Myopic Loss Aversion or Randomness in Choice? An Experimental Investigation*

Piotr Evdokimov
ITAM

First version: October 6, 2016
This version: February 11, 2017

Abstract

This paper reinterprets a well-known behavioral phenomenon, usually attributed to myopic loss aversion (MLA), through the lens of stochastic choice. First, it is argued that stochastic choice models can explain behavior in prior studies of MLA and reconcile some conflicting results in the literature. Second, the predictions of stochastic choice and MLA are contrasted in a new experiment based on simple choices between gambles and sure amounts. The results of this experiment show subjects to be more risk-averse with frequent feedback if the gamble is attractive, as in Gneezy and Potters (1997), and more *risk-seeking* with frequent feedback if the gamble is unattractive, as in Haisley et al. (2008). This pattern of results can be rationalized by a random utility or a random parameter model, but not MLA.

*The author thanks Simon Board, David Cooper, Joshua Miller, Pietro Ortoleva, and the brown bag participants at ITAM for helpful comments. The financial support of the Asociación Mexicana de Cultura A.C. is acknowledged.

1 Introduction

> *When faced with repeated choices between x and y, people often choose x in some instances and y in others. Furthermore, such inconsistencies are observed even in the absence of systematic changes in the decision makers taste which might be due to learning or sequential effects. It seems, therefore, that the observed inconsistencies reflect inherent variability or momentary fluctuation in the evaluative process. This consideration suggests that preference should be defined in a probabilistic fashion.*
>
> Tversky (1969, p. 31)

There is an element of randomness in choice. Moreover, randomness may bias decisions more in some situations than in others, which leads to patterns that can be misinterpreted by a hypothesis that doesn't take the random element into account.[1] The present paper argues that allowing for randomness in choice provides a more comprehensive account of the existing data than myopic loss aversion (MLA) in its widely accepted deterministic formulation (Benartzi and Thaler, 1995; Thaler et al., 1997; Gneezy and Potters, 1997; Benartzi and Thaler, 1999; Haigh and List, 2005). Importantly, the paper tests the stochastic choice explanation with a new experiment based on simple choices between lotteries and sure amounts. The results of this experiment are found to be consistent with a random utility model (RUM) or a random parameter model (RPM), but not MLA.

To understand how MLA has been studied in previous experiments, it is useful to consider the following anecdote from Samuelson (1963):

> [...] a few years ago I offered some lunch colleagues to bet each $200 to $100 that the side of a coin *they* specified would not appear at the first toss. One distinguished scholar - who lays no claim to advanced mathematical skills - gave me the following answer:
>
> ”I won't bet because I would feel the $100 loss more than the $200 gain. But I'll take you on if you promise to let me make 100 such bets.”

One theoretical foundation for the distinguished scholar's response is provided in Thaler et al. (1997). Assume that the decision maker (DM) has a utility function of the following

[1] See Regenwetter et al. (2011) for an illustration of this point in the context of reexamining violations of transitivity and Andersson et al. (2016) in the context of reexamining the relationship between risk aversion and cognitive skills.

form:[2]

$$u(x) = \begin{cases} x & \text{if } x \geq 0 \\ \lambda x & \text{if } x < 0 \end{cases} \tag{1}$$

Then, the expected utility of a single play of Samuelson's lottery is given by $U(\text{Lottery}) = (1/2)(200) - (1/2)\lambda(100) = 100 - 50\lambda$, while the expected utility of playing the lottery *twice* is $U(\text{Two lotteries}) = (1/4)400 + (1/2)(100) - (1/4)\lambda(200) = 150 - 50\lambda$. If $\lambda = 2.5$, then $U(\text{Lottery}) < 0$ and $U(\text{Two lotteries}) > 0$. Thus, if (i) lotteries are evaluated in an aggregated way[3] and (ii) the DM is *loss averse*, then the DM will appear more risk-averse when he makes risky decisions one at a time than when he makes them two at time. MLA consists of the combination of assumptions (i) and (ii) above.

Stochastic choice can provide an alternative explanation of the outcome of Samuelson's thought experiment. To see this, assume a risk-neutral RUM with $U(\text{Lottery}) = 50 + \epsilon$, $U(\text{Two lotteries}) = 100 + \epsilon$, $U(0) = \epsilon$, and ϵ independently drawn from the logistic distribution for every choice.[4] It follows that the probability of choosing a single lottery over a sure amount of zero is given by $\frac{1}{1+\exp(-50)}$, while the probability of choosing two lotteries is given by $\frac{1}{1+\exp(-100)} > \frac{1}{1+\exp(-50)}$. Intuitively, the stakes are higher in the decision involving two lotteries, which amplifies utility differences and makes it less likely that the random element ϵ influences the DM's decision.

Section 2 of the present paper shows that an argument along these lines can rationalize the data of several prior studies of MLA in the literature, while Sections 3, 4, and 5 of the paper compare the predictions of a wide class of RUMs and RPMs to those of MLA by means of a new laboratory experiment. The idea behind the experiment can also be illustrated using the lotteries of Samuelson. To this end, assume that the DM's preferences over final outcomes are as in Equation 1, and notice that $\lambda > 1$ if and only if $U(\text{Two lotteries}) > 2U(\text{Lottery})$. Let Decision H refer to a choice between a certain amount $c \in \{25, 75\}$ and a single play of Samuelson's lottery, and let Decision L refer to a choice between a certain amount $2 \times c$ and playing Samuelson's lottery two times. Assume that a certain amount c is accepted in Decision H. Then $U(\text{Lottery}) > c$, and hence $U(\text{Two lotteries}) > 2U(\text{Lottery}) > 2c$. Any DM with $\lambda > 1$ will be at least as likely to make the risky choice in Decision L *for both values of* $c \in \{25, 75\}$.[5] It is therefore not possible that a random sample of loss-averse subjects

[2] While we are assuming a reference point of zero here, the argument also holds under other assumptions (Section 4).

[3] This is referred to as *mental accounting* in the behavioral literature and reduction of compound lotteries in expected utility theory.

[4] Risk-neutrality is not necessary for this argument, and a RPM produces similar qualitative conclusions.

[5] In fact, a loss-averse decision maker will choose the safe option in both decisions when $c = 75$. Linearity

shows more risk-taking in Decision H when $c = 25$ and more risk-taking in Decision L when $c = 75$. This kind of behavior, however, is what we observe in the data.

This pattern of results can be rationalized by a RUM. To see this, let $U(\text{Lottery}) = 50 + \epsilon$ in Decision H, $U(\text{Two lotteries}) = 100 + \epsilon$ in Decision L, and $U(\text{Certain amount}) = \text{Certain amount} + \epsilon$ in both decisions, where ϵ is randomly drawn from an extreme value distribution for each choice. It follows that when $c = 25$,

$$P(\text{Lottery}) = \frac{1}{1 + \exp(25 - 50)} < \frac{1}{1 + \exp(2 \times (25 - 50))} = P(\text{Two lotteries}) \tag{3}$$

while when $c = 75$,

$$P(\text{Lottery}) = \frac{1}{1 + \exp(75 - 50)} > \frac{1}{1 + \exp(2 \times (75 - 50))} = P(\text{Two lotteries}) \tag{4}$$

The simple RUM argument can be generalized to situations where there is more than than two decisions to be made and where the non-random part of the utility function is risk- or loss-averse or even risk-seeking. For a wide class of utility functions, it will remain the case that extremely unattractive decisions (e.g., accepting a relatively small certain amount over a lottery or accepting a lottery over a relatively large certain amount) will be less likely when feedback is received infrequently.[6]

It has been argued that a RUM may lead to biased estimates of preference parameters, and that a RPM can provide a better alternative (Apesteguia and Ballester, 2016). While estimating preference parameters is not the goal of the present paper, it is not difficult to show that a RPM can predict the pattern of results in Equation 3 and Equation 4. To see this, assume that the DM's preferences are as in Equation 1, with the caveat that the λ parameter is subject to shocks. Assume for simplicity that the shocks follow the logistic distribution. Then, the probability that the lottery is chosen over c in Decision H is given

is not necessary for this argument. Assume, for example, that the utility function is concave for gains and convex for losses (Tversky and Kahneman, 1992):

$$u(x) = \begin{cases} x^{\alpha} & \text{if } x \geq 0 \\ -\lambda(-x)^{\alpha} & \text{if } x < 0 \end{cases} \tag{2}$$

When $c = 25$, the threshold λ for choosing the lottery remains greater in Decision L than in Decision H for any $\sigma \in (0, 1]$. When $c = 75$, the threshold λ is below one in both Decision H and Decision L for any $\sigma \in (0, 1]$, which means that a loss-averse decision maker would choose the safe option in both decisions.

[6]Here and elsewhere in this paper, "unattractive" means "unattractive when evaluated using the non-random component of the utility function."

by:

$$P(\text{Lottery}) = P(\lambda + \epsilon < \lambda(c, H)) = \frac{1}{1 + \exp(\lambda - \lambda(c, H)))},$$

where $\lambda(c, H)$ makes the decision maker indifferent between playing the lottery and choosing the sure amount. Similarly, the probability that the lottery is chosen over $2c$ in Decision L is:

$$P(\text{Two lotteries}) = P(\lambda + \epsilon < \lambda(2c, L)) = \frac{1}{1 + \exp(\lambda - \lambda(2c, L)))}.$$

It follows that, for a given c, $P(\text{Lottery}) > P(\text{Two lotteries})$ if and only if $\lambda(c, H) > \lambda(2c, L)$. Intuitively, if $\lambda(c, H) > \lambda(2c, L)$, there are less shocks to λ that can switch the decision to the safe option in Decision H than in Decision L. It's easy to check that $\lambda(25, H) = 1.5 < 2 = \lambda(50, L)$ and $\lambda(75, H) = 0.5 > 0 = \lambda(150, L)$. Hence, the DM acts more risk-averse in Decision H when $c = 25$ and less risk-averse in Decision H when $c = 75$. Notice that this argument holds even if $\lambda = 1$, i.e. if the DM is risk-neutral on average.

Because the literature on stochastic choice is large, providing an exhaustive survey is beyond the scope of this paper. Several studies have made the point that because experiments provide stochastic data, testing models of decision making requires an allowance for a stochastic element in choice. These include Iverson and Falmagne (1985), Busemeyer and Townsend (1992), Harless and Camerer (1994), Hey and Orme (1994), Carbone and Hey (2000), Loomes (2005), Regenwetter et al. (2011), Agranov and Ortoleva (2016), and Andersson et al. (2016). While none of these papers focused on MLA, Blavatskyy and Pogrebna (2010) made the point that data from some existing experiments, such as Gneezy and Potters (1997), can be better explained by a stochastic choice model than MLA. They stopped short of pointing out that stochastic choice can predict the pattern of greater risk taking with frequent evaluations. They also did not provide an experimental test of any stochastic choice model, as is done in the present paper.

2 Prior Experimental Studies of MLA

Stochastic choice provides a more comprehensive explanation of the existing experimental data than MLA. In what follows, Thaler et al. (1997), Gneezy and Potters (1997) (as well as related studies such as Haigh and List 2005 and Fellner and Sutter 2009), and Benartzi and Thaler (1999), are taken as prominent examples of studies showing evidence of (non-random) MLA while Langer and Weber (2001), Langer and Weber (2005), and Haisley et al. (2008) are taken as examples of studies the results of which are incompatible with the same hypothesis. For the sake of simplicity, we focus on the predictions of a RUM. As argued in the introduction, similar qualitative predictions (riskier behavior with low frequency feedback for unattractive lotteries and with high frequency feedback for attractive lotteries) can be obtained from a RPM.

2.1 Gneezy and Potters (1997) and Related Studies

In Gneezy and Potters (1997), the subjects decide how much to invest out of a 200 cent endowment. The investment is lost with probability 2/3 and otherwise returned with interest. Specifically, given an investment about x, the DM faces a lottery that earns $200 - x$ with probability 2/3 and $200 + 2.5x$ with probability 1/3, with expected value $200 + (1/6)x$. Notice that a risk-neutral investor would find the largest possible investment ($x = 200$) to be attractive, as the expected value is increasing in x.

In Treatment H, subjects make such investment decisions one at a time, with feedback following every decision. In Treatment L, they make these decisions three at a time, with the restriction that the investment is identical in each of the three decisions ($x_1 = x_2 = x_3$, where x_i is the investment in decision i). Thus, given an investment x, subjects face an investment that pays $600 - 3x$ with probability 8/27, $600 + 0.5x$ with probability 12/27, $600 + 4x$ with probability 6/27 and $600 + 7.5x$ with probability 1/27, with expected value $600 - (48/54)x + (12/54)x + (48/54)x + (15/54)x = 600 + (1/2)x$. Notice, again, that a risk neutral investor finds the largest possible investment ($x = 200$) to be optimal, although the stakes in Treatment H are higher than in Treatment L.

To see how a RUM can generate the prediction of smaller investments in Treatment H, assume that the underlying utility is risk-neutral and perturbed by a random error ϵ, so that $U_L(x) = 200 + (1/6)x + \epsilon$ and $U_H(x) = 600 + (1/2)x + \epsilon$. If ϵ follows an extreme value distribution, then:

$$P_L(x) = \frac{\exp(200 + (1/6)x)}{\sum_{x'} \exp(200 + (1/6)x')}$$

and

$$P_H(x) = \frac{\exp(600 + (1/2)x)}{\sum_{x'} \exp(600 + (1/2)x')}$$

It follows that the expected bet is $\sum_x xP_L(x) = 141.63$ (71% of the endowment) in Treatment H and $\sum_x xP_L(x) = 176.54$ (88% of the endowment) in Treatment L. A risk-neutral RUM generates the prediction of more risk-averse behavior with frequent feedback in Gneezy and Potters (1997). While the predicted bets are slightly higher than those observed in the Gneezy and Potters (1997) data, a better fit can be obtained by introducing a risk aversion parameter. Thus, with a CRRA utility function $u(x) = x^r$, a risk aversion parameter of $r = 0.91$, and a logistically distributed error term (as above), we get a predicted bet of 54% of the endowment in Treatment H and 71% of the endowment in Treatment L, much like the average bets of 54.7% and 71.9% observed in Rounds 7-9 of the Gneezy and Potters (1997) experiment.[7]

Several studies of MLA provide replications of Gneezy and Potters (1997) using the same parameters as Gneezy and Potters (1997) but different subject pools and, in the case of Fellner and Sutter (2009), additional experimental conditions. Examples of these include Fellner and Sutter (2009), Haigh and List (2005), and one of the treatments in Langer and Weber (2005). For these studies, the prediction of greater risk aversion with high frequency feedback can be obtained by the argument outlined above.

2.2 Thaler et al. (1997)

In Thaler et al. (1997), subjects choose how to allocate a portfolio of 100 shares between two investments. Fund A as mean return of 0.25 percent per period and a standard deviation of 0.177, while Fund B has a mean return of 1 percent per period with a standard deviation of 3.54 percent. The experiment was designed to simulate monthly and yearly investments, with monthly investments corresponding to a high frequency feedback condition in Gneezy and Potters (1997), and yearly ones corresponding to a low frequency of feedback one. Subjects made 200 decisions in the monthly condition and 25 decisions in the yearly one, with each decision binding for eight periods in the latter case. Notice that a risk neutral DM, whose decision is unaffected by the variance of the income-generating process, is predicted to allocate all of his endowment to Fund B in both conditions. Following the discussion above, assume that the utility function is risk-neutral with a random component following an extreme value distribution. Because the stakes are lower in the monthly condition than in the yearly one, the model predicts greater investment in Fund B in the latter case.

[7] It's worth pointing out that a CRRA utility function with any risk aversion parameter predicts greater risk taking in Treatment L than in Treatment H of Gneezy and Potters (1997) *with and without a stochastic component* (Iturbe-Ormaetxe et al. 2015). On the other hand, CRRA utility without randomness cannot account for the results of Haisley et al. (2008) and other experiments that failed to replicate the Gneezy and Potters (1997) results under different experimental conditions, including the experiment reported in the present paper.

2.3 Benartzi and Thaler (1999)

Benartzi and Thaler (1999) provides the results of several separate studies aimed at investigating MLA. Study 1 provides an experimental test of Samuelson's thought experiment and finds evidence of MLA among MBA students; see Section 1 for how this can be accommodated by a RUM. Study 2 tests the MLA hypothesis using simple binary gambles and therefore is the closest in spirit to the experiment reported in the present paper. It is unlike the present experiment, however, in that a risk-neutral DM would find the gambles in Study 2 of Benartzi and Thaler (1999) attractive. A random utility based model predicts greater risk-aversion in a high frequency condition *for any such gamble*. Studies 3 and 4 find evidence in favor of MLA using investment tasks similar to that in Thaler et al. (1997); their results can be rationalized by an argument analogous to that in the previous subsection.

2.4 Haisley et al. (2008)

In Haisley et al. (2008), subjects choose whether or not to buy state lottery tickets. They make these decisions one at a time in a high frequency treatent and several at a time in a low frequency treatment. Unlike the studies described above, Haisley et al. (2008) find subjects to be *less* risk-averse in a high frequency condition. The authors attribute this to over-weighting of small probabilities (of winning the lottery) and under-weighting of the cost of the lottery ticket. An alternative explanation of their results can be provided by a RUM. Thus, because a risk-neutral DM would find the lotteries in Haisley et al. (2008) unattractive, one can think of the decision to buy a lottery ticket as attributable to randomness. Such randomness has less of an influence in the low frequency condition, where the decisions are made several at a time and the stakes are higher.

2.5 Langer and Weber (2001, 2005)

Like Haisley et al. (2008), Langer and Weber (2001) find that subjects are sometimes more willing to take risks a high frequency treatment. Specifically, given a lottery that pays 400 units with probability 96% and subtracts 2100 with probability 4%, Langer and Weber (2001) find an acceptance rate of 20% for a single choice and an acceptance rate of 10% if subjects were shown an aggregated distribution associated with playing the lottery twice. The low acceptance rates in both conditions suggest that most subjects find this lottery unattractive, and this pattern of results can be rationalized by a RUM assuming that subjects are moderately risk-averse. Assume, for example, that

$$u(x) = \begin{cases} x^r & \text{if } x \geq 0 \\ -(-x)^{1/r} & \text{if } x < 0 \end{cases} \tag{5}$$

If r is low enough, the expected value of the gamble will be negative and smaller in the low frequency condition than in the high frequency one. The results in Langer and Weber (2005) can be explained by a similar argument. There, the authors use accepted parameters from the prospect theory literature to predict acceptance rates of 22.7% in the high frequency condition and 19.3% in the low frequency condition for a gamble that pays 15% interest with probability 0.9 and makes subjects lose their investment with probability 10%. These low acceptance rates suggest the gamble to be unattractive, and, indeed, it would be unattractive to a sufficiently risk-averse DM.

3 A New Experiment

The experiment described below was designed to be as simple as possible,[8] be relatable to existing MLA experiments in the literature, and provide reasonably high incentives to the subjects. It can be seen as an intersection of Gneezy and Potters (1997) and Benartzi and Thaler (1999). The experiment is similar to Benartzi and Thaler (1999) in that it uses simple binary gambles and similar to Gneezy and Potters (1997) in that subjects made nine decisions, one decision at a time, in Treatment H and three decisions at a time in Treatment L. Some additional methodological details that the experiment shares with Gneezy and Potters (1997) are described in detail below.

Each session in the experiment was randomly assigned into either Treatment H (high frequency of feedback) or Treatment L (low frequency of feedback). Both treatments had the following procedural features in common. After handing in their consent forms, the subjects were handed out their instructions. The instructions were then read out loud to the subjects. At this point, the subjects were allowed an opportunity to ask questions. After all the questions were privately answered, the subjects made practice decisions (three practice decisions in Treatment H and one practice decision in Treatment L). It was understood that any decisions made at this point would not count for the subjects' earnings. Following the practice decisions, subjects made decisions for real money. After all the decisions were made,

[8]Hence the use of binary gambles. In Gneezy and Potters (1997), for example, the subjects effectively choose between 100 lotteries in every decision (Section 2.1).

each subject was privately paid.

In **Treatment H**, each subject was randomly assigned a personal winning outcome (heads or tails) in the beginning of the session. They were then told that they will make several decisions between a sure amount and a lottery that pays 30 pesos with probability 50% and 0 pesos with probability 50%. At the time of the experiment, a 15km Uber ride from the author's house to the airport cost around 80 pesos.

In any given round, after every subject in the session made his or her decision, the researcher flipped a 10 peso coin. If any subject chose the lottery and the outcome of the coin flip matched his or her personal winning outcome, 30 pesos were added to the subject's earnings. If any subject chose the lottery and the outcome of the coin flip did not match his or her personal winning outcome, nothing was added to the subject's earnings for the decision. If any subject chose the certain amount, the certain amount was added to the subject's earnings. Every subject made three practice decisions with a certain amount of 15 before making decisions for real money.

When making decisions for real money, each subject was randomly assigned an order of the following four certain amounts: 10, 13, 16, 19. The subjects did not know what certain amounts they will face before making their decisions. They also made three decisions for each certain amount. For example, a subject with the order (13, 16, 10, 19) first made three choices between the fixed lottery and a certain amount of 13, then three choices between the fixed lottery and a certain amount of 16, then three choices between the fixed lottery and a certain amount of 10, and then three choices between the fixed lottery and a certain amount of 19. This block structure was borrowed from Gneezy and Potters (1997) to make Treatment H and Treatment L as similar as possible. Following every decision, each subject waited for all other subjects' decisions to be made. The researcher then flipped a coin, announced the outcome out loud, asked one of the participants in the room to verify the outcome, and entered the outcome of the coin flip into his computer. The computer software then calculated each subject's earnings and displayed it on the subject's screen.

In **Treatment L**, as in Treatment H, each subject was randomly assigned a personal winning outcome (heads or tails) in the beginning of the session. They were then told that they will make several decisions between a sure amount and *three plays* of the lottery in Treatment H. In any given round, after every subject in the session made his or her decision, the researcher flipped a 10 peso coin *three times*. As in Gneezy and Potters (1997), the three outcomes were announced together after all three coin flips were made. This was done to facilitate the evaluation of the three lotteries in an aggregated way. If any subject chose the certain amount, the certain amount was added to the subject's earnings. If any subject chose the three lotteries, he or she was paid 30 pesos for each instance of the researcher's coin flip agreeing with the subject's personal winning outcome. For example, if the outcome was "two heads, one tails" and the subject's personal winning outcome was heads, the subject was paid 60 pesos. Each subject made one practice decision with a certain amount of 45 before making decisions for real money.

When making decisions for real money, each subject was randomly assigned an order of the following four certain amounts: 30, 39, 48, 57. Note that dividing these certain amounts by three we obtain the certain amounts in Treatment H. The subjects did not know what certain amounts they will face before making their decisions. They also made one decisions for each certain amount. Following every decision, each subject waited for all other subjects' decisions to be made. The researcher then flipped three coins, announced the outcomes out loud, asked one of the participants in the room to verify the outcomes, and entered them into his computer. The computer software then calculated each subject's earnings and displayed it on the subject's screen.

4 Predictions

4.1 MLA

Given an investment decision x, subjects in Gneezy and Potters (1997) faced a lottery that paid $(200 - x)$ with probability 2/3 and $200 + 2.5x$ with probability 1/3. As pointed out in Harrison and Rutstrom (2008), if we consider the reference point to be zero, then subjects actually faced no losses in that experiment. The implicit assumption is that gains and losses are evaluated relative to the reference point of 200, leading to expected utility of zero from choosing the certain amount and $(1/3)(2.5x) - (2/3)(\lambda x)$ from the lottery.

Following this interpretation, one possibility is that subjects in our experiment evaluated gains and losses relative to the certain amount they faced in each decision. Assume, then, that the reference point is given by $c \in \{10, 13, 15, 16, 19\}$ for every decision in Treatment H and $3c \in \{30, 39, 45, 48, 57\}$ for every decision in Treatment, and let subjects' reference-dependent utility be defined as follows:

$$u(x|r) = \begin{cases} x - r & \text{if } x - r \geq 0 \\ \lambda(x - r) & \text{if } x - r < 0 \end{cases} \tag{6}$$

It follows that the expected utility from picking the lottery in Treatment H is:

$$U(\text{Lottery}|c) = \frac{30-c}{2} - \frac{\lambda c}{2} = \frac{30-c(1+\lambda)}{2}, \tag{7}$$

while the expected utility from picking the three lotteries in Treatment L is:

$$\begin{aligned} U(\text{Three lotteries}|3c) &= \frac{90-3c}{8} + \frac{3(60-3c)}{8} + \frac{3\lambda(30-3c)}{8} - \frac{\lambda(3c)}{8} = \\ &= \frac{270-12c+\lambda(90-12c)}{8}. \end{aligned} \tag{8}$$

Notice that $U(\text{Three lotteries}|3c) > 3U(\text{Lottery}|c)$ if and only if $\lambda > 1$. Assume that $\lambda > 1$ and the lottery is chosen in Treatment H. Then,

$$U(\text{Lottery}|c) > U(c|c) = 0 \Rightarrow U(\text{Three lotteries}|3c) > U(3c|3c) = 0.$$

Thus, loss aversion predicts the DM to be more risk-averse in Treatment H than in Treatment L for any value of c. It's worth noting that the linearity assumption is unnecessary for this qualitative prediction. Assume, for example, that the utility function is concave for gains and convex for losses (Tversky and Kahneman, 1992):

$$u(x|r) = \begin{cases} (x-r)^\alpha & \text{if } x-r \geq 0 \\ -\lambda\big(-(x-r)\big)^\alpha & \text{if } x-r < 0 \end{cases}.$$

Then the risky option is chosen in Treatment H if and only if $\left(\frac{30-c}{c}\right)^\alpha > \lambda$, and in Treatment L if and only if $\frac{(90-3c)^\alpha+3(60-3c)^\alpha}{(3c)^\alpha+3(3c-30)^\alpha} > \lambda$. For any $\sigma \in (0,1]$, the threshold in remains greater in Treatment L when $c \in \{10,13\}$, equal to one in both treatments when $c = 15$, and below one in both treatments when $c \in \{16,19\}$.

Because the lottery was held fixed in both Treatment H and Treatment L while the certain amounts changed, another possibility is that the reference point is given by the expectation of the lottery. To make the statement more general, assume that the reference point is $\bar{r}_H$ in Treatment H for all values of c and $\bar{r}_L = 3\bar{r}_H$ in Treatment L for all values of c. Then, following an argument similar to the one above, $U(\text{Three lotteries}|\bar{r}_L) > 3U(\text{Lottery}|\bar{r}_H)$ if and only if $\lambda > 1$. Moreover, $U(3c|r_L) = U(3c|3r_H) = 3U(c|r_H)$. Therefore, loss aversion predicts the DM to take on more risks in Treatment L for any value of c.

Yet another possibility is that the lottery itself served as a reference point, as in Kőszegi and Rabin (2006). In this case, we may assume that the utility of a gamble F given a referent lottery G is given by:

$$U(F|G) = \int\int u(x|r)dF(x)dG(r) \quad (9)$$

with $u(x|r)$ determined as in Equation 6. Let G_1 denote the lottery with a 50% chance of earning 30 pesos. Because G_1 was fixed throughout Treatment H as one of the choices while the certain amounts changed, assume that it served as a stochastic referent. It's easy to check that $U(G_1|G_1) = U(15|G_1)$ for any $\lambda > 1$.[9] Thus, the certainty equivalent of the referent lottery is 15. I.e., a loss averse DM is predicted to appear risk-neutral in Treatment H.

Now consider Treatment L, and denote by G_2 the lottery which pays 90 with probability 1/8, 60 with probability 3/8, 30 with probability 3/8, and 0 with probability 1/8. If we assume that the DM uses G_2 as an expectations-based reference point, it can be shown that for any $\lambda > 1$,

$$U(G_2|G_2) < U(45|G_2)$$

Thus, a loss-averse DM is predicted to appear more risk-averse in Treatment L than in Treatment H. Notice that this prediction holds for any $\lambda > 1$. While the prediction is qualitatively the opposite to that suggested by deterministic reference points, note that neither stochastic nor deterministic reference points can accommodate a result where loss-averse subjects appear more risk-averse in Treatment L for low values of c and more risk-averse in Treatment H for high values of c.

4.2 RUM

Assume that the DM's utility is given by $U = EU + \epsilon$, where EU is the expected utility of the prospect being evaluated and ϵ is an i.i.d. error term with a logistic distribution. It follows that:

$$P(\text{Lottery}) = \frac{1}{1 + \exp(EU(c) - EU(\text{Lottery}))}$$

and

$$P(\text{Two lotteries}) = \frac{1}{1 + \exp(EU(3c) - EU(\text{Three lotteries}))}$$

[9]See Sprenger (2015) for the general argument using binary gambles.

Assume that $EU(3c) - EU(\text{Three lotteries}) < EU(c) - EU(\text{Lottery})$ for low values of c and $EU(3c) - EU(\text{Three lotteries}) > EU(c) - EU(\text{Lottery})$ for high values of c, which holds for a wide class of utility functions. It holds, for example, if the decision maker is risk-neutral. Then, the probability of choosing the lottery is smaller in Treatment H for low values of c and smaller in Treatment L for high values of c.

4.3 RPM

Assume preferences of the form given in Equation 6. By the same argument as that in the introduction,

$$P(\text{Lottery}) = \frac{1}{1 + \exp(\lambda - \lambda(c, H))}$$

and

$$P(\text{Two lotteries}) = \frac{1}{1 + \exp(\lambda - \lambda(3c, L))}.$$

As in the introduction, $\lambda(\cdot, \cdot)$ represents the value of λ that makes the decision maker indifferent between choosing the lottery and the certain amount.[10] It follows that the lottery is more likely to be chosen in Treatment H if and only if $\lambda(c, H) > \lambda(3c, L)$. Regardless of whether the reference point is given by the certain amount or the expected value of the lottery, this condition holds when $c \in \{16, 19\}$. Likewise, $\lambda(c, H) < \lambda(3c, L)$ when $c \in \{10, 13\}$. Thus, we can use the RPM to predict greater risk-taking in Treatment L for low values of c and greater risk-taking in Treatment H for high values of c.

5 Results

Data was collected at ITAM in September 2016 over the course of six sessions (three in each treatment) from 87 subjects (46 in Treatment H and 41 in Treatment L). Average earnings were approximately 188 pesos, slightly less than 10 dollars at the time the experiment was run.[11] Each session lasted close to 30 minutes on average.

The results of the experiment are shown in Table 1.[12] Safe decisions were more likely

[10] If we assume that the reference point is given by the certain amount, the first argument of $\lambda(\cdot, \cdot)$ represents both the certain amount *and* the reference point.

[11] As noted above, this is a relatively high payment in terms of purchasing power.

[12] While decisions in the $c = 15$ condition were not incentivized, they are included in the analysis because

	$c = 10$	$c = 13$	$c = 15$	$c = 16$	$c = 19$
Treatment H	**17%**	**24%**	**29%**	45%	60%
	(23/138)	**(33/138)**	**(40/138)**	(62/138)	(83/138)
	$\vee$	$\vee^{**}$	$\vee$	$\wedge$	$\wedge^{**}$
Treatment L	8%	8%	22%	**49%**	**78%**
	(3/41)	(3/41)	(9/41)	**(20/41)**	**(32/41)**

Table 1: **Observed probabilities of making a safe decision.** *When c was low, subjects were more risk-averse in Treatment H. When c was high, subjects were more risk-averse in Treatment L.* ** *denotes significance at a* $P < 0.05$ *level in a Fisher's exact test.*

in Treatment H of the experiment when c was low, which mirrors the findings of previous MLA experiments where risky choices are attractive to a risk-neutral DM. When c was high, safe decisions were more likely in Treatment L, which mirrors the findings of previous MLA experiments with unattractive risky choices, such as Haisley et al. (2008). This pattern of results is not predicted by MLA.

To study the significance of these findings, we can compare the distributions of safe choices in Treatment H and Treatment L for each c using a Fisher's exact test. The test shows that subjects were significantly more likely to make a safe choice in Treatment H when $c = 13$ and significantly more likely to make a safe choice in Treatment L when $c = 19$, with $P < 0.05$ in both cases. Notice from Table 1 that sign of the effect is always negative (less safe choices in Treatment H) when $c \leq 15$'s and positive (more safe choices in Treatment H) when $c > 16$.

While the comparisons above assume independence of observations, the results are robust to allowing standard errors to be correlated within subjects. Table 2 reports the marginal effects of a Treatment L dummy for each value of $c \in \{10, 13, 15, 16, 19\}$ in a logit model where the probability of making a safe choice is modeled as a function of (i) a Treatment L dummy, (ii) the c variable, and (iii) an interaction between the Treatment L dummy and c, with standard errors are clustered at the subject level. The estimated marginal effects suggest that the effect of feedback frequency was significant and negative when $c = 10$ and $c = 13$ ($P < 0.05$ in both cases) and significant and positive when $c = 19$ ($P < 0.05$). These results can be summarized as follows:

Result 1. *In line with the stochastic choice predictions but not MLA, subjects were more risk-averse in Treatment H when the lottery was attractive and more risk averse in Treatment*

they fit into the overall pattern of the data both in Treatment H and Treatment L. The major conclusions of the paper are not altered if we exclude these decisions from the statistical analysis.

	Dependent variable: Prob. of safe choice
$c = 10$	-0.171**
	(0.0814)
$c = 13$	-0.145**
	(0.0654)
$c = 15$	-0.0615
	(0.0548)
$c = 16$	0.00235
	(0.0564)
$c = 19$	0.196**
	(0.0893)
Observations	895

Subject-clustered standard errors in parentheses
** $p < 0.05$

Table 2: **The estimated marginal effect of integrating the three lottery decisions for each value of c.**

L when the lottery was unattractive.

The data can also be explored structurally using maximum likelihood. To this end, assume that the underlying utility function is given by Equation 6, with the reference point given by c in Treatment H and $3c$ in Treatment L. Then the expected value of the lottery in Treatment H is captured by Equation 7, while the expected value of the three lotteries in Treatment L is captured by Equation 8. Assume that the probability of choosing the lottery is given by

$$\frac{\exp(\alpha EV_c^k(\lambda))}{1+\exp(\alpha EV_c^k(\lambda))}, \tag{10}$$

where $k \in \{H, L\}$ denotes the treatment, $c \in \{10, 13, 15, 16, 19\}$, $EV_c^k(\lambda)$ is the expected value of the lottery, and α is a noise parameter. Estimating this model using all observations in the experiment, we get $\hat{\lambda} = 0.669$ and $\hat{\alpha} = 0.245$. Using subject-clustered standard errors, we get that $\hat{\lambda}$ is significantly smaller than one ($P < 0.001$). I.e., the data suggests loss-loving behavior even though subjects' behavior is consistent with the MLA pattern for attractive lotteries. This is highlighted below:

Result 2. *Assuming (i) that sure amounts serve as reference points and (ii) that the non-random component of preferences does not differ across subjects, behavior in the experiment is consistent with loss seeking ($\lambda < 1$).*

To see Result 2 from a different angle, the model in Equation 10 can be estimated under the assumption that $\lambda = 1$. This is a logit model with a single explanatory variable equal to the difference between the expected value of the lottery (15 in Treatment H, 45 in Treatment L) and the sure amount (c in Treatment H, $3c$ in Treatment L).[13] Figure 1 plots the *predicted* probabilities of making the safe choice for each value of c in both Treatment H and Treatment L together with the *observed* probabilities reported in Table 1. Black bars represent the predictions of the model in Equation 10, while light grey bars represent the predictions of the loss-neutral model with $\lambda = 1$. Notice that the loss-neutral model consistently over-estimates the probability of making the safe choice (the light grey bars are above the white bars in nine out of ten cases), which is is consistent with Result 2.

The estimated parameters λ and α do not significantly differ across treatments. Thus, if we use Treatment H data to estimate λ_H and σ_H and Treatment L data to estimate λ_L and σ_L, we find no significant treatment effects ($P > 0.1$ for both statistical comparisons). The treatment-specific parameters can also be used to make out-of-sample predictions. Thus, the dark grey bars in Figure 1 are predicted probabilities of making a safe choice, computed for Treatment H (resp., L) using the estimates $\hat{\lambda}_L$ and $\hat{\sigma}_L$ (resp., $\hat{\lambda}_H$ and $\hat{\sigma}_H$). Notice that the out-of-sample predictions still suggest that subjects are more risk-averse in Treatment H if and only if c is low.

If λ and α are allowed to vary *across sessions*, we get an estimate of λ smaller than one *in every session*. Comparing λ's and α's across sessions reveals no pairwise comparisons significant at a $P < 0.05$ level and only one pairwise comparison significant at a $P < 0.1$ level. This provides evidence against session effects in the experiment, thereby justifying the use of each subject as an independent observation.

Additional evidence of randomness in choice can be obtained by looking at subject-level data. Behavior of subject i is consistent with a deterministic cutoff if there exists a c_i such that the subject always chooses the lottery for all $c < c_i$ and always chooses the safe option for all $c \geq c_i$. By this definition, behavior of only 6/46 (13%) subjects in Treatment H is consistent with a deterministic cutoff (see Figures 3, 5, and 7 in the appendix). For each c in Treatment L, each subject had only one opportunity to choose between the three lotteries and $3c$, which necessarily makes randomness more difficult to detect in this treatment at the

[13]This model results in the estimate $\hat{\alpha} = 0.196$.

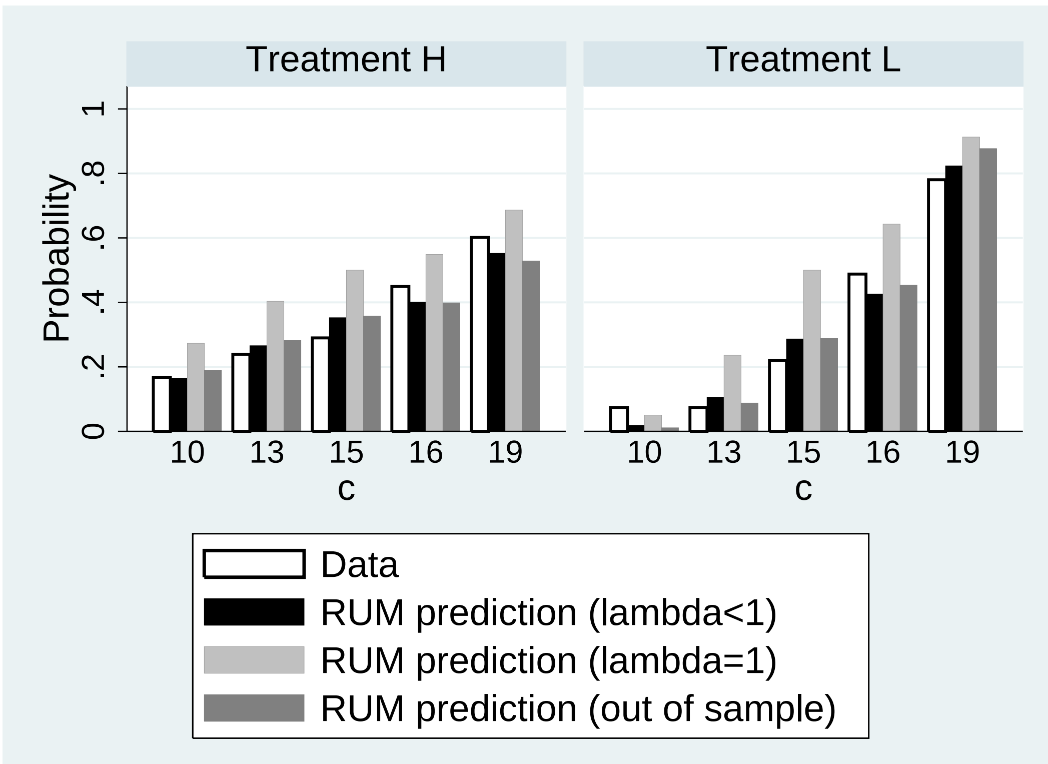

Figure 1: **Data and model predictions.** *When c was low, subjects were more risk-averse in Treatment H. When c was high, subjects were more risk-averse in Treatment L. This is reflected in the data (white bars), the predicted probabilities of a loss-averse RUM (black bars), the predicted probabilities of a risk-neutral RUM (light grey bars), and out-of-sample predicted probabilities (dark grey bars). The parameters of the RUMs are estimated using MLE. The treatment parameter c is on the horizontal axis.*

individual level. Moreover, because the stakes are higher, randomness should have less of an effect on behavior in Treatment L in theory. Nevertheless, we find that behavior of only 27/41 (66%) of subjects in Treatment L is consistent with a deterministic cutoff. Overall, approximately 38% of subjects show behavior consistent with a deterministic cutoff, which provides strong evidence of randomness in behavior.

6 Conclusion

A RUM or a RPM can better explain the results of the simple lottery choice experiment above than MLA. This suggests that the accepted interpretation of the effect of segregating lottery decisions on behavior should be reconsidered. While the present paper focuses on the evaluation period effect in MLA experiments, there are several promising avenues for future work. Other well-established behavioral regularities that have been given a deterministic explanation may possibly be reinterpreted through the lens of stochastic choice. Consider, for example, the following two problems from Kahneman and Tversky (1979):

Problem 7:
Option A: A payoff of 6000 with probability 45%,
Option B: A payoff of 3000 with probability 90%.

Problem 8:
Option A: A payoff of 6000 with probability 0.1%,
Option B: A payoff of 3000 with probability 0.2%.

Kahneman and Tversky (1979) find that subjects are more likely to choose B in Problem 7 and A in problem 8. This is inconsistent with expected utility because if $0.45u(6000) > 0.90u(3000)$, then $0.001u(6000) > 0.002u(3000)$. Because the stakes in Problem 7 are substantially higher than in Problem 8, a RUM can provide at least a partial account of these results. Thus, if there is a random component to the DM's utility function, this random component will have more of an influence in the latter problem. A risk-averse DM will therefore be more likely to choose Option B in Problem 7 than in Problem 8. The same intuition applies to Problems 3, 4, 5, and 6 in Kahneman and Tversky (1979). Future research should explore the extent to which stochastic choice can account for these and similar patterns of behavior in the judgment and decision making literature.

The message of this paper is that there is an element of randomness in behavior, and that a proper test of any model requires that the behavioral implications of the random element are accounted for. When we do account for these implications, we find that a loss averse utility function is unnecessary for rationalizing the typical pattern of results in the MLA literature. In Regenwetter et al. (2011), the authors state that Luce's *twofold challenge* (Luce, 1995) is to "(a) recast a deterministic theory as a probabilistic model (or a hypothesis) and (b) properly test that probabilistic model of the theory (or the hypothesis) on available data." An open question for future work is to determine how much of subjects' behavior is driven by the assumptions made in Kahneman and Tversky (1979) when prospect theory and expected utility are recast as probabilistic models.

References

Agranov, M. and P. Ortoleva (2016): "Stochastic choice and preferences for randomization," *Journal of Political Economy*, forthcoming.

Andersson, O., H. J. Holm, J.-R. Tyran, and E. Wengström (2016): "Risk aversion relates to cognitive ability: Preferences or Noise?" *Journal of the European Economic Association*, forthcoming.

Apesteguia, J. and M. A. Ballester (2016): "Monotone stochastic choice models: The case of risk and time preferences," *Journal of Political Economy*, forthcoming.

Benartzi, S. and R. H. Thaler (1995): "Myopic loss aversion and the equity premium puzzle," *The Quarterly Journal of Economics*, 110, 73–92.

——— (1999): "Risk aversion or myopia? Choices in repeated gambles and retirement investments," *Management science*, 45, 364–381.

Blavatskyy, P. R. and G. Pogrebna (2010): "Reevaluating evidence on myopic loss aversion: aggregate patterns versus individual choices," *Theory and decision*, 68, 159–171.

Busemeyer, J. R. and J. T. Townsend (1992): "Fundamental derivations from decision field theory," *Mathematical Social Sciences*, 23, 255–282.

Carbone, E. and J. D. Hey (2000): "Which error story is best?" *Journal of Risk and Uncertainty*, 20, 161–176.

Fellner, G. and M. Sutter (2009): "Causes, consequences, and cures of myopic loss aversion–An experimental investigation," *The Economic Journal*, 119, 900–916.

Gneezy, U. and J. Potters (1997): "An experiment on risk taking and evaluation periods," *The Quarterly Journal of Economics*, 631–645.

Haigh, M. S. and J. A. List (2005): "Do professional traders exhibit myopic loss aversion? An experimental analysis," *The Journal of Finance*, 60, 523–534.

Haisley, E., R. Mostafa, and G. Loewenstein (2008): "Myopic risk-seeking: The impact of narrow decision bracketing on lottery play," *Journal of Risk and Uncertainty*, 37, 57–75.

Harless, D. W. and C. F. Camerer (1994): "The predictive utility of generalized expected utility theories," *Econometrica: Journal of the Econometric Society*, 1251–1289.

Harrison, G. W. and E. E. Rutstrom (2008): "Risk aversion in the laboratory," *Risk Aversion in Experiments*, 12, 41.

Hey, J. D. and C. Orme (1994): "Investigating generalizations of expected utility theory using experimental data," *Econometrica: Journal of the Econometric Society*, 1291–1326.

Iturbe-Ormaetxe, I., G. Ponti, and J. Tomás (2015): "Some (mis) facts about myopic loss aversion," Working paper.

Iverson, G. and J.-C. Falmagne (1985): "Statistical issues in measurement," *Mathematical Social Sciences*, 10, 131–153.

Kahneman, D. and A. Tversky (1979): "Prospect theory: An analysis of decision under risk," *Econometrica: Journal of the econometric society*, 263–291.

Kőszegi, B. and M. Rabin (2006): "A model of reference-dependent preferences," *The Quarterly Journal of Economics*, 1133–1165.

Langer, T. and M. Weber (2001): "Prospect theory, mental accounting, and differences in aggregated and segregated evaluation of lottery portfolios," *Management Science*, 47, 716–733.

——— (2005): "Myopic prospect theory vs. myopic loss aversion: how general is the phenomenon?" *Journal of Economic Behavior & Organization*, 56, 25–38.

Loomes, G. (2005): "Modelling the stochastic component of behaviour in experiments: Some issues for the interpretation of data," *Experimental Economics*, 8, 301–323.

Luce, R. D. (1995): "Four tensions concerning mathematical modeling in psychology," *Annual Review of Psychology*, 46, 1.

Regenwetter, M., J. Dana, and C. P. Davis-Stober (2011): "Transitivity of preferences." *Psychological Review*, 118, 42.

Samuelson, P. (1963): "A model of reference-dependent preferences," *Scientia*, 108–113.

Sprenger, C. (2015): "An endowment effect for risk: Experimental tests of stochastic reference points," *Journal of Political Economy*, 123, 1456–1499.

Thaler, R. H., A. Tversky, D. Kahneman, and A. Schwartz (1997): "The effect of myopia and loss aversion on risk taking: An experimental test," *The Quarterly Journal of Economics*, 647–661.

Tversky, A. (1969): "Intransitivity of preferences." *Psychological review*, 76, 31.

Tversky, A. and D. Kahneman (1992): "Advances in prospect theory: Cumulative representation of uncertainty," *Journal of Risk and uncertainty*, 5, 297–323.

A Individual Behavior

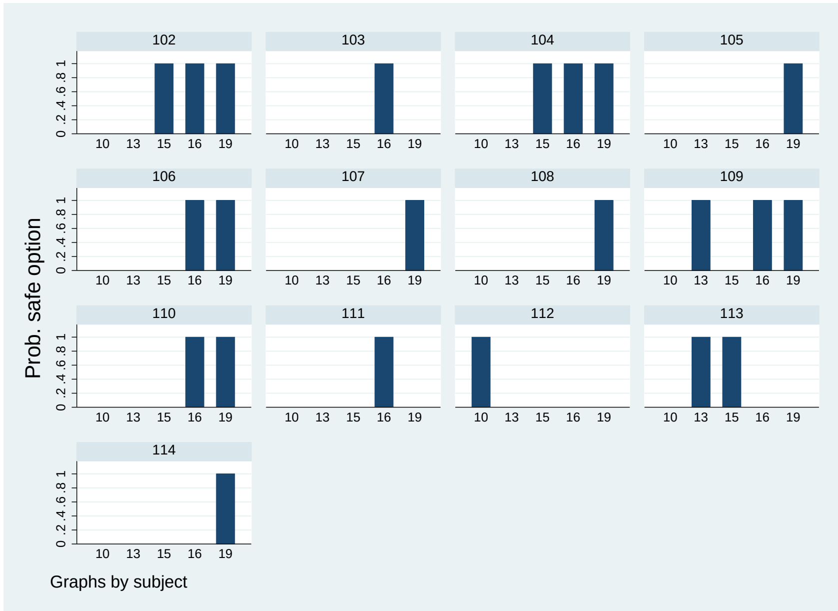

Figure 2: **Session 1.**

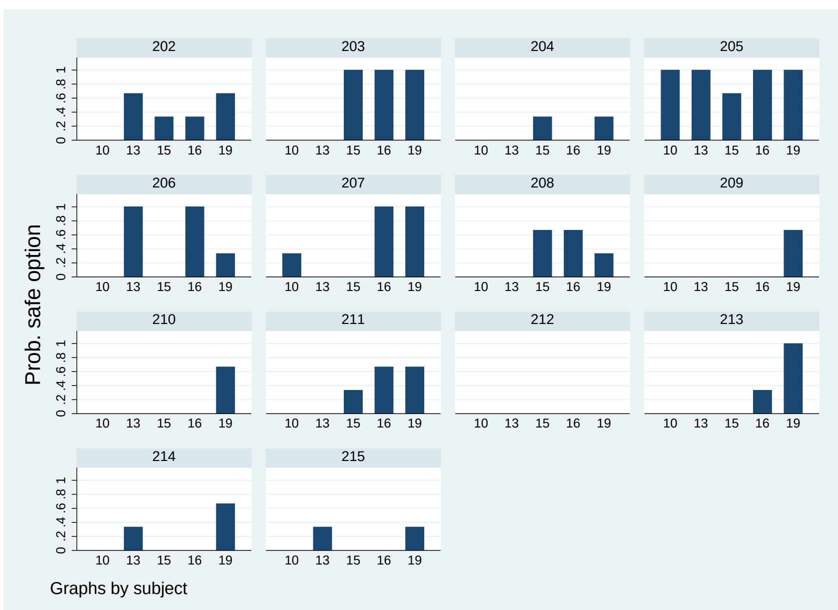

Figure 3: **Session 2.**

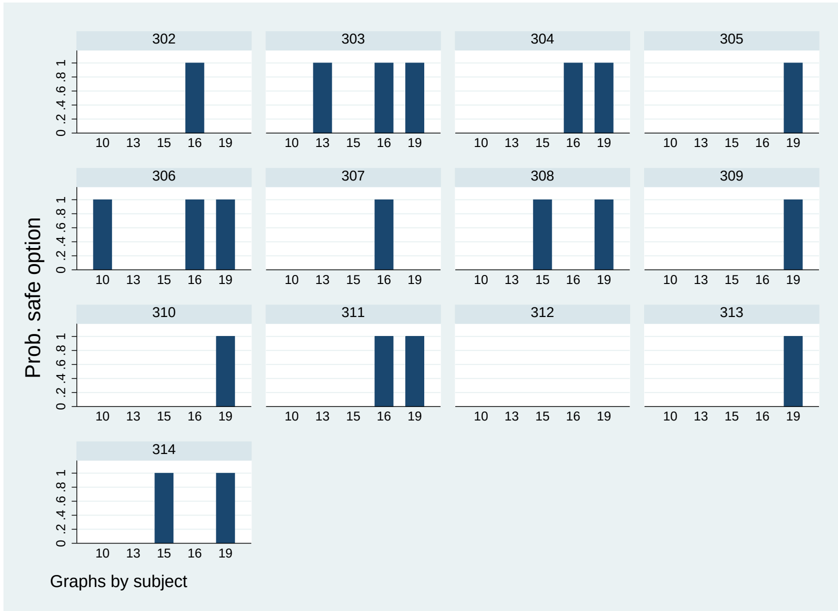

Figure 4: **Session 3.**

Figure 5: **Session 4.**

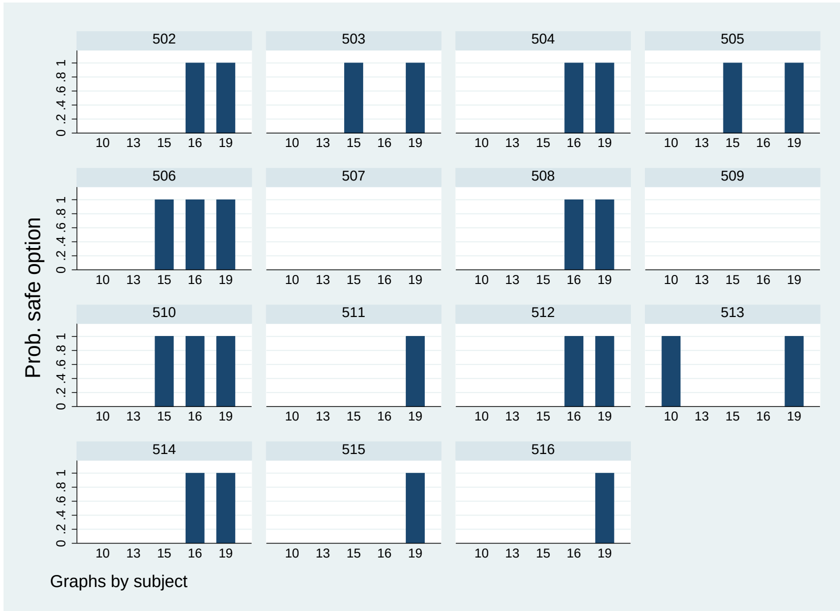

Figure 6: **Session 5.**

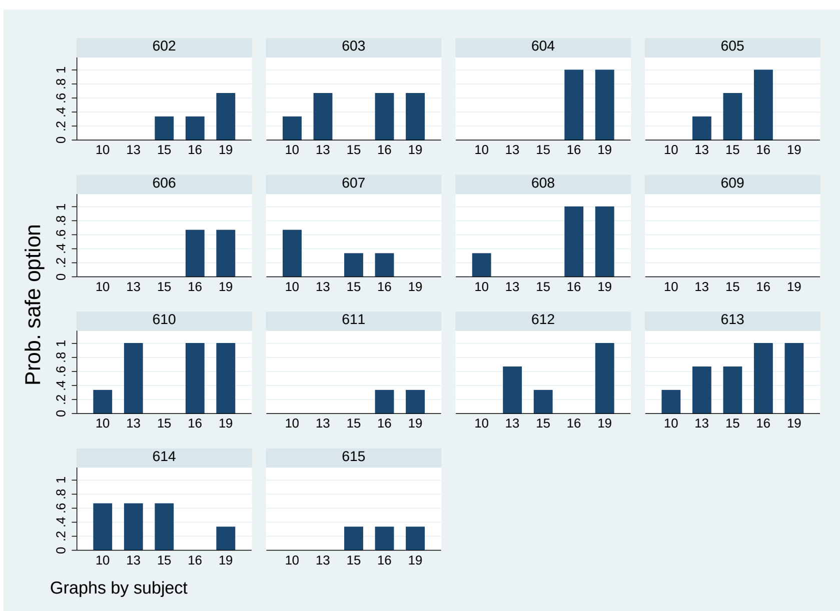

Figure 7: **Session 6.**

B Instructions (Treatment H)

Instructions

In this experiment, you will make a number of choices between a lottery and a certain amount. The lottery will always be the same: 30 pesos with probability 50% and 0 pesos with probability 50%. The certain amount will vary. For example, one possible choice may be between the lottery described above and a certain amount of 15 pesos.

Whenever you choose the lottery, we will determine your earnings based on an outcome of a coin flip. Every participant in this room has been assigned a winning outcome. Any coin flip will result in a 30 peso prize for you personally if its outcome is:

HEADS

Please note that your personal winning outcome, HEADS, will be fixed throughout the experiment.

Every time you make a decision, the researcher will flip the coin and show you the outcome. If you choose the certain amount, the certain amount will be added to your earnings. If you choose the lottery and the outcome of the coin flip matches your personal winning outcome, 30 pesos will be added to your earnings. Your overall earnings will always be displayed on your computer screen.

Example 1:

Your choice was between the lottery and a certain amount of 15, and you chose the certain amount. In this case, regardless of the outcome of the researcher's coin flip, 15 pesos will be added to your earnings.

Example 2:

Your choice was between the lottery and a certain amount of 15, and you chose the lottery. The researcher flipped the coin and the outcome of the coin flip was TAILS. In this case, 0 pesos will be added to your earnings.

Example 3:

Your choice was between the lottery and a certain amount of 15, and you chose the lottery. The researcher flipped the coin and the outcome of the coin flip was HEADS. In this case, 30 pesos will be added to your earnings.

Please raise your hand if you have any questions.

You will now play a practice trial in which the certain amount of 15. The practice trial is there to help you understand how the experiment works; the outcome of this trial will not count for your earnings.

C Instructions (Treatment L)

Instructions

In this experiment, you will make a number of choices between a certain amount and playing the same lottery three times. The lottery will always be: 30 pesos with probability 50% and 0 pesos with probability 50%. The certain amounts will vary. For example, one possible choice may be between playing the lottery above three times and a certain amount of 45 pesos.

Whenever you choose the lottery, we will determine your earnings based on an outcome of three coin flips. Every participant in this room has been assigned a winning outcome. Any coin flip will result in a 30 peso prize for you personally if its outcome is:

HEADS

Please note that your personal winning outcome, HEADS, will be fixed throughout the experiment.

After you make your decision, the researcher will flip three coins and show you their outcomes. If you choose the certain amount, the certain amount will be added to your earnings. If you choose playing the lottery three times and one of the coin flips matches your personal winning outcome, 30 pesos will be added to your earnings. If you choose playing the lottery three times and two of the coin flips match your personal winning outcome, 60 pesos will be added to your earnings. If you choose playing the lottery three times and three of the coin flips match your personal winning outcome, 90 pesos will be added to your earnings. Your overall earnings will always be displayed on your computer screen.

Example 1:

Your choice was between playing the lottery three times and a certain amount of 45, and you chose the certain amount. In this case, regardless of the outcome of the researcher's coin flip, 45 pesos will be added to your earnings.

Example 2:

Your choice was between playing the lottery three times and a certain amount of 45, and you chose the lottery. The researcher flipped the three coins and the outcome of every coin flip was TAILS. In this case, 0 pesos will be added to your earnings.

Example 3:

Your choice was between playing the lottery three times and a certain amount of 45, and you chose the lottery. The researcher flipped the three coins and the outcome of every coin flip was HEADS. In this case, 90 pesos will be added to your earnings.

Please raise your hand if you have any questions.

You will now play a practice trial in which the certain amount of 45. The practice trial is there to help you understand how the experiment works; the outcome of this trial will not count for your earnings.