuation/Volatility: Normally, which asset displays the highest fluctuations over time? (i) Savings accounts; (ii) Bonds; (iii) Stocks; (iv) DK; (v) Refuse.

13. Risk Diversification: When an investor spreads his money among different assets, does the risk of losing money: (i) Increase; (ii) Decrease; (iii) Stay the same; (iv) DK; (v) Refuse.