Chapter 2

Red Chip, Blue Chip, One Chip, Two Chip: Oh the Things People do that Don't Match BEU[1]

An Experiment

Kathleen Rodenburg

Abstract

The question as to how people judge the probabilities or likelihoods of uncertain events has been a major focus in behavioral decision research for a number of years. The fact that intuitive judgments often deviate from the laws of probability are widely accepted (Harrison & al. 2010, Holt & Smith 2009, Von Witherfeldt & Edwards 1986). However, controversy still exists surrounding both the identification and root cause of systematic deviations from optimal behaviour. In this laboratory experiment, I test how subjects behave in an individual binary choice decision task with the option to purchase or observe for free additional information before reaching a final decision. I find that subjects' behaviour over time converges toward optimal decisions prior to observing an imperfect information signal. However, when subjects observe an imperfect information signal prior to their terminal choice there is greater deviation from optimal behaviour. I find in addition to behaviour that is reflective of a risk-neutral BEU maximizer, evidence of status quo bias, over-weighing the informational value of the message received and past statistically independent outcomes influencing future choices. Charness & Levin (2005) suggested that subjects use different decision heuristics (i.e., reinforcement learning) when decision environments are more complicated, in particular when faced with harder updating tasks. To test this proposition, a sub-set of subjects when presented with an imperfect information signal were provided with the Bayes law calculation. These subjects performed no better relative to optimal decision theory than the subjects who were only provided with the parameter values necessary to calculate Bayes law. The findings from this study suggest that individuals adopt different decision rules depending on both personal attributes (i.e. skillset, gender, experience) and on the context and environment in which the decision task is conducted.

[1] Bayesian Expected Utility Theory

Introduction

The question as to how people judge the probabilities or likelihoods of uncertain events has been a major focus in behavioral decision research for a number of years. The fact that intuitive judgments often deviate from the laws of probability are widely accepted (Harrison & al. 2010, Holt & Smith 2009, Von Witherfeldt & Edwards 1986). However, controversy still exists surrounding both the identification and root cause of systematic deviations from optimal behaviour. The results from several experiments serve to both corroborate and refute the theoretical predictions of classic decision theory. As no broadly applicable model of learning behavior has emerged from these and other studies, this may imply that learning behavior depends in small or large part on the context and environment in which the decision making is conducted. It follows that the proper identification of the situational restrictions that should be applied to existing economic models would enhance their predictive power in both laboratory and real world settings, where a more representative model of decision making processes would serve to enhance economic policy development.

In this laboratory experiment I observe how subjects behave in an individual decision task involving the choices between two different actions (non-strategic 2-action binary decision task) with the option to purchase or observe for free additional information before reaching the final decision (terminal choice). In addition to testing whether subjects' choices follow the predictions of risk neutral Bayesian Expected Utility (BEU) theory, I also test whether they follow the predictions of a Reinforcement Learning (RL) model using a simple RL algorithm employed by Charness and Levin (2005)[2].

Bayesian Expected Utility theory and Reinforcement Learning models are different in how they presume learning progresses and as such, in some cases, the consequent outcomes. Bayesian learning is widely used in economic models and laboratory experiments where the focus is on comparing a subject's behavior with so-called Bayesian 'optimal' behavior. Agents who learn according to the BEU learning model use the laws of probability to update prior beliefs when observing new information and then use these new updated beliefs to make decisions that maximize their expected utility. In contrast, Reinforcement learning, widely discussed and tested

[2] See chapter 1 for detail surrounding the selection of this Reinforcement learning algorithm for this study.

in psychology literature, supposes that agents make decisions based on past positive and negative outcomes. That is, agents' actions that lead to good outcomes in the past are more likely to be repeated in the future, whereas agents' actions that lead to bad outcomes in the past are less likely to be repeated. While decisions based on past successful and unsuccessful outcomes may be relevant and potentially optimal when future and past decision choices are statistically inter-dependent, this may not be the case when future decisions are statistically independent.

Despite the numerous experiments conducted to test behavior in a binary-choice decision task environment, there are in both economic and psychology literature several important issues that warrant further investigation and are subsequently addressed in this study.

While there is pervasive evidence of reinforcement learning behavior in the strategic game context (Erev & Roth, 1998; Camerer & Ho, 1999; Feltovich, 2000), little has been done to investigate reinforcement learning relative to Bayesian learning in a simple non-strategic decision task in an economic laboratory context (as highlighted by Charness & Levin(2005)). It is possible that the incentive mechanisms used in the binary choice decision task may have excluded the prevalence of a subject's use of a reinforcement heuristic as past payoffs were not always realized and if realized not always known after each round of the decision task. It is interesting to note that in the strategic game context with opponents, a subject's winnings are a key benchmark for measuring performance relative to other players. Regardless of whether the subject is paid for each round of play or paid for just one round, the winnings are a critical element used to assess the strategic play of the opponent. This visual cue could be a key trigger to activating the reinforcement heuristic in a subject's behavior in a simple non-strategic binary-decision task context.

Of particular interest to this study is an experiment conducted by Charness & Levin (2005). In their study they test Bayesian updating with expected utility maximization (BEU) and Reinforcement learning (RL) using a simple RL heuristic design in an individual choice task. The objective of the experiment is to observe whether subjects continue to use Bayes rule when it is aligned or not aligned with the RL heuristic and whether 'the propensity' to use Bayes rule in either case is affected by the introduction of immediate reinforcement after the first decision and prior to the final decision. As an extension to their study they identify a need to understand the 'cross-over threshold' between simple and more complicated updating given the observation that subjects in their study use BEU when decision tasks are easy and RL when decision tasks become more difficult. Furthermore, their study looks specifically at the use of the RL heuristic by subjects in the context of a statistically inter-dependent decision choice. For purposes of this study, understanding whether subjects use the history from previous decision choices which are statistically independent is the main focus.

Furthermore, from the studies reviewed, there is inconclusive and insufficient evidence as to why individuals do not accurately apply Bayes law beyond the simple explanation that people 'lack the cognitive sophistication' to do the math. For the Bayesian Expected Utility Model, separating subjective probabilities from other key aspects of decision-making behavior, i.e. expected Utility theory, has been a difficult challenge for many researchers (starting as early as Wallesten, 1968).

Another disadvantage in past experimental designs is that the willingness to pay for additional information is measured in terms of a fixed specified cost per observation. How much a subject is willing to pay for information is determined by the number of samples purchased. As the experimenter sets the purchase price, the precise value that a subject places on the service is not known (i.e., only a lower bound can be established). While the Bayesian optimal benchmark for the value of new information can be identified, how subjects value new information under the option to buy mechanism presented in these papers cannot be assessed. Furthermore, under this payment scheme how the amount paid for the information influences the accuracy of the decisions made is not observable.

Finally, many of the studies and consequent findings specific to our decision task design[3] occurred prior to 1980 (Fried & Peterson, 1965; Green & Swets, 1966; Peterson & Ducharme, 1967; Edwards, 1968; Wallsten, 1968; Pitz, 1968; Hershman & Levine, 1970). It is questionable whether the payment structure for many of these studies was faithful to the theory they set out to test. For example, in some studies subjects were not paid for performance, were paid very small amounts (Hershman & Levine, 1970), or paid in a manner which may have induced incorrect behaviors

[3] The experiment design in this study is built from an old set of experiments conducted between 1965-1980. Although the Charness and Levin(2005) is most relevant as it compares the BEU decision in conjunction with the RL heuristic as in this study, the design for this experiment is more closely related to the earlier studies (i.e., 2 urns, two colour ball) and therefore, an important aspect of the literature review.

(Fried & Peterson,1965[4]; Wallsten, 1968[5]; Grether , 1980, 1992). Additionally, for these earlier studies, the sample size for the experiments were extremely small (i.e. Wallsten, 1968, n=14; Hershman & Levine, 1970, n=20) and several of the econometric procedures and techniques used today were not available.

Therefore, as augmentation to these previous studies, this experiment provides further insight to the following important questions: 1) Do subjects use Bayesian learning to maximize their expected payoffs (BEU) when making decisions? And if so, does a cross-over threshold exist where the task becomes too difficult for subjects to apply BEU decision rules? 2) Do subjects who are provided with the posterior probability calculation deviate less from BEU optimal decision theory than subjects who are left to calculate the posterior probabilities on their own? 3) Do subjects treat independent rounds of the decision task as interdependent events applying a Reinforcement Learning heuristic when making these decisions? 4) Will the paid for observations result in fewer deviations from BEU optimal or RL action choices than the observations which are provided to subjects for free? And, 5) are there systematic deviations from optimal BEU behaviour beyond the partial RL heuristic used in this study?

From this study, I find evidence of both the risk-neutral Bayesian Expected Utility (BEU) and the Reinforcement learning (RL) algorithm adapted from the Charness and Levin study (2005) reflecting subjects' decision choices, with the former being more prevalent. Furthermore, I find that subjects' action choices prior to receiving an imperfect information signal were non-optimal at the beginning of the experiment, even though the optimal action was associated with the first-order stochastically dominant (FOSD) lottery. Eventually, however, these subjects' first action choices converged on the optimal action through repeated rounds of the same decision task. Additionally, when subjects observed a relevant information signal prior to their terminal decision and the consequent lottery associated with the optimal action was no longer first-order stochastically dominant there was a greater deviation from optimal behaviour over all the rounds of this same decision task. There is evidence that suggests the existence of a threshold where subjects' behaviour is less reflective of the BEU model when tasks become more complex

[4] Reward if correct, larger penalty if wrong. Subjects could lose some of their own money
[5] Money illusion- received 1/6¢ per point

(Charness & Levin, 2005). However, in the Charness and Levin's paper they suggest that threshold is at a point where the BEU decision rule is too complicated to calculate. This study contradicts this explanation by simplifying the math component of the BEU decision rule for a sub-set of subjects and finds that there is no improvement toward optimal behaviour. Additionally, Charness and Levin found that when tasks became more complex the subjects' behaviour was more reflective of the RL heuristic. Results from this study, found that the likelihood of BEU decisions increased when the BEU and RL heuristics are aligned (i.e., the BEU decision choice is the same as the RL decision choice for the subject) and the likelihood of BEU decisions decreased when the two heuristics are not aligned (i.e., the BEU decision choice is not the same as the RL decision choice for the subject). Finally, I find two systematic deviations from optimal BEU decision theory in addition to the reinforcement learning model used in this study. First, in a two-action choice binary decision task, where the second action choice is dependent on the information received from an imperfect message, a subset of subjects behaviour reflects an under-weighing of the value of new information when it is contrary to their original choice. Subjects prefer to stay with their first action choice regardless of the message received.[6] Second, there is a smaller sub-set of subjects whose decision choices reflect a consistent over-weighing of the informational value of the message received.

Section I describes the experiment. In Section II the results and interpretations are presented. Section III provides conclusions and future research opportunities.

Section I-Experimental Design

I conducted 6 different treatments during 12 classroom sessions on the University of Guelph campus, Guelph, Ontario, with 180 students recruited by e-mail from the undergraduate Bachelor of Commerce student population. On average subjects earned $33.60 for a 90 minute session. Each classroom session consisted of approximately 15 students who participated in 24 rounds of an individual task consisting of two (2) binary-choice decisions per round; where the second binary choice decision occurred after observing an imperfect statistically relevant information signal. For a subset of the groups and rounds, subjects had a third decision choice that required them to

[6] Samuelson and Zeckhauser (1988) refer to this type of phenomenon as a 'status quo bias'.

specify their willingness to pay for this additional information, where the WTP amount specified determined whether a subject's first (before observing the message) or second (after observing the message) action choice was recognized for payment. Therefore, some subjects participated in two types of decision tasks; a decision task with a 'FREE' (FREE) message and a decision task with an 'OPTION TO PURCHASE' (OTP) a message (appendix 2). The remainder of the subjects were placed in the control group and participated in 24 rounds of the FREE message decision task only.

Upon arrival, participants were given a handout explaining the experiment set-up and detailed instructions. The facilitator read the instructions aloud and demonstrated the experiment (appendix 1). The subjects were told that the amount of money that they would earn depends both on their individual choices and on random chance. In addition, they were told that the objective of the experiment is to maximize their earnings. Each subject participated in a practice round for both the FREE message and the OTP message decision task prior to commencing the rounds designated for payment.

For the FREE and OTP message decision task, subjects are shown at the beginning of each round two opaque bags, each containing a combination of red and blue poker chips. The distribution of red to blue chips within the two bags is symmetric with one bag containing a greater proportion of red chips and one bag containing a greater proportion of blue chips. For example, if bag 1 contains 35 red and 15 blue chips, bag 2 will contain 15 red and 35 blue chips. Subjects are told and shown the precise number and combination of red and blue chips contained within each bag.

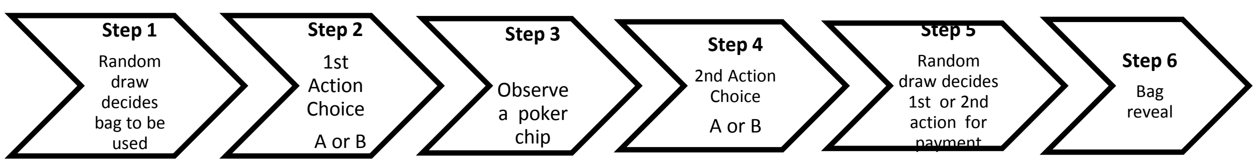
The step-by-step procedure for the FREE message decision task is outlined in Table 1 and described below.

In step 1, a subject is selected to perform a random draw that determines with equal probability which one of the two bags described above is selected for use during the round. All participants, including the subject performing the random draw, do not learn until the end of the round which bag has been chosen. In step 2, subjects are asked to choose one of two actions (action A or action B), where each action is associated with two different payoff amounts dependent on the bag that was randomly selected in step 1. In step 3, subjects are shown a sample draw of a poker chip (imperfect message) from the selected bag. In step 4 subjects can either maintain the action choice selected in step 2 BEFORE observing the sample draw or change their action choice selection AFTER observing the sample draw.

Table 2 provides the information that is shown and communicated to the subjects prior to taking their first and second action choice decisions for rounds 1-4 when performing the FREE message decision task.

In step 5 a random draw determines with equal chance whether the subjects' first or second action choice is used to calculate earnings. This payment mechanism incentivizes participants to apply effort to both action choices. In step 6, the bag that was used during the round is revealed.

Table 1: Sequential steps for the Free Message Task



The action that was selected (1st or 2nd) based on the random draw in step 5 determines the size of the payment received by the participant as outlined in table 3. From table 3 for rounds 1-4, if bag 1 is revealed as the bag selected in step 1 of the experiment, the participant will receive $2.00 if they selected action A and $0.50 if they selected action B. However, if bag 2 is revealed as the bag selected in step 1, the participant will receive $0.75 if they selected action A and $1.75 if they selected action B.

Table 2: FREE Message Task Exogenous Parameters

Rounds 1-4

BAG 1		BAG 2	
Red chips	35	Red chips	15
Blue chips	15	Blue chips	35
Total chips	50	Total chips	50

Potential Earnings

Pick Action A:	If the bag chosen by participant was bag 1 you receive	$2.00
	If the bag chosen by participant was bag 2 you receive	$0.75
Pick Action B:	If the bag chosen by participant was bag 1 you receive	$0.50
	If the bag chosen by participant was bag 2 you receive	$1.75

The step-by-step procedure for the OTP message decision task is outlined in Table 3 and is described below.

The first four procedural steps for the OTP message task are identical to the FREE message task. However, after observing the sample draw (imperfect message) in step 3 and selecting an action conditional on this draw in step 4, subjects in step 5 must indicate how much they would be willing to pay in order for their second versus their first action choice (the decision made prior to observing the sample draw) to be used for determining their payment. Once the willingness to pay (WTP) price has been specified, the experimenter in step 6 asks a participant to draw a random price from a box which contains 51 tokens each specifying a unique price point ranging from $0.00 to $0.50. The subjects are unaware of the range of prices contained within the box. The random price drawn determines the actual price required for using the second versus the first action choice to calculate earnings. In step 7, if the subject’s specified WTP is less than the randomly determined price, the initial action choice (first choice) will be used to calculate her earnings and there will be no price deducted from the payoff associated with this decision. However, if her specified WTP is greater than or equal to the randomly determined price, then the revised decision (second choice) will be used to calculate the earnings and the random price drawn will be deducted from the total earnings for the round for the subject. This WTP elicitation method is designed to be incentive compatible; thus ensuring that subjects reveal their truthful valuation of the information signal.

Table 3: Sequential steps for the OTP Message Task

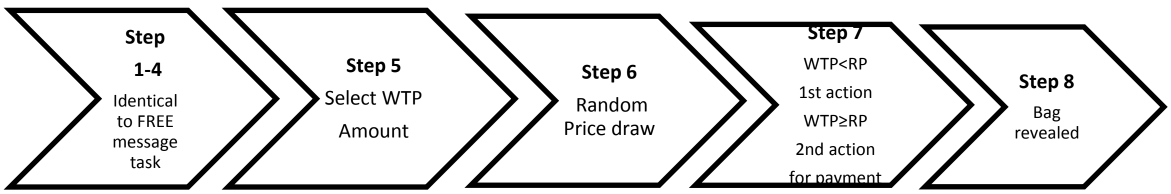
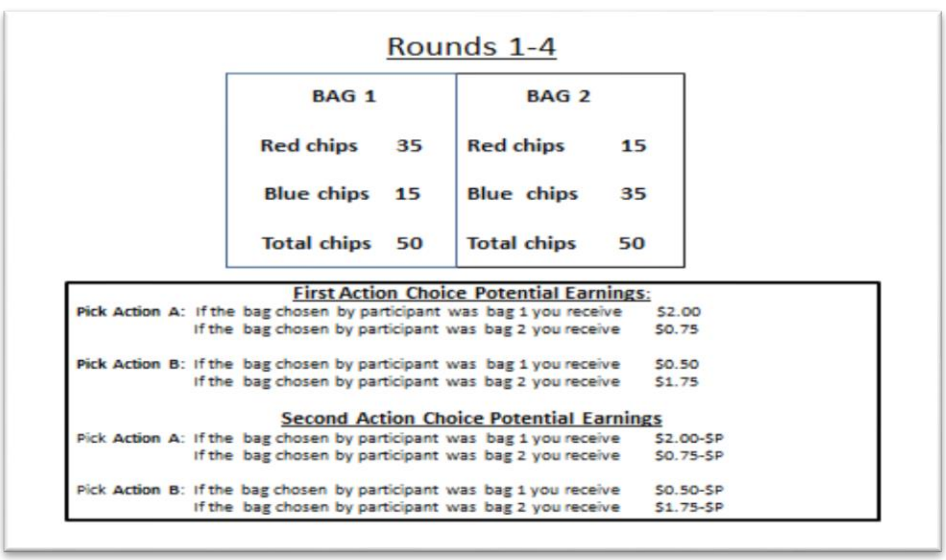


Table 4 provides the information shown and communicated to the subjects prior to making their first and second action choice decisions for rounds 1-4 when performing the OTP message decision task.[7]

Table 4: OTP Message Task Exogenous Parameters



Rounds 1-4

BAG 1		BAG 2	
Red chips	35	Red chips	15
Blue chips	15	Blue chips	35
Total chips	50	Total chips	50

First Action Choice Potential Earnings:

Pick Action A: If the bag chosen by participant was bag 1 you receive $2.00
If the bag chosen by participant was bag 2 you receive $0.75

Pick Action B: If the bag chosen by participant was bag 1 you receive $0.50
If the bag chosen by participant was bag 2 you receive $1.75

Second Action Choice Potential Earnings

Pick Action A: If the bag chosen by participant was bag 1 you receive $2.00-$P
If the bag chosen by participant was bag 2 you receive $0.75-$P

Pick Action B: If the bag chosen by participant was bag 1 you receive $0.50-$P
If the bag chosen by participant was bag 2 you receive $1.75-$P

[7] Note: During the 'Option to Purchase' (OTP) decision task, subjects must indicate how much they would be willing to pay in order for their second action choice to be used for payment. This is analogous to the following scenario: I book a trip to Florida, I observe that a hurricane is potentially pending; I book a trip to California but am informed that I must pay more to change my reservation. Based on my confidence in the weather forecast, how much would I be willing to pay to make this change? The rationale for this design is threefold: 1) it is easier to execute. All participants in a session observe the chip versus a design where only the subset of permitted participants can observe; 2) The Bayesian calculation for the optimal WTP amount is simplified. The added probability of receiving one of two messages is removed from the calculation providing subjects with an easier optimal WTP calculation; 3) More observations of subjects' second action choices are collected.

In step 8, the bag that was used during the round is revealed.

The action that was selected (1st or 2nd) based on the WTP amount in step 5 relative to the random price drawn in step 6 will determine whether participants will receive payment for their first versus their second action choice. From table 5 for rounds 1-4, if the subject's WTP is less than the random price drawn in step 6, they will receive payment on their first action choice. And if bag 1 is revealed as the bag selected during the round (for rounds 1-4), the participant will receive $2.00 if they selected action A and $0.50 if they selected action B; and, if bag 2 is revealed, the participant will receive $0.75 if they selected action A and $1.75 if they selected action B. However, if the subject's WTP amount is greater than or equal to the random price drawn in step 6, they will receive payment on their second choice. Let's assume the subject's WTP amount is $0.15 and the random price drawn was $0.10. Given that the WTP is greater than the random price drawn, the subject's second choice will be used to determine payment. If bag 1 is revealed as the bag selected during the round (for rounds 1-4), the participant will receive $2.00 minus the random price drawn ($2.00-$0.10), $1.90, if action A was selected and $0.40 ($0.50-$0.10) if action B was selected; and if bag 2 is revealed, the participant will receive $0.65 ($0.75-$0.10) if action A was selected and $1.65 ($1.75-$0.10) if action B was selected.

For both types of decision tasks (FREE and OTP) all subjects are informed each round of their earnings. Subjects are asked to record their first and second action choices, the results of each of the random draws, whether they received payment for their first or second action choice and their actual earnings for each round on the provided tracking sheet. The objective of the tracking sheet is to keep an account of each subject's history of events from past rounds to allow for the potential manifestation of reinforcement learning behaviour (appendix 3).[8]

The exogenous parameters, the distribution of red to blue chips contained within each bag and the payoffs associated with each action choice, change every four rounds and remain constant for 4 consecutive rounds. Table 5 provides the exogenous parameter values for the 24 rounds and the decision rule required to follow the risk- neutral BEU behaviour.

[8] In a strategic game subjects follow past successes or failures by observing whether they won or loss relative to an opponent. This tracking sheet provides a similar history of events within an individual decision task with no opponent.

Given the exogenous parameters for this experiment, the risk neutral (RN) optimal action taken prior to receiving an imperfect message is associated with a lottery that first-order stochastically dominates the alternative action's lottery for all rounds. Therefore, any expected utility maximizer with monotonic preferences should select the optimal first action regardless of risk preferences. The rationale for this design is to assist subjects in an easy optimal first choice, allowing for a cleaner assessment of subject behaviour when selecting a second action conditional on an imperfect information signal.

Similarly, the 2nd RN optimal action conditional on the red chip message is also associated with the lottery that first order stochastically dominates the alternative action's lottery for all rounds. Again in this case, any expected utility maximizer with monotonic preferences should select the optimal action regardless of risk preferences. On the other hand, there is no first or second order stochastic dominate lottery associated with either of the action choices conditional on a blue chip message. Although in this case it is now possible for risk preferences to influence choice, the optimal second choice for the risk neutral BEU maximizer continues to be the same optimal choice over a wide range of constant relative and absolute risk aversion utility curves.[9] Therefore, given this experimental design, when the message received is a blue chip versus a red chip, the consequent action choice is more suggestive of a subject's ability to follow the BEU decision rules. The willingness to pay (WTP) action to use the information to activate the subject's second action choice for payment results in changes to the lottery parameters, and as such risk aversion may influence the optimal WTP benchmark. This is discussed in more detail in chapter 3.

There is one final note on the choice of the risk neutrality assumption when establishing the BEU benchmark for comparison with subject behaviour. Arrow (1971) demonstrates in his Essays on the Theory of Risk Bearing that expected utility maximizers are (almost everywhere) arbitrarily close to risk neutral behaviour when stakes are arbitrarily small. This is later verified by the Rabin

[9] Under CRRA assumptions, ($u(c) = \frac{C^{1-\delta}}{1-\delta}$) when 0<$\delta < 2$, optimal second choice equals the RN BEU choice for rounds 1-8 & 13-20. Under CARA assumptions($u(c) = -e^{-\lambda C}$) when 0<λ<5, optimal second choice equals RN BEU choice for rounds 1-8 & 13-20. These parameter values, δ and λ are further relaxed during rounds 9-12 & 21-24.

Calibration (Rabin, 2000) which shows that the risk neutral prediction holds not only for small stakes but also for large and economically important stakes.[10]

Following the exogenous parameters identified in Table 5, suppose the subject behaves as a risk-neutral Bayesian Expected Utility maximizer. Each set of four rounds forces a new optimal decision. There are two possible states, represented by S_j, j ϵ {1,2}, where S_1 indicates bag 1 and S_2 indicates bag 2. A risk neutral BEU participant takes an initial action given the unconditional (prior) probability of either state with the objective of maximizing her expected earnings. Let the unconditional probability (initial belief) of playing in state j be, $prob\ (S_j)$, where, $\sum_j prob\ (S_j)$ = 1. Let C(***a***,S_j) be the payoff if action ***a*** is chosen conditional on the state (S_j), where ***a*** ϵ {A,B}. Without any additional information about the probability of the state being the bag with predominately red chips or the bag with the predominately blue chips, the initial decision to choose action A or B is based on the prior probabilities of being in either state, $prob\ (S_j)$, and the state contingent payoffs associated with each action, C(***a***,S_j). Specifically, the risk-neutral BEU will choose action A versus action B when:

$$EP_{action\ A} = prob\ (S_1)C(A, S_1) + \ prob\ (S_2)C(A, S_2)\ \geq EP_{action\ B} = prob\ (S_1)C(B, S_1) + \ prob\ (S_2)C(B, S_2)$$

Therefore, given the parameter values for rounds 1-4 & rounds 13-16 presented in Table 5, the initial BEU action choice in the absence of a message will be action A, as the expected payoff from action A is greater than that of action B.[11] The first action choice by the risk neutral BEU maximizer only requires the application of expected utility theory portion of the decision rule.

Next, the risk neutral BEU maximizer is provided with one of two possible messages in the form of a colour chip drawn from the randomly selected bag. Let the two possible messages be M_k, k ϵ {1,2}, where $M_1$is message 1(indicating a red chip message) and M_2 is message 2 (indicating a blue chip message). The participant is then required to propose a second action choice conditional on the message received. To do this the BEU maximizer will first, update her prior probabilities of being in either state to a new set of probabilities (posterior) using Bayes theorem. Second, she will

[10] Of course, there are others who argue these results using experimental data. However, these findings in addition to the exogenous parameter choices for this experiment provide good rational for the Risk neutral assumption when establishing this benchmark.

[11] From Eqn. 1, .5($2.00) + .5(.75) =$1.37 >.5(1.75)+.5(.50)=$1.125.

combine these updated probabilities to determine the expected payoff from taking either action and then choose the action with the highest expected payoffs.

Bayes theorem states that the posterior probability that a risk-neutral BEU maximizer should attach to the state after receiving a message, $prob\ (S_j|M_k),$ is:

$Prob\ (S_j|M_k) \equiv \frac{(prob\ S_j)(prob(M_k|S_j)}{prob(M_k|S_j)(prob\ S_j)+prob(M_k|S_{\neq j})(prob\ S_{\neq j})}$; j=1,2; j≠1,2; k=1,2; (Eqn. 1)

Where the $prob\big(M_k\big|S_j\big)$represents the likelihood of the message (M_k) conditional on state, S_j .

Note that regardless of the message received, one of two states must persist. Therefore,

$prob\ (S_j|M_k) + prob\ (S_{\neq j}|M_k) = 1$ (Eqn. 2)

Using Bayes theorem from Eqn. 1, the probability that the bag selected is bag 1(S_1) given that a red chip (M_1) was drawn is:

$$Prob(S_1|M_1) = \frac{prob(M_1|S_1)prob(S_1)}{prob(M_1|S_1)prob(S_1) + prob(M_1|S_2)prob(S_2)}$$

In short-form notation let,

$prob(S_j) \equiv\ \pi_j\ ; prob(M_k|S_j) \equiv q_{k.j};\ prob(S_j|M_k) \equiv\ \pi_{j.k}.$

Hence, the conditional probabilities of S_j given message M_1(red message) using short-form notation are:

$\pi_{1.1} = \frac{q_{1.1}\pi_1}{q_{1.1}\pi_1 + q_{2.1}\pi_2};\ \text{ and from Eqn. 2 }\ \pi_{2.1} = 1 - \pi_{1.1};$ (Eqns. 3 &4)

And, the conditional probabilities of S_j given message M_2(blue chip) are:

$\pi_{1.2} = \frac{q_{2.1}\pi_2}{q_{2.1}\pi_1 + q_{2.2}\pi_2}\ ;\ \text{ and from Eqn. 2 }\ \pi_{2.2} = 1 - \pi_{1.2};$ (Eqns. 5 & 6)

Therefore, from Eqns. 3 & 4, the expected payoff of choosing action A when message 1 (red chip) is received is:

$EP_{action\ A}|_{M_1} = \pi_{1.1}\ C(A, S_1) + \ \pi_{2.1}\ C(A, S_2)$ (Eqn. 7)

The expected payoff of choosing action B when message 1 (red chip) is received is:

$EP_{action\ B}|_{M_1} = \pi_{1.1}\ C(B, S_1) + \ \pi_{2.1}\ C(B, S_2)$ (Eqn. 8)

Given the red chip message (M_1), the risk neutral BEU maximizer will choose action A if the expected payoff is greater than choosing action B given the posterior probabilities conditional on the red chip message.

From Eqns. 7 & 8, the risk-neutral BEU will choose action A if:

$$EP_{action\ A}|_{M_1} = \pi_{1.1}\ C(A,S_1) + \pi_{2.1}\ C(A,S_2) \geq EP_{action\ B}|_{M_1} = \pi_{1.1}\ C(B,S_1) + \pi_{2.1}\ C(B,S_2).\ \text{(Eqn. 9)}$$

Table 5: Exogenous Parameters by Round Set

SET	1	2	3
Rounds	1-4 13-16	5-8 17-20	9-12 21-24
State Contingent Payoffs			
Action A **Bag 1 revealed, C(A,S_1)** **Bag 2 revealed, C(A,S_2)**	 \$2.00 \$0.75	 \$1.75 \$0.50	 \$1.00 \$0.50
Action B **Bag 1 revealed, C(B,S_1)** **Bag 2 revealed, C(B,S_2)**	 \$0.50 \$1.75	 \$0.75 \$2.00	 \$0.75 \$2.00
Initial Beliefs **Bag 1/Bag 2 (π_1/ π_2)**	 .5/.5	 .5/.5	 .5/.5
BEU decision rule prior to a message signal(chip draw)	**Action A**	**Action B**	**Action B**
State Characteristics **Total chips bag 1** **# red chips($q_{1.1}$)** **# blue chips ($q_{2.1}$)** **Total chips bag 2** **# red chips($q_{1.2}$)** **# blue chips ($q_{2.2}$)**	 50 35(.70) 15(.30) 50 15(.30) 35(.70)	 50 12(.24) 38(.76) 50 38(.76) 12(.24)	 50 20(.40) 30(.60) 50 30(.60) 20(.40)
Bayes Law Posterior Probabilities $\pi_{1.1}$ $\pi_{2.1}$ $\pi_{1.2}$ $\pi_{2.2}$	 **.70** **.30** **.30** **.70**	 **.24** **.76** **.76** **.24**	 **.40** **.60** **.60** **.40**
BEU decision rule After a message signal is received (chip draw)	**If Red : Action A** **If Blue: Action B**	**If Red : Action B** **If Blue: Action A**	**If Red : Action B** **If Blue: Action B**

Given the parameter values in Table 5 for rounds 1-4 & rounds 13-16, and given a red chip draw, one should choose action A, given that, $EP_{action\ A}|_{M_1} \geq EP_{action\ B}|_{M_1}$:

From Eqn. 7,

$$EP_{action\ A}|_{M_1} = \frac{.70(.5)}{.70(.5) + .30(.5)}(\$2.00) + \frac{.30(.5)}{.70(.5) + .30(.5)}(\$0.75) = \$1.625$$

From Eqn. 8,

$$EP_{action\ B}|_{M_1} = \frac{.70(.5)}{.70(.5) + .30(.5)}(\$0.50) + \frac{.30(.5)}{.70(.5) + .30(.5)}(\$1.75) = \$0.875$$

Similarly, if a blue chip is drawn, choose action B, given that, $EP_{action\ B}|_{M_2} > EP_{action\ A}|_{M_2}$:

$$EP_{action\ B}|_{M_2} = \frac{.70(.5)}{.70(.5) + .30(.5)}(\$1.75) + \frac{.30(.5)}{.70(.5) + .30(.5)}(\$0.50) = \$1.375$$

$$EP_{action\ A}|_{M_2} = \frac{.70(.5)}{.70(.5) + .30(.5)}(\$0.75) + \frac{.30(.5)}{.70(.5) + .30(.5)}(\$2.00) = \$1.125$$

For rounds 5-8 & rounds 17-20, following the math from above (see appendix 4), it is always optimal for the BEU decision maker to select action B as the first choice and action B as the second choice if a red chip is drawn and action A as the second choice if a blue chip is drawn. For rounds 9-12 & rounds 21-24, it is optimal to select action B as the first choice and also select action B as the second choice regardless of the colour chip drawn. The second action choice after observing the imperfect message requires the application of Bayes law in conjunction with expected utility theory in order to follow the BEU decision rule.

Prior to the announcement of the distribution of red to blue chips contained within each bag, and given the expected payoffs associated with each action choice, the critical values for the posterior probabilities, $\pi^c_{j.k}$, i.e., the switching rule where the BEU decision switches to the alternative action choice ($\mathrm{i.e. A\ to\ B\ or\ B\ to\ A}$), can be calculated.

For example, a BEU decision maker will switch her choice (from action A to B) conditional on observing a blue chip (M_2) if:

$$EP_{action\ B}|_{M_2} = \pi_{1.2}\ C(B,S_1) + \ \pi_{2.2}\ C(B,S_2)$$

$$\geq EP_{action\ A}|_{M_2} = \pi_{1.2}\ C(A,S_1) + \ \pi_{2.2}\ C(A,S_2) \quad \text{(Eqn. 10)}$$

The critical values of the posterior probabilities, $\pi^c_{1.2}$ & $\pi^c_{2.2}$, where a BEU decision maker will switch to the alternative action choice conditional on observing message 2 (blue chip) can be calculated by changing the weak inequality sign to an equality sign in Eqn. 10 and solving for $\pi_{1.2}$ & $\pi_{2.2}$.

$$EP_{action\ B}|_{M_2} = \pi_{1.2}\ C(B,S_1) + \ \pi_{2.2}\ C(B,S_2)$$

$$= EP_{action\ A}|_{M_2} = \pi_{1.2}\ C(A,S_1) + \ \pi_{2.2}\ C(A,S_2)$$

Noting that, $\pi_{1.2} = 1 - \ \pi_{2.2}$, and simplifying gives:

$$(1 - \pi_{2.2}\ )[C(B,S_1) - C(A,S_1)] + \ \pi_{2.2}\ [C(B,S_2) - \ C(A,S_2)] = 0 \quad \text{(Eqn. 11)}$$

$$\pi^c_{2.2} = \frac{C(B,S_1) - C(A,S_1)}{[C(B,S_1) - C(A,S_1)] + [C(A,S_2) - C(B,S_2)]} \qquad \text{(Eqn. 12)}$$

Given, $C(B,S_1) < C(A,S_1)$, $C(A,S_2) < \ C(B,S_2)$ then $\pi^c_{2.2}\ \in (0,1)$ and $\pi^c_{1.2}\ \in (0,1)$

Next, denote the equation for the difference in expected payoffs between action B and A (left-hand-side of Eqn. 11) by θ, and evaluate the partial derivative of θ with respect to $\pi_{2.2}$,

$$\frac{\partial\theta}{\partial\pi_{2.2}} = -[C(B,S_1) - C(A,S_1)] + [C(B,S_2) - C(A,S_2)] > 0;$$

Given, $C(B,S_1) - C(A,S_1) < 0 \; and \; C(B,S_2) - C(A,S_2) > 0$.

Since $\theta = 0$ at $\pi_{2.2} = \pi_{2.2}^c \in (0,1)$ and θ is monotonically increasing in $\pi_{2.2}$, it follows that $\theta > 0$ if $\pi_{2.2} > \pi_{2.2}^c$ and $\theta < 0$ if $\pi_{2.2} < \pi_{2.2}^c$.

Hence as $\pi_{2.2}$ increases, the expected payoff from taken action B increases. Conversely, as $\pi_{2.2}$ decreases the expected payoff from taking action B also decreases and when $\pi_{2.2} < \pi_{2.2}^c$ the expected payoff from selecting action B has decreased to a point where the greater expected payoff is now associated with selecting action A ($EP_{action\ A}|_{M_2} > EP_{action\ B}|_{M_2}$). Therefore, for any $\widehat{\pi_{2.2}} > \pi_{2.2}^c$ a BEU decision maker will switch to action B, otherwise she will remain with the initial action A.

As the state contingent payoffs for either action ($A\ orB$) are not symmetrical for any round, the critical values for the posterior probabilities $\pi_{j.k}^c$, where the BEU decision switches to the alternative action choice ($A\ to\ B\ or\ B\ to\ A$) will differ depending on the message received. As such the range of posterior probabilities where a BEU decision maker will switch to the alternative action conditional on the message will also be different.

Table 6 provides the ranges of posterior probabilities ($\pi_{j.k}$) where a BEU subject will switch to the alternative action choice conditional on the message received for each set of rounds that is governed by the same exogenous parameters.

Table 6: Posterior Probabilities: Critical Values and Ranges of Estimated Posterior Probabilities by Round

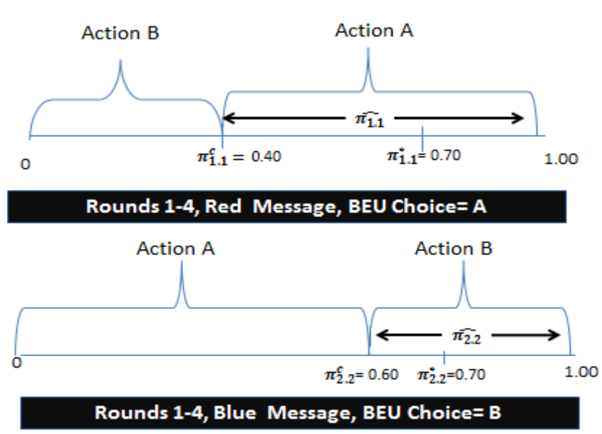
A BEU Participant Switches to the Alternative Action

Posterior Probabilities	Rounds 1-4/13-16			Rounds 5-8/17-20			Rounds 9-12/21-24		
$\pi_{j.k}$	Bayes	Critical Value (CV)	Range	Bayes	Critical Value (CV)	Range	Bayes	Critical Value (CV)	Range
$\pi_{bag1.red}$	.70	.40	[0.4, 1.0]	.24	.60	[0.0, 0.6]	.60	.857	[0.0, 0.86]
$\pi_{bag2.red}$	.30	.60	[0.0, 0.6]	.76	.40	[0.4, 1.0]	.40	.143	[0.14, 1.0]
$\pi_{bag1.blue}$	.30	.40	[0.0, 0.4]	.76	.60	[0.6, 1.0]	.40	.857	[0.86, 1.0]
$\pi_{bag2.blue}$	.70	.60	[0.6, 1.0]	.24	.40	[0.0, 0.4]	.60	.143	[0.0, 0.14]

Suppose that the posterior probability using Bayes law indicates that the optimal choice is to change to the alternative action, Table 6 illustrates that there exists a range of posterior probability estimates $(\widehat{\pi_{j.k}})$ where a subject may make a Bayesian updating error $(\widehat{\pi_{j.k}} \neq \pi^*_{j.k})$ yet still arrive at the correct BEU action choice. Additionally, it illustrates that in some cases this range is narrower than in others, allowing for less flexibility in posterior probability estimate $(\widehat{\pi_{j.k}})$ errors that still result in BEU action choices.

For example in rounds 1-4 and 13-16 from Table 6, a subject whose estimated posterior probability $(\widehat{\pi_{j.k}})$ falls within the range listed, assuming the subject is risk neutral and can perform the expected utility portion of the BEU problem, will also select the BEU action choice. However, the available range is reduced from a span of 0.6 points conditional on a red message to a span of 0.4 points conditional on the blue message (see also Figure 1).

Figure 1: Posterior Probability Estimate $(\widehat{\pi_{j.k}})$ Range for Each Action Choice Conditional on the Message Received Rounds 1-4 & 13-16



Note from figure 1, that the range for $\widehat{\pi_{j.k}}$ is narrower and the critical value boundary, $\pi_{j.k}^{c}$, where a subject should switch to the alternative action, is closer to the true $\pi_{j.k}^{*}$ when a blue message is received. Therefore, there is a greater likelihood of observing subject behaviour which reflects an under-weighing of the informational value of the blue message (i.e., in rounds 1-4, a risk-neutral subject capable of maximizing expected payoffs who under-weighs the informational value of the blue or red message and estimates a $\widehat{\pi_{1.1}} = \widehat{\pi_{2.2}} = .59$ would have behaviour reflective of BEU when a red message was received and would have non-BEU behaviour when the blue message was received). Furthermore, in an example where the critical value boundary $\pi_{j.k}^{c}$ for switching to the alternative action is greater than the true posterior probability $(\pi_{j.k}^{c} > \pi_{j.k}^{*})$ than we can potentially observe behaviour that is reflective of over-weighing the informational value of the message received. Specifically in this experiment, there is a greater likelihood of observing non-BEU decision choices conditional on a blue chip message that is reflective of under-weighing the informational value of the message for rounds 1-4(13-16) and 5-8(17-20) and over-weighing the informational value of the message for rounds 9-12 (21-24).

The Reinforcement Learner (RL) decision rule used in this study is based on the simple WIN-STAY, LOSE-SHIFT heuristic also used by Charness and Levin (2005). If a subject is successful in the first round of the experiment, she will STAY with this same action choice in the second round (WIN-STAY) and if the subject is unsuccessful in the first round, she will shift to the alternative action choice in the second round (LOSE-SHIFT); where, both RL actions are predicated on the subject experiencing the same past history dictated by both the fixed and random exogenous parameters set by the experiment.

Although the RL heuristic is simple in theory, difficulty arises when attempting to define past successful or unsuccessful action choices. The challenge arises when we attempt to use the financial gain or loss as the win or loss indicator. The necessary mechanisms required to ensure subjects exert best effort at each decision point and provide truthful valuation of a message service results in several possible reasons why a subject won or lost monetarily. Specifically, the reasons why a subject lost monetarily during the OTP decision task are: 1) the willingness to pay (WTP) price is less than the random price(RP) drawn forcing payment to be received on the subjects first action choice, the first choice incorrectly identified the higher payoff state (the payoff for this action given the state is less than the payoff for the alternative action given the state if payment was extended) and the second choice correctly identified the higher payoff state (the payoff for this action given the state is greater than the payoff for the alternative action given the state if payment was extended); 2) the WTP is less than the RP drawn forcing payment to be received on the subjects first action choice, and the first and second action choice incorrectly identified the higher payoff state; Or 3) the WTP is greater than the RP drawn forcing payment to be received on the second action choice and the second choice incorrectly identified the higher payoff state. Similarly, the reasons why subjects lose financially during the FREE decision task are attributed to: 1) pay draw was first, the first choice was incorrect and the second choice was correct; 2) pay draw was first, the first and second choice were incorrect ; 3) pay draw was second and the second choice was incorrect.

As a result of the above interpretations, I restrict the observation of the RL behavior to the second action choice. I assume that the subject will apply the WIN-STAY heuristic for a current round second action choice when the prior round second action choice correctly identified the state associated with the higher payoffs (WIN-guessed the right bag) and will apply the LOSE-SHIFT heuristic for a current round second action choice when the prior round second action choice incorrectly identified the state associated with the higher payoffs (LOSE-guessed the wrong bag), regardless of the actual amount of payment received. Additionally, the WIN-STAY or LOSE-SHIFT heuristic can only be applied in a current round if the exogenous parameter values experienced by the subject are the same as what was experienced in a prior round. Specifically, if the configuration of red to blue chips within each bag, the consequent payoffs conditional on the action choice taken and the message received in a prior round are the same as the current round.

For example, for the first round a RL participant is presented with a set of fixed exogenous parameters (bag configuration and consequent payoffs) that will remain constant for four consecutive rounds. Given this set of parameter values, regardless of the colour chip observed in round 1, there is no RL behaviour for the subject to follow. However, going forward (for 3 more rounds) subjects accumulate history from prior rounds. Let's assume the following outcomes for round 1: a red chip was observed, the subject took a second action that was associated with the higher payoff state once the true state was revealed (regardless of whether payment was extended or not) i.e. a WIN outcome. If the RL participant observes a red chip in the second round, she will use the past history gathered from round 1 to determine her second action choice for round 2. Hence, she will STAY with the same decision choice from round 1 in round 2 given that the choice in round 1 resulted in a WIN (WIN-STAY). However, if the RL participant instead observes a blue chip in the second round there is no blue chip history and therefore, no RL heuristic to apply. Now let's assume the following outcomes for round 1: a red chip was observed, the subject took a second action which was associated with the lower payoff state once the true state was revealed (regardless of whether payment was extended or not) i.e. a LOSE outcome. In this case, when a subject observes a red chip in the second round, the subject will SHIFT her second choice decision to the alternative action from the one chosen in round 1 (LOSE-SHIFT).[12]

Table 7 outlines the details of each of the six different treatment groups. Treatments 1, 2, 3 and 4 participated in 12 rounds of a decision task containing a FREE message (FREE) and 12 rounds of a decision task containing the OPTION TO PURCHASE (OTP) message. Treatments 1 & 2 received the FREE message task first and the OTP message task second. These decision tasks are reversed for treatments 3 & 4. Treatments 5 & 6 represent the control groups, participating in 24 rounds of the FREE message task only. Treatments 1, 3 & 5 are designated as un-informed. This group is provided with all the information required to apply Bayes law; however, they are not given the posterior probability calculations. Treatments 2, 4 & 6 are designated as informed and are provided with the posterior probability calculations described in terms of chances out of 100 that the colour chip drawn is either from bag 1 or bag 2.[13]

[12] Given the above interpretation, for the first round of each set (rounds 1, 5, 9, 13, 17 & 21) the histories from the prior sets are assumed irrelevant given the new set of parameters defined by the distribution of blue to red chips contained in each bag, and the change in payoffs associated with each action. A LOSE-SHIFT or WIN-STAY action choice only exists in the second round of each set (rounds 2, 6, 10,14,18 & 22) if the random chip observed is the same colour as in round 1, 5, 9, 13, 17 & 21, respectively (similar histories). Conversely, if the colour draw has never been observed in the prior round, there is no RL heuristic to be applied in rounds 2, 6, 10, 14, 18 & 22. For rounds 3 and 4 (15 and 16); 7 and 8 (19 and 20); 11 and 12 (23 and 24) a RL participant must observe the histories from rounds 1,2 and 1,2,3 (13, 14 and 13, 14, 15); 5,6 and 5,6,7 (17, 18 and 17, 18, 19); 9,10 and 9,10,11 (21, 22 and 21, 22, 23), respectively, in order to determine the second action RL decision choices. Note that when no same history exists, the RL observation is recorded as a blank observation in the data set.

[13] I.e. there are 70 chances out of 100 that the chip drawn came from bag 1 and therefore 30 chances out of 100 that the chip came from bag 2. This description of the posterior probabilities avoids any confusion associated with the term 'probability'.

Table 7-Treatment Group Specification

Treatment	1	2	3	4	5	6
No. Subjects (180)	30	30	30	31	29	31
Order: Round 1-12 : Round 12-24	Free OTP	Free OTP	OTP Free	OTP Free	Free Free	Free Free
Bayes Law	Uninformed[14]	Informed	Uninformed	Informed	Uninformed	Informed

In total I collected 4320 observations of subjects' first and second action choices and 1452 observations of the subjects' willingness to pay decisions.

Section II -Results

The data are analyzed using two different measurement criteria. First**,** subject behaviour is benchmarked relative to the action choices of a risk-neutral Bayesian Expected Utility maximizer and that of a Reinforcement Learner (RL) using the simple WIN-STAY, LOSE-SHIFT heuristic described in section II. In the data set, 'inconsistency rates' describe deviations from these two behavior types. Hence, for each subject in the experiment a Bayesian Expected Utility (BEU) first and second choice inconsistency rate (BDR1 & BDR2)[15] and a Reinforcement Learner (RL) second choice inconsistency rate (RDR2)[16] are calculated.

Second, for each round the subject's sequence of first and second action choices conditional on the imperfect message received are tracked and the proportion of subjects who follow each sequence is calculated. A subject can follow one of eight potential two action choice decision sequences for each round. The sequence that follows the BEU optimal action choice varies depending on the exogenous parameters of the experiment (Table 8).

Table 8: Eight Potential Two-Action Choice Sequences

Sequence	1st Choice	Message	2nd Choice	2nd Action BEU or Non BEU by Round Set					
of Action Choices	Action	Received	Action	Rds 1-4	Rds 5-8	Rds 9-12	Rds 13-16	Rds 17-20	Rds 21-24
S1	BEU	RED	STAY with 1st	BEU	BEU	BEU	BEU	BEU	BEU
S2	BEU	RED	Shift to Alternative	Non BEU	Non BEU	Non BEU	Non BEU	Non BEU	Non BEU
S3	BEU	BLUE	STAY with 1st	Non BEU	Non BEU	BEU	Non BEU	Non BEU	BEU
S4	BEU	BLUE	Shift to Alternative	BEU	BEU	Non BEU	BEU	BEU	Non BEU
S5	Non BEU	RED	STAY with 1st	Non BEU	Non BEU	Non BEU	Non BEU	Non BEU	Non BEU
S6	Non BEU	RED	Shift to Alternative	BEU	BEU	BEU	BEU	BEU	BEU
S7	Non BEU	BLUE	STAY with 1st	BEU	BEU	Non BEU	BEU	BEU	Non BEU
S8	Non BEU	BLUE	Shift to Alternative	Non BEU	Non BEU	BEU	Non BEU	Non BEU	BEU

[14] Uninformed: subjects given enough information to calculate Bayes law on their own
Informed: subjects provided with the Bayes Law calculation(posterior probabilities)
[15] BEU 1st and 2nd choice deviation rate
[16] RL 2nd choice deviation rate

For example from Table 8, when a red message is received sequence 1 (S1) is BEU optimal for all rounds, whereas, when a blue message is received sequence 3 (S3) is BEU optimal for rounds 9-12 and 21-24, and sequence 4 (S4) is BEU optimal for rounds 1-8 and 13-20. Additionally, the table highlights for each sequence whether a subject stays with their initial action or shifts to the alternative as their second choice.

To understand the causes of the observed BEU and RL inconsistency rates, I run logit regressions (both random and fixed effects) with the 1st and 2nd choice BEU inconsistency and 2nd choice RL inconsistency as the dependant variable to determine the odds ratio[17], log odds[18], and the marginal effects[19] of the independent variables on these three outcomes The dependant variable in equation (1) represents a 1st choice inconsistency from the risk neutral BEU decision by round and subject. The dependant variables in equation (2) & (3) represent a 2nd choice inconsistency from the risk neutral BEU decision and the 2nd choice inconsistency from the RL decision, respectively, by round and subject. In all three equations the dependent variable is a dichotomous outcome variable, where 0 represents consistency and 1 represents inconsistency relative to the designated behavior benchmarks. There are three types of variables used to explain the data. First, there is a group of explanatory variables that change over the rounds but are the same for all individuals in a given round. Second, there is a set of explanatory variables that vary both over the rounds and between subject and session. Finally, there are explanatory variables that vary between individuals but do not vary over the rounds.[20] The three equations are presented and described in Table 9.

[17] $\text{Odds Ratio} = (\text{Proportion of sucesses: positive dependent variable}(1)/\text{Proportion of failures: nonpositive dependent variable }(0))$

[18] The logarithm of the odds ratio

[19] Change in the probability of observing the dependent variable, if the independent variable changes by one unit

[20] The Breusch and Pagan Lagrangian multiplier test for all three equations established that individual effects are present in the data. The Hausman test cannot reject the null hypothesis that the coefficients for the fixed and random effects model are the same; implying that the random effects coefficients are not correlated with the individual error terms. As an additional test, I ran a GLS regression fixed and random effects model and performed the Hausman test and obtained the same result. Comparisons of the same coefficients from all models show the differences to be minimal; the signs and the statistical significance on the coefficients remain the same. Given the additional degrees of freedom, I report on the results from the random effects model and provide the fixed effects results in the appendix.

Table 9: Logit Regression Equations & Descriptive Summary of Variables

	Equation 1	Equation 2	Equation 2
Dependent Variable	**1^{st} Action Choice** **BEU Inconsistency** (Prior to observing a message) **1** Choice inconsistent **0** Choice consistent	**2nd Action Choice** **BEU Inconsistency** (After observing a message)	**2nd Action Choice** **RL Inconsistency** (After observing a message)
1. Explanatory Variables: vary over rounds, are the same for all individuals	**Experience** **1** second 12 rounds **0** first 12 rounds	**Experience**	**Experience**
	OTP **1** OTP message task **0** Free message task	**OTP**	**OTP**
	Informed **1** subjects given Posterior **Prob.** **0** subjects not given Posterior Probabilities	**Informed**	**Informed**
	Ex-ante Difference in expected POs for choosing Action A or B[21] -Continuous, changes every 4 rds.	**N/A**	**N/A**
	N/A	**Informative Power of chip draw**[22] -Continuous, changes every 4 rds.	**Informative Power of chip draw** .
2. Explanatory Variables: vary over rounds & between subjects	**N/A**	**Difference in Expected Payoffs between Action A and B conditional on the message received**[23] -continuous, changes every 4 rds. & conditional on chip draw	**Difference in Expected Payoffs between Action A and B conditional on the message received**
	N/A	**Shift Required from 1^{st} choice to be BEU optimal** 1 Shift 0 Stay	**Shift Required from 1^{st} choice to be BEU optimal**
	BEU action not consistent with the higher payoff state in prior round **1** Inconsistent **0** Consistent	**BEU action not consistent with the higher payoff state in prior round**	**BEU action not consistent with the higher payoff state in the prior round**
	Paid Second **1** payment on 2^{nd} action **0** payment on 1st action	**Paid Second**	**Paid Second**
3. Explanatory Variables: same over all rounds but vary by individual	**Female** **1** Female **0** Male	**Female**	**Female**
	English Second **1** English 2^{nd} language **0** English 1^{st} language	**English Second**	**English Second**
	Post Survey **1** Classified as RL **0** Classified as theorist	**Post Survey**	**Post Survey**
	Risk Aversion Continuous-Eckel-Grossman test 1 highest RA to 10 least RA	**Risk Aversion**	**Risk Aversion**
	Econ Math **1** Math/econ/optimization **0** non-math student	**Econ Math**	**Econ Math**

[21] $\text{Diff EPO} = \frac{[\pi_1 C(A,S_1) + \pi_2 C(A,S_2)] - [\pi_1 C(B,S_1) + \pi_2 C(B,S_2)]}{[\pi_1 C(A,S_1) + \pi_2 C(A,S_2)] + [\pi_1 C(B,S_1) + \pi_2 C(B,S_2)]}$

[22] [(# of red chips -#blue chips)/(#blue chips+#red chips)]

[23] $EP_{\text{action A}} > EP_{\text{action B}} \rightarrow \; [\pi_{1.1}\, C(A,S_1) + \pi_{2.1}\, C(A,S_2)] - [\pi_{1.1}\, C(B,S_1) + \pi_{2.1}\, C(B,S_2)]$
= difference between good and bad state conditional on the colour chip draw

i. Summary Results

Table 10 presents the mean values of first and second choice inconsistency rates relative to the BEU benchmark and the 2nd choice inconsistency rate relative to the RL benchmark for all six treatment groups (appendix 7 shows summary statistics for all variables).

Table 10: 1st and 2nd BEU and RL Inconsistency Rates by Treatment Group

Inconsisteny Rate		Treatment 1 FREE/OTP Uninformed	Treatment 2 FREE/OTP Informed	Treatment 3 OTP/FREE Uninformed	Treatment 4 OTP/FREE Informed	Treatment 5 FREE/FREE Uninformed	Treatment 6 FREE/FREE Informed	ALL
1st Choice BEU Inconsistency Rate	**All Rounds**	**12.5%**	**12.7%**	**14.2%**	**13.3%**	**13.4%**	**16.7%**	**13.8%**
	Rounds 1-12	14.7%	16.4%	17.5%	18.3%	18.4%	20.8%	(0.345)
	Rounds 13-24	10.3%	8.9%	10.8%	8.3%	8.3%	12.5%	
2nd Choice BEU Inconsistency Rate	**All Rounds**	**16.7%**	**14.9%**	**15.6%**	**15.4%**	**18.0%**	**16.0%**	**16.0%**
	Rounds 1-12	14.2%	14.4%	19.4%	17.4%	19.5%	20.8%	(0.367)
	Rounds 13-24	19.2%	15.3%	11.7%	13.4%	16.4%	11.1%	
2nd Choice RL Inconsistency Rate	**All Rounds**	**36.1%**	**46.9%**	**37.7%**	**36.6%**	**41.1%**	**44.8%**	**40.6%**
	Rounds 1-12	34.9%	49.7%	36.1%	33.9%	39.1%	42.8%	(0.491)
	Rounds 13-24	37.2%	44.0%	39.2%	39.2%	43.0%	46.7%	

Informed: subjects provided with the posterior probability calculations conditional on the chip draw
Uninformed: Subjects provided with sufficient information to calculate the posterior probabilities on their own

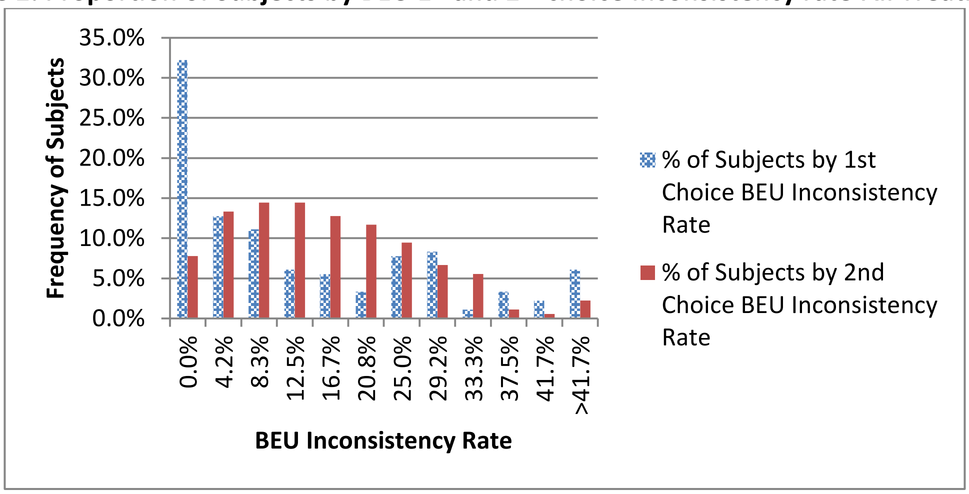
One sample t-tests comparing subject behaviour to 1st and 2nd choice BEU and 2nd choice RL benchmarks confirms that, in aggregate, subjects do not have an action choice (first or second) inconsistency rate relative to the BEU or the RL heuristic that is equal to zero (99% confidence interval). However, subjects' behaviour is less divergent from the risk neutral BEU benchmark.

The first action choice taken by subjects occurs prior to an imperfect message, where the lottery associated with the optimal action is first-order stochastically dominant (FOSD) to the alternative action lottery. Therefore, from the BEU inconsistency rate, it is not possible to conclude that subjects can or cannot maximize their expected utility, it may be that a subject's first action choice that is consistent with the BEU benchmark simply indicates the subject's ability to recognize the FOSD lottery. The second action choice BEU inconsistency rate gives insight into the subjects' ability to combine both Bayes law and Expected Utility Theory in particular when a blue chip message is observed. The second action choice caused greater diversion overall from the BEU benchmark as evidenced by the increase in subjects' BEU inconsistency rates from 13.8% for the first action choice to 16% for the second action choice[24]. Figure 2 shows that 58 subjects (32.2%)

[24] This difference between 1st and 2nd BEU inconsistencies is statistically significant at the 1% level

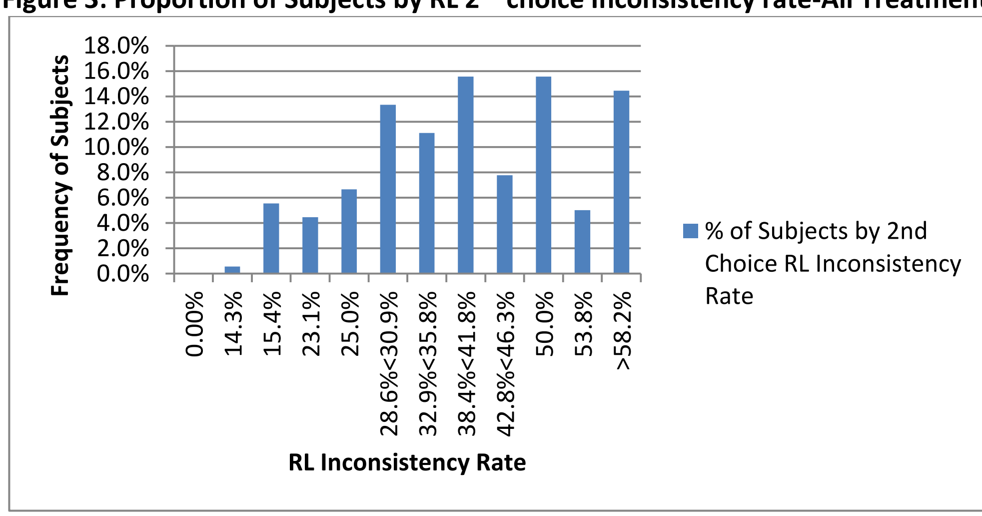
followed precisely the BEU decision rule for their first action choice for all 24 rounds. However, this number drops to 14 subjects (7.8%) who followed precisely the BEU decision rule for their second action choice. Relaxing the BEU decision rule assumptions to include subjects who followed the BEU model prediction within a 95% of the time (23 of 24 rounds are BEU consistent) increases the proportion of subjects following the action choices of a BEU decision maker to 81 subjects (45%) for the first action choice and 36 subjects (20%)for the second action choice.

Figure 2: Proportion of Subjects by BEU 1st and 2nd choice Inconsistency rate-All Treatments



On average, subjects' RL 2nd action choice inconsistency rate is 41% (Table 10). Figure 3 highlights that there are no subjects (0%) who followed precisely the simple RL heuristic for all 24 rounds. Unlike the BEU model, relaxing the RL decision rule model to include subjects who followed the RL model within a 95% confidence interval does not increase the proportion of subjects who followed the RL model predictions.

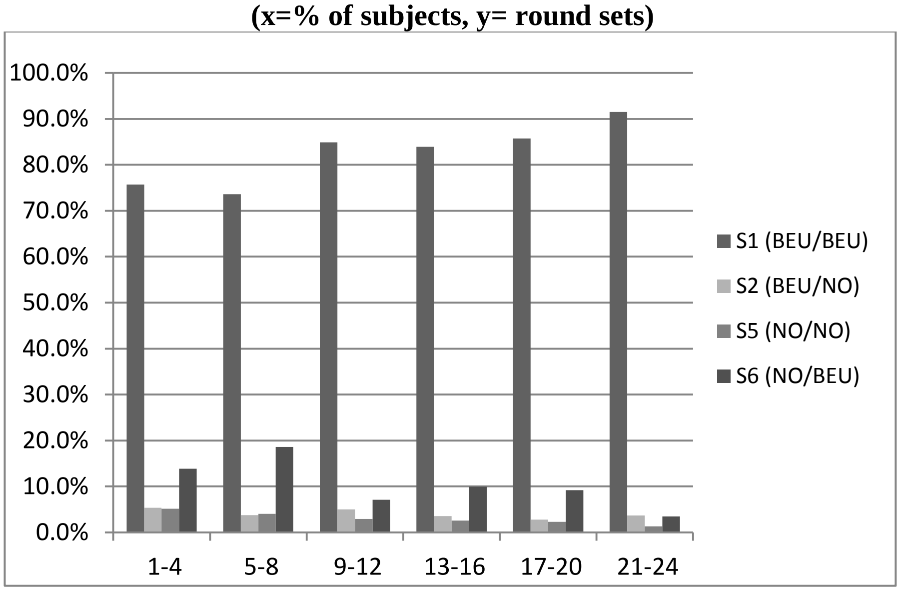
Figure 3: Proportion of Subjects by RL 2nd choice Inconsistency rate-All Treatments



The average financial payoff for subjects over the 24 rounds was $33.50, 4.2% less than the average BEU decision maker ($34.98) and 4.1% more than the average earned if all subjects were Reinforcement Learners ($32.19)[25].

Figures 4 & 5 identify the proportion of subjects who follow each of the 2 action sequence choices (described in table 8) for each round set conditional on observing a red (Figure 4) or blue (Figure 5) chip message. Recall, each round requires a subject to take an action choice prior to a colour chip draw and an action choice after observing a chip draw. These two choices per round conditional on the colour chip draw represent one sequence of decisions.

Figure 4: Proportion of Subjects who follow each of the 2-Action Sequence decision choices conditional on a Red Chip message by the sets of rounds sharing the same exogenous parameters. (x=% of subjects, y= round sets)

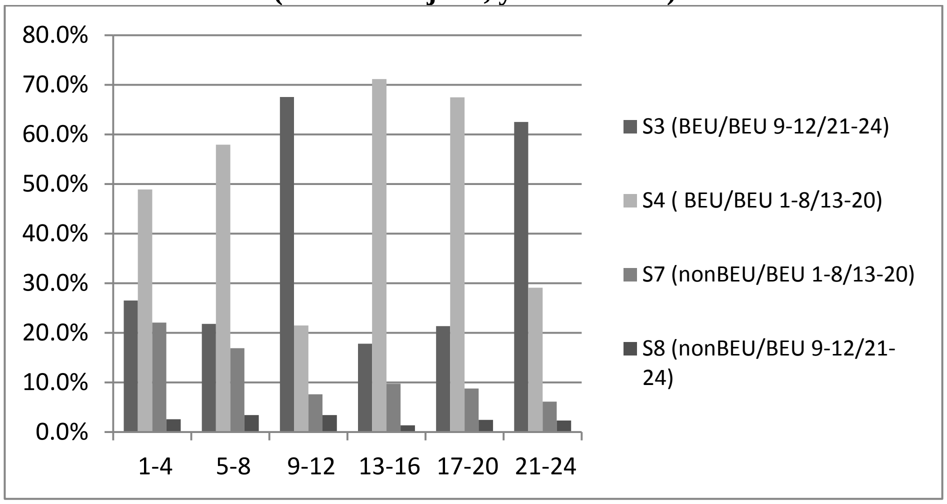


BEU/BEU- 1st Action follows the prediction of the BEU model/2nd Action follows the BEU model etc…

The red message accounted for 55.3% of all the imperfect messages observed by the subjects. Eighty-two percent (82%) of subjects who observed a red chip message chose, on average over the 24 rounds, the two-action sequence that equaled the BEU optimal sequence (see sequence 1, Figure 4).

[25] These payoffs are determined based on subject decision and random choice: 1) the decision choices of the subjects (i.e., first, second and WTP choices) and, 2) random draws (during the Free task a draw decides whether a subject will receive payment on their first or second action choice and during the OTP task a random price draw determines the price a subject will pay for the information observed). It is for this reason that payoffs are not used as an outcome variable in this analysis.

Figure 5: Proportion of Subjects who follow each of the 2-Action sequence decision choices conditional on a Blue Chip Message by the sets of rounds sharing the same exogenous parameters. (x=% of subjects, y= round sets)



BEU/BEU 9-12/21-24- 1st Action follows the prediction of the BEU model/2nd Action follows the prediction of the BEU model for rounds 9-12 & 21-24 and does not follow the BEU model otherwise etc…

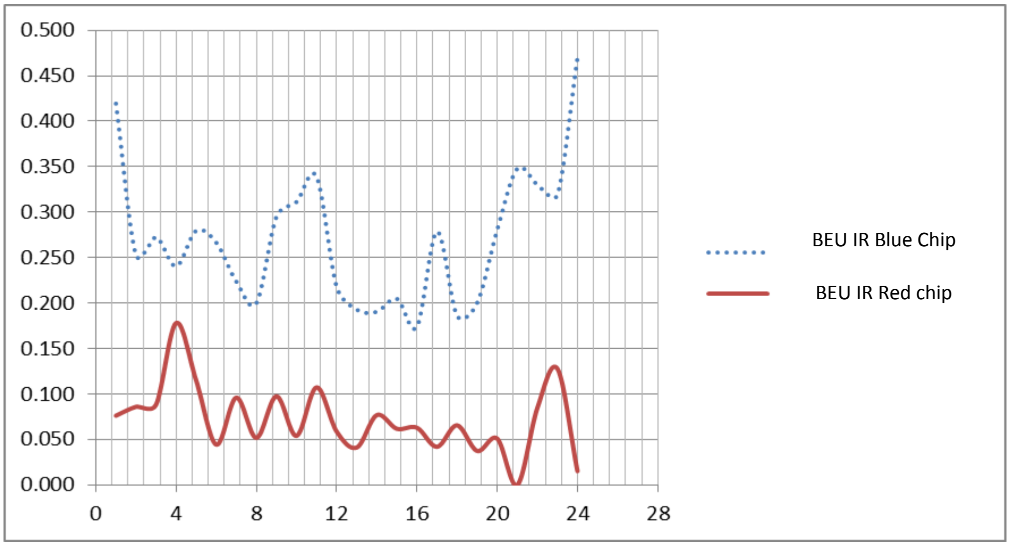
The blue chip message accounted for 44.7% of all the imperfect messages observed by the subjects. Sixty-three percent (63%) of subjects who observed a blue chip message chose, on average over 24 rounds, the two-action sequence that equaled the BEU optimal sequence (see sequence 3, rounds 9-12 and 21-24 & sequence 4, rounds 1-8 & 13-20, Figure 5).

On average the subjects' behaviour is reflective of the BEU sequence of decisions more often when the red message was received (7.1% BEU 2nd choice inconsistency rate). If subjects' behaviour is a consequence of BEU decision rules when observing the red chip, then they assigned probabilities of possible outcomes based on this message and calculated expected payoffs to value these outcomes. They then selected the action associated with the highest payoff. It follows that they should apply this same decision rule when they observe a blue chip. However, in contrast, although some subjects follow the BEU model predictions, there is a higher 2nd choice BEU inconsistency rate across all rounds of 27% when a blue chip is observed (see figure 6).

This difference in BEU inconsistency rates between the red versus the blue chip message may be the result of one or more of the following. First, it is possible that subjects did not follow BEU decision rules, but rather were capable of identifying the FOSD lottery associated with the optimal action given the red chip message. Therefore, as there was no FOSD lottery associated with the optimal action when a blue chip message was received, the decision environment became more complicated leading to more BEU inconsistencies. Second, it is possible that subjects in both decision environments (i.e., red and blue chip message) inappropriately estimated the informational value of the message received (i.e., the posterior probabilities associated with the message) when maximizing expected utility. In this case, the exogenous parameters of the experiment are such that when a red chip message is received their exists a broader range of posterior probability estimates where a subject may make a Bayesian updating error yet still arrive at the correct BEU action choice. On the other hand, the narrower range associated with the blue chip message for making a Bayesian updating error where a subject would still arrive at the correct BEU decision results in a greater likelihood of a BEU inconsistency. Finally, the second action choice conditional on the red chip message requires a BEU maximizer to stay with the initial optimal action chosen prior to receiving the message for all rounds. However, the second action choice conditional on the blue chip message requires the BEU maximizer to shift to the alternative action from the initial optimal action chosen prior to receiving the message for two-thirds of the rounds (16 of 24 rounds) and to stay with the original optimal action for one-third of the rounds. When the BEU decision rules changes from a shift to a stay with the initial pre-message action choice, the subject's behaviour is more suggestive of how they estimated the informational value of the message received particularly in the absence of a FOSD lottery choice.

When a blue message is observed, in addition to observing behaviour which is reflective of BEU decision rules, I observe during different round intervals behaviour that is reflective of both over-weighing and under-weighing the informational value of the message received.

Figure 6: BEU 2nd Choice Inconsistency rates conditional on chip colour observed by round (x= rounds and y= BEU inconsistency rate)



ii. Main Results

Table 11 provides the results from the 3 logit regressions with BEU 1st choice inconsistency (eqn. 1), BEU 2nd choice inconsistency (eqn. 2) and RL 2nd choice inconsistency (eqn. 3) as the dependent variable. I report the random effects results; odds ratios and marginal effects[26] ; in Table 11 and provide the fixed effects results in the appendix.[27]

[26] Marginal effects are calculated using Average marginal effects, that is, a marginal effect is computed for each case and the effects are then averaged.

[27] Appendix 8 & 9 shows the Fixed Effects model for the 1st and 2nd choice BEU Inconsistency rate and Appendix 12 shows the Fixed effects model for the 2nd choice RL inconsistency rate.

Table 11: Logit Regression Results from equation 1, 2 and 3 found in Table 9

	(1) BEU 1st Choice Inconsistency			(2) BEU 2nd Choice Inconsistency			(3) RL 2nd Choice Inconsistency		
Variables (see table 9)	**Marginal Effects**	**Odds Ratio**	**xtlogit, Re**	**Marginal Effects**	**Odds Ratio**	**xtlogit, Re**	**Marginal Effects**	**Odds Ratio**	**xtlogit, Re**
Experience	**-0.059*****	**0.451*****	**-0.797*****	**-0.028****	**0.778****	**-0.250****	**0.032**	**1.147**	**0.137**
	0.010	0.046	0.103	0.014	0.100	0.129	0.025	0.124	0.108
OTP	**0.015**	**1.226**	**0.204**	**0.062*****	**1.77*****	**0.554*****	**0.005**	**0.884**	**0.024**
	0.009	0.154	0.126	0.022	0.35	0.199	0.038	0.152	0.167
Experience*Paid	-	-	-	**0.002**	0.981	**0.018**	**-0.034**	**1.156**	**-0.148**
				0.029	0.255	0.260	0.044	0.225	0.194
Informed	**0.014**	**1.226**	**0.194**	**0.016**	**1.15**	**0.143**	**0.062*****	**1.311*****	**0.271*****
	0.020	0.322	0.265	0.018	0.188	0.163	0.025	0.142	0.108
Paid*Informed	-	-	-	**-0.065*****	**0.560*****	**-0.581*****	**-0.046**	**0.818**	**-0.201**
				0.023	0.118	0.210	0.044	0.157	0.193
Ex-ante Difference in Expected Payoffs Action A Vs. B	**-0.342***** 0.057	**0.009***** 0.006	**-4.66***** 0.616		-	-		-	-
Informative Power of the Chip Draw	-	-	-	**-0.476*****	**0.014*****	**-4.287*****		**0.099*****	**-2.31*****
				0.051	0.006	0.436		0.036	0.367
Difference in Expected Payoffs Action A Vs. B Conditional on message received	-	-	-	**-1.027***** 0.090	**0.000***** 0.000	**-9.25***** 0.769	**-0.713***** 0.158	**0.045***** 0.031	**-3.11***** 0.699
Shift from 1st action Required to follow BEU for 2nd action	-	-	-	**0.128***** 0.013	**3.160***** 0.366	**1.150***** 0.116	**0.034** 0.026	**1.159** 0.130	**0.148** 0.112
BEU action not consistent with the higher payoff state in prior round	**-0.005** 0.008	**0.932** 0.101	**-0.070** 0.109	**0.022***** 0.013	**1.230***** 0.120	**0.205***** 0.098	**-0.151***** 0.023	**0.519***** 0.053	**-0.656***** 0.103
Subject Paid on second action in Prior Round	**0.024***** 0.008	**1.377***** 0.146	**0.320***** 0.106	**-0.009** 0.011	**0.922** 0.091	**-0.081** 0.099	**-0.035** 0.022	**0.860** 0.082	**-0.151** 0.094
Female	**0.027**	**1.439**	**0.364**	**0.021**	**1.200**	**0.184**	**0.022**	**1.102**	**0.098**
	0.021	0.401	0.279	0.017	0.184	0.153	0.022	0.106	0.096
English 2nd	**0.066*****	**2.458*****	**0.900*****	**0.021**	**1.210**	**0.189**	**-0.001**	**0.997**	**-0.003**
	0.025	0.787	0.320	0.120	0.215	0.177	0.025	0.113	0.113
Post survey: RL	**-0.005**	**0.934**	**-0.069**	**0.011**	**1.110**	**0.103**	**0.063*****	**0.759*****	**-0.276*****
	0.022	0.276	0.295	0.018	0.179	**0.161**	**0.023**	0.076	0.099
Risk Aversion	**-0.003**	**0.959**	**-0.042**	**-0.002**	**0.979**	**-0.021**	**0.001**	**0.994**	**-0.006**
	0.004	0.046	0.048	0.003	0.026	0.026	0.004	0.017	0.017
Age	**-0.013*****	**0.845****	**-0.168****	**-0.001**	**0.989**	**-0.011**	**-0.005**	**0.979**	**-0.021**
	0.006	0.068	0.081	0.005	0.042	0.043	0.006	0.026	0.026
Econ Math Student	**-0.004**	**0.951**	**-0.04**	**-0.027**	**0.783**	**-0.245**	**0.032**	**1.148**	**0.138**
	0.023	0.295	0.310	0.019	**0.135**	0.172	0.024	0.120	0.105
Obs.		4320			4320			2265	
log likelihood/R-squared		-1451.85			-1630.61			-1472.77	

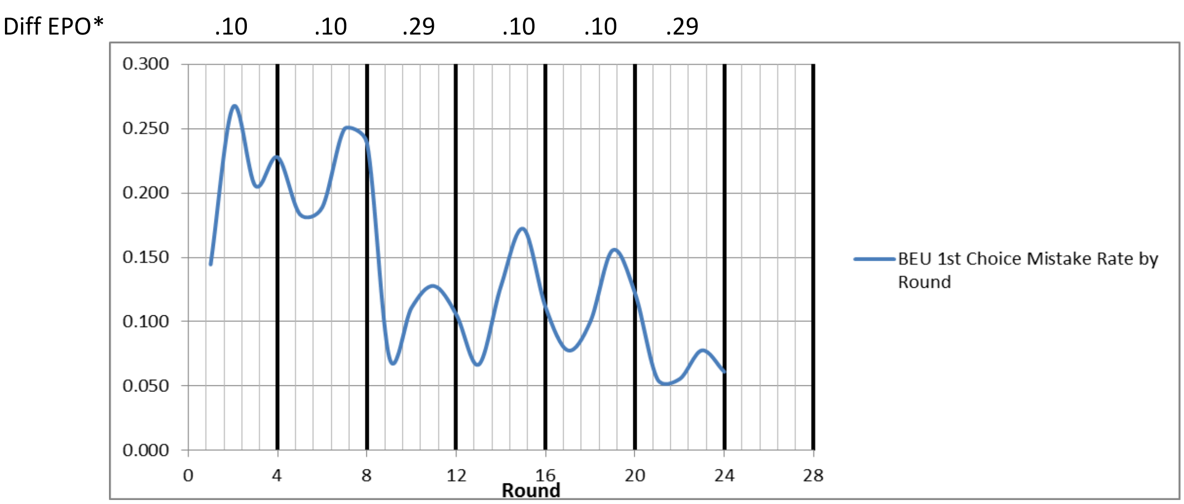
******p-value ≤.01 <**p-value ≤.05 < *p-value < .10***

Result 1: ***Subjects' decisions over time converge toward optimal choices prior to observing an imperfect message.***

A subject's first action choice is made in advance of observing an imperfect message, where there is an equal probability that the decision is being conducted using either bag 1 or bag 2. For all rounds the lottery associated with the BEU first action choice first-order stochastically dominates the lottery associated with the alternative action. Table 11, column 1, highlights the results from the Logit regression with the first choice BEU inconsistency as the dependent variable and provides some insight into the key aspects of subject decision behaviour when choosing a first action.

The predicted probability of a subject committing a first action choice BEU inconsistency is 5.9 percentage points (ppts) greater during the first 12 versus the last 12 rounds of the experiment. Through task repetition of the first decision choice, subjects' behaviour over-time converges on the optimal action. Additionally, as the difference between the FOSD lottery associated with the optimal action and the alternative action's lottery is exaggerated, the odds of a BEU inconsistency are less likely (significant at 1%)[28]. Figure 7 illustrates the subjects' first choice inconsistency rate relative to the BEU benchmark over the 24 rounds. In early rounds, subjects violate both BEU decision rules as well as first-order stochastic dominant choices and only converge on optimal decisions with practise and when the difference between the FOSD lottery and the alternative lottery are exaggerated. This first result does not necessarily imply irrational behaviour on behalf of the subject as the informational knowledge gained from the consequences of either action choice could also be considered a rational learning process.

Figure 7: BEU 1st Choice Inconsistency Rate by Round-All Subjects (x=1st Choice BEU Inconsistency Rate, y= round)



$$* \, Diff\ EPO = \frac{[\pi_1 C(A,S_1) + \pi_2 C(A,S_2)] - [\pi_1 C(B,S_1) + \pi_2 C(B,S_2)]}{[\pi_1 C(A,S_1) + \pi_2 C(A,S_2)] + [\pi_1 C(B,S_1) + \pi_2 C(B,S_2)]}$$

[28] The difference in ex-ante expected payoffs is the same and remains static during rounds 1-8 and 13-20 and increases by the same amount twice during the experiment; during rounds 9-12 and 21-24.

Result 2*: Subjects have a higher BEU inconsistency rate when the BEU decision rule requires the combination of both Bayes Law and Expected Utility theory to arrive at the optimal response. Additionally, this higher inconsistency rate is accentuated when subjects are performing the OTP versus the FREE message decision task.*

Subjects make a second decision, selecting either action A or B, after observing the colour of the chip drawn from the selected bag. To follow the 2nd choice BEU model predictions, a subject applies Bayes law in conjunction with Expected Utility theory. Table 11, column 2, highlights results from the Logit regression with the 2nd choice BEU inconsistency as the dependent variable and provides some insight into the key aspects of subject decision behaviour when choosing a second action. The action selected by subjects after observing an imperfect message results in a higher BEU inconsistency rate than the action selected prior to observing the message. Subjects' 2nd choice BEU inconsistency rate is 16%, 2.2% greater than the first choice BEU inconsistency rate of 13.8%.[29]

When performing the FREE and OTP message task, subjects take two actions; an action before and an action after observing an imperfect message. During the OTP message task subjects' make an additional decision and specify their willingness to pay in order to have their second versus their first action choice be used to calculate their earnings. Subjects' 2nd choice BEU inconsistency rate is 17.8% when subjects are performing the OTP message decision task, 2.6% greater than the BEU inconsistency rate when subjects perform the FREE message task.[30] Table 11, column 2, confirms this result and highlights that the predicted probability for a 2nd choice BEU inconsistency increases by 6.2 ppts when subjects are performing the OTP versus the FREE message task. Potential explanations for the higher 2nd choice BEU inconsistency rate associated with OTP message task follow.

It is of interest to investigate whether this result is due to a lack of commitment to the accuracy of the second choice on behalf of the subject when they are required to pay in order to have this choice used to calculate earnings. In this case, the lack of commitment in decision quality of second action choices could result in; 1) a WTP amount of zero or a relatively small WTP bid even though the subjects' second choice is different from their first choice, and 2) a BEU 2nd choice inconsistency rate which is higher for these subjects than for those who specified a more

[29] A two sample t-test identifies this difference to be statistically significant at the 1% level.
[30] See footnote 19

substantial WTP amount. Overall, the BEU 2nd choice inconsistency rate is 24.2% when the second action is different from the first action choice during the OTP message task. This inconsistency rate drops to 20.6% when the WTP is greater than \$0.05 and increases substantially to 41.7% when the WTP amount is less than \$0.05. This suggests that subjects with a higher willingness to pay amount are more BEU accurate. Although this result supports 'a lack of effort toward decision accuracy in the face of an additional cost' hypothesis, subjects who specified a WTP less than \$0.05 only accounted for 16% of the total observations where the first choice did not equal second choice actions.

Samuelson & Zeckhauser (1988) provide another potential explanation for the differences in the BEU inconsistency rate between the FREE and OTP message tasks. They demonstrate through a study of a series of decision making experiments that individuals have a tendency to maintain their previous decision choice even when new incomplete information is acquired that indicates that this decision choice is no longer optimal. They highlight that this 'status quo' bias in some cases leads individuals and firms to partake in fewer information searches than what is required to arrive at an optimal decision because they put greater importance on their original decision choice relative to the informational value provided by a new contradictory message. Table 11, column 2, show the predicted probability of a 2nd choice BEU inconsistency increases 12.8 ppts when subjects are required to change their initial action to the alternative action conditional on the chip draw in order to follow the BEU model predictions. To test whether subjects have a bias for the status quo and to determine whether this bias is more prevalent when performing the OTP versus the FREE message task, a comparison is made between the BEU inconsistency rates for both these tasks when the subjects' second action choice is the same as their first action choice (subjects maintaining the status quo). The 2nd choice BEU inconsistency rate when first and second action choices are aligned is 11.7% when subjects perform the FREE message task and 15% when subjects perform the OTP message task. A two sample t-test confirms that this difference is statistically significant at the 1% level. The logic, given this evidence, is that subjects maintain their first decision choice due to a status quo bias which is accentuated when there is an added cost (WTP decision requirement).

Another plausible explanation may be loss aversion. Tversky & Kahneman (1984) suggest that losses are two times more psychologically powerful than gains. Therefore, it may be that the subjects preferred to avoid the losses associated with the cost of using the information over the acquired gains from the informational knowledge the message provided. This is discussed in more detail in chapter 3.

Finally, is it possible that the change imposed by the additional step required to complete the OTP versus the FREE task created confusion or additional complexity to the decision environment? As a facilitator during the experiments I observed that the WTP decision task and the random price draw required for truthful elicitation of the WTP amount caused confusion and resulted in many requests by participants to repeat the instructions. Furthermore, during the OTP task, subjects relied on the posterior probability calculation (informed) as well as past outcomes to assist them with their decision choice.

The predicted probability of a 2nd choice BEU inconsistency is 6.5 ppts less when a subject is informed (provided the probability of being in either state conditional on the chip draw) versus uninformed when performing the OTP message task and this same coefficient has no significance when subjects are performing the FREE message task.[31] The finding that informed subjects perform no better relative to the BEU benchmark than subjects who were uninformed when performing the FREE message task appears contrary to the hypothesis that subjects do not follow the BEU decision rule because they lack the math skills or cognitive sophistication to perform the Bayes law component of the decision rule. A more plausible explanation may be that the procedural steps to complete the FREE task were easy to master allowing more clarity around odds estimations and less reliance on other cues when making decisions

According to the data, informed subjects performing the OTP message task have the same 2nd choice BEU inconsistency rate as subjects performing the FREE message task[32]. That is, the 2nd choice BEU inconsistency rate difference between the FREE and OTP message task is eliminated when subjects are provided with the Bayes law posterior probability calculation (informed). Additionally, subjects have a higher 2nd choice BEU inconsistency rate during the OTP rounds when the BEU 2nd choice in the prior round was inconsistent with the higher pay-off state for that round.

[31] To confirm this result, I add the interaction term, OTP*informed and show that the odds of a BEU inconsistency during the OTP message rounds are 1.77 (1/.571) times more likely when subjects are uninformed.

[32] To confirm this result, I test paid-paid*informed=0 and I cannot reject the null.

Given the application of different decision rules (informed and past history) and recognition that OTP task represented a more complex decision environment, these results suggest that learning behavior depends to some extent on the context and environment in which the decision making is conducted.

Result 3: There is evidence that suggests that the RL and BEU heuristics are complementary behaviours and when both are present can either enhance or diminish optimal decisions.

The predicted probability that a subject's behaviour will be reflective of RL model used in this study increases by 6.2 ppts when subjects are uninformed (column 3, Table 11) versus informed. Furthermore, the predicted probability that subjects will apply the RL heuristic in the current round is 15.1 ppts greater when the BEU 2nd action choice in the prior round was inconsistent versus consistent with the higher pay-off state for that round (column 3, Table 11).

There is evidence that suggests a subject's 2nd choice RL heuristic can improve and/or diminish behaviour that is reflective of optimal decision theory. The 2nd choice RL heuristic equals the 2nd choice BEU heuristic 39.8% of the time.[33] The BEU inconsistency rate is 13.6% when the RL and BEU heuristics are aligned, 18.9% when the heuristics clash and 16.5% when no RL heuristic exists (see Table 12). The differences between these BEU inconsistency rates are statistically significant at the 1% level (pairwise tests).[34] These results indicate that when a past BEU decision is rewarded (i.e. a WIN) a subject has a greater propensity to apply the BEU decision rule in the future resulting in fewer BEU inconsistencies (the RL and BEU heuristic are aligned). Additionally, if the BEU decision is not rewarded (i.e. LOSE), potentially creating a future decision environment where the subject's RL and BEU heuristic clash, optimal decision behaviour is compromised. However, it does not follow that when the heuristics clash, subjects' behaviour is more reflective of Reinforcement learning (see Table 12). For example, subject behaviour is more reflective of RL when they are uninformed (statistically significant at the 1% level), but is not more reflective of

[33] As a result the difference in expected payoffs between the action A and B conditional on the chip draw and the informative power of the chip draw (defined by the distribution of red to blue chips contained within each bag) coefficients found in table 12, column 3, have the same negative sign and statistical significance (at the 1% level) as the BEU Logit regression found in table 11, column 2. Removing the aligned BEU and RL heuristics from the RL inconsistency rate regression found in table 12 (see appendix 12) and observing only the RL heuristics where the BEU and RL are different results in a positive and statistical significant coefficient on the difference in payoffs between the good action and bad action state and the Informative Power of the chip draw coefficient losing its statistical significance.

[34] Subjects are 1.3 times less likely to be BEU inconsistent when their RL heuristic is the same as the BEU optimal choice versus when these two heuristics clash. Similarly, subjects are 1.25 times less likely to be BEU inconsistent when no RL heuristic exists (i.e. they have no prior history) versus when the two heuristics clash

BEU actions when informed (unless performing the OTP task). Similarly, although subjects have a higher 2nd choice BEU inconsistency rate when performing the OTP task, this same coefficient is not accompanied with a lower statistically significant 2nd choice RL inconsistency rate.

Charness and Levin (2005), suggest in their study the existence of a cross-over threshold where subjects' behaviour no longer reflect BEU decision theory when the task increases in complexity and becomes more reflective of this simple reinforcement learning model. However, the implication of the results from this study is that subjects do not substitute BEU with RL decision rules when tasks increase in complexity, but rather complement the BEU decision with the RL heuristic when they are aligned in an effort to achieve decision optimality in more complicated environments.

Table 12: BEU & RL Inconsistency Rates conditional on BEU and RL (Not) Alignment

	Aligned	**Not Aligned**	**No RL**
No. of Obs.	1361	905	2055
BEU 2nd Choice Inconsistency Rate	13.6%	18.1%	16.5%
RL 2nd Choice Inconsistency Rate	13.6%	81.9%	NA

Two additional findings of interest pertaining to Reinforcement Learning model utilized in this study: 1) The predicted probability of a first action choice BEU inconsistency is 2.4 percentage points greater when the subjects' second action choice versus their first actions choice receives payment in the prior round (see Table 11, column 1).[35]This suggests that subjects place more weight on optimal decision making at an action choice decision point (1st or 2nd) when this same action choice (1st or 2nd) was rewarded in the past, providing some additional evidence of reinforcement learning; And, 2)subjects who were categorized as reflective/reinforcement learners versus logical/theoretical learners, based on the condensed version of the Honey & Mumford (1986) personality type questionnaire (see appendix 5) conducted at the end of the experiment, were less likely to deviate from the RL benchmark.

[35] Recall, a first action choice will receive payment as a result of a random draw (FREE message task) or an unsuccessful specified WTP amount (OTP message task).

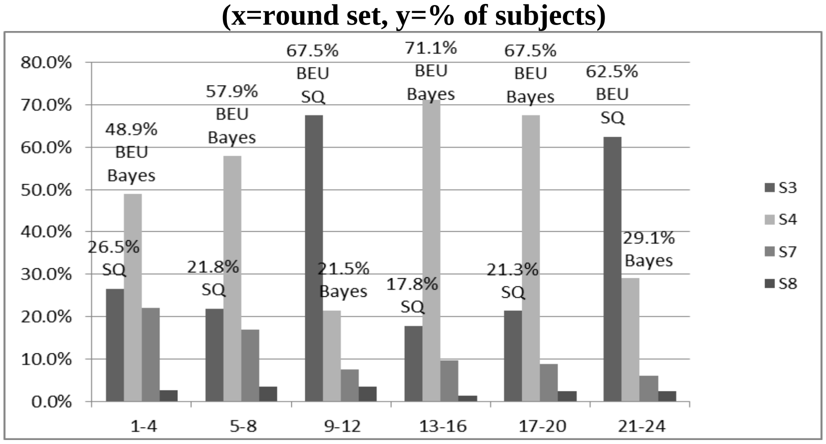
Result 4: There are two systematic decision behaviour patterns that deviate from optimal BEU decision theory that are not fully explained by the Reinforcement Learning model:

1. **Behaviour which is reflective of an over-weighing of the informational value of the message received (a.k.a. Over-weigh) ;**
2. **Behaviour which is reflective of an under-weighing of the informational value of the message received (a.k.a. Status Quo[36])**

The design of the experiment, specifically the asymmetric state contingent payoffs associated with either action (A or B), creates intervals of rounds when a blue message is received where behaviour either reflective of over-weighing or under-weighing of the informational value of the message received can be observed separate from BEU decision rule. In contrast, when a red message is received, subjects who display either of these two behaviour patterns look identical in behaviour to subjects who follow the BEU model.

Figure 8 highlights the proportion of subjects who follow each of the four potential sequence of action choices (1st and 2nd) conditional on observing a blue chip message. The proportion of subjects who followed sequence 3 during rounds 9-12 and 21-24 and sequence 4 during rounds 1-8 and 13-20 provide the main evidence for the two identified alternative behaviour types; Over-weigh (S3) and Status Quo bias (S4), respectively. [37] Note that the subjects who followed the non-optimal decision, sequence 3, during rounds 9-12 are not the same subjects who followed the non-optimal decision, sequence 4, during rounds 1-8 and 13-20.

Figure 8: Proportion of subjects by 2-Action sequence decision choice conditional on a Blue Chip message by sets of rounds sharing the same exogenous parameter values. (x=round set, y=% of subjects)



[36] Subjects display a bias towards staying with original action choices regardless of the message received.

[37] While sequence 3 is BEU optimal for rounds 9-12 and 21-24, it is not BEU optimal for rounds 1-8 and 13-20. Similarly, while sequence 4 is BEU optimal for rounds 1-8 and 13-20, it is not BEU optimal for rounds 9-12 and 21-24.

It is assumed that a subject who exhibits either the Over-weigh or the Status Quo decision pattern during the rounds where a blue message is received also applies these same decision rules when a red message is received; even though the behaviour pattern during these rounds is not observable.[38]

To determine group or individual characteristics that contribute to a subject's behaviour which is reflective of either BEU, an over-weighting of the informational value of the message received or a status quo decision rule, I run 3 OLS regressions with the proportion of decisions by subjects that reflect each behaviour pattern as the dependent variable and determine the marginal effects of the independent variables on these three outcomes. For all three regressions, the dependant variable is a continuous outcome variable, where the outcome represents the percentage of decisions where subjects followed this behaviour pattern (BEU, Status quo or Over-weight). The explanatory variables include a subset of the variables detailed in Table 9, specifically, the variables that are the same over all rounds but vary by individual (see no. 3 Table 9). Two additional independent variables are added to capture which treatment group that the subject belonged. All three regressions can be found in table 13.

Subject behaviour is more likely to be reflective of the BEU sequence of decisions when they are performing the FREE message task in advance of the OTP task (4.7ppts increase), when they are male (4.1ppts increase) and English is their language of origin (4.6ppts increase).

Subject behaviour is most likely to be reflective of the Status quo decision rule when they are not math or economics students (3.7ppts increase) and classified as a Reinforcement Learner based on the post experiment survey (4.2ppts increase).

Although a proportion of subjects have behaviour reflective of over-weighing the informational value of the message received, there are no characteristics that are statistical significant contributing to this behaviour type.

[38] Recall, that subjects who overweigh or play according to status quo rules will look identical to a BEU maximize when a red message is received.

Table 13: OLS Regression Results[39]

Variables	BEU OLS	Status Quo OLS	Overweight OLS
Free followed by OTP	**0.047*** **	**0.005**	**0.016**
	0.021	0.020	0.029
OTP followed by Free	**0.012**	**-0.024**	**0.033**
	0.021	0.020	0.030
Informed	**-0.006**	**-0.023**	**0.032**
	0.017	0.016	0.024
Female	**-0.042*** **	**-0.012**	**0.031**
	0.019	0.017	0.026
English 2nd	**-0.046*** **	**0.002**	**-0.015**
	0.022	0.020	0.031
Reinforcement Learner Survey	**0.010**	**0.042*** **	**-0.004**
	0.019	0.017	0.027
Risk Aversion	**0.002**	**0.002**	**-0.007**
	0.003	0.003	0.005
Age	**-0.006**	**0.005**	**-0.007**
	0.005	0.005	0.007
Econ Math Student	**0.003**	**-0.037** **	**-0.005**
	0.020	0.019	0.028
Obs.	180	180	180
Adjusted R-squared	0.0702	0.041	0.009

III. Conclusions

Subjects performing a relatively simple binary-decision task are adept at selecting optimal choices over time prior to observing additional statistically relevant information. Although this may provide evidence that suggests that subjects are capable of maximizing expected utility, it is also possible based on the lottery choices associated with each action, that subjects choose optimally simply by properly ranking the action associated with the FOSD lottery. When subjects observe a relevant information signal in the absence of a FOSD lottery and are required to combine Bayes law with expected utility theory in order to follow the BEU model predictions, there is greater deviation from this optimal behaviour. Furthermore, when the decision environment changes requiring subjects to perform a decision task which requires an additional step, optimal decision behaviour is further compromised. Specifically, when the decision environment changes from the

[39] Appendix 12 shows the complete Logit regressions for all independent variables by behaviour type.

FREE to the OTP message task subjects rely on past outcomes of success or failure and when available the information provided by the Bayes Law calculation (informed) to assist them with their decision choices.

Although the results are not sufficient evidence to confirm or refute the existence of a cross-over threshold where subjects no longer apply BEU decision rules due to task complexity (Charness and Levin, 2005), it does lend further support to the notion that learning behavior depends in small or large part on the context and environment in which the decision making is conducted. Grether (1989) suggests that in environments of uncertainty individuals use different decision rules in different decision situations. Furthermore, psychologists have also identified this finding and refer to these different decision rules used in different environments as the 'contingent judgement'[40] hypothesis (Payne, Bettman, and Johnson 1992). As observed in this study, when there is no FOSD lottery associated with the optimal action, potentially creating a more difficult decision environment, subjects' behaviour reflects the use of a different set of decision rules.

In addition to Rational decision theory (BEU) and the simple 'Reinforcement Learning' model used in this study, two alternative behaviour patterns emerged: 1) a sub-group of subjects put greater informational value on the message received; And, 2)a sub-group of subjects apply a status quo decision rule; under-weighing the value of new information when it is contrary to their original choice, preferring to stay with their first action choice regardless of the message received.

Samuelson and Zeckhauser (1988) identify regret avoidance and a taste for consistency (Charness and Levin, 2005) as a possible reason for this status quo behaviour. Individuals may have a desire to justify previous commitments, wish to avoid feelings of regret and have a need to feel in control. Kahneman and Tversky (1982), identified that individuals 'feel strong regret from bad outcomes that are the consequence of new actions taken than for similar bad consequences resulting from inaction'. It is reasonable to conclude that a subject's lack of math skills may create a more risky and uncertain decision environment leading to greater incidences of status quo behaviour.

[40] Subjects will substitute different decision rules contingent on the decision environment.

Risk aversion is ruled out as a possible determinant of the subjects' first and second action choice non-optimal behaviour. When selecting an action prior to observing an imperfect message or selecting an action conditional on receiving a red chip message, the optimal action is always associated with the first-order stochastically dominant lottery. Hence, any subject with monotone utility preferences should select the optimal action regardless of risk preferences. Additionally, although the optimal second action conditional on receiving a blue chip message does not entail a first order stochastically dominant lottery, the differences in the expected payoffs from either action allows for a broad range of CRRA and CARA utility curves (see footnote 9). An attempt was made to determine the risk attitudes of the subjects at the end of the experiment through the administration of the Eckel-Grossman risk task (Eckel-Grossman, 2002; see appendix 6). Regardless of the subjects risk preferences (as defined by the Eckel-Grossman risk test), results from the random and between effects model found in Table 11 & 13, show no statistical significance on the risk attitude coefficient; suggesting that risk preferences did not influence subject behaviour.

The findings from this study suggest that individuals adopt different decision rules depending on both personal attributes (i.e. skillset, gender, experience) and on the context and environment in which the decision task is conducted. Of most interest is understanding whether these deviations from BEU persist in other decision making environments more representative of a real world market setting. As such, further research should be conducted to determine which individual characteristics have a higher propensity for a certain behaviour type and then determine how and when changes to the decision environment influence the choices of the individuals with these characteristics. These changes could also assist in identifying the threshold where subjects no longer apply BEU optimal decision rules but apply different decision making criteria. It may also help identify whether this threshold varies depending on the individual type.

Given the above, the ability to identify individual type, the degree of risk and uncertainty within the decision environment and how individuals behave in these changing environments would be essential in determining the proper mechanism necessary to ensure optimal choices.

References

Antoniou, C., Harrison, G., Lau, M. & Read, D. 'Subjective Bayesian Beliefs.' Working Paper, January 2010.

Arrow, K. J., 'Essays on the Theory of Risk Bearing'. Publisher: Markham Pub. Co., Chicago, 1971.

Börgers, T., Sarin, R. ' Learning Through Reinforcement and Replicator Dynamics.' Journal of Economic Theory, 1997, 77, p.p. 1-14.

Camerer, C.F. '*Individual Decision Making*' in: Kagel, J.H. & Roth, A.E. (Eds.): The Handbook of Experimental Economics, 1995, 588–703.

Charness, Gary, and Dan Levin,. 'When Optimal Choices Feel Wrong: A laboratory Study of Bayesian Updating, Complexity and Affect.' American Economic Review, 2005, 95, pp. 1300-1309.

Davis, D., Holt, C., 'Chapter 8.' In: Experimental Economics, 1993, University Press, Princeton, New Jersey.

Ducharme, W. 'Response Bias Explanation of Conservative Human Inference.' Journal of Experimental Psychology, 1970, 85, p.p. 66-74.

Eckel, C., Grossman, P. ' Forecasting risk attitudes: An experimental study using actual and forecast gamble choices.' Journal of Economic Behaviour and Organization, Volume 68, 2008, 68, p.p. 1-17.

Edwards, W., 'Conservatism in human information processing.' In B. Kleinmuntz (Ed.), Formal representation of human judgment , 1968, pp. 17-52. New York: Wiley.

Edwards, Ward, 'Optimal Strategies for Seeking Information: Models for Statistics, Choice Reaction Times, and Human Information Processing', Journal of Mathematical Psychology, 1965, 2, 312-329.

Edwards, W. & Slovic, P. Seeking information to reduce the risk of decisions. *American Journal of Psychology,* 1965, 78, 188-197.

Erev, Ido & Roth, Alvin E. 'Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria.' Amercian Economic Review, 1998, 88(4), pp. 848-81.

Erev, Ido & Roth, Alvin E. 'Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term.' Games and Economic Behavior, 1995, 8, pp. 164-212

Feltovich, N. *'Reinforcement-Based vs. Belief-Based Learning Models in Experimental Asymmetric-Information Games',* Econometrica, 2000, Vol. 68, No. 3, pp. 605-641.

Fried, L. S. & Peterson, C. R. 'Information seeking: Optional versus fixed stopping.' *Journal of Experimental Psychology,* 1969, 80, 525-529.

Grether, David M. 'Testing Bayes Rule and the Representativeness Heuristic: Some Experimental Evidence.' Journal of Economic Behavior and Organization, 1992, 17(1), pp. 31-57.

Grether, David M. 'Bayes Rule as a Descriptive Model: The Representativeness Heuristic.' Quarterly Journal of Economics, 1980, 95(3), pp. 537-57.

Griffin, D., Tversky, A. 'The Weighing of Evidence and the Determinants of Confidence.' Cognitive Psycology, 1992, 24, pp. 411-35.

Hammerton, M., 'A case of radical probability estimation.' Journal of Experimental Psychology ,1973, 101, 252–254.

Heath C., and A. Tversky. 'Preference and belief: Ambiguity and competence in choice under uncertainty. Journal of Risk and Uncertainty', 1991, 4:5-28.

Hershman, R. L. & Levine, J. R. 'Deviations from optimum information-purchase strategies in human decision-making.' *Organizational Behavior and Human Performance,* 1970, 5, 313-329.

Holt, C., & Smith, A.,M. 'An Update on Bayesian Updating.' Journal of Economic Behaviour & Organization, 2009, 69, pp. 125-134.

Holt, C.A., Laury, S.K., 'Risk aversion and incentive effects.' American Economic Review, 2002, 92, 1644–1655.

Hopkins, Ed, *'Two Competing Models of How People Learn in Games'*, Econometrica, 2002, Vol. 70, No. 6, pp. 2141-2166.

Kahneman, D. & Tversky, A. 'Subjective Probability: A Judgment of Representativeness.' Cognitive Psychology, 1972, 3(3), pp. 430-54.

Kahneman, D. & Tversky, A. 'On the psychology of prediction.' *Psychological Review,* 1973, 80, 237-251.

Kahneman, D. & Tversky, A. 'Rational Choice and the Framing of Decisions.' The Journal of Business, 1986(1984), 59, p.p. S251-S258.

Loomes, G., 'Some Lessons from Past Experiments and some Challenges for the Future.' The Economic Journal, 1999, 109, F35-F45.

Malcolm, D., Lieberman, B., ' The Behavior Responsive Individuals play a two-person, zero sum game Requiring the use of Mixed Strategies.' Psychonomic Sci., 1965, 12, p.p. 373-74.

Ochs, J. 'Simple Games with Unique Mixed Strategy Equilibrium: An Experimental Study.' Games and Economic Behavior, 1995, 10(1), p.p. 202-17.

Offerman, T., Sonnemans, J. 'What's causing overreaction? An experimental investigation of recency and the hot-hand effect. Scandinavian Journal of Economics, 2004, 106, 533–553.

Offerman, T., Sonnemans, J., van de Kuilen, G.,Wakker, P.P., 'A truth-serum for non-Bayesians: correcting proper scoring rules for risk attitudes.' Working paper, 2007, University of Amsterdam (CREED).

Peterson, C. R. & DuCharme, W. M. 'A primacy effect in subjective probability revision.' *Journal of Experimental Psychology,* 1967, 73, 61-65.

Pitz, G. F. Information seeking when available information is limited. *Journal of Experimental Psychology,* 1968, 76, 25-34.

Pitz, G. F. ‘The influence of prior probabilities on information seeking and decision making.’ *Organizational Behavior and Human Performance,* 1969, 4, 213-226.

Pitz, G. F., Reinhold, H. & Geller, E. S. ‘Strategies of information seeking in deferred decision making.’ *Organizational Behavior and Human Performance,* 1969, 4, 1-19.

Phillips, L.D., Edwards, W. ‘Conservatism in a simple Probability Inference Task.’ Journal of Experimental Psychology, 1966, 72, p.p. 346-354.

Rabin, Mathew, ‘Risk Aversion and Expected Utility Theory: A Calibration Theorem.’ Econometrica, Vol. 68, No. 5, September 2000, p.p. 1281-1292.

Rapoport, A., Boebel, R. B. ‘Mixed Strategies in Strictly Competitive Games: A further test of the minmax hypothesis.’ Games and Economic Behavior, 1992, 4, p.p. 261-83.

Rapoport, A., Tversky, A., ‘Choice Behavior in an Optimal Stopping Task.’ Organizational Behavior and Human Performance, 1970, 5, p.p. 105-20.

Samuelson, William and Richard Zeckhauser, ‘Status Quo Bias in Decision Making’. Journal of Risk and Uncertainty, 1:7-59, 1988

Sarin, R., Vahid, F. ‘Predicting How People Play Games: A Simple Dynamic Model of Choice.’ Games and Economic Behavior, 2001, 34, p.p. 104-22.

Slembeck, T., ‘A Behavior Approach to Learning in Economics: Toward an Economic Theory of Contingent Learning.’ European Economic Association Congress, September 1998, Berlin.

Smith, V.L., Walker, J.M., ‘ Monetary Rewards and Decision Cost in Experimental Economics.’ Economic Inquiry, 1993, 31:2, p.p. 245-61.

Suppes, Patrick and Atkinson, Robert C. ‘Markov learning models for multiperson interaction.’ Palo Alto: Stanford University Press, 1960.

Swets, J. A. and Green, D. M. ‘Sequential observations by human observers of signals in noise.’ In C. Cherry (Ed.) *Information theory.* London: Butterworths Scientific Publications, 1961 (in press).

Tversky, A., & Edwards, W. ‘Information vs. reward in binary choices.’ *Journal o] Experimental Psychology,* 1966, 71, 680-683.

Tversky, A. & Kahneman, D. ‘Advances in Prospect Theory: Cumulative Representation of Uncertainty.’1992,5(4), pp.297-323.

Tversky, A. & Kahneman, D. ‘Availability: A heuristic for judging frequency and probability.’ *Cognitive Psychology,* 1973, 5, 207-232.

Tversky, A. & Kahneman, D. ‘Belief in the law of small numbers.’ *Psychological Bulletin,* 1971, 76, 105-110.

Von Winterfeldt ,D. ,Edwards, W. ‘ Decision Analysis and Behavioral Research’ Cambridge University Press, New York, 1986.

Wallsten, T. ‘ Failure of Predictions from Subjectively Expected Utility Theory in Bayesian Decision Task.’ Organizational Behavior and Human Performance, 1968, 3, p.p. 239-52.

Appendix 1

Instructions

Thank you for participating today!

By participating in this experiment, you will have the opportunity to earn money. The actual amount of money you will earn depends both on your choices and on random chance. During this session, we will ask you to make a series of decisions. Please make sure that you completely understand the instructions for each part of the experiment before making any decisions in that part of the experiment. If you have any questions at any point or need clarification, please raise your hand and the experimenter will come to you and answer your question.

In this session there are two rounds that are practice rounds for you to get familiar with two different decision tasks and a series of experimental rounds which will be used to calculate your earnings in a manner to be described in the workbook. You are not allowed to use a calculator, but may write down anything you may need to make your decision on the yellow tracking sheet provided by us.

You will have a ***Workbook*** that will contain the instructions for each round of the game. You will also use the workbook to record all your decision choices for each round of the game. To ensure confidentiality, your workbook is identified only by a participant number, which is never connected to your actual name.

You will also be asked at the end of the session to complete a short questionnaire. Please respond to this questionnaire truthfully and as accurately as possible. The questions provide the experimenter with important data that is of enormous help in organizing and interpreting your decisions. Your decisions and answers to the questionnaires are confidential and will not be revealed to anyone other than the experimenter. The data will only be identified by the participant code assigned to you and will not at any point be connected to your name in any way.

Please make sure that you completely understand the instructions for the experiment. It is important not to make any noises that might disturb others around you. If you have any questions, raise your hand and we will answer your questions individually.

Appendix 2

Practice Round -The Decision task with a 'FREE' Message Signal

As this is a practice round the potential earnings highlighted on the next page will not actually be paid. The intention of this round is to allow you to become familiar with the decision task with a "FREE" message signal.

In all free message signal rounds, there are two bags, each containing a combination of red and blue poker chips. There are 50 poker chips in each bag. However the number of chips that are red and the number of chips that are blue differs between these two bags. You will be informed of the number of red and the number of blue chips contained within each bag. In step 1, a random draw will decide which bag will be selected for use during the round. There is an equal chance that we will be playing the round using bag 1 or bag 2. However, you will not know until the end of the round which bag has been randomly selected for play during the round. In Step 2, you are asked to choose one of two actions and told the financial consequences of taking each of these two actions. In step 3, you are shown a sample draw of a poker chip from the bag. This chip is replaced back into the bag once it has been observed. You can either maintain the decision choice made in step 2 <u>BEFORE</u> observing the sample draw or change your decision choice AFTER observing the sample draw. In step 4, the bag that was used during this round of play will be revealed. A random draw will be made to determine whether your step-2 action choice or your step-4 action choice will be used to calculate your earnings. Therefore, there is an equal chance of receiving earnings calculated based on the action choice you made <u>BEFORE</u> observing the sample draw in step 2 or of receiving earnings based on the action choice you made <u>AFTER</u> observing the sample draw in step 4. You will be informed of your earnings for the round. Please record these earnings on the tracking sheet provided. Since this is a practice round, the earnings for this round will not actually be paid.

For this practice round, the bags contain the following number of red and blue poker chips:

Bag 1	Bag 2
Red chips 35	Red chips 15
Blue chips 15	Blue chips 35
Total chips 50	Total chips 50

Step 1:

A random draw will determine the bag to be used for this round. The procedure is as follows. The experimenter will show you the contents of both bags to verify the number and colour of the poker chips contained within each bag. The number 1 will be pinned to the inside of bag 1 and the number 2 will be pinned to the inside of bag 2. From the exterior of the bag it will be impossible for you to tell which bag is designated 1 or 2. Both of these bags will be placed in a large cardboard box. For each round, a participant will be selected to come forward and reach into the box and select a bag. You will be unable to identify which bag has been selected. There is an equal chance that the round is being played using either bag 1 or bag 2.

Appendix 2 –Continued: Free message signal Task

Step 2:

<u>BEFORE</u> observing the sample draw, please circle below whether you wish to take action A or action B based on the following potential earnings:

Pick **Action A**: If the bag chosen by the participant was bag 1 you receive $2.00
If the bag chosen by the participant was bag 2 you receive $0.75

Pick **Action B**: If the bag chosen by the participant was bag 1 you receive $0.50
If the bag chosen by the participant was bag 2 you receive $1.75

Action Choice <u>BEFORE</u> observing a Sample Draw

Circle either Action A or Action B

Action A Action B

Rip off this sheet and place it on the corner of your desk for the research assistant to collect.

Step 3:

The experimenter will ask one participant to draw one poker chip from the bag, show you the colour of the chip and replace it back into the bag.

After observing the sample draw, please circle below whether you wish to take action A or action B based on the following potential earnings:

Pick **Action A**: If the bag chosen by participant is bag 1 you receive $2.00
If the bag chosen by participant is bag 2 you receive $0.75

Pick **Action B**: If the bag chosen by participant is bag 1 you receive $0.50
If the bag chosen by participant is bag 2 you receive $1.75

Action Choice <u>AFTER</u> Observing a Sample Poker chip Draw

Circle either Action A or Action B

Action A Action B

Rip off this sheet and place it at the corner of your desk for the research assistant to collect.

Step 4:

A random draw will take place to determine whether you will be paid based on your initial choice as indicated in Step 2 or your revised choice as indicated in Step 3. Therefore, there is an equal chance of receiving payment for either the action choice you made BEFORE observing the sample draw or the action choice you made AFTER observing the sample draw. The experimenter will reveal the bag used for this round. You will then be informed of your earnings for the round. Please record these earnings on the tracking sheet provided.

Practice Round - A Decision Task with an "OPTION TO PURCHASE" a Message Signal

As this is a practice round the potential earnings highlighted on the next page will not actually be paid. The intention of this round is to allow you to become familiar with the decision task with an "OPTION TO PURCHASE" a message signal.

In all the 'option to purchase' a message signal rounds, there are two bags, each containing a combination of red and blue poker chips. There are 50 poker chips in each bag. However the number of chips that are red and the number of chips that are blue differs between these two bags. You will be informed of the number of red and the number of blue poker chips contained within each bag. In step 1, a random draw will decide which bag will be selected for use during the round. There is an equal chance that we will be playing the round using bag 1 or bag 2. However, you will not know until the end of the round which bag has been randomly selected for play during the round. In Step 2, you are asked to choose one of two actions and told the financial consequences of taking each of these two actions. In step 3, you are shown a sample draw of a poker chip from the bag. This chip is replaced back into the bag once it has been observed. You can either maintain the decision choice made in step 2 BEFORE observing the sample draw or change your decision choice AFTER observing the sample draw. However, in step 4, in order to determine whether this revised decision will be used to calculate your earnings, you must indicate how much you would be willing to pay in order for it to be so used.. In step 5, the experimenter will ask a participant to draw a random price from a random price box which will determine the actual price of using your revised decision rather than your initial decision to calculate your earnings. If your specified willingness to pay is less than the randomly determined price, your initial decision will be used to calculate your earnings. Therefore, your earnings will be based on the action choice you made BEFORE you observed the sample draw. However, if your specified willingness to pay is greater than or equal to the randomly determined price, your revised decision will be used to calculate your earnings. Therefore, your earnings will be based on the action choice made AFTER you observed the sample draw and your earnings for this round will be reduced to include the randomly determined price of using the new information. In step 6, the bag that was used during this round of play will be revealed and you will be informed of your earnings for the round. Please record these earnings on the tracking sheet provided.

Step 1 to 3 is identical to the Free message signal task(subjects were walked through these steps again…condensed for the appendix)

Step 4:

Please indicate on the next page the amount of money you would be willing to pay so that the revised action choice you made AFTER observing the sample draw rather than the initial action choice you made BEFORE observing the sample draw is used in order to calculate your earnings.

Your earnings for this round will be determined as follows:

Once you have indicated your willingness to pay to use your revised action choice, a random draw will determine the actual price that you must pay for the action choice you made AFTER observing the sample draw to be used to calculate your earnings. The procedure is as follows. The experimenter will ask a participant to choose a price from a box that contains many possible prices, some low prices and some high prices. If the random price drawn is less

than or equal to your specified willingness to pay, the action choice you made AFTER observing the sample draw will be used to calculate your earnings. You will pay the random price selected. Therefore, the action choice used to calculate your earnings will be the one made AFTER you observed the sample draw. The cost of using your revised decision will be subtracted from your earnings. , And, you will earn:

Pick **Action A**:	If the bag chosen by participant is bag 1 you receive	\$2.00-\$P
	If the bag chosen by participant is bag 2 you receive	\$0.75-\$P
Pick **Action B**:	If the bag chosen by participant is bag 1 you receive	\$0.50-\$P
	If the bag chosen by participant is bag 2 you receive	\$1.75-\$P

Appendix 2-Continued: OTP a message signal decision task

Otherwise, if the random price drawn is greater than your specified willingness to pay, the action choice used to calculate your earnings will be the one made <u>BEFORE</u> you observed the sample draw (i.e., the action choice made in step 2). And, you will earn:

Pick **Action A**:	If the bag chosen by participant is bag 1 you receive	\$2.00
	If the bag chosen by facilitator is bag 2 you receive	\$0.75
Pick **Action B**:	If the bag chosen by participant is bag 1 you receive	\$0.50
	If the bag chosen by participant is bag 2 you receive	\$1.75

Please indicate the amount that you are willing to pay in order to use the revised action choice made AFTER the sample draw rather than the initial action choice made <u>BEFORE</u> the sample draw to calculate your earnings.

$______________

Rip off this sheet and place it on the corner of your desk for the research assistant to collect.

Step 5:
The experimenter will now ask a participant to draw a random price from the box.
If the randomly chosen price $P is less than or equal to your specified willingness to pay, then you are paid based on the action choice you made <u>AFTER</u> you observed the sample draw.

If the randomly chosen price $P is greater than your specified willingness to pay, then you are paid based on the action choice you made <u>BEFORE</u> you observed the sample draw.

Step 6:
The experimenter will reveal the bag used for this round and you will be informed of your earnings. Please record your earnings for the round on the tracking sheet provided.

Appendix 3

Participant #__________

Participant Tracking Sheet

Round #	(F) First Action Choice A or B	Chip Colour r or b	(S) Second Action Choice A or B	Action choice for Payment F or S	Bag revealed 1 or 2	Earnings $
Practice						
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						

Round #	(F) First Action Choice A or B	Chip Colour r or b	(S) Second Action Choice A or B	Your Willingness to Purchase Price $WTP	Random Price drawn $P	Action choice for Payment If $P >WTP=F If $P≤WTP=S	Earnings $
Practice							
13							
14							
15							
16							
17							
18							
19							
20							
21							
22							
23							
24							

Appendix 4

First BEU Action Choice

Without any additional information about the probability of the state being S_1 or S_2, the initial decision to choose action A or B is based on prior probabilities, π_s. Specifically, the risk-neutral BEU will choose action A versus action B when:

$$\pi_1 C(A,S_1) + \pi_2 C(A,S_2) \geq \pi_1 C(B,S_1) + \pi_2 C(B,S_2)$$

Given the parameter values from table 4 and prior to a message signal

For rounds 1-4 & 13-16 the initial BEU action choice will be action A, as the expected payoff from action A is greater than that of action B.

.5($2.00) + .5(.75) =$1.37 >.5(1.75)+.5(.50)=$1.125

For rounds 5-8 & 17-20 the initial BEU action choice will be B

.5($2.00) + .5(.75) =$1.37 >.5(1.75)+.5(.50)=$1.125

For rounds 9-12 & 21-24 the initial BEU action choice will be B

.5($2.00) + .5(.75) =$1.37 >.5(1.00)+.5(.50)=$0.75

Second BEU Action Choice

Bayes theorem states: $\boldsymbol{\pi_{S.M} = \frac{j_{S.M}}{q_M}}$; Where, $\boldsymbol{\pi_{S.M}}$ is the conditional (posterior) probability of state S given the message M; $\mathbf{j_{S.M}}$ is the joint probability of state S and the message M; and q_M is the unconditional probability of receiving message M. Therefore, given message 1(red chip), the BEU players chooses action A if the expected payoff is greater than choosing action B given the posterior probabilities associated with message 1. The expected payoff when choosing action A when message 1(red chip) is received is:

$$EP_{action\ A} = \pi_{1.1}\ C(A,S_1) + \pi_{2.1}\ C(A,S_2) = \frac{q_{1.1}\pi_1}{q_{1.1}\pi_1 + q_{1.2}\pi_2}\ C(A,S_1) + \frac{q_{1.2}\pi_2}{q_{1.1}\pi_1 + q_{1.2}\pi_2}\ C(A,S_2),$$

And the risk-neutral BEU will choose action A if:

$$EP_{action\ A} = \pi_{1.1}\ C(A,S_1) + \pi_{2.1}\ C(A,S_2) > EP_{action\ B} = \pi_{1.1}\ C(B,S_1) + \pi_{2.1}\ C(B,S_2).$$

Therefore, given the parameter values in Table 4 for **rounds 1-4 & 13-16**, given red chip draw, **the RN BEU picks action A**,

$$EP_{action\ A} = \frac{.70(.5)}{.70(.5) + .30(.5)}(\$2.00) + \frac{.30(.5)}{.70(.5) + .30(.5)}(\$0.75) = \$1.625$$

$$EP_{action\ B} = \frac{.70(.5)}{.70(.5) + .30(.5)}(\$0.50) + \frac{.30(.5)}{.70(.5) + .30(.5)}(\$1.75) = \$0.875$$

For **rounds 1-4&13-16**,

if a blue chip is drawn, the risk-neutral BEU will choose action B, given that, $EP_{action\ B} > EP_{action\ A}$:

$$EP_{action\ B} = \frac{.70(.5)}{.70(.5) + .30(.5)}(\$1.75) + \frac{.30(.5)}{.70(.5) + .30(.5)}(\$0.50) = \$1.375$$

$$EP_{action\ A} = \frac{.70(.5)}{.70(.5) + .30(.5)}(\$0.75) + \frac{.30(.5)}{.70(.5) + .30(.5)}(\$2.00) = \$1.125$$

For **rounds 5-8 & 17-20 and a red chip draw RN BEU picks Action B**

$$EP_{action\ A} = \frac{.76(.5)}{.76(.5) + .24(.5)}(\$0.50) + \frac{.24(.5)}{.76(.5) + .24(.5)}(\$1.75) = \$0.80$$

$$EP_{action\ B} = \frac{.76(.5)}{.76(.5) + .24(.5)}(\$2.00) + \frac{.24(.5)}{.76(.5) + .24(.5)}(\$0.75) = \$1.70$$

For **rounds 5-8 & 17-20 and a blue chip draw RN BEU picks Action A**

$$EP_{action\ A} = \frac{.76(.5)}{.76(.5) + .24(.5)}(\$1.75) + \frac{.24(.5)}{.76(.5) + .24(.5)}(\$0.50) = \$1.45$$

$$EP_{action\ B} = \frac{.76(.5)}{.76(.5) + .24(.5)}(\$0.75) + \frac{.24(.5)}{.76(.5) + .24(.5)}(\$2.00) = \$1.05$$

For **rounds 9-12 & 21-24 and a red chip draws RN BEU picks Action B**

$$EP_{action\ A} = \frac{.60(.5)}{.60(.5) + .40(.5)}(\$0.50) + \frac{.40(.5)}{.60(.5) + .40(.5)}(\$1.00) = \$0.70$$

$$EP_{action\ B} = \frac{.60(.5)}{.60(.5) + .40(.5)}(\$2.00) + \frac{.40(.5)}{.60(.5) + .40(.5)}(\$0.75) = \$1.50$$

For **rounds 9-12 & 21-24 and a blue chip draws RN BEU picks Action B**

$$EP_{action\ A} = \frac{.60(.5)}{.60(.5) + .40(.5)}(\$1.00) + \frac{.40(.5)}{.60(.5) + .40(.5)}(\$0.50) = \$0.80$$

$$EP_{action\ B} = \frac{.60(.5)}{.60(.5) + .40(.5)}(\$0.75) + \frac{.40(.5)}{.60(.5) + .40(.5)}(\$2.00) = \$1.25$$

Appendix 5

Post-experiment Questionnaire

Participant Code_______.

We would appreciate it if you could provide the following information. Your responses will remain confidential, and you may decline to answer any question if you wish.

1. What is you gender? (Please mark one circle with an X) Male O Female O

2. What is your age? ____________________.

3. Where were you born?
   (Specify one response only, according to present boundaries)
   Born in Canada (Specify province or territory): _________________
   Born outside Canada (Specify country): _________________

4. What language do you speak most often at home? (Please mark one circle with an X)
   O English O French
   O Other – Specify: ___________, ___________, ___________

5. What program are you in at the University? ________________________________

6. What year are you currently in? (1st, 2nd etc..)________________________________

7. At what grade levels did you take English courses at High School? (please mark with an X)
   O Grade 9 O Grade 10
   O Grade 11 O Grade 12

8. Which Math courses did you take at High School? If you attended High School outside of Ontario, please mark from the list below the courses that are approximately equivalent to those you studied. (please mark with an X)
   O Grade 9 O Grade 10
   O Grade 11 O Grade 12- Relations & Functions
   O Grade 12-Calculus O Grade 12- Data Management

9. Have you taken any Math or Statistics courses in university (list)? ________________
   ____

10. Have you participated in an Economics or Psychology experiment before? (check one)
    __ Yes __ No

For the questions 11-22 (Honey & Mumford, 1986),

Please circle the degree to which you agree or disagree with the following statements

	Agree Strongly … Disagree Strongly
11. I tend to solve problems using a step-by- step approach.	1 2 3 4 5 6 7
12. I take pride in doing a thorough job.	1 2 3 4 5 6 7
13. What matters most is whether something works in practice.	1 2 3 4 5 6 7
14. I like to relate my actions to a general principle.	1 2 3 4 5 6 7
15. I find it difficult to produce ideas on impulse.	1 2 3 4 5 6 7
16. I prefer to have as many sources of information as possible- the more data to think over the better.	1 2 3 4 5 6 7
17. I am keen to reach answers via a logical approach.	1 2 3 4 5 6 7
18. In discussions I enjoy watching the maneuverings of other people.	1 2 3 4 5 6 7
19. I get along on best with logical, analytical people and less well with spontaneous people.	1 2 3 4 5 6 7
20. I accept and stick to laid down procedures so long as I regard them as an efficient way of getting the job done.	1 2 3 4 5 6 7
21. I am keen on exploring the basic assumptions, principles and theories underpinning things and events.	1 2 3 4 5 6 7
22. In discussions with people I often find I am the most dispassionate and objective.	1 2 3 4 5 6 7

23. How would you describe your strategy for making choices in this study?

____

____

____

Please write any comments you may have:

____

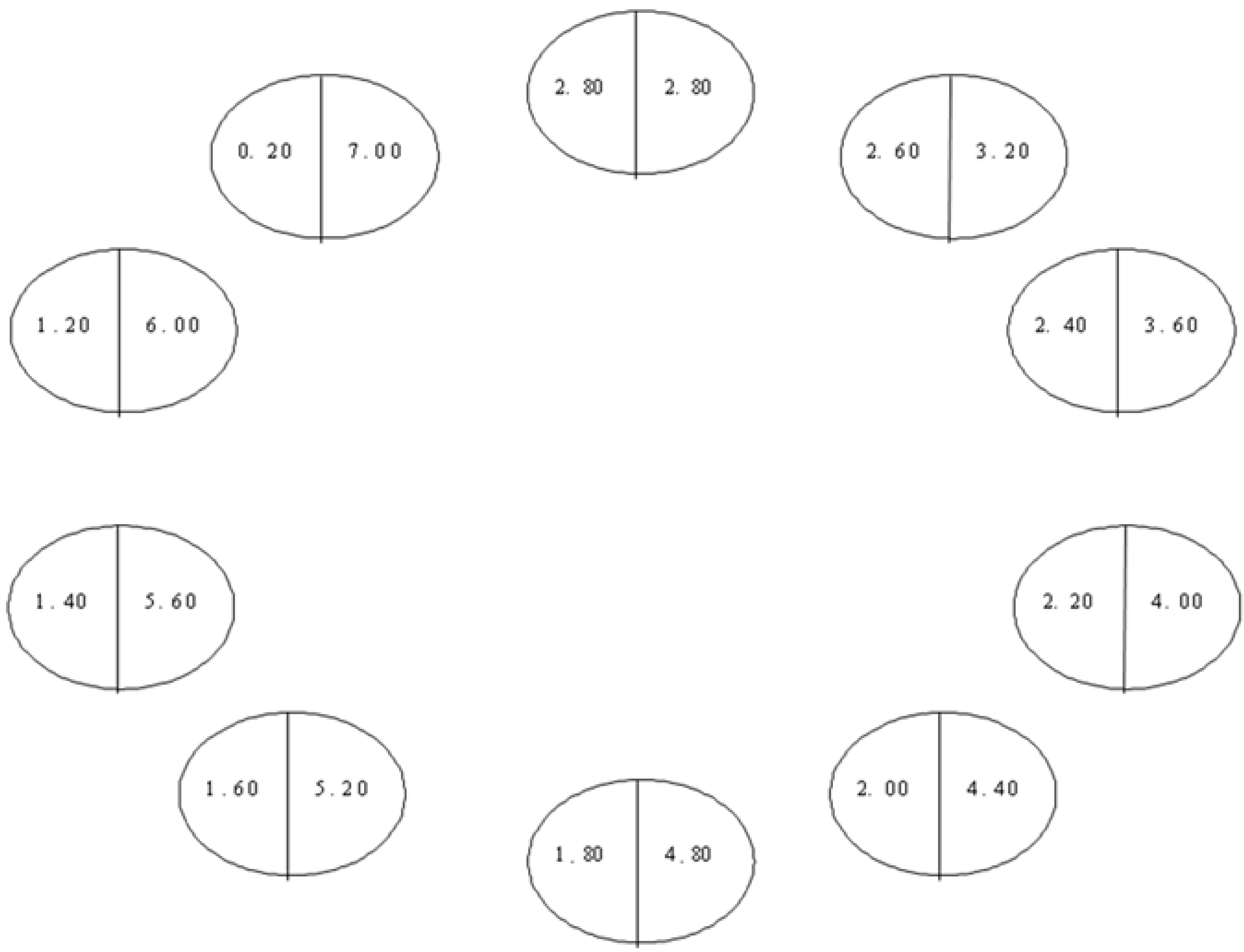
____

Appendix 6

Eckel-Grossman, Test for Risk Aversion (2002).

Participant Code_______.

24. Each of these ten circles represents a lottery with two possible prizes. The lotteries are played by flipping a coin. You will earn the amount on the left side of a circle if the flip is a HEAD, while you will the amount on the right side of a circle if the flip is a TAIL. **Please choose the lottery you most prefer by placing an X over it. You will play the lottery you choose for cash.**



Appendix 7

		BEU First Choice Deviations				BEU Second Choice Deviations				RL Second Choice Deviations		
ORDER	Obs.	**All**	**Un-inform**	**Inform**	**Obs.**	**All**	**Un-inform**	**Inform**	Obs.	**All**	**Un-inform**	**Inform**
Free 1-12	720	15.5%	14.7%	16.4%	720	14.3%	14.2%	14.4%	386	42.2%	34.9%	49.7%
		(0.363)	(0.354)	(0.371)		(0.350)	(.349)	(0.352)		(0.495)	(0.478)	(0.501)
OTP 13-24	720	9.6%	10.3%	8.9%	720	17.2%	19.2%	15.3%	378	40.7%	37.2%	44.0%
		(0.299)	(0.313)	(0.285)		(0.378)	(0.394)	(0.360)		(0.492)	(0.484)	(0.498)
Free/OTP	1440	**12.6%**	**12.5%**	**12.7%**	1440	**15.8%**	**16.7%**	**14.9%**	764	**41.5%**	**36.1%**	**46.9%**
OTP 1-12	732	17.9%	17.5%	18.3%	732	18.4%	19.4%	17.4%	374	35.0%	36.1%	33.9%
		(0.384)	(0.380)	(0.387)		(0.387)	(0.396)	(0.380)		(0.478)	(0.482)	(0.475)
Free 13-24	732	9.6%	10.8%	8.3%	732	12.6%	11.7%	13.4%	374	39.4%	39.2%	39.2%
		(0.294)	(0.311)	(0.277)		(0.332)	(0.321)	(.342)		(0.489)	(0.490)	(0.490)
OTP/Free	1464	**13.8%**	**14.2%**	**13.3%**	1464	**15.5%**	**15.6%**	**15.4%**	748	**37.2%**	**37.7%**	**36.6%**
Control 1-12	708	19.6%	18.4%	20.8%	708	20.2%	19.5%	20.8%	372	40.9%	39.1%	42.8%
		(0.397)	(0.387)	(0.407)		(0.402)	(0.397)	(0.407)		(0.492)	(0.489)	(0.496)
Control 13-24	708	10.4%	8.3%	12.5%	708	13.8%	16.4%	11.1%	389	45.0%	43.0%	46.7%
		(.306)	(0.277)	(0.331)		(0.345)	(0.371)	(0.315)		(0.498)	(0.496)	(0.500)
Control	1416	**15.0%**	**13.4%**	**16.7%**	1416	**17.0%**	**18.0%**	**16.0%**	761	**43.0%**	**41.1%**	**44.8%**
Total	4320	**13.8%**	**13.3%**	**14.2%**	4320	**16.0%**	**16.6%**	**15.4%**	2265	**40.6%**	**38.3%**	**42.8%**
		(0.345)	(.341)	(.349)		(0.367)	(0.374)	(0.361)		(0.491)	(0.428)	(0.495)

Aggregate Summary Statistics By Round Type

	All			Rounds 1-4			Rounds 5-8			Rounds 9-12		
Variable	**Obs.**	**Mean**	**Std. Dev.**	**Obs.**	**Mean**	**Std. Dev.**	**Obs.**	**Mean**	**Std. Dev.**	**Obs.**	**Mean**	**Std. Dev.**
Informed	4320	0.506	0.500	1440	0.506	0.500	1440	0.506	0.500	1440	0.506	0.500
Paid	4320	0.336	0.472	1440	0.336	0.473	1440	0.336	0.473	1440	0.336	0.473
Free 2nd	4320	0.503	0.500	1440	0.503	0.500	1440	0.503	0.500	1440	0.503	0.500
Female	4320	0.467	0.499	1440	0.467	0.499	1440	0.467	0.499	1440	0.467	0.499
Eng. 2nd	4320	0.200	0.400	1440	0.200	0.400	1440	0.200	0.400	1440	0.200	0.400
No HS Math	4320	0.028	0.164	1440	0.028	0.164	1440	0.028	0.164	1440	0.028	0.164
No U Math	4320	0.150	0.357	1440	0.150	0.357	1440	0.150	0.357	1440	0.150	0.357
Reinf Survey	4320	0.683	0.465	1440	0.683	0.465	1440	0.683	0.465	1440	0.683	0.465
Risk Aversion	4320	7.383	2.772	1440	7.383	2.772	1440	7.383	2.772	1440	7.383	2.772
age	4320	19.972	1.733	1440	19.972	1.734	1440	19.972	1.734	1440	19.972	1.734
year	4320	2.316	1.213	1440	2.317	1.214	1440	2.317	1.214	1440	2.317	1.214
Blue Draw	4320	0.447	0.497	1440	0.396	0.489	1440	0.441	0.497	1440	0.504	0.500
Degree of Informativeness	4320	0.373	0.132	1440	0.400	0.000	1440	0.520	0.000	1440	0.200	0.000
WTP	1452	0.404	3.887	484	0.546	4.723	484	0.550	4.768	484	0.117	0.485
Bayes WTP	1452	0.103	0.215	484	0.104	0.190	484	0.177	0.269	484	0.027	0.140
Diff WTP & BWTP	1452	0.302	3.878	484	0.443	4.723	484	0.373	4.753	484	0.090	0.470
Yes Bayes WTP	1452	0.352	0.478	484	0.403	0.491	484	0.442	0.497	484	0.211	0.408
WTP Restricted at $2	1452	0.171	0.384	484	0.197	0.423	484	0.219	0.407	484	0.099	0.304
Diff Restricted WTP & BWTP	1452	0.069	0.375	484	0.092	0.430	484	0.042	0.380	484	0.072	0.306
Random Price draw	1452	0.235	0.147	484	0.233	0.131	484	0.245	0.152	484	0.228	0.157
Exp Bayes PO	4320	1.435	0.212	1440	1.458	0.209	1440	1.473	0.263	1440	1.374	0.125
Paid on 2nd Choice	2868	0.446	0.497	956	0.448	0.498	956	0.447	0.497	956	0.442	0.497
Bag 2 Revealed	4320	0.523	0.500	1440	0.502	0.500	1440	0.552	0.518	1440	0.526	0.500
BR not Bayes Prediction	4320	0.363	0.481	1440	0.356	0.479	1440	0.257	0.437	1440	0.474	0.500
PO	4320	1.396	0.621	1440	1.415	0.611	1440	1.445	0.615	1440	1.329	0.632
Bayes PO	4320	1.457	0.608	1440	1.440	0.606	1440	1.525	0.586	1440	1.407	0.624
Diff PO	4320	-0.061	0.479	1440	-0.026	0.511	1440	-0.079	0.537	1440	-0.078	0.373
% PO maximized	4320	0.534	0.499	1440	0.536	0.499	1440	0.556	0.497	1440	0.510	0.500
% PO Bayes Maximized	4320	0.557	0.497	1440	0.541	0.498	1440	0.603	0.489	1440	0.526	0.499
% lose	4320	0.415	0.493	1440	0.410	0.492	1440	0.367	0.482	1440	0.467	0.499

Appendix 8: 1st Choice BEU Inconsistency Regression

Variables	(1) BEUD1 1st Choice Deviation xtlogit, Re	(2) BEUD1 1st Choice Deviation xtlogit, Fe	(3) BEUD1 1st Choice Deviation GLS, Re	(4) BEUD1 1st Choice Deviation GLS, Fe
Experience	**-7970***** (.1039)	**-.7873***** (.1022)	**-.0754***** (.0095)	**-.0753***** (.0095)
Free followed by OTP	**-.0101** (.2768)	**(omitted)**	**-.0121** (.0237)	**(omitted)**
OTP	**.2038*** (.1260)	**.2467**** (.1299)	**.0191*** (.0116)	**.0202*** (.0119)
Informed	**.1942** (.2652)	**(omitted)**	**.0124** (.0224)	**(omitted)**
Difference in ex-ante Payoffs	**-4.665***** (.6159)	**-4.585***** (.6096)	**-.4090***** (.0527)	**-.4085** (.0528)
Bag Reveal in Prior round not equal to BEU prediction	**-.0699** (.1060)	**-.0760** (.1082)	**-.0044** (.0102)	**-.0051** (.0102)
Subject Paid on second action in Prior Round	**.3200***** (.1060)	**.3162***** (.1060)	**.0334***** (.0103)	**.0330***** (.0104)
Female	**.3629** (.2806)	**(omitted)**	**.0319** (.0240)	**(omitted)**
English 2nd	**.8994***** (.3200)	**(omitted)**	**.0828***** (.0282)	**(omitted)**
Reinforcement Learner Survey	**-.0678** (.2956)	**(omitted)**	**-.0110** (.0248)	**(omitted)**
Risk Aversion	**-.0421** (.0479)	**(omitted)**	**-.0027** (.0042)	**(omitted)**
Age	**-.1682***** (.0808)	**(omitted)**	**-.0142***** (.0065)	**(omitted)**
Econ Math Student	**-.0510** (.3103)	**(omitted)**	**.0015** (.0263)	**(omitted)**
Obs.	4320	2928	4320	4320
log likelihood/R-squared	-1451.84	-1041.51	.1060	.0339

Appendix 9: Second choice BEU inconsistencies (RE and FE) and Marginal Effects

Variables	(5a) BEUD2 2nd Choice Deviation xtlogit, Re	(5b) BEUD2 2nd Choice Deviation Marginal Effects	(6) BEUD2 2nd Choice Deviation xtlogit,Fe	(7) BEUD2 2nd Choice Deviation GLS,Re	(8) BEUD2 2nd Choice Deviation GLS,Fe
Experience	**-0.257*** **	**-.0285*** **	**-0.261*** **	**-0.025*** **	**-0.025*** **
	0.093	0.010	0.093	0.010	0.010
Free followed by OTP	**-0.167**	**-0.018**	(omitted)	**-0.020**	(omitted)
	0.153	0.016		0.017	
OTP	**0.577*** **	**.0341*** **	**0.682*** **	**0.056*** **	**0.067*** **
	0.151	0.013	0.1636	0.017	0.018
Informed	**0.147**	**-0.008**	**(omitted)**	**0.007**	**(omitted)**
	0.162	0.016		0.017	
Paid*Informed	**-0.5739*** **		**-0.655*** **	**-0.055*** **	**-0.063*** **
	0.210		0.227	0.024	0.025
Difference in Payoffs Good vs. Bad state conditional on chip draw	**-3.570*** **	**-0.397*** **	**-3.665*** **	**-0.360*** **	**-0.369*** **
	0.296	0.035	0.294	0.033	0.033
Shift Required to follow BEU	**1.114*** **	**0.133*** **	**1.049*** **	**0.126*** **	**0.117*** **
	0.117	0.015	0.115	0.014	0.014
Degree informative	**-3.025*** **	**-0.336*** **	**-2.896*** **	**-0.235*** **	**-0.224*** **
	0.425	0.048	0.422	0.045	0.045
Bag Reveal in Prior round not equal to BEU prediction	**0.218*** **	**0.025*** **	**0.213*** **	**0.025**	**0.025*** **
	0.098	0.011	0.098	0.011	0.011
Subject Paid on second action in Prior Round	**-0.078**	**-0.008**	**-0.037**	**-0.018**	**-0.016**
	0.099	0.011	0.101	0.011	0.012
Female	**0.167**	**0.019**	**(omitted)**	**0.014**	**(omitted)**
	0.153	0.017		0.017	
English 2nd	**0.190**	**0.022**	**(omitted)**	**0.014**	**(omitted)**
	0.177	**0.021**		0.020	
Reinforcement Learner Survey	**0.112**	**0.012**	**(omitted)**	**0.014**	**(omitted)**
	0.161	**0.017**		0.017	
Risk Aversion	**-0.021**	**-0.002**	**(omitted)**	**-0.003**	**(omitted)**
	0.026	0.003		0.003	
Age	**-0.011**	**-0.001**	**(omitted)**	**-0.001**	**(omitted)**
	0.043	**0.005**		0.005	
Econ Math Student	**-0.258**	**-0.029**	**(omitted)**	**-0.027**	**(omitted)**
	0.172	**0.019**		0.018	
Obs.	4320	4320	3984	4320	4320
log likelihood/R-squared	-1631.05		-1224.486	0.099	0.094

Appendix 10: Logit regression Free Vs. OTP and OTP uninformed Vs. Informed

xtlogit, Re	(9)	(10)	(11)	(12)	(13)
	BEUD2	BEUD2	BEUD2	BEUD2	BEUD2
	2nd Choice Deviation	2nd Choice Deviation	2nd Choice Deviation	2nd Choice Deviation	2nd Choice Deviation
Variable	**All**	**Free**	**OTP**	**OTP Uninformed**	**OTP Informed**
Experience	**-0.221*****	**-0.315*****	**-0.361**	**-0.264**	**-0.466**
	0.091	0.143	0.235	0.358	0.346
Free followed by OTP	**-0.125**	**-0.277**	**(omitted)**		
	0.149	0.195			
Informed	**-0.07**	**0.134**	**-0.534*****		
	0.142	0.159	0.235		
Difference in Payoffs Good vs. Bad state conditional on chip draw	**-4.418*****	**-4.431*****	**-4.650*****	**-4.757*****	**-4.435*****
	0.271	0.338	0.486	0.630	0.782
Degree informative	**-0.934*****	**-0.708***	**-1.5226*****	**0.670**	**-4.143*****
	0.345	0.432	0.602	0.841	0.921
Bag Reveal in Prior round not equal to BEU prediction	**0.162***	**0.052**	**0.441*****	**0.594*****	**0.146**
	0.095	0.121	0.167	0.227	0.257
Subject Paid on second action in Prior Round	**-0.119**	**-0.074**	**0.034**	**0.011**	**0.180**
	0.095	0.117	0.194	0.295	0.267
Female	**0.185**	**0.175**	**0.255**	**0.173**	**0.473**
	0.151	0.167	0.255	0.387	0.363
English 2nd	**0.189**	**0.246**	**0.064**	**0.029**	**-0.091**
	0.175	0.191	0.300	0.392	0.504
Reinforcement Learner Survey	**0.16**	**0.045**	**0.366**	**0.311**	**0.407**
	0.159	0.176	0.272	0.384	0.397
Risk Aversion	**-0.018**	**-0.020**	**-0.014**	**0.006**	**-0.018**
	0.026	0.028	0.045	0.067	0.063
Age	**-0.0137**	**-0.015**	**0.001**	**0.085**	**-0.041**
	0.042	0.047	0.068	0.099	0.097
Econ Math Student	**-0.269**	**-0.308***	**-0.197**	**0.337**	**-0.657***
	0.17	0.192	0.272	0.401	0.382
Obs.	**4320**	**2868**	**1452**	**720**	**732**
log likelihood/R-squared	**-1687.076**	**-1084.140**	**-593.102**	**-306.188**	**-276.043**

Appendix 11- 2nd Choice Reinforcement Inconsistency Rate Logit Fixed and Random Effects

	(14a)	(14b)	(15)	(16)	(17)
	RLD2	RLD2	RLD2	RLD2	RLD2
	2nd Choice Deviation	2nd Choice Deviation	2nd Choice Deviation	2nd Choice Deviation	2nd Choice Deviation
Variables	xtlogit, Re	Marginal Effects	xtlogit,Fe	GLS,Re	GLS,Fe
Experience	**0.185*** **	**0.042*** **	**0.199*** **	**0.041*** **	**0.043*** **
	0.089	0.020	0.089	0.020	0.020
Free followed by OTP	**0.132**	**0.030**	(omitted)	**0.029**	(omitted)
	0.098	0.023		0.022	
OTP	**-0.088**	**-0.043**	**-0.010**	**-0.020**	**-0.003**
	0.141	0.023	0.162	0.032	0.036
Informed	**0.272*** **	**0.048*** **	(omitted)	**0.061*** **	(omitted)
	0.108	0.021		0.025	
Paid*Informed	**-0.194**		**-0.181**	**-0.045**	**-0.037**
	0.193		0.223	0.044	0.051
Difference in Payoffs Good vs. Bad state conditional on chip draw	**-1.207*** **	**-0.277*** **	**-1.222*** **	**-0.276*** **	**-0.288*** **
	0.271	0.061	0.285	0.063	0.066
Shift Required to follow BEU	**0.144**	**0.033**	**0.022**	**0.032**	**0.002**
	0.114	0.026	0.119	0.026	0.028
Degree informative	**-1.901*** **	**-0.436**	**-1.844*** **	**-0.429*** **	**-0.42*** **
	0.380	0.085	0.383	0.087	0.088
Bag Reveal in Prior round not equal to BEU prediction	**-0.653*** **	**-0.146*** **	**-0.815*** **	**-0.147*** **	**-0.188*** **
	0.103	0.022	0.107	0.023	0.024
Subject Paid on second action in Prior Round	**-0.153**	**-0.035**	**-0.186**	**-0.036**	**-0.038**
	0.095	0.022	0.100	0.022	0.023
Female	**0.104**	**0.024**	(omitted)	**0.025**	(omitted)
	0.096	0.022		0.022	
English 2nd	**0.000**	**0.000**	(omitted)	**0.000**	(omitted)
	0.113	0.026		0.026	
Reinforcement Learner Survey	**-0.284*** **	**-0.066*** **	(omitted)	**-0.065**	(omitted)
	0.100	0.023		0.023	
Risk Aversion	**-0.007**	**-0.002**	(omitted)	**-0.002**	(omitted)
	0.017	0.004		0.004	
Age	**-0.023**	**-0.005**	(omitted)	**-0.005**	(omitted)
	0.026	0.006		0.006	
Econ Math Student	**0.141**	**0.032**	(omitted)	**0.032**	(omitted)
	0.1048	0.024		0.0242	
Obs.	2265	2265	2265	2265	2265
log likelihood/R-squared	-1472.553		-1124.277	0.0486	0.0382

Appendix 12: Logit Regressions by Behaviour Type

Obs	35		28		35		18	
% of sample	19.4%		15.6%		19.4%		10.0%	
	BEU		Status Quo		Overweight		Non BEU	
Variables	Logit	OR	Logit	OR	Logit	OR	Logit	OR
Free followed by OTP	**-0.236**	**0.790**	**0.568**	**1.765**	**0.090**	**1.094**	**-0.654**	**0.520**
	0.488	0.393	0.535	0.929	0.510	0.556	0.651	0.347
OTP followed by Free	**-0.238**	**0.788**	**-1.244***	**0.288***	**0.440**	**1.553**	**-0.169**	**0.844**
	0.545	0.394	0.643	0.181	0.477	0.748	0.603	0.509
Informed	**-0.485**	**0.615**	**-1.012****	**0.363****	**-0.128**	**0.880**	**0.014**	**1.014**
	0.412	0.255	0.512	0.181	0.399	0.350	0.541	0.535
Female	**-0.535**	**0.586**	**0.198**	**1.219**	**0.879****	**2.408****	**-0.349**	**0.705**
	0.492	0.267	0.549	0.640	0.447	1.046	0.545	0.398
English 2nd	**-1.138***	**0.320***	**0.819**	**2.268**	**-0.232**	**0.793**	**0.988***	**2.685***
	0.674	0.213	0.519	1.264	0.522	0.399	0.574	1.510
Reinforcement Learner Survey	**-0.398**	**0.672**	**1.524*****	**4.591*****	**-0.295**	**0.745**	**-0.270**	**0.763**
	0.448	0.285	0.733	2.966	0.443	0.332	0.526	0.442
Risk Aversion	**0.021**	**1.021**	**0.195****	**1.215****	**-0.116***	**0.891***	**-0.126**	**0.882**
	0.076	0.080	0.103	0.121	0.067	0.062	0.094	0.080
Age	**0.120**	**1.127**	**0.158**	**1.171**	**0.012**	**1.012**	**-0.192**	**0.825**
	0.122	0.129	0.134	0.159	0.109	0.115	0.177	0.140
Econ Math Student	**0.922****	**2.513****	**-1.745*****	**0.174*****	**0.037**	**1.037**	**0.500**	**1.648**
	0.451	0.218	0.842	0.148	0.481	0.498	0.532	0.9336
Obs.	180		180		180		180	
log likelihood/R-squared	-79.419		-61.184		-83.427		-53.808	

******p-value ≤.01 <**p-value ≤.05 < *p-value < .10***

Robust standard errors