PERUVIAN ECONOMIC ASSOCIATION

Measuring efficiency and risk preferences in dynamic portfolio choice

Jacopo Magnani

Jean Paul Rabanal

Olga A. Rud

Yabin Wang

Working Paper No. 165, April 2020

Measuring efficiency and risk preferences in dynamic portfolio choice

Jacopo Magnani [*] Jean Paul Rabanal[†] Olga A. Rud [‡]

Yabin Wang[§]

April 2, 2020

Abstract

This paper uses non-parametric methods to study the efficiency (Dybvig, 1988) and risk-profile (Varian, 1988) of dynamic portfolio choices. We design an experiment which varies the number of states (complexity), and includes an equivalent static Arrow-Debreu problem. The results suggest that complexity reduces efficiency, as does lower cognitive ability. Efficiency is also lower in the static problem, and in the dynamic task it is mostly driven by a form of stop-loss strategy. Further, we find that a representative agent exhibits decreasing absolute risk aversion and constant relative risk aversion, despite significant individual heterogeneity.

[*]Department of Finance, EM Lyon.

[†]Department of Banking and Finance, Monash University. jeanpaul.rabanal@monash.edu (Corresponding author)

[‡]Department of Economics, RMIT.

[§]Hong Kong Monetary Authority. The views expressed are those of the author and do not necessarily reflect the official positions of the Hong Kong Monetary Authority.

1 Introduction

Understanding the sources of inefficiency in dynamic portfolio choices has important implications for theory, policy-making and market participants. Inefficient portfolio choices have a negative effect on financial well-being, and may enhance the disparity in performance across investors. Sound financial advice, on the other hand, can help mitigate some of these inefficiencies. To date, there is no systematic way to identify the size and source of possible inefficiencies, though some studies point to irrational behaviours, such as naïve diversification. A major hurdle in identifying the (in)efficient strategies is the inability to specify the underlying risk preferences of investors. In this paper, we propose a novel experiment which can simultaneously quantify the efficiency loss of dynamic portfolio strategies, using the method of Dybvig (1988), and measure risk preferences, applying the techniques of Varian (1988).

In our experiment, we ask subjects to allocate their endowed wealth between a risky and a risk-free asset. In the dynamic task, the allocation decision is made for each period, and every possible state of a binomial tree, to effectively create a contingency plan. Consumption occurs only in the final period, where each terminal state has an equal probability of being drawn. In order to evaluate whether a portfolio strategy is efficient, we follow Dybvig (1988) who defines a choice as efficient if the final wealth in the terminal states is non-increasing with respect to state-prices.[1] Inefficiency is measured as the extra cost above the minimum expenditure necessary to achieve the chosen distribution of final wealth. This approach requires no strong parametric assumptions about the utility function, nor knowledge of investor risk preferences.[2]

[1]If this condition is violated, then it is possible to propose an allocation that first-order stochastically dominates (FOSD) the chosen one. This holds under the assumption of equiprobable states, which we implement in our design.

[2]Parametric models of portfolio choice may fail to separate true sub-optimal choices from misspecified investor preferences. Standard preferences cannot explain inefficiencies identified using Dybvig (1988), who shows that a portfolio strategy is efficient if, and only if, it maximizes a strictly increasing von Neumann-Morgenstern or Expected Utility function over terminal wealth. Bernard et al. (2014)

This study complements both theoretical and empirical literature in two ways. First, measuring inefficiency is difficult in the field because it requires knowledge of state contingent choices, which are rarely observable. Instead, one needs to rely on the distribution of returns observed in markets.[3] Second, data on state contingent choices, allows us to analyze the patterns and sources of inefficient strategies. In this regard, we focus on two main sources of inefficiencies: (i) the length of investment horizon, where complexity increases with the number of states, and (ii) path-dependent strategies, such as stop-loss.[4] To evaluate the impact of the former, we implement a dynamic portfolio problem with three and four periods, resulting in eight and 16 terminal states, respectively. To evaluate the impact of the latter, we implement two analogous static Arrow-Debreu problems, with eight and 16 states, respectively. This setup is formally equivalent to the dynamic portfolio decision in our experiment, but without the channel for path-dependence, and thus allows for a clean comparison across treatments.

Our results show that most of the choices made by subjects are efficient, and therefore consistent with the maximization of expected utility (EU). Among inefficient choices, losses are higher when subjects (i) face a greater array of terminal states, (ii) make allocation decisions in the static Arrow-Debreu environment, and (iii) have a lower score on the cognitive reflection test (CRT). While the first result confirms our initial hypothesis regarding the effect of increasing complexity, the second result suggests that breaking down the investment problem into intermediate steps can improve the efficiency of portfolio choices, relative to the static environment. Our result on the effect of cognitive skills is consistent with previous evidence that higher CRT scores are

generalizes the approach of Dybvig (1988) and Cox and Leland (2000) by studying state-dependent constraints.

[3]The approach of Dybvig (1988) has been applied empirically to study the performance of hedge funds (Amin and Kat, 2003). There is a relatively recent literature that studies portfolio choices using the criterion of stochastic dominance, mentioned in the next section.

[4]Stop-loss strategy prescribes a sale of assets when the price falls below some predetermined level. Dybvig (1988) shows stop-loss strategy is inefficient.

positively correlated with performance in laboratory asset markets (such as Corgnet et al., 2018) and choice under risk (Choi et al., 2014). Additional analyses on the pattern of strategies from the dynamic environment point to a form of stop-loss inefficiency, at a particular node in the tree.

Using the data on individual portfolio allocations, we can also categorize investor risk profiles, following the techniques of Varian (1988), who abstains from making strong assumptions about the nature of the utility function. On aggregate, we find that a representative agent's risk preference is consistent with decreasing absolute risk aversion (DARA) and constant relative risk aversion (CRRA). At the individual level, there is significant heterogeneity. An important share of subjects displays non-monotonic risk-aversion, in contrast to the monotonic assumption frequently used in standard parametric models. Thus, testing efficiency under CARA or hyperbolic absolute risk aversion utility (HARA) may lead us to overlook individual heterogeneity that is present in portfolio choices.

Our findings also have important implications for industry. For example, diversification, promoted as a golden rule of investment practice, may not be enough to improve the financial well-being of individuals. We show that when decisions are dynamic, people sometimes embrace a form of stop-loss strategy which leads to sub-optimal outcomes. However, despite the known inefficiency of stop-loss orders, some financial advisors still promote their use.[5]

2 Related literature

There exists a long tradition of experiments and empirical studies seeking to analyze portfolio efficiency with non-parametric methods. These studies use lab experiments

[5] For example, see https://www.schwab.com/resource-center/insights/content/be-defensive-use-stop-orders. Extracted on March 3rd, 2020.

(Baltussen and Post, 2011, Levy et al., 2004, Kroll et al., 1988, Kroll and Levy, 1992, Sundali and Guerrero, 2009, Biais et al., 2017), surveys (Bateman et al., 2016) and market data (Post, 2003, Post and Levy, 2005, Kopa and Post, 2009, Amin and Kat, 2003). Most papers in this literature adopt the static portfolio choice framework. To the best of our knowledge, we are the first to use Dybvig (1988)'s non-parametric methods to evaluate dynamic portfolio choices.

Our paper also contributes to the literature that elicits risk preferences through financial allocation decisions in the lab (see Gneezy and Potters, 1997, Choi et al., 2007, Friedman et al., 2019, Goldstein et al., 2008, Kaufmann et al., 2013). Most of these papers use a static allocation task with two state-contingent securities, whereas our experiment provides a richer data set. We are also the first to use Varian (1988)'s method to non-parametrically estimate risk aversion. Consistent with previous findings, our results suggest that the format of the elicitation task matters.[6] When subjects face a binomial tree diagram in the dynamic task of our experiment, they exhibit less risk aversion relative to the price list environment in the static task. Indeed, we find that many subjects hedge their allocation decision in the static task, behavior consistent with the equal allocation or $1/n$ heuristic across funds (Huberman and Jiang, 2006; Benartzi and Thaler, 2001) as well as with other economic decisions (e.g. Sonnemann et al., 2013, and Rud et al., 2019).[7]

In terms of literature on risk preferences, perhaps the most closely related paper to our work is that of Brocas et al. (2019), where they implement a dynamic portfolio problem in the lab and employ a parametric approach (HARA) to evaluate the risk preferences of subjects. Their results show that while subjects are quite heterogeneous,

[6] Friedman et al. (2019) also compare different risk elicitation tasks, including a jar interface, a budget line in a two-state economy, and a multiple price list (Holt and Laury, 2002), and find that (i) while violations of FOSD are not infrequent, they are generally quite small, and (ii) standard EU preferences cannot explain the observed variations in preferences across tasks.

[7] Task complexity can also lead to different behavioral rules (e.g. see d'Acremont and Bossaerts, 2008).

more than half exhibit decreasing absolute risk aversion and increasing relative risk aversion (DARA-IRRA) preferences.[8] At the aggregate level the results are mixed: Brocas et al. (2019) find evidence of CRRA, but this occurs because subjects who exhibit DRRA accumulate more wealth than those who exhibit IRRA and thus have more impact.[9]

Finally, there are a number of lab experiments exploring specific anomalies related to dynamic portfolio choices, notably the disposition effect (e.g. Weber and Camerer, 1998, Frydman et al., 2014, Magnani, 2015). Fischbacher et al. (2017) show that stop-loss strategies may help overcome the disposition effect, but do not analyze the portfolio inefficiency that such strategies can create.

3 The environment

The experimental design focuses on the portfolio allocation decision in two equivalent environments which yield the same terminal wealth but have different presentation formats. We refer to the two tasks as (i) dynamic, and (ii) static.

3.1 Dynamic portfolio choice

In the dynamic task, each investor is endowed with an initial wealth $\bar{w} > 0$ which they must allocate between a risky asset and a risk-free asset for every contingent state s in periods $t = 1, \dots T - 1$. The rate of return of the risky asset $R_{s,t}$ evolves according

[8]DARA at the subject level is consistent with the prudence measure (a convex marginal utility) of Kimball (1990), which appears frequently in high-order risk-aversion tests (e.g. see Ebert and Wiesen, 2011, Deck and Schlesinger, 2014, Noussair et al., 2014 and Trautmann and van de Kuilen, 2018 for a recent overview). The decrease in risk-aversion at the low-priced state is also consistent with the behavior documented in Thaler and Johnson (1990).

[9]Similarly, Rapoport et al. (1988), who allow subjects to invest in risky assets and one safe asset show evidence of IRRA at the aggregate level. In lottery choices, IRRA is found in Holt and Laury (2002) and Noussair et al. (2014), however Harrison et al. (2007) fail to reject existence of CRRA using a controlled field experiment with Danish subjects.

to a binomial process described below. The safe asset pays a risk-free rate of return, R_f. We denote the investor's wealth in state s and period t as $W_{s,t}$, the amount of wealth invested in the risky asset as $a_{s,t}$ and the amount of wealth invested in the risk-free asset as $b_{s,t}$. The objective of an investor is to maximize the expected Von Neumann-Morgensten utility on terminal wealth $u(W_{s,T})$:

$$\max_{\{a_{s,t}\}} E[u(W_{s,T})] \tag{1}$$

subject to

$$W_{s,t+1} = W_{s,t}R_f + a_{s,t}(R_{s,t} - R_f), \tag{2}$$

where the risk-free investment can be computed as

$$b_{s,t} = W_{s,t} - a_{s,t}, \tag{3}$$

In our implementation we assume the risk-free return is $R_f = 1$ and the risky asset return $R_{s,t}$ follows

$$R_{s,t} = \begin{cases} R_u = 2 > R_f = 1 \text{ w. prob. } 0.5; \\ R_d = .5 < R_f = 1 \text{ w. prob. } 0.5. \end{cases} \tag{4}$$

Further, the investor is allowed to borrow and short-sell with the constraint that wealth is greater than or equal to zero in the following period: $W_{s,t+1} \geq 0$.

3.2 Static portfolio choice

From option pricing theory we know that the dynamic portfolio model can be reformulated as a static model using Arrow-Debreu securities. For each of the $S = 2^T$ terminal states we can define an Arrow-Debreu security as a state-contingent asset that pays out one if that state is realized and zero otherwise. The payoff to any portfolio strategy in the dynamic problem can be replicated by an appropriate portfolio of Arrow-Debreu securities. Moreover, this property can be used to derive the prices of the Arrow-Debreu securities, or state-prices. If u denotes the number of times the price of the risky asset increases prior to reaching state s in the binomial tree, and d the number of times the price decreases, then the price of Arrow-Debreu security s using the parameters of the risky asset return in equation (4) is:

$$p_s = \left(\frac{1}{3}\right)^u \left(\frac{2}{3}\right)^d.$$

We can also define the state-price density q_s as:

$$q_s := p_s/\pi_s = \left(\frac{2}{3}\right)^u \left(\frac{4}{3}\right)^d. \tag{5}$$

where π_s is the probability of final state s, which equals $\frac{1}{S}$, where S is the number of terminal states.[10] Hence, the investor's maximization problem can then be specified as

$$\begin{aligned} \max_{\{c_s\}} \quad & \sum_{s=1}^{S} \pi_s u(c_s), \\ \text{subject to} \quad & \sum_{s=1}^{S} p_s c_s = \bar{w}, \\ & c_s \geq 0 \ \forall s. \end{aligned} \tag{6}$$

where c_s denotes the amount of an Arrow-Debreu security s in the investor's portfolio.

[10] π_s can be re-written in terms of the number of ups and downs, or $\pi_s = \left(\frac{1}{2}\right)^u \left(\frac{1}{2}\right)^d$. In our experiment, the probability of up (down) is 1/2. Dividing p_s, known as fair price or risk-neutral price, by π_s, we obtain q_s. The risk-neutral price is computed by setting $c_u = 2a + b$ equal to $c_d = a/2 + b$. Solving for a and b, we obtain $a + b = (1/3) \cdot c_u + (2/3) \cdot c_d$.

3.3 Efficiency measures

Reformulating the original dynamic portfolio problem into the static Arrow-Debreu problem allows us to evaluate whether an investment strategy is efficient. Dybvig (1988) provides an important theoretical result:

Proposition 1. *The following are equivalent:*

1. *The portfolio is optimally chosen by some agent who has strictly increasing and concave von Neumann-Morgenstern (or Expected Utility) preferences over terminal wealth.*

2. *Final wealth c_s is non-increasing in the terminal state-prices p_s.*

3. *The portfolio minimizes the expenditure $\sum p_s c_s$ required to achieve the chosen distribution of final wealth.*

Proof: See Theorem 1 in Dybvig (1988).

By solving the static problem, and using the first order condition where marginal utility $u'(.)$ is proportional to state price density q

$$u'(c_s^*) = \lambda q_s, \tag{7}$$

we can show that (1) implies (2) in Proposition 1. Equation (7) also implies cyclic monotonicity[11] due to the concavity of the utility function

$$c_1^* \geq c_2^* \geq \cdots \geq c_{2^T}^* \Rightarrow q_1 \leq q_2 \leq \cdots \leq q_S. \tag{8}$$

[11]For a textbook analysis of the monotonicity condition using risk-neutral probabilities, refer to Chapter 13 of Danthine and Donaldson (2014) which includes a formal derivation of the main proposition of Dybvig (1988).

Since each state is equiprobable, we can rewrite the above condition as

$$c_1^* \geq c_2^* \geq \cdots \geq c_{_{2^T}}^* \Rightarrow p_1 \leq p_2 \leq \cdots \leq p_{_S}. \tag{9}$$

Proposition 1 shows that cyclic monotonicity is not only a necessary, but also sufficient condition for EU maximization. Moreover, Proposition 1 guarantees that only allocations that are non-increasing in prices are efficient, in the sense that they minimize the expenditure required to achieve the chosen distribution of wealth. If an agent picks a consumption profile that does not satisfy cyclic monotonicity, then it is possible to re-order the consumption choices in such way that the agent is able to achieve the same distribution of consumption and pay less. The savings can be spent on increasing wealth in any state, resulting in a distribution that first-order stochastic dominates the initial distribution of wealth.[12]

The theoretical result leads to a measure of inefficiency: if an investor chooses a distribution of c_s that is not non-increasing in p_s, then we can derive a dominating distribution by reordering the c_s to be non-increasing in p_s. Denoting the dominating distribution as c_s^*, Dybvig (1988)'s measure of efficiency loss (L) can then be specified as

$$\text{efficiency loss (L)} := \sum p_s c_s - \sum p_s c_s^* = \sum p_s (c_s - c_s^*). \tag{10}$$

In experimental and field data, we cannot observe actual preferences of agents. Rather L in equation (10) represents a tight lower bound on the amount of wealth an agent would pay to switch from c to an optimal strategy that satisfies the cyclical monotonicity condition in equation (8).

[12] Recall that a lottery A (strictly) first-order stochastically dominates lottery B iff $F_A(x) \leq F_B(x)$ for all x, with strict inequality for some x. $F_A(x)$ is the cumulative distribution function, or the probability that the realized wealth in profile A is no greater than x.

3.4 Measuring risk-preferences from portfolio choices

If a portfolio strategy induces a distribution of final wealth that satisfies the cyclical monotonic condition in equation (8), then it follows that the observed choices satisfy a maximization process according to EU. Next, we present the non-parametric techniques originally suggested by Varian (1988) to study the risk profiles revealed by portfolio choices.

We characterize absolute risk aversion using portfolio choice data and the Arrow-Pratt measure of risk defined as,

$$A(c) := -\frac{u''(c)}{u'(c)} = -\frac{d\log(u'(c))}{dc}. \tag{11}$$

$A(c)$ is the absolute value of the slope of the log of the marginal utility relative to c, and it follows that an investor will have decreasing (increasing) absolute risk aversion if and only if the log of marginal utility is convex (concave):

$$A'(c) < 0 \Leftrightarrow \frac{d^2\log(u'(c))}{dc^2} > 0 \tag{12}$$

While we do not observe marginal utility directly, we can infer it from portfolio choices of subjects. By choosing an appropriate affine transformation of the utility function, the first order condition of the static problem

$$u'(c_s) = q_s, \tag{13}$$

suggests that we can use state-prices in place of marginal utilities. If we plot $\log(q_s)$ versus c_s, we obtain a piece-wise linear demand curve, as illustrated in Figure 1. A steep (or more inelastic) demand curve represents higher risk-aversion. If the demand curve is vertical (perfectly inelastic) then an agent will invest all her wealth in the risk-free asset, and thus perfectly hedge. Moreover, if $\log(q_s)$ is convex (concave) in c_s, then we can conclude that the investor has a decreasing (increasing) absolute risk aversion. Varian (1988) shows that under decreasing absolute risk aversion it is possible to establish the following bounds for the coefficient $A(c)$ at each wealth level c_s, so that

$$\frac{\log(q_s) - \log(q_{s-1})}{c_{s-1} - c_s} \leq A(c_s) \leq \frac{\log(q_{s+1}) - \log(q_s)}{c_s - c_{s+1}}, \tag{14}$$

where states have been re-numbered so that $c_1 < c_2 < ... < c_S$ and $q_1 > q_2 > ... > q_S$. We will refer to this as Varian's *ratio condition*. In the case of increasing absolute risk aversion, the inequalities in (14) are reversed.

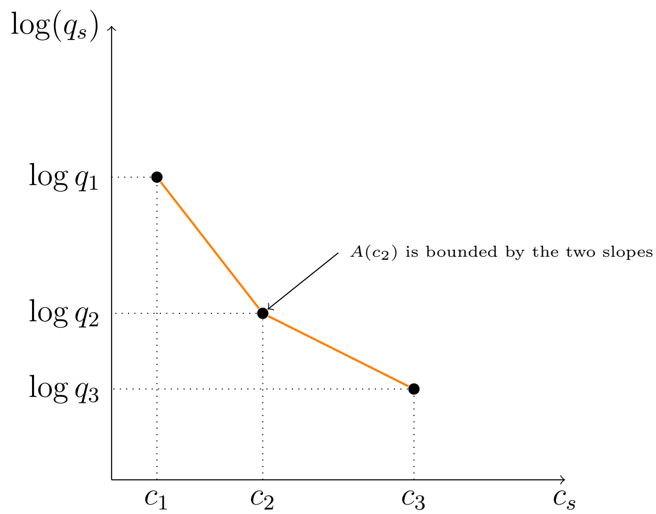

Figure 1: Illustration of Varian (1988) method.

A similar argument can be made to analyze relative risk aversion, defined as

$$R(c) := -\frac{c \cdot u''(c)}{u'(c)} = -\frac{d\log(u'(c))}{d\log(c)}. \tag{15}$$

If $\log(q_s)$ is convex (concave) in $\log(c_s)$, then we can conclude that the investor has a decreasing (increasing) relative risk aversion. Furthermore, under decreasing relative risk aversion the coefficient $R(c)$ at each wealth level c_s has the following lower and upper bounds:

$$\frac{\log(q_s) - \log(q_{s-1})}{\log(c_{s-1}) - \log(c_s)} \leq R(c_s) \leq \frac{\log(q_{s+1}) - \log(q_s)}{\log(c_s) - \log(c_{s+1})}, \tag{16}$$

where states have been numbered so that $c_1 < c_2 < ... < c_S$ and $q_1 > q_2 > ... > q_S$. As before, inequalities are reversed for the case of increasing relative risk aversion.

4 Hypotheses

The first environment in our experiment reproduces the dynamic portfolio problem (1). To test whether inefficiency is affected by the complexity of the portfolio problem, we implement two versions of problem, one with $T = 3$ periods and one with $T = 4$ periods, resulting in eight and 16 terminal states respectively. In addition to the dynamic portfolio problem, we include a static portfolio problem (6) which allows us to test whether efficiency is affected by the use of path-dependent strategies, such as stop-loss. To allow for a clean comparison with the dynamic portfolio problems, we similarly employ two versions of the static portfolio problem, one with eight states and one with 16 states.

Hypothesis 1: *Under the assumption of perfect rationality, efficiency loss is zero in both dynamic and static tasks, and across the number of terminal nodes.*

If we assume perfect rationality, then there should be no path-dependent behavior, and the task format and number of terminal nodes is non-consequential. Thus, we expect similar efficiency loss L (equation 10) and risk-aversion across all treatments.

Alternatively, under path-dependence (e.g. stop-loss, or lock-in strategies) we expect that efficiency loss L will be higher in the dynamic task, with the effect stronger when there are 16 terminal nodes compared to the eight terminal nodes. Similarly, it is more likely that a boundedly-rational player, as measured by the CRT score, will make more mistakes ($L > 0$).

Hypothesis 2: *Under the assumption of expected utility maximization, subjects will exhibit non-increasing absolute risk aversion in both static and dynamic tasks.*

We expect that $A(c)$, as defined in equation (11), is non-increasing in terms of final consumption or wealth, or $A'(c) \leq 0$ for both tasks.

Hypothesis 3: *Under the assumption of expected utility maximization, subjects exhibit non-decreasing relative risk aversion in both static and dynamic tasks.*

Following the results of Brocas et al. (2019), who estimate parameters of a HARA utility function (under EU assumptions) in a dynamic portfolio experiment, we expect that $R(c)$, as defined in equation (15), is non-decreasing in terms of terminal wealth, or $R'(c) \geq 0$.

5 Laboratory procedures

Our experiment consists of four treatments, labeled $D8$, $D16$, $S8$ and $S16$, where the letter denotes either a dynamic (D) or a static (S) environment, and the number pertains to the terminal states in the task. This is a within-subject design, where each subject participates in all four versions of the task ($D8$, $D16$, $S8$ and $S16$). In some sessions, subjects first complete the dynamic tasks followed by static tasks, while in other sessions, subjects first complete the static tasks and then dynamic tasks.

While the static/dynamic task order was randomized, the first task always had eight terminal states.[13] To ensure comprehension, the subjects answered control questions and practiced each task with four terminal nodes for two rounds. We also include a survey regarding the field of study and gender, and an incentivized cognitive reflection test (Frederick, 2005) at the end of the experiment. The subject's payoff is calculated using the wealth in an independently drawn terminal state, which is randomly selected for all four versions of the task ($D8$, $D16$, $S8$ and $S16$). Payoffs are revealed only after the subjects complete all tasks.

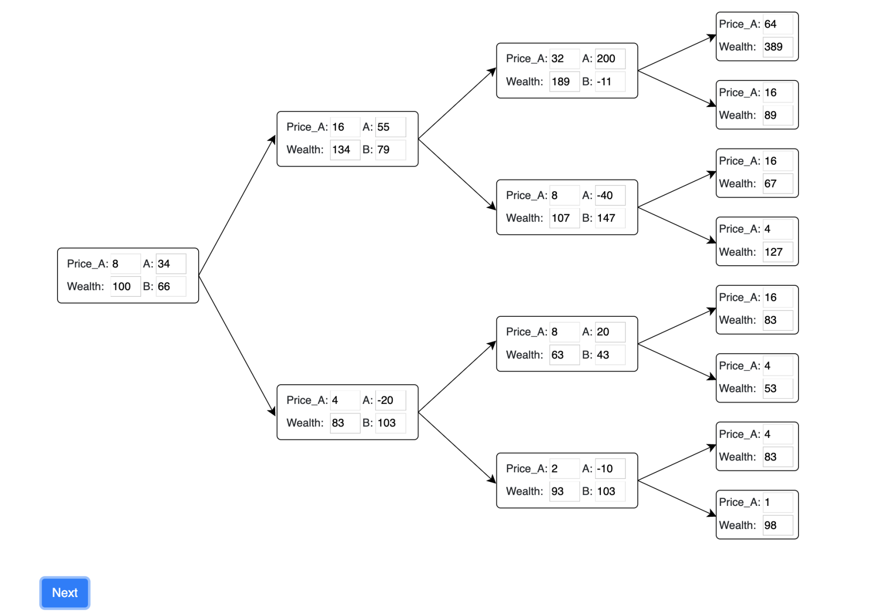

Figure 2: User-interface of task D8.

The user interface in the dynamic task presents the price of the risky-asset in every node in the tree. An example of the interface, designed in oTree (Chen et al., 2016), for $D8$ is depicted in Figure 2. Each box in the tree displays the price of the risky asset, the current wealth and the amount of wealth allocated to the risky asset A and to the risk-free asset B. The initial price of the risky asset A is eight, and follows the process

[13] We opted for a lower complexity task first to dampen the effect of complexity in the experiment with experience.

described by equation (4). In the interface, an upward-pointing (downward-pointing) arrow represents a price increase (decrease). Subjects can decide how much to invest in the risky asset ($a_{s,t}$) for all $2^T - 1$ nodes of the tree until the terminal period T. The user-interface for $D16$ is similar to the one presented in Figure 2, but with 16 terminal nodes and 15 asset allocation decisions.

Subjects always begin with an initial wealth of $\bar{w} = 100$ or 2 AUD. After the subjects select the risky allocation $a_{s,t}$ at initial node, the interface computes (i) the risk-free investment $b_{s,t}$ following equation (3), and (ii) the wealth $W_{s,t+1}$ in equation (2) for the next two nodes, in which the asset price either increases or decreases. Recall that the subjects are allowed to short-sell the risky asset (i.e., take a negative position on $a_{s,t}$) or borrow from the risk-free bond (i.e., take a negative position on $b_{s,t}$) with the constraint that $W_{s,t+1} \geq 0$. As the subject fills out the investment plan, the software immediately updates the wealth levels in the subsequent boxes. In Figure 2, we entered an initial value of $a = 34$ and the software computed $b = 100 - 34 = 66$, and the subsequent wealth $134 = 2 \times 34 + 66$ in the up state and $83 = 1/2 \times 34 + 66$ in the down state. When the investment plan is complete (all nodes have an allocation, and the terminal wealth across all states is non-negative), it can be submitted. Following submission, the software randomly draws a price path and the subjects final wealth is realized (without feedback, as subjects learn of outcomes only after completing all tasks).

In the static task, subjects also begin with an initial wealth $\bar{w} = 100$ or 2 AUD, and allocate their budget to buy tickets across eight or 16 different lotteries, depending on the treatment. The payoff of a subject is calculated as the number of tickets bought for a randomly selected lottery. The price of each lottery represents the price of the Arrow-Debreu security at each of the terminal nodes of the dynamic task. The interface is a large table with either 8 or 16 rows, where each row represents a lottery.

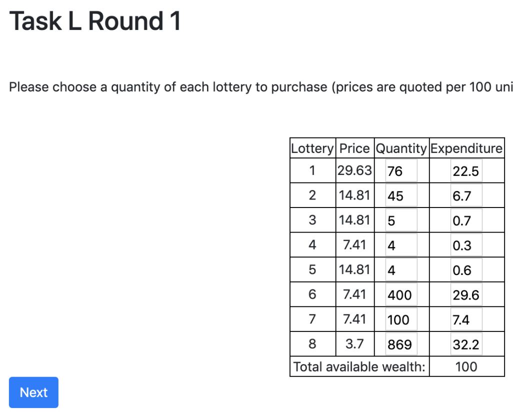

Figure 3: User-interface of task S8.

Figure 3 presents the user-interface for the 8 lottery case. The first column contains the lottery id,[14] and the second column posts the lottery ticket prices in terms of 100 units. Lottery prices are computed as state-prices for each of the final states, using the method described in section 3. The third column is where subjects enter how many tickets they would like to purchase, and the last column provides the expenditure on each lottery after the subjects enter the desired quantity. The interface automatically fills in the quantity of tickets for the last lottery (the lowest-priced lottery) so that the budget constraint is satisfied. Subjects can purchase only non-negative amounts of each lottery. If the number of tickets is negative for any lottery, then subjects have to adjust their allocation across lotteries in order to proceed to the next screen. In our example in Figure 3, we entered 76 for the first lottery, which is an expenditure of 22.5 (76×0.2963). Once quantities for lotteries 2-7 are entered, and the software computes the quantity that satiates the budget of 100 for the last lottery.

In all, six sessions were conducted at Monash University (MonLEE lab), with 119 subjects, recruited from all fields of study using the software SONA. Subjects earned on average 23 AUD, including a show-up fee of 10 AUD. The exchange rate was 100

[14] The lotteries are sorted in the reverse order of the final states in the dynamic task tree.

points for 2 AUD. Each correct answer in the CRT task yielded a payoff of 50 points. From a total of three CRT questions, the average score was 2/3. The participants were 48 percent female, 66 percent declared their field of study as STEM, 21 percent as Business and Finance, 7 percent as Art and Humanities, with the rest classified as Social Science or not a student.

6 Results

We begin our discussion of results with a look at the efficiency loss L, as defined in equation (10), for each of the four tasks. Recall that a value of zero for L means that the subject's consumption choice satisfies the monotonicity condition presented in equation (8). Table 1 presents the mean and median values of L, as well as the percentage of choices where $L \leq 2$, and where $L = 0$.

Table 1: Observed and Predicted Inefficiency L

	Data				Predictions			
Task	Mean	Median	$\% \leq 2$	$\% = 0$	stop-loss	lock-in	random	worst
D16	2.25	0.00	74	57	1.23	4.94	23.57	93.75
D8	1.22	0.00	81	70	0.00	0.00	19.48	87.85
S16	9.92	1.88	51	26	na	na	23.57	93.75
S8	7.35	0.19	58	47	na	na	19.48	87.85

Note:
a. The statistics are computed across subjects for each treatment, and thus the percentage is in terms of the total number of subjects. b. The initial asset price is equal to 8. The stop-loss strategy switches to the risk-free bond when the asset price drops to 4. The lock-in strategy switches to the risk-free bond when the asset price increases to 16.
c. Random strategy: selects a random allocation of terminal wealth from a flat Dirichlet distribution on the S-dimensional simplex defined by the budget constraint. The values of L reported are the averages over 5,000 iterations.
d. Worst strategy: assumes that the investor puts all wealth in the most expensive state. The inefficiency is then computed against switching the wealth to the lowest-priced state

Result 1a: *The inefficiency of portfolio choices, measured by L in equation (10), is small in both dynamic and static tasks. Most of the subjects choose an efficient port-*

folio allocation.

Overall, we find that median value of L is zero in the dynamic task, and close to zero for the static task (0.19 for $S8$ and 1.88 for $S16$). If we look at the percentage of choices per subject where $L \leq 2$, then we find a large percentage of subjects selecting efficient portfolios in the dynamic task (81 percent for $D8$ and 74 percent for $D16$). For $S8$ and $S16$, the percentage of subjects with efficient choices is 58 percent and 51 percent, respectively.

Result 1b: *The inefficiency is higher in the tasks with a larger number of terminal states and higher in static tasks relative to dynamic tasks.*

The inefficiency is greater in the static task relative to the dynamic task, and is increasing in complexity. The mean value of L is about seven and 10 in $S8$ and $S16$, respectively, and about one and two in $D8$ and $D16$, respectively.

We test whether inefficiency L varies systematically across treatments using Wilcoxon tests. The p-values of two-sided tests are summarised in Table 2. We find that tasks with a greater number of terminal states have higher inefficiency on average. This difference is statistically significant for both static and dynamic tasks (p-values of 0.043 and 0.009, respectively). We also find that for a given number of terminal states static tasks have significantly higher inefficiency.

Table 2: Wilcoxon Test: L (p-values)

	D16	D8	S16
D16			
D8	$<$ (0.009)		
S16	$>$ (0.000)	$>$ (0.000)	
S8	=	$>$ (0.000)	$<$ (0.043)

To have a better understanding of the L values observed in our data, we perform simulations using four portfolio selection strategies. First, we consider two inefficient dynamic strategies originally featured in the analysis of Dybvig (1988) and which may be salient for subjects: a stop-loss strategy, and a lock-in strategy. The stop-loss strategy involves investing all wealth in the risky asset at the initial node, and moving all wealth to the risk-free asset as soon as the risky asset price falls to 4. The lock-in strategy involves investing all wealth in the risky asset at the initial node, and moving all wealth to the risk-free asset as soon as the risky asset price rises to 16. Next, we consider a random strategy which draws an allocation of terminal wealth from a flat Dirichlet distribution on the S-dimensional simplex defined by the budget constraint. To compute the inefficiency of the random strategy we simulate 5,000 portfolio choices and calculate the average L. Finally, we consider the worst strategy, which allocates all wealth to the most expensive state. The value of L reported under the worst-case strategy is computed by switching the consumption from the most expensive state to the cheapest one. The simulation results of efficiency loss L for all four strategies are depicted in Table 1.

Across all treatments, subjects outperform the worst-case strategy, which results in high values of L (about 87 for $D8$ and $S8$ and 94 for $D16$ and $S16$). The level of inefficiency generated by the random strategy is also much larger than the average inefficiency of subjects' choices in our data. For the stop-loss and lock-in strategies, we only encounter inefficiencies in the treatment with 16 terminal states. The value of L in $D16$ is 1.23 for the stop-loss strategy, and 4.94 for the lock-in strategy. This suggests that forgoing consumption in up states is more costly compared to the down states. By design, the static environment does not allow path-dependent strategies, and thus in Table 1 there are no values to report for $S8$ and $S16$. The inefficiency generated by path-dependent strategies is closer in magnitude to the actual inefficiency we observe in our data. We elaborate on whether subjects implement these strategies during the experiment further below. First, we discuss whether subject characteristics can explain the observed variation in inefficiency.

Result 1c: *Subjects with higher CRT scores make more efficient portfolio choices.*

In order to study whether subject characteristics can help predict the efficiency loss L observed in our sessions, we run OLS regressions with L as a dependent variable. The results are summarized in Table 3, and include the following as explanatory variables: (i) *Small* takes the value of one if the task has 8 terminal nodes, and zero otherwise, (ii) *Dynamic* takes the value of one if the task is $D8$ or $D16$, and zero otherwise, (iii) *Man* takes the value of one if the subject is a self-reported male, and zero otherwise, and lastly, (iv) CRT can take the value of $\{0, 1, 2, 3\}$, and represents the correct number of questions answered in the cognitive reflection test. Thus, the intercept of 14.83 for specification (I) in Table 3 captures the inefficiency L of a female subject in the $S16$ treatment with a CRT score of zero (the average CRT score in our data is 2).

Inefficiency is 7.66 points lower in the dynamic tasks relative to static tasks. Inefficiency is also 2.57 points lower in tasks with 8 terminal states relative to tasks with 16 terminal states. The smallest inefficiency is observed in the $D8$ treatment ($= 14.83 - 2.57 - 7.66$). We do not find any statistically significant interaction effects between treatments, with the coefficient *Dynamic* $\times$ *Small* not statistically different from zero. Furthermore, the inefficiency is also smaller for subjects with higher CRT scores. An additional correct answer decreases L by 2.43. Lastly, there are no gender differences in the efficiency of portfolio choices. Specification (II) in Table 3 includes the same set of independent variables as specification (I), but controls for the field of studies declared. While we do not find any major differences across the two specifica-

Table 3: OLS Regressions. Inefficiency L

	(I)	(II)
Intercept	14.83^{***}	13.06^{***}
	(2.85)	(2.63)
Dynamic	-7.66^{***}	-7.66^{***}
	(1.66)	(1.67)
Small	-2.57^{**}	-2.57^{**}
	(1.19)	(1.19)
Dynamic × Small	1.54	1.54
	(1.34)	(1.35)
Man	0.14	0.59
	(1.33)	(1.35)
CRT	-2.43^{***}	-2.12^{**}
	(0.84)	(0.82)
Field of Study (FE)	No	Yes
R^2	0.13	0.15
N	476	476

Notes:
a. L is defined in equation (10).
b. Standard errors are clustered at the individual level.
$^{***}p \leq .01$, $^{**}p \leq .05$, $^{*}p \leq .1$

tions, we do want to note that the field of study reduces the significance of the CRT from one to five percent, which suggests that these two variables may be correlated.

Result 1d: *In $D8$ and $D16$, the most common mistakes leading to inefficiency can be explained by a form of stop-loss, and insufficient risk-taking following early declines in the price of the risky asset.*

To study the sources of inefficiency, we assess whether subjects make common mistakes in their allocation decisions. Recall that inefficient strategies violate cyclic monotonicity. In other words, if a strategy is inefficient, then there is at least a pair of terminal states i and j where $p_i > p_j$ and the strategy allocates more wealth to state i than state j ($c_i > c_j$). Focusing on inefficient strategies in the dynamic tasks, we analyze the frequency of different violations of monotonicity, which we refer to as mistakes below.

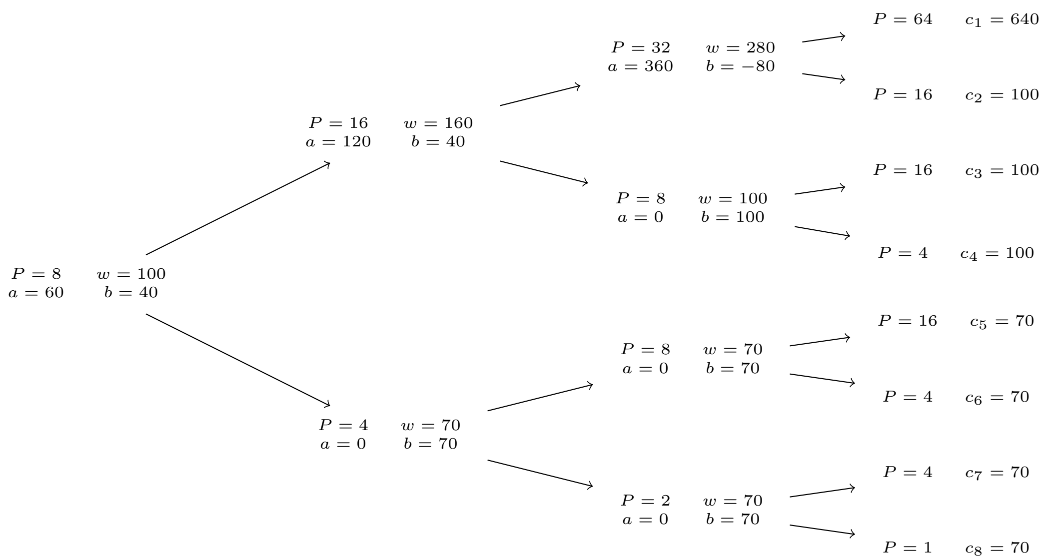

Figure 4: Example of inefficient strategy in $D8$

Note: P denotes the price of the risky asset, w wealth, a the amount of wealth invested in the risky asset, b the amount of wealth invested in the risk-free asset and c_i denotes final wealth in state i. The inefficiency arises from $c_4 > c_5$ at the terminal period. The Arrow-Debreu price at state 4 is 0.1481 while at state 5 is 0.0741. Therefore, the inefficiency L can be computed as $L = 2.22 = (100 - 70) \times (0.1481 - 0.0741)$. The subject stops ($a = 0$) when the price drops to $P = 4$ in period 1 and when the price drops to $P = 8$ in period 2 after having increased to 16 in period 1.

To begin, we sort the states according to the position of each terminal node in the binomial tree diagram, from high ($s = 1$) to low ($s = 8$ for $D8$, $s = 16$ for $D16$).[15] Table 4 presents the count of each type of mistake for $D8$ and $D16$, and the median value of the efficiency loss, L, among subjects who make the same mistake. $D16$ data is divided into two sections, upper (nodes 1-8) and lower tree (nodes 9-16). Out of the 35 subjects that have a value of L greater than zero in $D8$, 12 subjects made a mistake where $c_4 > c_5$, and with a median value of L equal to 2.11.[16] A form of stop-loss can explain this mistake, where the market participant leaves the market once the price falls below the initial level. An example of such strategy, taken from one of the subjects, is illustrated in Figure 4. The second most common mistake in $D8$ involves a misallocation of final wealth in states $s = 5$ and $s = 7$. Subjects who

[15]Note that as in the binomial tree diagram, states are not ordered according to state-prices.
[16]We use median instead of mean so our measures are not biased by extreme mistakes.

commit this mistake seem to disproportionately increase their risk tolerance following two consecutive declines in the price of the risky asset.

Table 4: Mistakes in D8 ($N = 35$ subjects) and D16 ($N = 51$ subjects)

D8			D16 upper tree			D16 lower tree		
Mistake	N	L (median)	Mistake	N	L (median)	Mistake	N	L (median)
$c_2 > c_1$	1	5.33	$c_2 > c_1$	2	43.32	$c_{10} > c_9$	0	–
$c_5 > c_1$	1	2.78	$c_5 > c_1$	3	34.58	$c_{13} > c_9$	4	19.16
$c_4 > c_3$	4	8.88	$c_4 > c_3$	1	34.58	$c_{12} > c_{11}$	0	–
$c_4 > c_5$	12	2.11	$c_4 > c_5$	10	4.55	$c_{12} > c_{13}$	10	2.40
$c_6 > c_2$	2	1.63	$c_6 > c_2$	2	43.32	$c_{14} > c_{10}$	3	34.58
$c_7 > c_3$	5	2.44	$c_7 > c_3$	2	43.32	$c_{15} > c_{11}$	16	5.71
$c_7 > c_5$	7	2.44	$c_7 > c_5$	4	18.95	$c_{15} > c_{13}$	2	3.43
$c_8 > c_7$	3	3.70	$c_8 > c_7$	2	33.36	$c_{16} > c_{15}$	3	14.66
						$c_8 > c_{11}$	20	2.82

Notes:
a. The terminal nodes are sorted from high (one) to low (eight for D8 and 16 for D16).
b. N is the count of subjects.
c. L is defined in equation (10).
d. The tree for D16 is split in two. The upper tree corresponds to nodes 1-8 and the lower tree to nodes 9-16.

The stop-loss strategy also explains many of the mistakes observed in the $D16$. In this task, 51 subjects incur some level of inefficiency. In the upper tree, we observe ten subjects making a mistake similar to the one described above in the context of the $D8$ task, but now the mistakes are costlier given the higher growth of the asset values (4.55). A similar pattern appears in the lower tree, at nodes 12 and 13. Ten subjects make this mistake, with the median value of $L = 2.40$. The most common mistake in $D16$ involves states 8 and 11. Although state 8 has a higher state-price than state 11, 20 subjects opt for a higher terminal consumption in state 8 than state 11, resulting in a median value of $L = 2.82$. This mistake can also be interpreted as a form of stop-loss strategy. We illustrate one example of this behavior in Figure 5. The stop-loss strategy implemented in this case is not triggered by the initial price decrease, but by the decline in later periods. The stop-loss strategy presented in our simulations in Table 1, is implemented by one subject, who obtains a terminal wealth of 85, which required investing 30 in the risky asset and 70 in the risk-free bond at the initial node.

The same subject also used a lock-in strategy, which yields a wealth of 130 for the top eight terminal nodes. Using the Arrow-Debreu prices of 0.0988 and 0.0247 for states 8 and 9, respectively, we are able to calculate the inefficiency generated by these choices as $L = 3.33 = (130 - 85) \times (0.0988 - 0.0247)$.

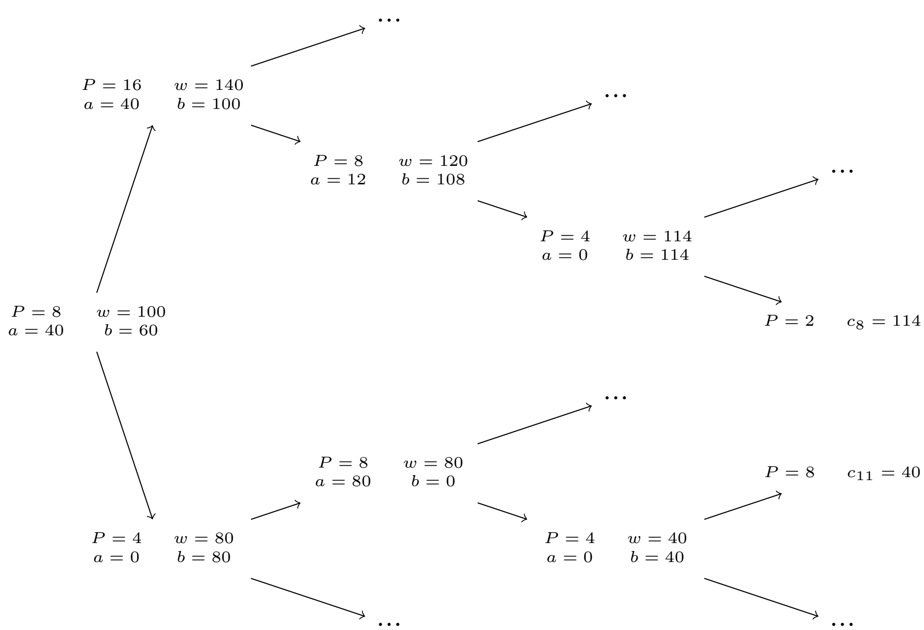

Figure 5: Example of inefficient strategy in $D16$

Note: P denotes the price of the risky asset, w wealth, a the amount of wealth invested in the risky asset, b the amount of wealth invested in the risk-free asset and c_i denotes final wealth in state i. The inefficiency arises when $c_8 > c_{11}$ at the terminal period. The Arrow-Debreu price at state 8 is 0.0988 while at state 11 is 0.0494. Therefore, the inefficiency L due to this mistake can be computed as $L = 3.458 = (114 - 40) \times (0.0988 - 0.0494)$. The subject stops ($a = 0$) when the price drops to $P = 4$ in period 3.

The second most common mistake (16 subjects) in the lower tree of $D16$ involves a misallocation of the final wealth in states 11 and 15. The inefficiency can be explained by the more risk-taking at the upper nodes, which yields less units available for consumption at node 11. This results in the median consumption of 38 at node 11 and 85 at node 15.[17]

[17] We also replicate Table 4 using data from the static task. The table is presented in Appendix A. We find that the most common mistake occurs in the first node. 31 out of 62 subjects with $L > 0$ in $S8$ opt for $c_2 > c_1$. In $S16$, we find 49 subjects (out of 88) with $L > 0$. The first node in the static task is filled by the software after the subjects enter the $S - 1$ choices. This implies that many

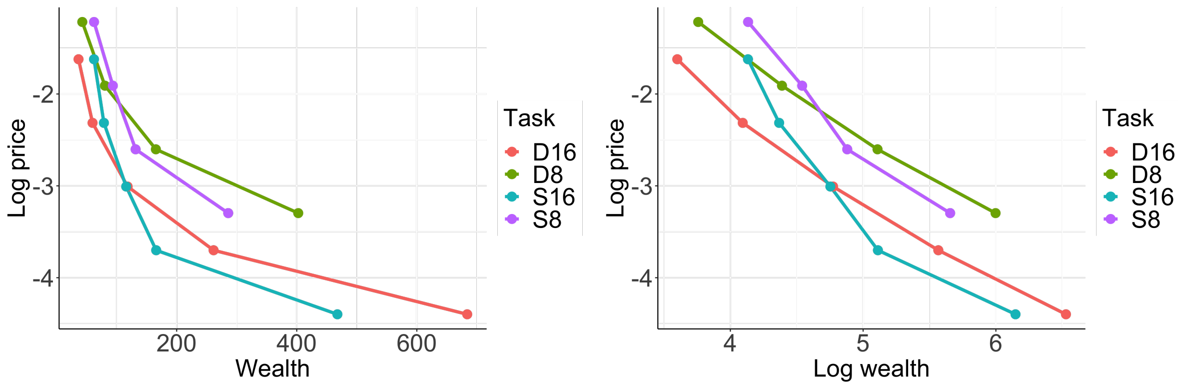

Figure 6: Piecewise linear function of log prices and final wealth for the representative agent (left panel), and in terms of log prices and log final wealth (right panel).

Result 2a: *The representative agent exhibits decreasing absolute risk aversion.*

Efficient portfolio strategies, rationalized by an EU preference, can be further analyzed using the techniques discussed in Section 3.4. To achieve this, we restrict our analysis to efficient portfolios only and evaluate whether these portfolios are consistent with decreasing, constant or increasing risk aversion. Prior to discussing our results, we first explain how we adapt the methods from Section 3.4 to our data. Each observation in our data consists of a vector of final wealth levels, $(c_1, ..., c_S)$ and state-prices $(p_1, ..., p_S)$. Some of the final states have the same state-price since different asset price paths lead to the same total number of ups and downs. The techniques require that we work with linearly independent states. Therefore, we aggregate states with the same state-price and compute the average wealth per subject in each independent state. For $D8$ and $S8$ we have four linearly independent states, and for $D16$ and $S16$, five linearly independent states.

subjects do not revise their choices even though they have the option to re-allocate their consumption before submitting. We also observe significant mistakes at other terminal nodes. For example, 17 (16) subjects pick $c_8 > c_7$ for $S8$ ($S16$).

To characterize the demand function of the representative agent, we aggregate subject consumption choices (wealth) at each terminal state price. The ratio condition in equation (14) requires that we work with log prices and levels for final wealth (consumption) in order to measure absolute risk aversion, $A(c)$. The left panel of Figure 6 shows the constructed piecewise linear function, or the demand function, from four treatments. All four demand curves in the left panel of Figure 6 are clearly convex. Thus, according to Varian (1988) the choices of a representative agent are consistent with a decreasing absolute risk aversion (DARA). Moreover, the demand curves for $D8$ and $D16$ are flatter than the curves for $S8$ and $S16$, especially when prices prices are low, which implies a lower degree of absolute risk aversion in dynamic tasks relative to static tasks.

Table 5: Ratio of $|\Delta \log p/\Delta w_T|$ for the representative agent

Ratios (slopes)	D16	S16	D8	S8
I	0.030	0.042	0.018	0.022
	[0.448]	[0.834]	[0.018]	[0.892]
II	0.012	0.019	0.008	0.018
	[0.067]	[0.896]	[0.000]	[0.018]
III	0.005	0.014	0.003	0.005
	[0.000]	[0.011]	–	–
IV	0.002	0.002	–	–
	–	–		
R^2	0.57	0.10	0.71	0.14
N	595	595	476	476

Notes:
a. The coefficients are estimated using a piecewise linear regression, in which the dependent variable is the terminal wealth, and the independent variables are state prices (in logs). We report the inverse of the coefficients. The slopes are sorted from high to low prices, or from low to high terminal wealth. There are four slopes in $D16$ and $S16$, and three in $D8$ and $S8$.
b. p-values are in brackets. The null hypothesis states that the difference of ratios at j and $j+1$ (marginal effect) is equal to zero.

The ratio $\left|\frac{\Delta \log q}{\Delta w_T}\right|$, constructed using the points from the left panel of Figure 6, is presented in Table 5. The ratio values (or the slopes of the demand curve) reported are estimated using a piece-wise linear regression.[18] In brackets, we present the p-value of the test where the null hypothesis states that the marginal effect, or the difference of the current slope with respect to the one below, is equal to zero. Table 5 presents slope values, instead of the marginal effects, to help convey the upper and lower bounds of the Arrow-Pratt measure $A(w_T)$. We sort the data from high to low prices (or low to high in terms of terminal wealth). Consistent with DARA, the slopes are non-increasing with respect to wealth for all treatments. As observed in Figure 6, we find that the demand curves for $S16$ and $S8$ are steeper compared to $D16$ and $D8$, respectively. The representative agent is therefore more risk-averse in the static tasks. The low correlation between prices and terminal wealth in the static tasks leads to a lower R^2 in Table 5.

Result 3a: *The representative agent exhibits constant relative risk aversion.*

To study the measure of relative risk aversion, $R(w_T)$ in equation (16), we follow the same procedure as in our study of $A(w_T)$ above, except that we work with log prices and log wealth, so that the ratio is $\left|\frac{\Delta \log q}{\Delta \log w_T}\right|$. The piecewise linear function for the representative agent is depicted in the right panel of Figure 6, and the piecewise linear regression is presented in Table 6. The representative agent exhibits CRRA across our treatments, with the static task showing levels of higher risk-aversion. According to Table 6, the value of R in $D16$ (0.868) and $D8$ (0.848) is smaller than in $S16$ (1.720) and $S8$ (1.475), respectively.[19]

[18] The piecewise linear regression includes the terminal wealth (w_T) as a dependent variable, and the state prices (q_j, where $j = 1, 2, 3,$ or 4) as independent variables. Table 5 presents the p-values for the regression coefficients, and the ratio $\left|\frac{\Delta \log q}{\Delta w_T}\right|$ is presented as the inverse of the coefficients since wealth is the dependent variable in the regression.

[19] The R^2 in Table 5 (for $R(w_T)$) is higher than in Table 6 (for $A(w_T)$) because the logs smooth the data, and improve the goodness of fit. Note that in Table 6, we have less observations (N) than in Table 5 because we cannot take logs where terminal wealth equal is to zero. Overall, we lose 10

Table 6: Ratio of $|\Delta \log p/\Delta \log w_T|$ for the representative agent

Ratios (slopes)	D16	S16	D8	S8
I	1.142	3.748	1.074	2.166
	[0.204]	[0.315]	[0.416]	[0.682]
II	0.888	1.883	0.958	2.684
	[0.839]	[0.554]	[0.324]	[0.155]
III	0.920	2.641	0.848	1.475
	[0.738]	[0.432]	–	–
IV	0.868	1.720	–	–
	–	–		
R^2	0.75	0.23	0.78	0.25
N	585	565	454	454

Notes:
a. The coefficients are estimated using a piecewise linear regression, in which the dependent variable is the log of terminal wealth, and the independent variables are state prices (in logs). We report the inverse of the coefficients. The slopes are sorted from high to low prices, or from low to high terminal wealth. There are four slopes in $D16$ and $S16$, and three slopes for $D8$ and $S8$.
b. p-values in brackets. The null is that the difference of ratios at j and $j+1$ (marginal effect) is equal to zero. For example, for D16 we fail to reject that ratios at III and at IV are equal.

Result 2b: *At the subject level, DARA is prevalent in both dynamic and static tasks, but non-monotonic behavior is also important, particularly for the static task.*

At the subject level, we work with the average subject wealth at each state price, and exclude upward sloping demand curves, and terminal wealth that is equal to zero from the data, since subsequent analysis requires us to take logs of wealth.[20] Table 7 presents the fraction of subjects from the restricted sample, and who are characterized by (i) decreasing, (ii) increasing, (iii) constant or (iv) non-monotonic absolute risk aversion, $A(w_T)$. Since we cannot use a statistical test to assess whether a slope change is different from zero at the individual observation level, we use a threshold.

observations in $D16$, 30 observations in $S16$, and 22 observations in both $D8$ and $S8$.

[20] An upward sloping demand means that subjects are making inefficient portfolio choices, or $L > 0$.

Specifically, we impute a change in the slope as equal to zero if the actual change is below a 10 percent threshold.[21] Thus, if the percentage changes in the slopes of the demand curves are all smaller than or equal to 10 percent in absolute value, then the profile is consistent with constant absolute risk aversion (CARA). If the slopes are non-increasing (and with at least one slope change greater than 10 percent), then we say that the profile is consistent with DARA. If the slopes are non-decreasing (and with at least one slope change greater than 10 percent), then we conclude that the profile is consistent with increasing absolute risk aversion (IARA). For the other cases, we classify the profiles as consistent with non-monotonic absolute risk aversion.

Table 7: Risk averse types

	Decreasing	Constant	Increasing	Non-monotonic
			D16	
A	0.83	0.02	0.00	0.15
R	0.33	0.08	0.15	0.44
			S16	
A	0.45	0.00	0.00	0.55
R	0.39	0.00	0.04	0.57
			D8	
A	0.81	0.03	0.00	0.16
R	0.42	0.11	0.18	0.29
			S8	
A	0.39	0.27	0.00	0.35
R	0.33	0.20	0.08	0.39

Note: A denotes absolute risk aversion while R refers to relative risk aversion. Treatments $D8$, $S8$, $D16$ and $S16$ use 0.89, 0.82, 0.92 and 0.77 of total observations, respectively. We drop observations due to inefficient choices and/or zero final wealth.

Table 7 shows that DARA is quite dominant in the treatments $D16$ and $D8$. Around 80 percent of the subjects are classified as DARA, while the rest of subjects are classified primarily as non-monotonic. This means that the slopes of the demand curve are

[21]Qualitatively, we obtain the same if, instead we work with a threshold of 5, or 20 percent. Since the slope is not defined when $\Delta w_T = 0$, we replace Δw_T with a small number. The purpose of this replacement is to gain observations without affecting the risk profiles.

decreasing for some segments, but increasing for others. In treatment $S16$, 45 percent of subjects are classified as DARA, with the rest classified as non-monotonic. For the treatment $S8$, 39 percent are classified as DARA, 27 percent as CARA, and the rest as non-monotonic. The main reason of CARA behavior in $S8$ is due to a constant terminal wealth across states, or hedging behavior. Furthermore, due to higher inefficiencies in the static task compared to the dynamic task, we retain more observations in $D16$ (92 percent of total) and $D8$ (89 percent) compared to $S16$ (77 percent) and $S8$ (82 percent), respectively.

Result 3b: *Subjects exhibit increasing and decreasing relative risk aversion.*

Classification of relative risk aversion (see R in Table 7) at the individual level is heterogeneous. We find that increasing and decreasing relative risk aversion frequently appear in the data, and that the combination of both results in a non-monotonic relative risk aversion. In treatment $D16$, we observe that 44 percent of subjects are classified as non-monotonic, 33 percent as DRRA, and 15 percent as IRRA. Treatment $S16$ follows a similar profile for DRRA, but the fraction of non-monotonic behavior increases to 56 percent, and the fraction of IRRA drops to four percent. DRRA is also common in $D8$ (42 percent) and $S8$ (33 percent), followed by non-monotonic behavior, with 29 percent in $S8$ and 39 percent in $D8$. In the dynamics tasks, $D8$ and $D16$, IRRA is observed more frequently than in the static tasks, $S8$ and $S16$. CRRA is observed most frequently in $S8$ (20 percent), and also in $D8$ (11 percent).

Result 4: *The dynamic task leads to a higher Sharpe ratio relative to the static task.*

More efficient portfolio choices and lower risk-aversion in the dynamic task lead to a higher Sharpe ratio, relative to the static task. Figure 7 presents the average terminal wealth across the states (μ) on the y-axis and the standard deviation of the terminal wealth on the x-axis. We compute the expected value μ and standard deviation σ of the final wealth allocation as follows:

$$\mu := \frac{\sum_s c_s}{S}$$

and

$$\sigma := \sqrt{\frac{\sum_s (c_s - \mu)^2}{S}}$$

In Figure 7, pairs σ, μ are shown for each subject, and each treatment. There is a significant mass of subjects (about 15 percent) at $(0, 100)$ in $S8$ and $S16$, suggesting perfect hedging, with a risk-free rate of return. The hedging decision is not present at other points where $\sigma >> 0$. Figure 7 shows that $D16$ and $D8$ align quite well, following a similar trend, and that expected wealth appears to be lower in $S16$ and $S8$. The Sharpe ratios, defined as the risk-premium ($\mu - 100$) divided by σ, are 0.44 for $D16$ and 0.26 for $S16$, which are significantly different (using a Wilcoxon test, with a p-value of 0.001). For $D8$, the Sharpe ratio is 0.41 and for $S8$, 0.29 (with a p-value of 0.001).[22] If we exclude choices that result in a negative return ($\mu < 100$), then the Sharpe ratio for $S16$ goes up to 0.41, which is statistically equivalent to $D16$, and to 0.36 for $S8$, which is statistically equivalent to $D8$. The observed Sharpe ratio is smaller than the maximal Sharpe ratio, defined as $[\sum_{s=1}^{S} p_s^2/\pi_s - 1]^{1/2}$, of 0.61 for $S = 8$ and 0.72 for $S = 16$.[23]

[22]The Sharpe ratio is not defined when $\sigma = 0$, and hence we omit these observations from our analysis.

[23]In a complete market, the maximal Sharpe ratio can be computed by minimizing the mean squared return subject to an arbitrary return with a cost of zero. See Appendix A of Goetzmann et al. (2007).

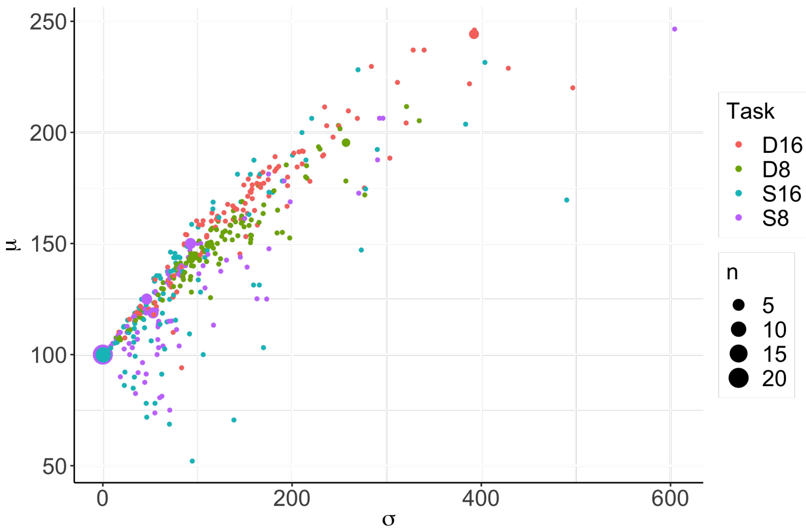

Figure 7: Expected value and volatility of returns

7 Discussion

In this paper we use non-parametric methods to study how individuals make dynamic portfolio choices in a laboratory experiment. Specifically, we employ a measure of efficiency proposed by Dybvig (1988) to test the optimality of portfolio choices, without specifying a parametric model of the investor's utility function, and then characterize the risk preferences of efficient choices using the techniques developed by Varian (1988).

Our results show that portfolio strategies of most subjects are efficient. Efficiency losses, when they occur, are greater when subjects face a large number of terminal states, when the task is static, and when subjects have a lower cognitive ability score. The latter complements the work of Grinblatt et al. (2011), who find that higher-IQ investors achieve higher Sharpe ratios. The correlation between cognitive skills and portfolio choice efficiency has important policy implications, suggesting that improving household financial literacy can yield large economic benefits. Further analysis of subject portfolio choices reveals that a form of stop-loss strategy generates some of the observed inefficiencies. We also find significant individual heterogeneity in the patterns of risk aversion. In particular, an important share of subjects display non-monotonic risk-aversion.

This paper provides important insights for how researchers should interpret heterogeneity in returns with respect to financial wealth. It is a well-known fact that there is a large variability in returns at the individual level. For instance, using data from Norway, Fagereng et al. (2020) show that individual returns on wealth have a standard deviation larger than 20 percent. The standard approach in financial economics is to interpret this heterogeneity as a result of differences in risk preferences and in individual exposure to background risk. While our lab experiment does not provide information on the latter, it does allow us to observe a large degree of heterogeneity in risk preferences and quantify the relative contribution of risk preferences and inefficiency to the overall heterogeneity in returns. Adjusting returns for inefficiency explains a small part of the variation in expected returns but most of the variation in risk-adjusted returns (Sharpe ratios).[24] We hope that these findings encourage other researchers and practitioners to apply the methods illustrated in this paper to elicit risk, and to measure the efficiency of investment choices.

8 Acknowledgements

We are grateful for the insightful discussions and comments received from Carole Bernard, Peter Bossaerts, Zachary Breig, Stephen Brown, Brice Corgnet, Abe de Jong, Dan Friedman, Ying Liang, Sébastien Pouget, Josh Shemesh, Stefan Trautmann, Jin Yu and participants at different seminars and talks. This project was approved by the human subjects committee at Monash University.

[24] Adjusting for inefficiency lowers the standard deviation of rates of return (across individuals) from 0.44 to 0.42 and lowers the standard deviation of Sharpe ratios from 0.23 to 0.09.

Appendix A: Inefficiency in static task

Table 8: Inefficiencies in S8 ($N = 62$ subjects) and S16 ($N = 88$ subjects)

S8			S16 upper nodes			S16 lower nodes		
Mistake	N	L (median)	Mistake	N	L (median)	Mistake	N	L (median)
$c_2 > c_1$	31	5.93	$c_2 > c_1$	49	2.43	$c_{10} > c_9$	23	17.78
$c_5 > c_1$	28	9.26	$c_5 > c_1$	49	2.58	$c_{13} > c_9$	22	23.73
$c_4 > c_3$	23	26.66	$c_4 > c_3$	14	29.11	$c_{12} > c_{11}$	21	31.35
$c_4 > c_5$	18	26.48	$c_4 > c_5$	22	20.97	$c_{12} > c_{13}$	23	21.56
$c_6 > c_2$	22	23.70	$c_6 > c_2$	23	21.56	$c_{14} > c_{10}$	11	19.75
$c_7 > c_3$	19	22.95	$c_7 > c_3$	23	20.60	$c_{15} > c_{11}$	17	31.35
$c_7 > c_5$	14	21.48	$c_7 > c_5$	26	20.49	$c_{15} > c_{13}$	15	26.86
$c_8 > c_7$	17	17.18	$c_8 > c_7$	16	32.34	$c_{16} > c_{15}$	21	26.67
						$c_8 > c_{11}$	22	26.77

Notes:
a. The terminal nodes are sorted from high (one) to low (eight for S8 and 16 for S16).
b. N is the count of subjects.
c. L is defined in equation (10).
d. The nodes for S16 is split in two. The upper nodes are 1-8 and the lower nodes are 9-16.

References

Amin, Gaurav S and Harry M Kat, "Hedge fund performance 1990–2000: Do the money machines really add value?," *Journal of financial and quantitative analysis*, 2003, *38* (2), 251–274.

Baltussen, Guido and Gerrit T Post, "Irrational diversification: An examination of individual portfolio choice," *Journal of Financial and Quantitative Analysis*, 2011, *46* (5), 1463–1491.

Bateman, Hazel, Christine Eckert, John Geweke, Jordan Louviere, Stephen Satchell, and Susan Thorp, "Risk presentation and portfolio choice," *Review of Finance*, 2016, *20* (1), 201–229.

Benartzi, Shlomo and Richard H Thaler, "Naive diversification strategies in defined contribution saving plans," *American economic review*, 2001, *91* (1), 79–98.

Bernard, Carole, Phelim P Boyle, and Steven Vanduffel, "Explicit representation of cost-efficient strategies," *Finance*, 2014, *35* (2), 5–55.

Biais, Bruno, Thomas Mariotti, Sophie Moinas, and Sébastien Pouget, "Asset pricing and risk sharing in a complete market: An experimental investigation," 2017.

Brocas, Isabelle, Juan D Carrillo, Aleksandar Giga, and Fernando Zapatero, "Risk aversion in a dynamic asset allocation experiment," *Journal of Financial and Quantitative Analysis*, 2019, *54* (5), 2209–2232.

Chen, Daniel L, Martin Schonger, and Chris Wickens, "oTreeAn open-source platform for laboratory, online, and field experiments," *Journal of Behavioral and Experimental Finance*, 2016, *9*, 88–97.

Choi, Syngjoo, Raymond Fisman, Douglas Gale, and Shachar Kariv, "Consistency and heterogeneity of individual behavior under uncertainty," *American economic review*, 2007, *97* (5), 1921–1938.

_ , Shachar Kariv, Wieland Müller, and Dan Silverman, "Who is (more) rational?," *American Economic Review*, 2014, *104* (6), 1518–50.

Corgnet, Brice, Mark Desantis, and David Porter, "What makes a good trader? On the role of intuition and reflection on trader performance," *The Journal of Finance*, 2018, *73* (3), 1113–1137.

Cox, John C and Hayne E Leland, "On dynamic investment strategies," *Journal of Economic Dynamics and Control*, 2000, *24* (11-12), 1859–1880.

d'Acremont, Mathieu and Peter Bossaerts, "Neurobiological studies of risk assessment: a comparison of expected utility and mean-variance approaches," *Cognitive, Affective, & Behavioral Neuroscience*, 2008, *8* (4), 363–374.

Danthine, Jean-Pierre and John B Donaldson, *Intermediate financial theory*, academic press, 2014.

Deck, Cary and Harris Schlesinger, "Consistency of higher order risk preferences," *Econometrica*, 2014, *82* (5), 1913–1943.

Dybvig, Phillip H, "Inefficient dynamic portfolio strategies or how to throw away a million dollars in the stock market," *The Review of Financial Studies*, 1988, *1* (1), 67–88.

Ebert, Sebastian and Daniel Wiesen, "Testing for prudence and skewness seeking," *Management Science*, 2011, *57* (7), 1334–1349.

Fagereng, Andreas, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri, "Heterogeneity and persistence in returns to wealth," *Econometrica*, 2020, *88* (1), 115–170.

Fischbacher, Urs, Gerson Hoffmann, and Simeon Schudy, "The causal effect of stop-loss and take-gain orders on the disposition effect," *The Review of Financial Studies*, 2017, *30* (6), 2110–2129.

Frederick, Shane, "Cognitive reflection and decision making," *Journal of Economic perspectives*, 2005, *19* (4), 25–42.

Friedman, Daniel, Sameh Habib, Duncan James, and Brett Williams, "Varieties of risk elicitation," Technical Report, Working Paper 2019.

Frydman, Cary, Nicholas Barberis, Colin Camerer, Peter Bossaerts, and Antonio Rangel, "Using neural data to test a theory of investor behavior: An application to realization utility," *The Journal of finance*, 2014, *69* (2), 907–946.

Gneezy, Uri and Jan Potters, "An experiment on risk taking and evaluation periods," *The Quarterly Journal of Economics*, 1997, *112* (2), 631–645.

Goetzmann, William, Jonathan Ingersoll, Matthew Spiegel, and Ivo Welch, "Portfolio performance manipulation and manipulation-proof performance measures," *The Review of Financial Studies*, 2007, *20* (5), 1503–1546.

Goldstein, Daniel G, Eric J Johnson, and William F Sharpe, "Choosing outcomes versus choosing products: Consumer-focused retirement investment advice," *Journal of Consumer Research*, 2008, *35* (3), 440–456.

Grinblatt, Mark, Matti Keloharju, and Juhani Linnainmaa, "IQ and stock market participation," *The Journal of Finance*, 2011, *66* (6), 2121–2164.

Harrison, Glenn W, Morten I Lau, and E Elisabet Rutström, "Estimating risk attitudes in Denmark: A field experiment," *scandinavian Journal of Economics*, 2007, *109* (2), 341–368.

Holt, Charles A and Susan K Laury, "Risk aversion and incentive effects," *American economic review*, 2002, *92* (5), 1644–1655.

Huberman, Gur and Wei Jiang, "Offering versus choice in 401 (k) plans: Equity exposure and number of funds," *The Journal of Finance*, 2006, *61* (2), 763–801.

Kaufmann, Christine, Martin Weber, and Emily Haisley, "The role of experience sampling and graphical displays on one's investment risk appetite," *Management science*, 2013, *59* (2), 323–340.

Kimball, Miles, "Precautionary Saving in the Small and in the Large," *Econometrica*, 1990, *58* (1), 53–73.

Kopa, Miloš and Thierry Post, "A portfolio optimality test based on the first-order stochastic dominance criterion," *Journal of Financial and Quantitative Analysis*, 2009, *44* (5), 1103–1124.

Kroll, Yoram and Haim Levy, "Further tests of the separation theorem and the capital asset pricing model," *The American Economic Review*, 1992, *82* (3), 664–670.

_ , _ , and Amnon Rapoport, "Experimental tests of the separation theorem and the capital asset pricing model," *The American Economic Review*, 1988, pp. 500–519.

Levy, Haim, Moshe Levy, and Natalie Alisof, "Homemade Leverage," *The Journal of Portfolio Management*, 2004, *31* (1), 84–93.

Magnani, Jacopo, "Testing for the disposition effect on optimal stopping decisions," *American Economic Review*, 2015, *105* (5), 371–75.

Noussair, Charles N, Stefan T Trautmann, and Gijs Van de Kuilen, "Higher order risk attitudes, demographics, and financial decisions," *Review of Economic Studies*, 2014, *81* (1), 325–355.

Post, Thierry, "Empirical tests for stochastic dominance efficiency," *The Journal of Finance*, 2003, *58* (5), 1905–1931.

_ **and Haim Levy**, "Does risk seeking drive stock prices? A stochastic dominance analysis of aggregate investor preferences and beliefs," *The Review of Financial Studies*, 2005, *18* (3), 925–953.

Rapoport, Amnon, Rami Zwick, and Sandra G Funk, "Selection of portfolios with risky and riskless assets: Experimental tests of two expected utility models," *Journal of economic psychology*, 1988, *9* (2), 169–194.

Rud, Olga A, Jean Paul Rabanal, and Manizha Sharifova, "An experiment on the efficiency of bilateral exchange under incomplete markets," *Games and Economic Behavior*, 2019, *114*, 253–267.

Sonnemann, Ulrich, Colin F Camerer, Craig R Fox, and Thomas Langer, "How psychological framing affects economic market prices in the lab and field," *Proceedings of the National academy of Sciences*, 2013, *110* (29), 11779–11784.

Sundali, James A and Federico Guerrero, "Managing a 401 (k) account: An experiment on asset allocation," *The Journal of Behavioral Finance*, 2009, *10* (2), 108–124.

Thaler, Richard H and Eric J Johnson, "Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice," *Management science*, 1990, *36* (6), 643–660.

Trautmann, Stefan T and Gijs van de Kuilen, "Higher order risk attitudes: A review of experimental evidence," *European Economic Review*, 2018, *103*, 108–124.

Varian, Hal R, "Estimating risk aversion from Arrow-Debreu portfolio choice," *Econometrica: Journal of the Econometric Society*, 1988, pp. 973–979.

Weber, Martin and Colin F Camerer, "The disposition effect in securities trading: An experimental analysis," *Journal of Economic Behavior & Organization*, 1998, *33* (2), 167–184.