



AUSTRALIAN
SCHOOL OF BUSINESS™
THE UNIVERSITY OF NEW SOUTH WALES

www.business.unsw.edu.au

The University of New South Wales
Australian School of Business

Australian School of Business Research Paper No. 2010ACTL08

Investment Risk Framing and Individual Preference Consistency

Hazel Bateman, Christine Ebling, John Geweke, Jordan
Louviere, Stephen Satchell & Susan Thorp

This paper can be downloaded without charge from
The Social Science Research Network Electronic Paper Collection:
<http://ssrn.com/abstract=1664869>

Investment risk framing and individual preference consistency*

PRELIMINARY: DO NOT QUOTE WITHOUT PERMISSION OF AUTHORS

A/Prof Hazel Bateman* Dr Christine Ebling# Prof John Geweke#

Prof Jordan Louviere# Prof Stephen Satchell**

Dr Susan Thorp#†

*Centre for Pensions and Superannuation, University of New South Wales

#Centre for the Study of Choice, University of Technology, Sydney, Australia.

*Trinity College, University of Cambridge

August 24, 2010

*We thank Pure Profile and CenSoC for their generous assistance with the development and implementation of the internet survey. Frances Terlich and Edward Wei gave excellent research assistance. We acknowledge financial support under ARC DP1093842. Part of this work was completed while Bateman visited the School of Finance and Economics at the University of Technology, Sydney.

†Contact author: susan.thorp@uts.edu.au

Abstract

Here we test the usefulness of a discrete choice experiment (DCE) for identifying individuals who consistently exhibit concave utility over returns to wealth, despite variations in the framing of risk. At the same time, we test the relative strengths of nine standard descriptions of investment risk. We ask a sample of 1200 retirement savings account holders to select their most and least preferred investment strategies from a menu of a safe (zero risk) savings account, a risky growth asset portfolio and a 50:50 share of both. We identify respondents who fail to conform with expected utility and test whether this behavior is predictable across different risk frames. Tests confirm that the DCE can help isolate individuals whose preferences violate global risk aversion despite variation in risk presentation. We also identify frames linked to significantly more consistent behavior by respondents. These are frames which simultaneously specify upside and downside risk. Frames that present risk as a frequency of failures or successes against a zero returns benchmark are more likely to generate violations of risk aversion.

Keywords: investment risk; household finance; financial literacy; context and framing effects; retirement savings

JEL Classification: G23; G28; D14

1 Introduction

Defined contribution schemes are now the dominant model for retirement income provision in many economies. As a result, investment decisions that were once the province of wealthier households and their advisors are now the norm for almost everyone. One of the most difficult of these decisions, particularly for unsophisticated investors, is choosing a portfolio for retirement savings, a choice where an appreciation of investment risk is crucial.

Conventional theoretical approaches to investment decisions assume that investors maximize globally concave expected utility functions, usually of the constant relative risk aversion (CRRA) form. In simple cases, random returns are assumed to be lognormally distributed, so that the optimal allocation of wealth among investments depends on the first two central moments of the returns distribution and the risk preference parameter. More generally, individuals with any globally concave utility function will prefer portfolios that have returns distributions exhibiting second order stochastic dominance over alternatives. Individuals who are risk averse will consistently penalize risk and, if given enough information about the distribution of returns, will reject investments that are dominated at the second order.

Researchers have employed an array of methods to measure risk preferences. Academic experimental studies have asked lottery questions, both unincentivized (e.g., Kimball, Sahn and Shapiro 2008) and incentivized (e.g., Holt and Laury 2002). Empirical investigations have inferred risk parameters from observed portfolio allocations (e.g., Friend and Blume 1975). More recent methods include interactive interfaces such as distribution builders (Goldstein, Johnson and Sharpe 2008) or experience sampling, which gives feedback about risky choices through web-based platforms (Haisley, Kaufmann and Weber 2010). The financial planning industry has long relied on questionnaires and risk profiling instruments, such as that implemented by Finametrica, to guide investors to the best mix of risk and return. All these methods assume that inferred preferences are stable and generalizable, a view not always shared by behavioral researchers.

Tversky and Kahneman (1981, 1986) established the importance of framing in eliciting

ing choices. Framing effects occur where changing the statement of a logically equivalent scenario can reverse one's choices. Numerous studies have demonstrated the sensitivity of preferences to lottery outcomes framed as gains or losses, but investigation of framing effects have widened to include many additional modes of presentation in many different contexts, including retirement savings choices returns. (See, among others, Benartzi and Thaler 1999, Rubaltelli, Rubichi, Savadori, Tadeschi and Ferretti 2005, Anagol and Gamble 2008 on asset allocation decisions, and Agnew, Anderson, Gerlach and Szykman 2008 and Brown, Kling, Mullainathan and Wrobel 2008 on annuitization decisions.) Sensitivity to frame presents special challenges to any experimental study or survey. If stated choices always change when frame changes, then little certainty or predictability should be ascribed to inferences from experiments or surveys. Every opinion and choice potentially can be manipulated, directed or reversed. If, on the other hand, individual preferences are more or less consistently exposed by choices, despite changes to presentation, then survey responses can be used for prediction.

Here we test the usefulness of a discrete choice experiment (DCE) for identifying individuals who consistently exhibit concave utility over returns to wealth, despite variations in the framing of risk. At the same time, we test the relative strengths of nine standard descriptions of investment risk. We ask a sample of 1200 retirement savings account holders to select their most and least preferred investment strategies for retirement accumulations from a menu of three. The three strategies are a safe (zero risk) savings account, a risky growth asset portfolio and a 50:50 share of both. Respondents repeat this choice under nine different descriptions of investment risk. In addition, the underlying distribution of investment returns to the risky account varies over four mean-preserving spreads. Each respondent sees three of the nine frames at four risk levels, generating 12 sets of most and least preferred investment strategies per respondent.

Our null hypothesis is that investors are globally risk averse, and this implies a series of restrictions over choices. First, the properties of concavity imply that risk averse individuals

will never choose the 50:50 portfolio as least preferred. This result is independent of the level of risk presented to respondents and represents a strong restriction on preferences. Second, respondents should not increase exposure to the risky asset as returns volatility rises, at a fixed expected return. This is a weak restriction because it is based on the further assumption that the respondent perceives, within each frame and across risk levels, that higher risk levels are mean preserving spreads of lower levels (i.e., a higher risk level is stochastically dominated by a lower level at the second order). We test these restrictions at an aggregate level and frame by frame.

If the DCE helps identify respondents who comply with expected utility, then breaches of concavity under one presentation of risk should predict breaches under the others. Likewise with sensitivity to mean-preserving spread. In both cases we find that the behavior of individual respondents is significantly predictable. Our first contribution is thus to confirm that the DCE can help isolate individuals whose preferences violate global risk aversion despite variation in risk presentation.

Our second contribution is to isolate from within the sets of nine frames, those that lead to significantly more consistent behavior by respondents. These are frames which include information about both upside and downside risk, assigning probabilities to high and low returns. Frames that present risk as a frequency of failures or successes against a zero returns benchmark are more likely to generate violations of risk aversion.

Regulators responsible for consumer financial decision making and the financial services industry now communicate with a mass market, and are investigating how to convey risk information in modes that are intelligible, standardized and easily comparable across products. Australian regulators, for example, have recently introduced regulations to require a short-form (four-page) financial product disclosure document which will present the basic information required to make pension fund and pension asset allocation decisions (Minister for Financial Services 2010) and the UK Financial Services Authority is working to improve insurance and pension disclosures to assist consumers to make informed decisions (Andrews

2009). Cross country coordination of simplified disclosures is being initiated by multilateral organisations such as the European Commission and the OECD (European Commission 2009a, b, OECD 2008). The two contributions of this study expose the tension faced by regulators and financial service providers. On the one hand individuals will to some degree ‘see through’ risk presentations and make decisions consistent with their underlying preferences. On the other hand, not all descriptions of risk are equal, and financial institutions and supervisors can influence retirement saving decisions by using risk framing that encourages people to choose consistently.

In the next section we outline the theoretical underpinnings and design of the discrete choice experiment. Results and discussion follow in section three and section four concludes.

2 Experimental design

Consider an investor with utility over wealth w determined by a constant relative risk aversion utility function

$$u(w) = w^{1-\alpha}/(1-\alpha) \tag{1}$$

where α is the risk aversion parameter. The value $\alpha = 0$ implies risk neutrality; $\alpha = 1$ is $u(w) = \log(w)$; when $\alpha < 1$, $u(w)$ is bounded below by zero and unbounded above; when $\alpha > 1$ $u(w)$ is unbounded below and bounded above by zero. We treat income from now until retirement as exogenous so that lifetime utility is the sum of utility before retirement and utility after retirement, which is proportional to retirement wealth. Investment portfolio allocation decisions are independent of consumption decisions, and the problem of maximizing retirement utility reduces to the single period maximization problem when returns are iid. In what follows, U denotes realized utility and can be written as a function of the gross return to wealth.

Risky asset returns, y , are identically and independently lognormally distributed

$$\log(y) \sim N(\mu, \sigma^2). \quad (2)$$

The investor can allocate current and future retirement savings wealth to one of three accounts (funds).

1. A safe account, S , with risk free return x .
2. A risky account, R , with return y that $E(y) = \exp(\mu + \frac{1}{2}\sigma^2)$, $var(y) = \exp(2\mu + \sigma^2) [\exp(\sigma^2) - 1]$.
3. A medium account, M , made up of equal proportions of risk-free and risky assets with return $z = (x + y)/2$, $E(z) = 0.5[x + E(y)]$, $var(z) = 0.25var(y)$.

In the DCE set out below, investors give a complete ranking of these three accounts by indicating their most and least preferred ('best' and 'worst') retirement savings allocations.

2.1 Survey Questionnaire

We recruited a sample of 1220 individuals aged 18-65 from the PureProfile online web panel of over 600,000 Australians.¹ PureProfile filtered the sample to ensure that all subjects held a current retirement savings ('superannuation') account, that genders were equally represented and that the age distribution did not deviate far from population proportions. Of this sample, 1199 fully completed the survey. Table 1 compares the full sample and Australian population demographics. The survey asked subjects to make hypothetical decisions, but we informed respondents that this was a university project designed to inform policy makers and industry participants. PureProfile paid participants a flat rate of \$3AUD (\$2.70USD) for completing the survey.

[Insert Table 1 here]

We presented a four-part questionnaire:

¹Under Australia's Superannuation Guarantee, all working Australians aged 18-65 who earn at least 8% of average earnings participate in the mandatory retirement savings system, contributing a minimum percentage of their earnings into a (usually privately managed) regulated 'superannuation' fund. Most workers are members of defined contribution funds.

- Introductory questions about subjects’ retirement savings, including the name of their superannuation fund and the aggregate amount in their superannuation accounts;
- 21 questions to measure numeracy and financial literacy skills, as well as self-assessed knowledge of finance, access to financial education, use of financial advice and confidence in stock market recovery;
- The hypothetical asset allocation task for retirement savings; and
- Demographic questions relating to marital status, work status, occupation, industry/business, education, income, assets, household make-up and number in household.

We provide access to the entire survey at

<http://survey.confirmit.com/wix/p1250911674.aspx>

2.2 Task and Risk Presentation

For the hypothetical asset allocation task, we asked the subjects to think about choosing an investment option for their retirement account in the context of possible government simplification of the arrangements. (Simplifications to the retirement savings system have been proposed and enacted by government and industry over the past few years and subjects would find the prospect of regulated change plausible.) Retirement plan providers would offer a menu of three investment accounts, each differing by expected rate of return and investment risk. The three investment choices were 100% bank account (S), 50% bank account and 50% growth (M) and 100% growth (R). The scenario included a verbal explanation of differences in expected real rates of return to the three accounts and an explanation of variability of returns to options R and M . Appendix A reproduces the on-line instructions to the subjects.

Each choice menu provided two key pieces of information for each of the three options: (1) the average rate of return above inflation; and (2) a single aspect of investment risk, for example the frequency of negative returns or return at risk, as detailed in Table 2. Using this information, the respondents were asked to nominate their ‘most likely’ (best) option and their ‘least likely’ (worst) option.

We explained investment risk to respondents in nine different ways (frames) and these are set out in Table 2. Eight of the frames described investment risk in text (with a combination of words and numbers) and one presented the risk information in a graphical format. There is no universally accepted method for presenting investment risk to retirement savers but these nine frames are drawn from standard financial service provider prospectuses in Australia, Europe and the United States as well as related studies (Vlaev, Chater and Stewart 2009). We presented a different aspect of investment risk in each of the nine frames while holding the expected rate of return to each account constant. By design, the aspects were all consistent with the same underlying return distribution, but since no single aspect of investment risk conveys the entire distribution, survey subjects could rationally impute different distributions across frames.

[Insert Table 2 here]

Frames 1, 2 and 9 presented investment risk as a range of returns, and therefore provided information about both the upper and lower tails of the returns distribution. Frame 1 described investment risk as the chance of a return *between* the 5th and 95th quantiles of the M and R returns distributions, while frame 2 is expressed in terms of a return *outside* these quantiles. Frame 9 is a graph illustrating the returns between the 5th and 95th quantiles for each investment option.

In the introductory scenario, subjects were reminded that ‘more risk of high returns also means more risk of low returns’ but the remaining frames provide information on just one side of the distributions of returns. Frames 3 and 4 present investment risk as the chance of a return above (below) the 95th (5th) quantile of the y and z distributions; frames 5 and 6 present investment risk as the number of years in 20 (on average) of positive (negative returns) for M and R , while frame 7 and 8 present investment risk as the number of years in 20 (on average) when returns to M and R will be above (below) the return to the bank account.

So that we could gauge survey subjects’ response to changes in risk, we increased the

standard deviation of y (and therefore of z) in four steps from 12-28% p.a. (Table 3). These four configurations for the lognormal risky asset we denote y_1, y_2, y_3, y_4 , and we denote the corresponding cumulative distributions functions $F_{y1}, F_{y2}, F_{y3}, F_{y4}$. Let $z_i = (x + y_i) / 2$ ($i = 1, 2, 3, 4$). Every subject made choices across all four risk levels in each investment risk frame they viewed. Numerical values (quantiles or frequencies) that we entered into the frames' wording or graphs thus took four increasingly risky levels, derived from simulated cumulative densities of the corresponding lognormal random variable y_i . The number of random draws was $N = 100,000$ and the simulated probability of a return less than a fixed value \bar{r} was $\Pr(Y_{i,n} \leq \bar{r})$

$$\Pr(Y_{i,n} \leq \bar{r}) = \frac{1}{N} \sum_{n=1}^N 1(Y_{i,n} \leq \bar{r}) \quad (3)$$

where $1(\cdot)$ is the indicator function, returning the value 1 when $Y_{i,n} \leq \bar{r}$ and zero otherwise. The draws were generated as $Y_{i,n} = \exp(\mu_i + \sigma_i z_n)$ where z_n is a random draw from a standard normal density.

[Insert Table 3 here]

Table 4 details the simulated values that we entered into the frames for $\{y_i\}_{i=1}^4$.

[Insert Table 4 here]

2.2.1 Order of choice sets

We arranged the 9 investment risk frames into four sets (denoted 'framesets'), each made up of three frames with the visual range graph frame (9) common to each set. That is, Frameset A (frames 1, 2 and 9) is the three 'range' frames including the visual range; Frameset B (frames 3, 4 and 9) is the two 'tail' frames and the visual frame; Frameset C (frames 5, 6 and 9) is the two frames describing risk in terms of years of positive/negative returns and the visual frame; while Frameset D (frames 7, 8 and 9) includes the two frames describing risk relative to returns above/below the bank account and the visual frame. We implemented a between-subjects design that presented one quarter of the sample with one of the four framesets (A, B, C and D). The whole survey was completed by 1199 of 1220 subjects

recruited: 300 subjects completed choices with Frameset A; 299 with Frameset B; 297 with Frameset C; and 303 with Frameset D. Each completing respondent made a total of 12 best and worst selections from the S , M , and R retirement accounts, choosing best and worst accounts at each of four risk configurations in each of 3 frames (one frameset). Table 5 illustrates a representative investment choice task. In this investment choice task, investment risk is presented as Frame 3 ("There is a 1 in 20 chance of a return above $x\%$.") and the underlying returns distribution y_1 has a standard deviation of 12% p.a.

[Insert Table 5 here]

2.3 Concave utility over returns to wealth: strong and weak restrictions

Each of the 12 investment choices made by a survey respondent gives his or her complete ranking of the three accounts, conditioning on risk level and frame. Six rankings are possible:

1. SMR
2. MSR
3. MRS
4. RMS
5. SRM
6. RSM

We consider two restrictions over choices that result from expected utility maximization. First, if a subject makes a choice on the basis of expected utility with a concave utility function, then rankings 5 and 6 are impossible. This follows from the fact that if realized utility U is concave,

$$U(z) = U\left(\frac{x+y}{2}\right) > \frac{U(x) + U(y)}{2}$$

and so

$$E[U(z)] > \frac{U(x) + E[U(y)]}{2}$$

and hence either $E[U(z)] > E[U(y)]$ or $E[U(z)] > U(x)$, or both. In other words, subjects with concave utility should never rank the 50:50 account (M) as worst. We describe this restriction as a strong restriction of expected utility theory because it does not require any assumption about survey subjects' inferences from the risk information we present.

The second restriction depends on subjects' inferences from risk and return information presented in the survey. Given the lognormal distributions of the asset returns y_1, y_2, y_3, y_4 , then y_j is a mean preserving spread of y_i if $i < j$ and $F_{y,i}$ stochastically dominates $F_{y,j}$ at the second order. (See Appendix B for proof.) It follows at once that for $i < j$, z_j is a mean preserving spread of z_i . So long as the respondent perceives that y_j is a mean preserving spread of y_i for all $i < j$, and if U is concave then for all $i < j$,

$$E[U(y_i)] > E[U(y_j)], \quad E[U(z_i)] > E[U(z_j)].^2$$

Within each frame, these conditions place restrictions on the orderings that can be observed across choice sets when subjects conform with concave utility summarized in Table 6:

TABLE 6:
RANKING RESTRICTIONS UNDER EXPECTED UTILITY

Rankings	Rankings using riskier asset y_j			
using y_i	(1) SMR	(2) MSR	(3) MRS	(4) RMS
(1) SMR	*	(5)	(4), (5)	(4), (5), (6)
(2) MSR	*	*	(4)	(4), (6), (6)
(3) MRS	*	*	*	(6)
(4) RMS	*	*	*	*

Note: Table shows account rankings that do not conform with expected utility when risk increase from y_i to y_j . Account S is 100% safe assets, R is 100% risky assets and M is 50:50 safe and risky assets. Ranking pairs marked with an asterisk conform with expected utility.

Because $E[U(y_i)] > E[U(y_j)]$,

$$U(x) - E[U(y_j)] > U(x) - E[U(y_i)]. \quad (4)$$

This eliminates the cells marked (4) in Table 6. Because $E[U(z_i)] > E[U(z_j)]$,

$$U(x) - E[U(z_j)] > U(x) - E[U(z_i)]. \quad (5)$$

This eliminates the cells marked (5) in Table 6. We also conjecture that

$$E[U(z_j)] - E[U(y_j)] > E[U(z_i)] - E[U(y_i)].^3 \quad (6)$$

This eliminates the cells marked (6) in Table 6. Across questions within a frame, ranking comparisons must fall in the cells indicated with an asterisk (*) in Table 6. We summarize these restrictions as follows. Consider an investment account choice made with risky return y_i produce ranking t_i , and let a choice with riskier asset y_j produce ranking number t_j . where $t_i, t_j = 1, \dots, 6$, then it must be the case that $t_j \leq t_i$, if the choices conform with expected utility maximization under concave utility. This is a weak restriction because it is based on the further assumption that the respondent perceives, within each frame and across risk levels, that riskier asset returns are mean preserving spreads of returns of less risky assets.

2.4 Testing the restrictions of expected utility theory

Here we set out tests designed to investigate whether there is substantial variation over survey respondents in their propensity to violate the strong and weak restrictions of expected utility, and more especially, to investigate whether the risk framings we present are useful

in identifying these respondents. We propose the following test: each respondent makes choices over three risk frames, two from frames 1-8, and frame 9. If a respondent violates a restriction in their first frame, does this imply that the respondent is more likely to violate that restriction in the next or final frame they saw? Were it to turn out that violations in their first frame have no implications for violations in the second and third, then this DCE is unlikely to be useful for eliciting attitudes toward risk. We might also call into question the value of many experimental surveys that condition on particular framings of investment risk.

2.4.1 Strong restriction test

For each respondent we have a 3×4 layout of responses in the form of investment account rankings: three different frames, and four different risk levels within each frame. In any of these 12 cells the respondent can violate the strong concavity restriction by ranking M last. We code the variable v_{fir} as being 1 if respondent r ranked M last in frame f at risk level i and 0 otherwise. There are $2^{12} = 4096$ possible outcomes for each individual so it is infeasible to model all possibilities. Instead we pick a cutoff j , where j is 1, 2, 3 or 4, and define $w_{fr}^{(j)} = 1$ if $\sum_{i=1}^4 v_{fir} \geq j$ and $w_{fr}^{(j)} = 0$ otherwise. Since there are three frames per respondent, r , there are eight possible outcomes represented by strings of zero and one indicators of concavity violations.

Our alternative hypothesis is that respondents' violations of concavity are not independent events across the three risk frames they saw. Under this alternative, the model assigns probabilities to each of the eight possible patterns of concavity violations. Explicitly,

$$P \left[w_{1r}^{(j)} = w_1, w_{2r}^{(j)} = w_2, w_{3r}^{(j)} = w_3 \right] = \rho^{(j)}(w_1, w_2, w_3) \quad (7)$$

where $w_f = 0$ or 1 ($f = 1, 2, 3$). This model permits the kind of variation one would expect to see our DCE approach is effective in identifying individuals who do not behave consistently

with concavity. In the extreme of a complete failure of independence we would see that only $\rho^{(j)}(0, 0, 0)$ and $\rho^{(j)}(1, 1, 1)$ have positive values. That is, if a respondent does (not) violate concavity in the first risk frame they see then the respondent will surely (not) violate concavity in their second and third frames.

The null hypothesis asserts that violations of concavity by a respondent are independent events moving from the first to the third frame. This model has a single probability parameter for each of the three frames. Explicitly,

$$P \left[w_{1r}^{(j)} = w_1, w_{2r}^{(j)} = w_2, w_{3r}^{(j)} = w_3 \right] = \rho_1^{*(j)}(w_1) \cdot \rho_2^{*(j)}(w_2) \cdot \rho_3^{*(j)}(w_3) \quad (8)$$

This model is nested in (7) and so the hypothesis can be tested with a conventional asymptotic chi-square statistic based on log-likelihoods, with $7 - 3 = 4$ degrees of freedom.

We also extend this procedure to include further information about respondents' age, retirement savings wealth and financial literacy via the conventional multinomial logit model, in the case of (7), and via the conventional logit model, in the case of (8), each with the same covariates and no restrictions on the logit coefficients. With these $k = 3$ covariates and an intercept, the test statistic will have $4(k + 1) = 16$ degrees of freedom.

2.4.2 Weak restriction test

The weak restriction of expected utility depends on respondents perceiving mean-preserving spreads in the risky asset distribution. The four-step increase in risk occurs once per frame so within each frame each respondent is either consistent with the weak restriction or not. We code the variable w_{fr} as 1 if respondent r is inconsistent with the weak restriction in frame f , and 0 otherwise. The test now proceeds precisely as for the strong restriction but with one test per frame rather than four. The interpretation is also the same. Specifically, a failure to reject the null hypothesis indicates that inference about preferences from DCEs are susceptible to framing and not informative about respondents' consistency with the weak

restriction of expected utility theory.

3 Results

3.1 Are violations of concavity predictable?

We compute the likelihood of a violation of the strong restriction of expected utility by estimating joint (frameset level) and independent (frame by frame) probabilities that a respondent chooses the 50:50 portfolio as worst at least j times. As noted above, there are eight possible patterns of violations of this condition across the set of three frames. Consequently we estimate the probability that a respondent fits one of these eight patterns using a multinomial logit model, with three covariates: age group, reported retirement savings, and score on a set of basic numeracy questions. Table 7 sets out two subsets of survey questions from which we draw covariates for these models. The first set of questions relate to several aspects of the respondent’s awareness of their own retirement savings account (superannuation account), and financial education. Responses to question P2 on superannuation amount were included in the MNL and logit models after recoding into four groups: less than \$19999, \$20,000-\$79999, \$80,000-\$499999, \$500,000 and over. (The median Australian retirement savings account (superannuation) balance is around \$70,000). The numeracy questions are drawn from Gerardi et al (2010) and are designed to test basic concepts such as fractions, percentages, division, multiplication and simple probability. We fitted a factor model to these questions, and used the fitted factor loadings to create a numeracy score for each respondent. (The survey asked further questions designed to test financial literacy, but we did not find them relevant to the models reported here.) Ages were coded into three groups (18-34 years, 35-54 years and over 55 years.).

Table 8 reports likelihood ratio tests of the restriction of independence of concavity violations across frames. In all but one test (the category ‘at least four violations for frameset 1,2,9’) the restriction of independence is rejected in favour of the alternative hypothesis of

predictability. In other words, respondents who choose the 50:50 portfolio as worst tended to do so across all risk frames which they observed, and similarly, respondents who did not violate concavity were unlikely to do so when risk was presented differently in later frames. Of the covariates used in estimating the joint MNL model, few were significant except that a poor numeracy is significant predictor of a high level of concavity violations.

[Insert Table 8 here]

3.2 Do respondents predictably fail to perceive mean-preserving spreads?

We repeat the testing process for violations of the weak restriction of expected utility, that is, we record those respondents who increase their retirement account risk exposure as the standard deviation of the risky return increases. These respondents do not conform with the prediction that expected utility maximizers prefer (second order) stochastically dominant returns distributions. Again, there are eight possible patterns of violations of this condition across the sets of three frames viewed by each respondent. We test whether the probabilities of the each of the eight possible outcomes are equal to the product of the independent (frame by frame) probabilities. Since respondents can violate this weak restriction only once per frame, we need to estimate one MNL model for the eight outcomes. Table 9 reports likelihood ratio tests of the restriction of independence against the alternative model. All tests clearly reject independence. Again we find that poor numeracy is a significant predictor of high and consistent violations of this weak restriction of expected utility.

[Insert Table 9 here]

Overall, we confirm that respondents who choose inconsistently with standard risk averse preferences do so predictably across the three risk frames presented in the survey. Likewise, respondents whose choices conform with concave, risk averse utility choose consistently across frames. We note a large variation in the degree to which respondents conform to these extremes, and we discuss this variation in more detail below.

3.3 Does consistency with expected utility vary across frames?

We are interested in finding if any risk presentations from within the set of nine frames lead to significantly more consistent behavior by respondents. Regulators responsible for consumer financial decision making and the financial services industry are investigating how to convey risk information in modes that are intelligible, standardized and easily comparable across products. As noted earlier, the frames used in this survey are derived from product disclosure statements of financial service providers and include upside and downside quantiles, frequencies of exceeding or failing benchmarks and graphs. The Australian Prudential Regulation Authority (APRA) recently released a guide on risk presentation to superannuation providers that specified risk framing in terms of the expected frequency of negative returns over a 20 year period. The APRA (2010) recommendation closely matches the wording of risk frame 6 in this survey. Frame 5 presents the same information but focussing on the frequency of positive returns. In this section we compare the nine risk frames to assess whether any, including frame 6, are associated with significantly more or less violations of standard risk averse preferences.

To begin, we test whether the proportions of choices going to the safe, medium and risky accounts vary significantly across frames. Table 10 presents the percentages of all best and worst choices given to each account, across the whole survey ('all') and frame by frame. The final column reports that test statistic for the chi-square test of joint equality of the frame percentages with the aggregate percentage given in the first ('all' column). We also tested each individual frame percentage for equality with the aggregate percentage. Those frames where the percentage of best or worst choices significantly deviated from the aggregate percentage at 10% or less are marked with an asterisk.

[Insert Table 10 here].

The null hypothesis of joint equality is rejected in 5 of 6 rows with most variation associated with frames 4 and 5. A significantly larger number of violations of the strong restriction of expected utility (where the medium account was chosen as worst) occurred when respon-

dents viewed frame 5, which presented risk in terms of the expected number of years from 20 where returns would exceed zero. Frameset B was comprised of frames 3, 4 and 9. This frameset produced significant reversals in allocation patterns. Best choices going to the safe account were significantly lower than expected in frames 3 and 9, and significantly higher than expected in frame 4. These choices shift in the direction of the information offered in these frames: frame 3 presented the chance of returns above the 95th quantile, frame 4 presented the chance of returns below the 5th quantile and frame 9 presented the graph.

A detailed analysis of individual choices allows us to make an exhaustive allocation of respondents to four mutually exclusive groups which are increasingly consistent with CRRA expected utility. Table 11 reports the number of respondents who fell into each group for each frame. We see that around 20% of respondents displayed violations of concavity ('concave inconsistent'), around 17% were consistent with concavity but did not perceive mean-preserving spreads and chose inconsistently with second order stochastic dominance ('concave consistent'). Around 32% satisfied the strong and weak restrictions of expected utility but did not match CRRA functional form ('ssd consistent') and around the the same proportion conformed to CRRA preferences. The final column in Table 11 reports a test for the joint equality of the proportions of respondents under each frame with the proportion found under frame 9, for each group. We use frame 9 as the reference level for this test since all respondents saw frame 9. Joint equality is rejected for the concave inconsistent group and the CRRA consistent group. Other cells marked with an asterisk indicate significant difference from frame 9 when each frame is tested individually. Higher concavity inconsistency is significantly more likely under frames 5, 6, and 8, and lower consistency with second order stochastic dominance under frame 5. Respondents are more likely to conform with standard CRRA preferences when presented with lower tail events (frame 4) but we noted earlier a higher level of choice reversals for respondents viewing the 3,4,9 frameset.

[Insert Table 11 here]

Finally, we consider the number of respondents who show consistency or inconsistency

across all three frames. Table 12 shows the number of respondents who saw each frameset, and then reports the number of those who conformed to increasingly stringent tests of expected utility consistency across their whole frameset. We note that lower overall consistency is linked to frameset C, which presented risk as expected years out of 20 when returns would exceed or fall below zero. Significantly higher perception of mean-preserving spreads emerged under the range frames which were presented in frameset A.

[Insert Table 12 here]

4 Conclusion

Changes to retirement saving systems mean that investment decisions involving risky financial assets are the responsibility of ordinary individuals. The financial services industry and academic researchers have recognised the key role of investment risk, and have devised many techniques to measure risk preferences, some for a mass market. Standard methods, such as risk profiles or lottery tasks assume that preferences are measurable, stable and predictable. At the same time, studies of framing and choice architecture have reinforced the power of presentation mode to direct (or even reverse) decisions, particularly decisions with uncertain outcomes.

The discrete choice experiment reported here indicates that framing is influential but not decisive. Respondents to the survey repeat the same simple retirement savings investment task 12 times, under three different modes of risk presentation, at four increasingly variable levels for returns risk. We isolate survey respondents whose choices do not conform with very basic features of risk aversion, that is concavity and perception of increasing risk. Failures to conform with concavity are predictable across frames, as are failures to conform with perceptions of mean-preserving spreads. In other words, we demonstrate that this discrete choice experiment, and most likely other similar experimental surveys, can identify individuals with specific preference patterns. Framing does not entirely determine portfolio

allocation choices.

Second, we expand this finding by showing that not all frames are linked with equally consistent choices. Around one fifth of respondents violated both concavity and mean-preserving spread perception, and around two thirds conformed with both, but we detect significant variation in these patterns by frame. Of the frames we investigate, we find the least consistent choices for the frames worded ‘On average, positive (negative) returns occur (20-x) (x) years in every 20’, a common mode of presentation by financial service providers and fund managers. We also demonstrate that respondents’ choices are more likely to switch from risky to conservative when they are given a sequence of frames that switch between tail quantiles, that is, those frames worded ‘There is a 1 in 20 chance of a return above (below) y%’. We find that most consistency is linked to range information that numerically specifies, or graphs, both upside and downside risks.

References

Agnew, J., Anderson, I., Gerlach, J., Szykman, L., 2008. Who chooses annuities? An experimental investigation of the role of gender, framing and defaults. *American Economic Review* 98(2), 418-22.

Anagol, S., Gamble, K.J. 2008. Presenting investment results asset by asset lowers risk taking. Yale University Working paper.

Andrews, P. 2009. Did life and pensions 'disclosure' work as expected? Occasional Paper Series No. 32, UK Financial Services Authority, April 2009, London, UK.

APRA 2010. Good Practice Guidance - Investment Risk Description. Letter to trustees of APRA Regulated Superannuation Funds, 29 June 2010, Australian Prudential Regulation Authority, 400 George Street, Sydney, Australia.

Arnold, B.C. 1980. *Majorization and the Lorenz Order: Brief Introduction (Lecture Notes)*. Springer Verlag, Berlin.

Benartzi, S., Thaler, R.H. 1999. Risk aversion or myopia? Choices in repeated gambles and retirement investments. *Management Science* 45(3), 364-81.

Brown, J., Kling, J., Mullainathan, S., Wrobel, M., 2008. Why don't people insure late life consumption? A framing explanation of the under-annuitization puzzle. *American Economic Review* 98(2).

European Commission 2009a. UCITS Disclosure Testing Research Report, prepared for the European Commission by IFF Research and YouGov, June 2009.

European Commission 2009b. Update on Commission Work on Packaged Retail Investment Products, December 2009.

Friend, I. Blume, M.E. 1975. The demand for risky assets. *American Economic Review* 65(5), 900-922.

Gerardi, K., Goette, L., and Meier, S., 2010. Financial literacy and subprime mortgage delinquency: evidence from a survey matched to administrative data. Federal Reserve Bank of Atlanta Working Paper Series No 2010-10, Atlanta, GA.

Goldstein, D G, Eric J J., Sharpe, W.F. 2008. Choosing outcomes versus choosing products: consumer-focused retirement investment advice, *Journal of Consumer Research* 35, 440-456.

Haisley, E., Kaufmann, C., Weber, M. 2010. How much risk can I handle? The role of experience sampling and graphical displays on one's risk appetite. Paper presented at the First Boulder Conference on Consumer Financial Decision Making, Boulder, Colorado, June 2010.

Holt, C.A., Laury, S.K. 2002. Risk aversion and incentive effects. *American Economic Review* 92(5),1644-1655.

Ingersoll, J. 1987. *Theory of Financial Decision Making*, Rowman and Littlefield, Lanham, USA.

Kimball, M. S., Sahm, C. R., Shapiro, M. D. 2008. Imputing risk aversion from survey responses. *Journal of the American Statistical Association* 103(483), 1028-1038.

Minister for Financial Services 2010. Explanatory Statement, Select Legislative Instrument 2010 (No.135), Corporations Amendment Regulations 2010 (No.5), Canberra, Australia.

Organisation for Economic Cooperation and Development 2008. Improving financial education and awareness on insurance and private pensions, OECD, Paris.

Rubaltelli, E., Rubichi, S., Savadori,L., Tadeschi,M., Ferretti, R. 2005. Numerical information format and investment decisions: Implications for the disposition effect and status quo bias. *Journal of Behavioral Finance* 6(1), 19-26.

Tversky, A., Kahneman, D. 1981. The framing of decisions and the psychology of choice. *Science* 211(4481), 453-458.

Tversky, A., Kahneman, D. 1986. Rational Choice and the Framing of Decisions. *Journal of Business* 2(2), 73-85.

Vlaev, I., Chater, N., Stewart, N. 2009. Dimensionality of risk perception: Factors affecting consumer understanding and evaluation of financial risk. *Journal of Behavioral Finance* 10(3), 158-181.

Appendix A: Instructions to survey subjects.

The Australian Government is concerned about the complexity of superannuation arrangements and is looking for ways of simplifying superannuation investment choices. One possibility is to offer only three investment options for all superannuation accounts. Each investment option has a different expected rate of return (the average rate at which your investment will grow each year), and a different amount of investment risk (year to year UPSIDE and DOWNSIDE variation in the return to your investment).

The options are

Option A: All (100%) of your superannuation account is invested in a guaranteed bank deposit with a fixed rate of interest paid each year.

Option B: Your superannuation account will be divided half and half (50%/50%) between the bank account and growth assets. You can anticipate that savings in this option will grow faster than the bank deposit (Option A), but will grow more slowly and be less risky than only choosing growth assets (Option C).

Option C: All (100%) of your superannuation account is invested in assets like shares and property. On average, you can anticipate that savings in this option will grow at a faster rate than in the bank deposit (Option A) but without a guarantee. There is some risk that your account will grow faster or slower than average if you choose this option.

We are going to show you 12 sets of these options for investing your superannuation. Each set includes 3 investment options like the ones described above. Each investment option has

a average rate of return and investment risk. The average rates of return stay the same in each of the twelve sets; only the risk will change. Remember that more risk of high returns also means more risk of low returns.

What we want you to do is simple. There are two questions to ask about each set of options::

1. If these superannuation options were available for you to invest your money today, which one of the three would you be most likely to choose?

2. If these superannuation options were available for you to invest your money today, which of the three would you be least likely to choose?

Your choices will inform government and industry about better ways to simplify superannuation arrangements.

Appendix B: Proof of SSD in mean-preserving spreads of lognormal random variables.

If $y_i \succ y_j$, meaning that y_i second order stochastically dominates y_j , then $cy_i \succ cy_j$ for any constant $c > 0$. Let y_i and y_j be lognormal random variables with the same mean, so that $E(y_i) = E(y_j)$. Consider transformations of y_i and y_j , $\tilde{y}_i = y_i/E(y_i)$ and $\tilde{y}_j = y_j/E(y_j)$ which will have identical means of 1 and variances $var(\tilde{y}_i) = \exp(\sigma_i^2 - 1)$ and $var(\tilde{y}_j) = \exp(\sigma_j^2 - 1)$ where σ_k^2 is the variance of $\ln(y_k)$ $k = i, j$.

Since $\ln(\tilde{y}_k) \sim N(0, \sigma_k^2)$, we know from Arnold (1980, problem 10 p.42), that \tilde{y}_k has a Lorenz curve of the form $L = \Phi(\Phi^{-1}(s) - \sigma_k)$ where Φ is the standard normal cumulative density function. Let ϕ be the normal probability density function.

Differentiating L with respect to σ_k gives

$$\frac{dL}{d\sigma_k} = -\phi(\Phi^{-1}(s) - \sigma_k)$$

which is negative, so that an increase in σ_k leads to a lowering of the Lorenz curve. By Arnold (1980, p85, Theorem 6.9), if A has a lower Lorenz curve than B then $A/E(A)$ is second order stochastically dominated by $B/E(B)$.

In our problem, we hold $E(y_i) = \exp(\mu_i + 1/2\sigma_i^2) = \exp(\mu_j + 1/2\sigma_j^2) = E(y_j)$. We also have $\ln(y_i) \sim N(\mu_i, \sigma_i^2)$ and $\ln(y_j) \sim N(\mu_j, \sigma_j^2)$ by lognormality of y_i and y_j .

Define $Z_i = (\ln(y_i) - \mu_i) \sim N(0, \sigma_i^2)$ and $Z_j = (\ln(y_j) - \mu_j) \sim N(0, \sigma_j^2)$ where $\sigma_j^2 > \sigma_i^2$. From above we have Lorenz domination for $\exp(Z_i)$ over $\exp(Z_j)$. This implies second order stochastic dominance of $\exp(Z_i)/E[\exp(Z_i)]$ over $\exp(Z_j)/E[\exp(Z_j)]$. By back-substitution, we see that $\exp(Z_i)/E[\exp(Z_i)] = y_i/E(y_i)$. Hence $y_i/E(y_i)$ second order stochastically dominates $y_j/E(y_j)$ and since $E(y_i) = E(y_j)$ it follows that $y_i \succ y_j$.

Table 1

SAMPLE DEMOGRAPHIC CHARACTERISTICS ^a

	Survey respondent population (%)	General Australian population (%)		Survey respondent population (%)	General Australian population (%)
<i>Gender</i>			<i>Industry</i>		
Male	49.9	50.1	Agriculture, forestry & fishing	0.98	3.17
Female	50.1	49.9	Mining	2.07	1.21
<i>Age (as % of 18-65 year pop'n)</i>			Manufacturing	4.79	10.74
18-34 years	35.8	37.4	Electricity, gas, water & waste services	1.20	1.01
35-54 years	43.2	43.6	Construction	5.01	8.00
55-65 years	21.1	18.9	Wholesale trade	2.18	4.47
<i>Marital status</i>			Retail trade	9.80	11.65
Not living with long term partner	42.94	46.72	Accommodation & food services	3.70	6.49
Married or living with long term partner	57.06	53.28	Transport, postal & warehousing	4.79	4.82
<i>Work status</i>			Information media & telecommunications	5.45	1.99
Employed full-time	51.72	40.79	Financial & insurance services	7.19	3.93
Employed part-time	23.52	18.79	Rental, hiring & real estate services	1.20	1.74
Unemployed	3.44	3.53	Professional, scientific & technical services	6.86	6.79
Not in the labour force	21.31	36.89	Administrative & support services	5.56	3.23
<i>Occupation</i>			Public administration & safety	3.70	6.86
Clerical and administrative worker	20.81	15.00	Education & training	11.22	7.87
Community and personal service worker	3.59	8.81	Health care & social assistance	11.55	10.78
Labourer	5.66	10.46	Arts & recreation services	0.98	1.44
Machinery operators and drivers	3.49	6.64	Other services	11.76	3.81
Manager	10.89	13.21			
Professional	31.15	19.84			
Sales worker	7.63	9.84			

Technicians and trades worker	7.41	14.38			
Other	9.37	1.82			
<i>Highest year of school completed</i>			<i>Number of people living in household</i>		
Year 12 or equivalent	70.49	46.87	1	10.98	24.36
Year 11 or equivalent	9.10	11.08	2	34.67	34.10
Year 10 or equivalent	17.13	25.36	3	22.95	15.79
Year 9 or equivalent	2.13	7.74	4	19.92	15.73
Year 8 or below	1.07	7.98	5	7.62	6.88
Did not go to school	0.08	0.96	6 or more	3.85	3.13
<i>Highest non-school qualification</i>			<i>Number of people in family fully/partially financially supported^b</i>		
Postgraduate or equivalent	13.59	6.58	None	45.66	50.18
Graduate Diploma and Graduate Certificate from University or equivalent	8.43	3.64	1	23.28	17.24
Bachelor Degree or equivalent	30.77	29.33	2 or more	31.06	32.58
Advanced Diploma and Diploma from University/TAFE equivalent	20.65	18.01	<i>Net wealth (individual)</i>		
Certificate or equivalent	26.55	42.43	Under \$10,000	13.93	-
<i>Annual total household gross income (before tax)</i>			\$10,000 - \$99,999	27.54	18.21
Less than \$18,200 pa (i.e. \$350 a week)	3.28	4.72	\$100,000 - \$999,999	35.00	62.44
\$18,200-\$72,799 pa (i.e.\$499-1,399 a week)	34.33	39.49	\$1,000,000 or over	6.80	19.35
\$72,800-\$129,999 pa (i.e. \$1,400-\$2,499 a week)	31.64	28.44	Prefer not to answer	16.72	-
\$130,000 pa (i.e. \$2,500 a week) or more	16.88	14.93			
Prefer not to answer	13.87	12.42 ^a			
<i>Household make-up</i>					
Couple family with no children	25.49	25.67			
Couple family with children	37.46	31.20			
One parent family	6.48	10.87			
Other family household	3.44	1.18			
Single person household	13.77	23.38			
Group household (i.e. shared)	13.36	7.68			

a Source for population statistics: Australian Bureau of Statistics Census of Population and Housing & Household Wealth and Wealth Distribution, Australia, 2005-2006

Table 2

FRAMES FOR INVESTMENT RISK

F1 : There is a 9 in 10 chance of a return between x% and y%.

F2 : There is a 1 in 10 chance of a return outside x% and y%.

F3 : There is a 1 in 20 chance of a return above y%.

F4 : There is a 1 in 20 chance of a return below x%.

F5 : On average, positive returns occur (20-x) years in every 20.

F6 : On average, negative returns occur x years in every 20.

F7 : On average, returns above the bank account occur (20-x) years in every 20.

F8 : On average, returns below the bank account occur x years in every 20.

F9 : Three options are shown in the chart below.

- Option A: 100% bank account, the rate of return is always exactly x% (black dot)
- Option B: 50% bank account & 50% growth asset, there is a 9 in 10 chance of a rate of return within the light blue box
- Option C: 100% growth asset, there is a 9 in 10 chance of a rate of return within the dark blue box

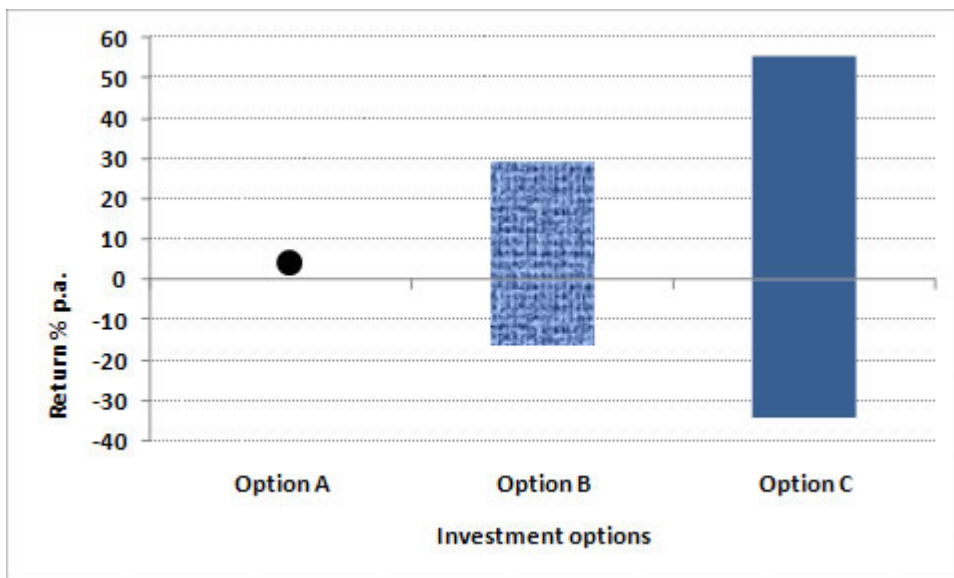


Table 3

INVESTMENT ACCOUNT RATES OF RETURN AND RISK LEVELS.

Level	Annual rates of return		Volatility	Lognormal parameters	
	Safe: E(x)	Risky: E(y)	Std dev(y)	μ	σ
1	0.02	0.045	0.12	.038	0.11
2	0.02	0.045	0.16	.032	0.15
3	0.02	0.045	0.20	.026	0.19
4	0.02	0.045	0.28	.009	0.26

Table shows expected rate of return to safe and risk accounts and standard deviation to risky returns distribution over four levels, and the corresponding lognormal parameters. The 50:50 account had an expected return of 3.25% and a standard deviation of 0.5(Std dev (y)).

Table 4

NUMERICAL VALUES FOR RISK FRAMES, LEVELS 1-4.

Frame	Option	Lower (L)/ Upper(U)	Standard deviation of returns distribution			
			12%	16%	20%	28%
			%	%	%	%
1	50:50 Account	L	-6	-9	-11.5	-16.5
		U	14	17.5	21	29
	Risky Account	L	-14	-19.5	-25	-34.5
		U	25.5	32.5	40	55.5
2	50:50 Account	L	-6	-9	-11.5	-16.5
		U	14	17.5	21	29
	Risky Account	L	-14	-19.5	-25	-34.5
		U	25.5	32.5	40	55.5
3	50:50 Account	U	14	17.5	21	29
	Risky Account	U	25.5	32.5	40	55.5
4	50:50 Account	L	-6	-9	-11.5	-16.5
	Risky Account	L	-14	-19.5	-25	-34.5
5	50:50 Account	-	14	13	12	11
	Risky Account	-	13	12	11	10
6	50:50 Account	-	6	7	8	9
	Risky Account	-	7	8	9	10
7	50:50 Account	-	11	10.5	10	9.5
	Risky Account	-	11	10.5	10	9.5
8	50:50 Account	-	9	9.5	10	11.5
	Risky Account	-	9	9.5	10	11.5

Table 5

ILLUSTRATIVE INVESTMENT CHOICE TASK

Features of Options	Option A	Option B	Option C
Option type	100% Bank account	50% Bank account & 50% Growth assets	100% Growth assets
Average annual rate of return (above inflation)	2%	3.25%	4.5%
Level of investment risk	No risk	There is a 1 in 20 chance of a rate of return <u>above</u> 14%	There is a 1 in 20 chance of a rate of return <u>above</u> 25.5%

If these superannuation options above were available for you to invest your money today

1. Which one of the three would you be **most likely** to choose?

- Option A
- Option B
- Option C

2. Which one of the three would you be **least likely** to choose?

- Option A
- Option B
- Option C

Table 7

PRELIMINARY SUPERANNUATION AND NUMERACY QUESTIONS

P1: Which fund manages your main superannuation account in Australia? (Responses: Please specify name of fund; Don't know.)

P2: Which of the following ranges best describes the total amount you currently have in all your superannuation accounts in Australia? (Responses: 13 ranges from 'Under \$10,000 to \$5,000,000 or over.)

P3: On a scale of 1 to 7, where 1 means very low and 7 means very high, how would you assess your understanding of finance?

P4: How much of your financial education was devoted to financial education, such as commerce, business studies, finance or economics? (Responses: A lot; Some; A little; Hardly at all.)

P5: Did any of the firms you have worked for (including your current employer) offer financial education programs, for example retirement seminars? (Responses: Yes; No; Not applicable.)

P6: Have you paid for professional financial advice about your superannuation over the past twelve months? (Responses: Yes; No.)

Numeracy skills

Q1: In a sale, a shop is selling all items at half price. Before the sale, a sofa costs \$300. How much will it cost in the sale? (Answers: \$150; \$300; \$600; Do not know; Refuse to answer.)

Q2: If the chance of getting a disease is 10 per cent, how many people out of 1,000 would be expected to get the disease? (Answers: 10; 100; 1000; Do not know; Refuse to answer.)

Q3: A second hand car dealer is selling a car for \$6,000. This is two-thirds of what it cost new. How much did the car cost new? (Answers: \$4,000; \$6,600; \$9,000; Do not know; Refuse to answer.)

Q4: If 5 independent, unrelated people all have the winning numbers in the lottery and the prize is \$2 million, how much will each of them get? (Answers: \$40,000; \$400,000; \$500,000; Do not know; Refuse to answer.)

Q5: If there is a 1 in 10 chance of getting a disease, how many people out of 1,000 would be expected to get the disease? (Answers: 10; 100; 1000; Do not know; Refuse to answer.)

Table 8

LIKELIHOOD RATIO TEST OF INDEPENDENCE OF STRONG RESTRICTION VIOLATIONS ACROSS
FRAMES

	Likelihood ratio test statistics, $\chi^2(16)$			
	Concavity violations per frame			
	at least 1	at least 2	at least 3	at least 4
Frameset 1,2,9	187.79	54.87	24.93	0.21
Frameset 3,4,9	104.74	42.59	50.37	40.43
Frameset 5,6,9	120.35	59.83	49.93	40.21
Frameset 7,8,9	123.64	97.15	63.62	68.80

Table reports $\chi^2(16)$ statistics from the test of independence of concavity violations across frames. Concavity is violated when a respondent chooses the 50:50 portfolio as worst. Joint probabilities (probability that a respondent displays one of eight possible patterns of concavity violations over three frames) are computed from multinomial logit models estimated with covariates, age, retirement savings amount, and numeracy score. Independent probabilities of violations by frame are computed from logit models with common covariates where the binary variable indicates concavity violation by a respondent at level $j=1, \dots, 4$.

Table 9

LIKELIHOOD RATIO TEST OF INDEPENDENCE OF WEAK RESTRICTION VIOLATIONS ACROSS
FRAMES

	Likelihood ratio test statistics, $\chi^2(16)$
Frameset 1,2,9	138.81
Frameset 3,4,9	55.27
Frameset 5,6,9	109.86
Frameset 7,8,9	88.86

Table reports $\chi^2(16)$ statistics from the test of independence of failures to perceive mean-preserving spreads of the risky returns distribution across frames. Joint probabilities (probability that a respondent displays one of eight possible patterns of violations over three frames) are computed from multinomial logit models estimated with covariates, age group, retirement savings amount, and numeracy score. Independent probabilities of violations of this restriction by frame are computed from logit models with common covariates where the binary variable indicates concavity violation by a respondent.

Table 10

PERCENTAGES OF BEST AND WORST CHOICES PER ACCOUNT BY FRAME

	all	F1	F2	F3	F4	F5	F6	F7	F8	F9	$\chi^2(9)$
best	%	%	%	%	%	%	%	%	%	%	
S	24	27	31	16*	50*	28	22	21	33	16*	40.45*
M	46	51	48	54	35	34*	44	52	41	47	9.19
R	30	22	21	30	15*	38	33	27	26	37	17.30*
worst											
S	49	43	39	57	28*	52	50	49	44	56	14.80*
M	9	6	8	8	7	20*	13	10	13	7	19.22*
R	42	51	53*	36	65*	29*	37	41	42	37	23.50*

Table reports percentage of best and worst choices allocated to the safe, medium and risky accounts over the entire survey ('all') and by 9 frames as described in Table 2. The final column reports the test statistic for the chi-square test of the joint equality of percentages across frames for each row, where the reference level is the aggregate percentage ('all'). Individual percentages marked with an asterisk are significantly different from the aggregate percentages according to a chi-square test with 1 d.f.

Table 11

NUMBER OF RESPONDENTS VIOLATING STRONG AND WEAK RESTRICTIONS, BY FRAME

	F1	F2	F3	F4	F5	F6	F7	F8	F9	$\chi^2(8)$
total	300	300	299	299	297	297	303	303	1199	
respondents										
concave	46	55	40	44	93*	75*	60	74*	177	35.1*
inconsistent										
concave	60	37	65	35	47	52	53	48	210	5.2
consistent										
ssd	106	105	116	73	63*	93	115	90	396	8.9
consistent										
crra	88	103	78	147*	94	77	75*	91	416	15.0*
consistent										

Table groups respondents into four mutually exclusive categories depending on conformity with expected utility. The final column reports the test statistic for the chi-square test of the joint equality of proportions across frames for each row, where the reference level is the percentage for frame 9. Individual cells marked with an asterisk are significantly different from the frame 9 percentage according to a chi-square test with 1 d.f.

Table 12

NUMBER OF RESPONDENTS CONFORMING TO STRONG AND WEAK RESTRICTIONS ACROSS
ALL THREE FRAMES

	Frameset A	Frameset B	Frameset C	Frameset D	$\chi^2(3)$
total respondents	300	299	297	303	
concave consistent	226	230	174*	203	9.55*
ssd consistent	152*	123	102*	120	10.34*
crra consistent	48	25*	34	40	7.7*

Table reports numbers of respondents who satisfied the restrictions of expected utility for all three frames in their frameset. The final column reports the test statistic for the chi-square test of the joint equality of proportions across framesets for each row, where the reference level is the average for that row. Individual cells marked with an asterisk are significantly different from the average.