

https://dspace.org/

A test taker's gamble: an analysis of the framing effect on a student's degree of risk aversion under a negative marking regime

Kika, Jesal Chandrakant

2015
http://hdl.handle.net/10539/19172

Downloaded from a DSpace repository. https://dspace.org/

A Test Taker's Gamble: An Analysis of the Framing Effect on a Student's Degree of Risk Aversion under a Negative Marking Regime

by

Jesal Chandrakant Kika

Student number: 354175

MASTER OF ECONOMIC SCIENCE (CC011)

in the

SCHOOL OF ECONOMIC AND BUSINESS SCIENCES

at the

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

Supervisors: Dr Aylit Romm & Mr Volker Schöer

Date of submission:

27 February 2015

SCHOOL OF ECONOMIC AND BUSINESS SCIENCES

Declaration Regarding Plagiarism

I (full names & surname):	Jesal Chandrakant Kika
Student number:	354175

Declare the following:

1. I understand what plagiarism entails and am aware of the University's policy in this regard.
2. I declare that this assignment is my own, original work. Where someone else's work was used (whether from a printed source, the Internet or any other source) due acknowledgement was given and reference was made according to departmental requirements.
3. I did not copy and paste any information directly from an electronic source (e.g., a web page, electronic journal article or CD ROM) into this document.
4. I did not make use of another student's previous work and submitted it as my own.
5. I did not allow and will not allow anyone to copy my work with the intention of presenting it as his/her own work.

Signature _____ Date _____

Acknowledgements

I would like to express my sincere gratitude to my supervisors, Dr Aylit Romm and Mr Volker Schöer for their invaluable support, guidance and reciprocal altruism. In addition, I gratefully acknowledge the funding for this dissertation from the African Micro-Economic Research Unit (AMERU), and the support and assistance from both the research team at AMERU and my fellow Masters classmates. Lastly, a special thank you goes out to my family and friends for their moral support and selfless encouragement.

Abstract

A test or an examination in an educational setting may be considered as a situation in which individuals face a series of decisions under risk. The experiment conducted in this study replicates an actual test situation in which students approach a multiple choice test having accumulated known scores from previous assessments, each contributing towards an aggregate. It permits the analysis of the framing effect on a student's degree of risk aversion, based on their tendency to guess, under the adoption of a negative marking regime. This enables an evaluation of whether or not the behaviour of students, influenced by the framing effect of a scoring rule as a potential gain or loss relative to a reference point, is consistent with the behaviour initiated by prospect theory. Furthermore, this study investigates whether a deviation of points further into the gain and loss domain yields increasingly or decreasingly risk averse behaviour. In general, the findings reveal that individuals exhibit a higher degree of risk aversion in the gain domain in comparison to the loss domain. In addition, individuals tend to become less risk averse/more risk seeking as a deviation of points away from the reference in the loss domain such that the lowest degree of risk aversion is displayed at the extreme end of the loss domain. Similarly, for a deviation of points further away from the reference in the gain domain, individuals in financial need i.e. those who are recipients of a lower monthly income tend to become increasingly risk averse, while those who are recipients of a higher monthly income tend to become decreasingly risk averse.

Table of Contents

Acknowledgements iii

Abstract iv

List of Figures vi

List of Tables vi

1 Introduction 1

2 Literature Review 3

3 Experiment 8

4 Analysis and Discussion 12

4.1 Treatment Effect 12

4.2 Effect of the treatment on participants who are willing to receive a monthly wage of R500 16

4.3 Effect of the treatment on participants who are NOT willing to receive a monthly wage of R500 17

4.4 Gender specific treatment effects 18

4.5 Gender specific treatment effects for participants who are willing and not willing to receive a monthly wage of R500 21

5 Regression Analysis 23

5.1 Methodology 23

5.2 Regression Results 23

6 Conclusion 27

7 References 29

A Supplementary Appendix 31

Test Questionnaire: 31

Solutions: 34

List of Figures

Figure 1: A Hypothetical Value Function 6
Figure 2: A Hypothetical Value Function in the Context of the Experiment 9

List of Tables

Table 1: Percentage values of observable characteristic for each treatment group..... 10
Table 2: Treatment effects for the first two groups who are simply in the gain and loss domain 13
Table 3: Treatment effects for the second two treatment groups with a deviation of points further into the loss and gain domain..... 14
Table 4: Treatment effects for the full sample..... 15
Table 5: Treatment effects for participants who are willing to accept a monthly wage of R500 17
Table 6: Treatment effects for participants who are not willing to accept a monthly wage of R500 ... 18
Table 7: Gender specific treatment effects 20
Table 8: Gender specific treatment effects for participants willing to accept a monthly wage of R500 21
Table 9: Gender specific treatment effects for participants who are not willing to receive a monthly wage of R500..... 22
Table 10: Regression results..... 24
Table 11: Effects of possible treatment heterogeneities 26

1 Introduction

There are many situations in which individuals are required to make a series of decisions in the presence of risk. A typical situation is that of a student during a multiple choice examination, under a negative marking regime. This negative marking regime, architected to discourage guessing, presents students who are not able to certainly identify the correct answer to a question with a gamble. Either the problematic question is omitted, ensuring that no points are gained or lost, or it is answered, manifesting the possibility of scoring points for a correct response, but inviting the risk of being penalised for an incorrect one. Hence, the performance on such a test is dependent upon whether or not the student chooses to answer a question when confronted with such a gamble. Expected utility theory has been a preponderant approach in the analysis of decision making under risk, in which explicit information pertaining to the potential consequences of choices made by economic agents and their associated probabilities are known. In addition, the expected utility approach has been viewed as a descriptive model of rational choice; however, the rationality of an individual is limited by the cognitive limitations of their minds, the information that they are presented with, and the time period that they have in order to make a decision.

The behavioural characteristics that influence risk taking by individuals may also have an effect on the decisions made by students during a test or an examination. Ordinarily, a rational examinee is defined as a student whose aim is to maximise his/her expected score; however, extensive literature on behavioural decision making advocates that individuals may behave in a non-rational manner, condemning the rational prescriptions of expected utility theory. The framing effect[1] (Tversky & Kahneman, 1981) of scoring rules as a potential gain or loss relative to a reference point has been proposed as a possible explanation of why students do not maximise their expected score and can be elucidated using prospect theory. Prospect theory - developed by Kahneman & Tversky (1979), is based on the premise that individuals are not always rational in their decision making. The authors postulate that individuals are risk averse when their outcomes are perceived as gains and risk seeking when perceived as losses.

In a decision making experiment conducted by Bereby-Meyer et al. (2002), participants – who were first-year, Israeli, undergraduate, social science students, were presented with a

[1] The "framing effect" is an example of a cognitive bias which occurs when the description of a set of options in terms of gains and losses elicits systematically different choices which affects a decision maker's risk tolerance.

hypothetical multiple choice test situation under different scoring rules and were asked to indicate the degree of certainty, ranging from 0 to 100, that they judged was necessary in order to answer each question presented to them. A gain and loss domain had been created as some participants were told to anticipate a grade above the passing grade required for the course, while others were told to anticipate a grade below the passing grade which was therefore assumed to be the reference point (Bereby-Meyer, Meyer, & Flascher, 2002). The authors conducted a further experiment which resembled an actual test situation aimed at replicating the results from the prior synthetic experiments of their study in a more realistic setting. Since this experiment was conducted as part of a course requirement, participants were told that they would receive half an hour of experimental credit in which they will be given an additional half an hour to complete any future tests for the course under consideration, if their final score reached a value above 75 points, thus serving as the new reference point. In this instance, a gain and loss frame had been created as one group was told that the average score obtained by students who had previously taken the test was above the 75 point reference, whereas the second group was told that the average score of these examinees was below the 75 point reference.

The experiment in this study is designed to analyse the effects of framing in a multiple choice test setting on the tendency to guess under a negative marking rule by replicating an actual test situation in which students enter a test or examination with a certain aggregate based upon their previous test scores or examination results. In addition, this study aims to assess whether the behaviour of students under such circumstances is consistent with the behaviour predicted by prospect theory. According to prospect theory, students should be risk averse in the domain of gains, rendering them reluctant to guess; and risk seeking in the domain of losses, encouraging more guessing. It therefore finds itself embedded in the realm of behavioural economic literature. As a departure from the experiment conducted by Bereby-Meyer et al. (2002), a gain and loss domain which reflects the average test scores of students from previous tests or examinations is replicated by assigning differential points randomly to different groups before they commence with the experimental test. Participants gain and lose points according to the negative marking rule adopted in this study[2] in addition to the points that they enter the test with. The payoff to the participants are in monetary terms on a Rand (R1) per point basis. As such, the students were supplied with an incentive to participate in the experiment. Participating

[2] A detailed description of the negative marking rule adopted in this study will be discussed in Section 3.

students were promised a bonus of 50 Rand (R50) if they attained a score above 50 points after completion of the test. Thus, a score of 50 points serves as a reference which is reflective of a pass mark for most courses in a realistic test setting.

A further contribution of this study is investigating whether the distance into the gain and loss domain as a deviation of marks further away from the reference point yields significantly different results. This, to our knowledge, has previously not been investigated in any empirical literature pertaining to prospect theory. In particular, the experiment seeks to investigate whether individuals are increasingly or decreasingly risk averse in the gain domain; and similarly, whether individuals are increasingly or decreasingly risk seeking in the loss domain. For a greater deviation of points in the gain domain, students may either exhibit increased risk aversion as they prefer to remain with certainty at a particular score and are thus reluctant to guess and lose the points which they have accumulated; or they may exhibit decreasingly risk averse behaviour in which their willingness to guess increases since they have more points to take a gamble with. Similarly, for a greater deviation in the loss domain, students may display an increase in their willingness to guess as it is beneficial for them to take a gamble and potentially receive a higher score from each correct guess; or they may display decreasingly risk seeking behaviour as they are reluctant to lose the few points that they have accumulated.

The remainder of this study will proceed as follows: Section 2 reviews literature pertaining to scoring rules and the framing effect, and provides a brief description of prospect theory. A description of the experiment conducted in this study is presented in Section 3 with an analysis of the data and a discussion of the corresponding results is presented in Section 4. Section 5 provides a description of the regression methodology and the accompanying results, and Section 6 concludes this paper.

2 Literature Review

The use and grading of multiple choice questions is a well-established and reliable method of assessing knowledge in standardised tests and examinations within the education space. These multiple choice tests are advantageous to both the instructor and the student. From an instructor's perspective, these tests offer increased accuracy and reliability in scoring (Walstad & Becker, 1994) as well as objectivity of the grading process (Becker & Johnston, 1999). In addition, Buckles & Siegfried (2006) reveal that these tests enable instructors to cover a wide range of subject material, and facilitate the availability of comparative statistical analysis.

From a student's perspective, the objectivity of the grading process is welcomed, since it ensures consistent scoring, and eliminates instructor bias (Kniveton, 1996). Furthermore, students can increase the probability of guessing the correct answer to a question by eliminating the unlikely choices (Bush, 2001), especially under scoring rules in which these random guesses are not subject to a penalty.

The basic scoring rule adopted by many examiners is the Number of Right (NR) rule in which the test score is the number of correct answers multiplied by some constant, and there is no penalty for an incorrect response. For a student who wishes to maximise his/her score on a test, answering every question, even though some answers will have to be chosen at random, will be his optimal strategy (Lord, 1975). A drawback of this rule is the deleterious effect of the guessing of responses on the reliability of a test. Given that there is some probability that a guessed response will be correct, an examinee may gain points for questions in which they have no knowledge of the correct solution. In the presence of different response strategies between two students with the same level of knowledge, in which one student displays a tendency towards guessing, and the other does not - the examinee who chooses to guess subsequently receives a higher observed score for the test. Thus, the guessing strategy adds random error to the variance of test scores which undermines the reliability of the assessment (Bereby-Meyer, Meyer, & Flascher, 2002).

Formula scoring rules, also known as negative marking, are frequently adopted as a means to discourage guessing by subtracting points for incorrect responses; and unanswered questions are neither penalised nor rewarded (Holt, 2006). The penalty for an incorrect response serves to augment the reliability of tests through a reduction in the measurement errors induced by guessing. A formula scoring rule of this nature will be adopted in this study. A basic property of such formula scoring rules is that the expected value of a pure guess is the same as the expected value from omitting a response (Budescu & Bar-Hillel, 1993). However, Davis (1967) asserts that a limitation of the formula scoring rule is the failure to take into account the partial knowledge of examinees, which enables them to eliminate one or more solution options. As students with partial knowledge eliminate one or more solution options, the expected value of guessing exceeds the expected value of omitting a response which therefore results in guessing as being the optimal strategy. The findings of Bliss (1980) reveals that risk averse students are more likely to omit questions for which they only have partial knowledge. Regardless of the positive expected payoff, these examinees underestimate their partial knowledge and are at a

disadvantage compared to other students. This approach fails to maximise the expected score, and the bias against risk averse students under a negative marking regime may vary systematically, depending on the demographic characteristics of the student (Hartford & Spearman, 2014).

Behavioural literature pertaining to gender differences and risk attitudes advocates that females exhibit a greater degree of risk aversion in comparison to their male counterparts. Within a multiple choice test framework among economics students in Spain in which students were penalised for incorrect responses, Marin and Rosa-Garcia (2011) report evidence of a higher degree of risk aversion for female students, measured by the number of unanswered questions. Consequently, the test scores of more risk averse students are lower since they behave in a non-rational manner; however, the differences among test scores between less risk averse students and more risk averse students are marginal. The Economics Department at the University of the Witwatersrand implemented a decision in 2013 to discontinue negative marking in multiple choice test questions for the first year economics course. Hartford and Spearman (2014) viewed this decision as an opportunity to conduct a natural experiment with the aim of investigating the gender bias effect under a negative marking regime. The assessment results were obtained in an actual test setting and were therefore free from any framing biases associated with controlled experiments. The authors' findings provide evidence of a gender bias against female students when a negative marking rule is adopted. In addition, female students in higher performance quantiles exhibit greater risk aversion and omit a larger number of questions than those in lower performance quantiles, and are thus adversely affected by negative marking (Hartford & Spearman, 2014).

According to Bereby-Meyer et al. (2002), the framing of scoring rules serves as a possible explanation of why students do not maximise their expected score. In particular, when individuals' choices are framed positively in terms of gains and negatively in terms of losses in the face of equivalent descriptions of a decision problem, this results in contradictory choices among them. Tversky and Kahneman (1981) suggest that the frame adopted by a decision maker is dependent partly upon the way in which the problems are presented, and partly upon the norms, habits and personal characteristics of the decision maker. According to decision theory, the framing effect may be defined in both a "strict" sense and a "loose" sense. Tversky and Kahneman (1981) are proponents of this definition in a strict sense which is related to the manipulation of the wording of identical problems. A classic example proposed by the authors is the "Asian Disease Problem" which reveals that decision makers tend to be risk averse when

an outcome is viewed as a gain and risk seeking when an outcome is viewed as a loss. Kuhberger (1998) is an advocate of framing in the loose sense in which a situation is referred to as an internal event induced by other contextual features aside from semantic manipulation; as well as individual factors. An example of the framing effect in a loose sense is provided by Aquino, Steisel & Kay (1992) in which differential responses to social dilemmas that result in a collective disaster as a consequence of individually reasonable decisions, are presented as give-some versus take-some dilemmas.

Bereby-Meyer et al. (2002) base their analysis of test-taking on prospect theory which is the most well-known formal theory, providing a descriptive model of choice under risk developed by Kahneman & Tversky (1979).

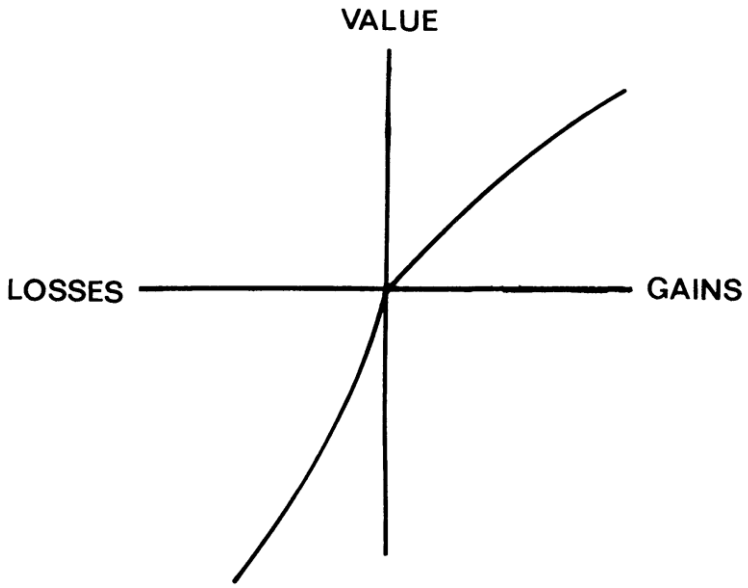


***Figure 1*: A Hypothetical Value Function**

For simplicity, Kahneman & Tversky (1979) restrict the formal treatment to simple gambles or lotteries involving stated numerical probabilities and quantitative outcomes. Consider a simple prospect $(x, p; y, q)$ that yields an outcome x with probability p, an outcome y with probability q, and preserves the status quo, which is usually an individual's reference point, with probability $(1 - p - q)$. According to prospect theory, a decision weight $\pi(p)$ reflects the impact of p on the overall value of a prospect, and a value v assigned to each outcome, such that $v(.)$ reflects the subjective value of a particular outcome. The overall value of a prospect is therefore given by $\pi(p)v(x) + \pi(q)v(y)$ where $v(0) = 0,\ \pi(0) = 0$ and $\pi(1) = 1$. In addition, Kahneman &Tversky (1981) propose an S-shaped valuation function defined on deviations from a reference point of which a hypothetical illustration is depicted in Figure 1

above. The decision weight function $\pi(p)$ is monotonic in p and tends to overweight highly unlikely outcomes or low probabilities, and underweight moderate to high probabilities as a result of the limited ability of individuals to comprehend and evaluate extreme probabilities.

The aforementioned features result in preferences that depend on the way in which problems are framed. Bereby-Meyer et al. (2002) suggest that the framing effect has implications for the guessing of solutions in a multiple choice test. In the loose sense of the term, framing in such a situation denotes the way in which students perceive themselves as a function of their individual factors. These factors include the current situation that they are faced with, their abilities, previous achievements and aspirations. Prospect theory predicts a concave and shallow value function for students who perceive themselves to be in the gain domain. A student may choose to omit an answer and remain with certainty at some position, or engage in a gamble and choose to guess, in which both a gain and a loss are possible outcomes. The shape of the valuation function in the gain domain implies that the value associated with expected gains is of lower magnitude than the value associated with an equally sized expected loss, thus omitting an answer seems more attractive than guessing, and students tend to be risk averse (Bereby-Meyer, Meyer, & Flascher, 2002). In contrast, prospect theory predicts a convex and relatively steep value function for students in the loss domain. The shape of the valuation function implies that the value associated with the expected gain is larger in magnitude than the value associated with the expected loss, thus guessing an answer is more attractive than omission, and students tend to be risk seeking in the loss domain.

Krawczyk (2011) makes use of a field experiment to manipulate the framing of simple decisions under risk. The framing of a microeconomics test question as an opportunity for either a gain or a loss is examined when a missing response is scored higher than an incorrect one, and the impact of a test-takers' likelihood to answer the question is evaluated. A gain and loss treatment is created, in which each treatment group receives different exam instructions regarding the number of points per item. Within the gain treatment, a student can never lose a point, whereas in the loss treatment, a student can never gain a point. Whether a solution to a question is correct, incorrect or omitted, a student in the gain treatment is awarded 3 points more as a baseline than a student in the loss treatment. Under this regime, in an examination comprising of 30 multiple choice questions, any solution combination would automatically yield 90 points more within the gain domain in comparison to the loss domain. The results reveal that the framing effect has no impact on the likelihood of answering a question. The experimental subjects were generally risk averse in both the gain and loss domain, and this

deviation from the pattern predicted by Kahneman & Tversky (1979) may be attributed to the underlying reward medium involving academic success instead of a monetary payoff (Krawczyk, 2011) such as the one which will be used in this study.

3 Experiment

A sample of 102 undergraduate Economics students from the University of the Witwatersrand voluntarily enrolled to participate in the experiment. Participating students were randomly allocated to four different treatment groups; however, an attempt was made in order to stratify allocation by gender to ensure that there were a sufficient number of male and female students within each treatment group. This stratification permitted the analysis of possible treatment heterogeneities along gender lines.

The decision making experiment was conducted in a classroom setting. A multiple choice test (see supplementary appendix) consisting of 10 questions was presented to the students, and each question was comprised of 3 possible solution options, of which the student was made to choose 1. The test consisted of a number of questions that could not be answered, whereby none of the alternative solutions provided were correct, and a response to such a question was indicative of a pure guess (Slakter, 1969). In order to create the perception of legitimacy of the multiple choice test, the first question was solvable and included a correct solution as an option; however, this solvable question was not considered in the final analysis of this study. A negative marking rule had been enforced. Specifically, each student received 2 points for a response that was predetermined by the research team[3] as the "correct" answer, and 1 point was to be subtracted for each response which was predetermined as an "incorrect" response. No points were to be gained or lost for each omitted response.

Each point awarded to a participant equated to a reward of R1. Furthermore, participants were informed that a bonus of R50 would be given to each student who reached a score above 50 points after completion of the test. As such, in this study, the threshold score of 50 points served as the reference point.

Each participant was allocated to one of the four experimental groups. These groups received equal treatment, aside from the points assigned to the students in each group, as a preliminary "base" score. These points were assigned prior to the commencement of the multiple choice

[3] African Micro-Economic Research Unit, University of the Witwatersrand

test, and were unique to each of the four treatment groups. The first and second groups were comprised of students who entered the experiment with 53 points (perceived as a gain) and 47 points (perceived as a loss) respectively. Similarly, the third and fourth groups were comprised of students who entered the experiment with 35 points and 65 points respectively, in which the treatment group entering the experiment with 35 points represented a deviation further into the loss domain, and the group entering with 65 points represented a deviation further into the gain domain.

The range of achievable results for the multiple choice assessment spanned from a minimum 25 points to maximum of 85 points. This encompasses the points that were either lost or gained during the test itself, in addition to the points that the students entered the experiment with as a function of their experimental group. The monetary payoff served as an incentive for students to participate in the experiment, as well as to induce relevant and appreciable consequences, since the payoff each student received was dependent upon their performance in the test. With the aforementioned R1 per point, and a bonus of R50 for reaching 50 points, the payoff limits ranged from a low of R25 to a high of R135[4].

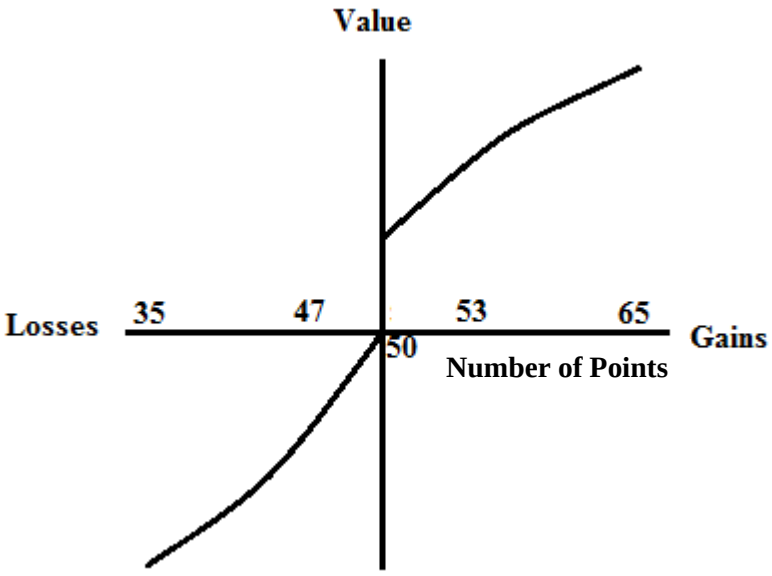


***Figure 2*: A Hypothetical Value Function in the Context of the Experiment**

A diagrammatic representation of a hypothetical valuation function in the context of the design of the experiment is presented in Figure 2 above. The horizontal axis denotes the number of

[4] Calculated as the final result in Rand terms plus an additional R50 for all students who obtain 50 points or more as their final result. Therefore, the highest possible monetary payoff of R135 is achieved by students who receive R85 for a final score of 85 points, plus the additional R50 for obtaining a final score of 50 points or above.

points relative to the 50 point reference, and the vertical axis represents the value obtained by participants which we assume was derived from the monetary payout corresponding to their performance in the multiple choice test. In comparison to the standard S-shaped valuation function proposed by Kahneman & Tversky (1979), which is continuous through the reference point, the valuation function in the context of the experiment conducted in this study is discontinuous at the 50 point reference due to the inclusion of the premium that each student receives for crossing into, or remaining within the gain domain.

Table 1 below documents the percentage of participants with particular observable characteristics after the assignment into each treatment group, inclusive of gender, measured by the percentage of female representation within each group, and the percentage of participants willing to accept a monthly wage of R500, which may thus be representative of financial need.

***Table 1*: Percentage values of observable characteristic for each treatment group**

Variable	Group 1	Group 2	Group 3	Group 4
Gender	46.43	44	45.83	48
Students Willing to Accept a Monthly Wage of R500	53.57	48	54.17	44

The experimental sample consists of 102 participating students who were not allowed to self-select themselves into a particular treatment group. The first two treatment groups in close proximity to the 50 point reference consist of 28 and 25 students respectively; and similarly, the third and fourth treatment groups at the extreme ends of the loss and gain domain were individually comprised of 24 and 25 participating students.

The four treatment groups in which students enter the experiment with differential points are approximately balanced in term of the gender variable and the percentage of participants willing to accept a monthly wage of R500, taking into consideration the relatively small sample sizes within each treatment group. The balance of the observable characteristics in this study supports the proposition that the assignment of students to the four different treatment groups was completely random.

In accordance with the University's ethical policy, participation in this study was absolutely voluntary and participation/non-participation did not affect the students' academic performance in any credit bearing course. Furthermore, the identity of participants has remained completely anonymous in the final analysis of this study since the findings are reported in an aggregate format only. In addition, participating students were made aware of the full range of possible financial payoffs before their decision to participate.

The use of a Randomized Control Trial (RCT) in this study as opposed to Items Response Theory (IRT)[5] is due to the fact that the behavior of students with an average class mark (before they sit for the final examination) above or below 50% cannot necessarily be compared. Such a comparison necessitates that students above and below the 50% pass mark are not systematically different, especially with respect to their content knowledge. Thus, differences in a student's behavior under a test-like situation, as captured by observational data from an actual examination, might be reflective of factors other than the position of the student in the gain or loss domain relative to the 50% pass mark threshold.

The experimental nature of this study enables a causal link to be established between the treatment in terms of the initial amount of points received by each student and the tendency to guess. Firstly, this relationship can be attributed to the exclusion of content knowledge as a covariate – as participants cannot reduce the number of possible solution options through their partial knowledge or obtain the correct answer to the question, since none of the alternative solutions provided are correct; and secondly, due to the random allocation of participating students into the four treatment groups which ensures that no other observable and unobservable characteristics explicates differences in the test behaviour of students. We can therefore assume that none of the four treatment groups consist of participants with particular characteristics that would make them respond differently to the multiple choice test questions; for example, due to a higher innate level of risk aversion or increased content knowledge. Thus, the utilisation of the RCT design for the experiment in this study singularly tests the guessing behaviour of students who enter a test or an examination with different average scores (marks) relative to a reference point, and any differences in the answering behaviour of the students can therefore be attributed to the random allocation into one of these four treatment groups.

[5] IRT provides a model based linkage between the latent characteristics of a test taker and the item responses assessed by a test i.e. it models the response of an examinee of a given ability to each item in a test or examination.

The expected points from guessing are computed using the equation $EP = \frac{1}{C}G + \frac{C-1}{C}L$, in which *G* and *L* represents the number of points gained for a correct response and the number of points lost for an incorrect response respectively; and *C* denotes the number of possible solution options. Since each question is comprised of 3 possible solution options and the student gains 2 points for a correct response, and loses 1 point for an incorrect response, the expected points from a random guess will be zero (*EP* = 0). Furthermore, when the expected score of a guess is zero, rational respondents should always guess, since guessing is at least as good as omitting a response (Bereby-Meyer, Meyer, & Flascher, 2002).

4 Analysis and Discussion

4.1 Treatment Effect

The focus of the analysis is embedded in each participant's decision to answer or omit a response for the nine unsolvable questions in the multiple choice test. A response to each unsolvable question is indicative of an individual's tendency to guess under the negative marking regime, corresponding to the manner in which their choices are framed. The analysis proceeds with a comparison of the first and second treatment groups, which are comprised of students who enter the experiment with 53 and 47 points respectively. The average and median number of responses for the nine unsolvable questions; the percentage of students responding to all or none of the unsolvable items; as well as the percentage of participants willing to respond to a sufficient number of questions, thus enabling them to cross the 50 point threshold in either direction (from the loss domain into the gain domain or vice versa), are documented for the first two treatment groups in Table 2 below.

An immediate observation is that 100% of participants entering the experiment with 47 points respond to two or more questions which facilitates crossing the 50 point threshold. A response to two or more questions are carried out by participants under the assumption that their answers are indeed correct. This in itself, is suggestive of the level of importance attached to traversing from the loss domain into the gain domain. However, 85.17% of participants in the treatment group entering the experiment with 53 points take a gamble of responding to four or more questions and risk falling below the 50 point threshold. The importance of crossing into the gain domain is therefore greater than the aversion or avoidance expressed by participants towards crossing into the loss domain.

***Table 2*: Treatment effects for the first two groups who are simply in the gain and loss domain**

Treatment Group	Average Number of Guesses	Median Number of Guesses	Percentage of Students that Respond to 0 Questions	Percentage of Students that Respond to 9 Questions	Percentage of Students Willing to Respond to a Sufficient Number of Questions to Cross Over the 50 Point Reference in either Direction
Group 1: 53 points	5.39	6	7.14	17.86	85.17
Group 2: 47 points	6.28	7	0	16	100

Prospect theory alludes to the fact that a gain is valued more than an equal loss is disvalued in the loss domain, and a loss is disvalued more than an equal gain is valued in the gain domain. However, it provides no indication of what seems to occur with regard to crossing into either the gain or loss domain. The introduction of the R50 premium assigns greater value by participants towards receiving it, implying that greater importance is attached to traversing from the loss domain into the gain domain, in comparison to how much they disvalue losing it. These dynamics, however, could be altered through the introduction of a higher premium.

The average number of guesses by participants in the gain domain (Group 1) is 5.39 in comparison to an average of 6.28 in the loss domain (Group 2), which advocates that participants in the gain domain display a lower tendency to guess than those in the loss domain. These results are commensurate with prospect theory in which individuals in the gain domain are more risk averse in comparison to those in the loss domain.

An equivalent conclusion arises through an inspection of the median number of guesses which are less sensitive to outliers. Participants in the gain domain have a median of 6 guesses, therefore exhibiting increased risk averse behaviour in comparison to those in the loss domain, with a median of 7 guesses. An analysis of the percentage of students that respond to none of the questions accentuates that participants in the gain domain are risk averse with 7.14% of the 28 participants in Group 1 choosing to respond to none of the questions presented to them. Participants in the loss domain are less risk averse in comparison to those in the gain domain, in which 100% respond to at least one question in the experimental assessment. The percentage of students that respond to all nine unsolvable items are relatively similar for the two treatment groups under consideration, therefore, no inference can be made regarding the difference in the degree of risk aversion expressed by these participants.

Groups 3 and 4 are comprised of students who enter the experiment with 35 and 65 points respectively. The average and median number of guesses; the percentage of students that respond to all or none of the nine unsolvable questions; and the percentage of students willing to respond to a sufficient number of questions to cross the 50 point reference, either from the loss domain into the gain domain or vice versa, are documented for the second two treatment groups in Table 3 below.

Table 3: **Treatment effects for the second two treatment groups with a deviation of points further into the loss and gain domain**

Treatment Group	Average Number of Guesses	Median Number of Guesses	Percentage of Students that Respond to 0 Questions	Percentage of Students that Respond to 9 Questions	Percentage of Students Willing to Respond to a Sufficient Number of Questions to Cross Over the 50 Point Reference in either Direction
Group 3: 35 points	7.42	9	0	54.17	66.67
Group 4: 65 points	6.84	8	0	32	100

Participants in each of the two extreme treatment groups respond to at least one of the nine unsolvable items for a deviation of points further into the gain and loss domain. 54.17% of the 24 participants in the group entering the experiment with 35 points, and 32% of the 25 participants in the group entering the experiment with 65 points respond to all nine questions.

Participants in the treatment group entering the experiment with 35 points need to correctly answer eight or more questions in comparison to two or more questions by participants entering the experiment with 47 points. 66.67% of participants in the treatment group entering the experiment at the extreme end of the loss domain i.e. those commencing the experiment with a preliminary score of 35 points, respond to eight questions or more, therefore, providing an indication of the level importance attached to crossing over the 50 point threshold into the gain domain.

Notwithstanding the fact that participants entering the experiment with 65 points are already in the gain domain, they have an added advantage such that an incorrect response to all nine questions will not result in a fall below the 50 point reference. In this instance, no inference can be made, since 100% of participants not taking a gamble of crossing into the loss domain

will not have any implications on the outcome of interest, due to the impossibility of falling below the 50 point threshold.

In addition to the aforementioned results, the average and median number of guesses suggests that participants entering the experiment with 35 points are more risk seeking in comparison to those entering with 65 points. Both the average and median number of guesses are significantly higher for the group commencing with 65 points.

***Table 4*: Treatment effects for the full sample**

Treatment Group	Average Number of Guesses	Median Number of Guesses	Percentage of Students that Respond to 0 Questions	Percentage of Students that Respond to 9 Questions	Percentage of Students Willing to Respond to a Sufficient Number of Questions to Cross Over the 50 Point Reference in either Direction
Group 1: 53 points	5.39	6	7.14	17.86	85.71
Group 2: 47 points	6.28	7	0	16	100
Group 3: 35 points	7.42	9	0	54.17	66.67
Group 4: 65 points	6.84	8	0	32	100

The findings for the full sample in Table 4 above, which is simply an amalgamation of Tables 2 and 3, advocates that as the number of points a participant enters the experiment with increases in the gain domain, they exhibit an increase in the tendency to guess. This is indicative of decreasingly risk averse behaviour. An inference can therefore be made that higher performing students with a higher class mark on average, before commencing with an examination, exhibit an increased tendency to guess in comparison to their lower performing counterparts. This is in accordance with the standard S-shaped valuation curve proposed by prospect theory. Similarly, as the number of points a student enters the experiment with decreases in the loss domain, participants become more risk seeking or less risk averse. It may be beneficial for these participants to guess more and potentially receive a higher score from each correct response.

4.2 Effect of the treatment on participants who are willing to receive a monthly wage of R500

In each of the four treatment groups, participants were asked whether or not they would be willing to accept a monthly wage of R500. This serves as a proxy for each individual's income status or level of financial need, in which participants willing to accept the R500 monthly wage are considered to be students who are recipients of a lower monthly income. In addition, this question provides an indication of the level of importance attached to the reference point i.e. how much participants' value being in the gain domain over being in the loss domain. Since the payoff in this study is in monetary terms, the financial need argument may be interpreted as the students' response to the level of importance assigned to passing any credit bearing course (remaining above the 50% pass mark). Therefore, students who are required to pass a course in order to proceed to the next level will attach a higher value to the R50 premium and hence, to remaining in the gain domain. For example, students who are registered for an Economics course as a Major, assign greater importance towards passing the course in comparison to those registered for the course as an elective. Table 5 below documents the treatment effects for participants in financial need allocated to the four treatment groups respectively.

The percentage of participants that respond to none of the nine unsolvable questions for the group commencing the experiment with 53 points, demonstrates that some of the students assign a higher value to the R50 premium and therefore, to remaining in the gain domain. Hence, these participants have no incentive to take a risk, and avoid guessing at all costs.

For a deviation of points further into the gain domain, lower income participants exhibit an increase in risk averse behaviour. The number of guesses decreases from an average of 5.47 for participants entering the experiment with 53 points to an average of 5.45 for participants entering the experiment with 65 points. However, despite the difference in the average number of guesses between the two treatment groups under consideration being infinitesimal, the median number of guesses decreases from 6 to 5, justifying the increase in risk averse behaviour. Furthermore, a mere 8% of lower income participants that enter the experiment with 65 points respond to all nine questions, emphasising the increasingly risk averse behaviour demonstrated for a deviation of points further into the gain domain. These results are contrary to the findings of the full sample and the behaviour postulated by the hypothesised valuation function initiated by Kahneman and Tversky (1979).

***Table 5*: Treatment effects for participants who are willing to accept a monthly wage of R500**

Treatment Group	Average Number of Guesses	Median Number of Guesses	Percentage of Students that Respond to 0 Questions	Percentage of Students that Respond to 9 Questions	Percentage of Students Willing to Respond to a Sufficient Number of Questions to Cross Over the 50 Point Reference in either Direction
Group 1: 53 points	5.47	6	13.33	10.71	86.67
Group 2: 47 points	5.5	5	0	8	100
Group 3: 35 points	7.54	9	0	29.17	69.2
Group 4: 65 points	5.45	5	0	8	100

Lower income participants exhibit an increase in risk seeking behaviour for a deviation of points further into the loss domain. This is in accordance to the findings revealed by the full sample of participating students, elucidated through an increase in the average and median number of guesses, and an increase in the percentage of participants who are willing to respond to all nine unsolvable items.

4.3 Effect of the treatment on participants who are NOT willing to receive a monthly wage of R500

The treatment effects for participants who are not willing to receive a monthly wage of R500 are represented in Table 6 below. In this instance, the proxy for income status is indicative of a higher monthly income and potentially a lower value attached to receiving the R50 premium.

Contrary to the findings revealed by lower income participants, all participating students not willing to accept a monthly wage of R500 – amongst those entering the experiment with 53 points, respond to at least one or more questions. The 7.14% of participants responding to none of the questions in the first group under consideration, for the full sample, may therefore be wholly attributed to the behaviour of the recipients of a lower monthly income. Hence, for individuals in financial need who value the R50 premium to a large extent, greater importance is attached to remaining in the gain domain in comparison to those who are not in financial need.

Table 6: **Treatment effects for participants who are not willing to accept a monthly wage of R500**

Treatment Group	Average Number of Guesses	Median Number of Guesses	Percentage of Students that Respond to 0 Questions	Percentage of Students that Respond to 9 Questions	Percentage of Students Willing to Respond to a Sufficient Number of Questions to Cross Over the 50 Point Reference in either Direction
Group 1: 53 points	5.31	5	0	15.38	100
Group 2: 47 points	7	7	0	15.38	100
Group 3: 35 points	7.27	9	0	54.55	63.64
Group 4: 65 points	7.93	8	0	42.86	100

For a deviation of points further into the gain domain, recipients of a higher monthly income exhibit a decrease in risk averse behaviour. These findings are contrary to those exhibited by lower income participants; however, it resonates with the findings of the full sample and the standard S-shaped valuation function proposed by prospect theory. Furthermore, for a deviation of points further into the loss domain, higher income participants become increasingly risk seeking. These findings are concurrent with the higher tendency to guess displayed by both lower income, and the full sample of participating students.

4.4 Gender specific treatment effects

The stratification of each treatment group according to gender permits the analysis of the treatment effects along gender lines. Table 7 below documents the average and median number of guesses among both males and females, as well as the percentage of gender representation in each treatment group, willing to respond to a sufficient number of questions in order to cross above or below the 50 point reference.

The results reveal that in close proximity to the reference i.e. for participants entering the experiment with 53 and 47 points respectively, the difference between both the mean and median number of guesses for male and female participants are marginal and not statistically significant. As the number of points a student enters the experiment with increases in the gain domain, both males and females become decreasingly risk averse. Similarly, as the number of

points a student enters the experiment with decreases in the loss domain, both male and female participants exhibit an increase in risk seeking behavior. The average and median number of guesses for the treatment groups entering the experiment with 35 and 65 points respectively, are higher for males in comparison to females.

It can therefore be deduced that gender specific treatment effects do not prevail in close proximity to the reference point; however, for a deviation of points further into the gain and loss domain, the gender differences in the treatment effects become more distinct. In particular, females display a greater degree of risk aversion in comparison to males for the treatment groups entering the experiment with 35 and 65 points respectively.

Treatment Group	Average Number of Guesses among Females	Median Number of Guesses among Females	Percentage of Females Willing to Respond to a Sufficient Number of Questions to Cross Above or Below the 50 Point Reference	Average Number of Guesses among Males	Median Number of Guesses among Males	Percentage of Males Willing to Respond to a Sufficient Number of Questions to Cross Above or Below the 50 Point Reference
Group 1: 53 points	5.77	6	69.23	5.07	6	66.67
Group 2: 47 points	6.27	6	100	6.29	7	100
Group 3: 35 points	6.82	8	54.55	7.92	9	76.92
Group 4: 65 points	6.25	7.5	100	7.38	8	100

***Table 7*: Gender specific treatment effects**

4.5 Gender specific treatment effects for participants who are willing and not willing to receive a monthly wage of R500

***Table 8*: Gender specific treatment effects for participants willing to accept a monthly wage of R500**

Treatment Group	Percentage of Males Willing to Accept a Monthly Wage of R500 (%)[6]	Average Number of Guesses among Males	Median Number of Guesses among Males	Percentage of Females Willing to Accept a Monthly Wage of R500 (%)	Average Number of Guesses among Females	Median Number of Guesses among Females
Group 1: 53 points	28.57	4.25	4	25	6.86	7
Group 2: 47 points	16	5	5.5	32	5.75	5
Group 3: 35 points	37.5	7.89	9	16.67	6.75	7.5
Group 4: 65 points	16	6	6.5	28	5.14	4

Table 8 above presents the percentage of both male and female participants who are willing to accept a monthly wage of R500 within each treatment group and both the average and the median number of guesses amongst them. For participating students who enter the experiment with 53 and 47 points respectively, the average number of guesses suggests that males display a lower tendency to guess in comparison to females and are therefore more risk averse. However, the median number of guesses for participants simply in the loss domain proposes that females are more risk averse in comparison to males. Thus, caution needs to be exerted when interpreting these results due to the smaller sample sizes within each treatment group.

For a deviation of points further into the gain domain, lower income males exhibit decreasingly risk averse behaviour in comparison to the increasingly risk averse behaviour demonstrated by

[6] The sample sizes on which the average and median number of guesses are being evaluated are getting smaller, therefore, caution needs to be exerted regarding the interpretation of these results.

lower income females. A deviation of points further into the loss domain results in both male and female participants becoming increasingly risk seeking. At the two extreme frames in which participants enter the experiment with 35 and 65 points respectively, lower income females exhibit a greater degree of risk aversion in comparison to their male counterparts.

The percentage of male and female participants who are not willing to accept a monthly wage of R500, and the average and median number of guesses among them is presented in Table 9 below. In close proximity to the 50 point reference, it is evident that females exhibit a greater degree of risk aversion in comparison to males. For a deviation of points further into the gain domain, both males and females exhibit increasingly risk averse behaviour. However, a deviation of points further into the loss domain, results in male participants displaying increasingly risk seeking behaviour, and females, decreasingly risk seeking behaviour. For participants entering the experiment with 35 and 65 points respectively, females who are not willing to receive a monthly wage of R500 are more risk averse than their male counterparts; however, when the analysis is based on the median number of guesses, the reverse holds for the group that enters the experiment with 35 points. Yet again, caution needs to be exerted when interpreting these findings.

***Table 9*: Gender specific treatment effects for participants who are not willing to receive a monthly wage of R500**

Treatment Group	Percentage of Males Not Willing to Accept a Monthly Wage of R500 (%)	Average Number of Guesses among Males	Median Number of Guesses among Males	Percentage of Females Not Willing to Accept a Monthly Wage of R500 (%)	Average Number of Guesses among Females	Median Number of Guesses among Females
Group 1: 53 points	25	6	6	21.43	4.5	3
Group 2: 47 points	40	6.8	7	12	7.67	9
Group 3: 35 points	16.67	8	8.5	29.17	6.86	9
Group 4: 65 points	36	8	9	20	7.8	8

5 Regression Analysis

5.1 Methodology

The objective of the regression analysis is to assess whether or not the assignment of participants into the four treatment groups produces a statistically discernable difference in the guessing behaviour of students, particularly; when the influence of other covariates are held constant. These covariates are held constant in order to control for possible imbalances in observable characteristics. Given that the experiment in this study makes use of a Randomized Control Trial, the analysis will be conducted using an Ordinary Least Squares (OLS) approach. The average number of guesses for the nine unsolvable questions in the multiple choice assessment will serve as the dependent variable (*Y)* in one instance, and the median number of guesses in another. Explanatory variables are comprised of a dummy variable for the different treatment groups (*T*) which enter the experiment with differential points, and a vector (*X)* inclusive of characteristics such as gender and participants willingness to accept a monthly wage of R500. The treatment group that enters the experiment with 53 points will serve as the baseline group to which the remaining three groups will be compared, such that:

$$Y = \beta_0 + \beta_1 T + \beta_2 X + \varepsilon \quad (1)$$

where ε represents the disturbance term.

The regression represented in equation 1 correspondingly permits the analysis of possible treatment heterogeneities along the lines of gender and income status, through an evaluation of interaction effects, utilising dummy variables, such that:

$$Y = \beta_0 + \beta_1 T + \beta_2 T * D + \varepsilon \quad (2)$$

where *T*D* is the interactive term between the treatment and a dummy variable representing either gender or a participant's willingness to accept a monthly wage of R500.

5.2 Regression Results

The regression analysis proceeds by adopting a simple Ordinary Least Squares (OLS) approach for the full sample of participating students. The empirical results for regression with the average number of guesses as the dependent variable are presented in column 1 of Table 10 below, with the robust standard errors in parenthesis. For the group that enters the experiment with 53 points serving as the baseline for comparison purposes, it is evident that the treatment

groups entering the experiment with 47, 35 and 65 points respectively are willing to guess more on average than the group of participants simply in the gain domain. The effect of the treatment in which participants enter the experiment with 47 points on the number of guesses is insignificant, even at a 10% level. However, the effect of the groups that enter with 35 and 65 points are significant at a 1% and 10% level respectively. Participants in Group 3 have an average of 2.02 more guesses in comparison to those in Group 1. Similarly, participants in Group 4 have an average of 1.37 more guesses than those in close proximity to the reference in the gain domain. The gender effects reveal that females display a lower tendency to guess in comparison to male participants; however, the effect is insignificant.

Table 10: **Regression results**

	OLS	**OLS**
Variables	**Number of Guesses (1)**	**Number of Guesses (2)**
Treatment (Group 2): 47 Points	0.880 (0.697)	1.000 (1.669)
Treatment (Group3): 35 Points	2.022*** (0.708)	3.000** (1.458)
Treatment (Group 4): 65 Points	1.372* (0.733)	2.000 (1.623)
Gender	-0.312 (0.493)	-6.90e-18 (1.137)
Constant	5.538*** (0.621)	6.000*** (1.247)
Observations (N)	102	102

*** Significance at 1% level **Significance at 5% level *Significance at 10% level

Robust standard errors in parenthesis

Similarly, column 2 represents the quantile regression coefficients which estimates the change in the median number of guesses. The effects of the treatment groups that enter the experiment

with 47 and 65 points on the median number of guesses are insignificant; however, the effect of the third treatment group entering the experiment with 35 points is significant at a 5% level.

Table 11 documents the mean regression results of the possible treatment heterogeneities evaluated in this study with the number of guesses per student serving as the dependent variable. Columns 1 and 2 represent the gender effects for participants who are willing and not willing to accept a monthly wage of R500 respectively. For participants who are willing to accept a monthly wage of R500, the effects of the treatment groups that enter the experiment with 47, 35 and 65 points results in a higher tendency to guess in comparison to the group that enters the experiment with 53 points, with the effects of the third group entering the experiment with 35 points being significant at a 5% level, and having 3.64 more guesses on average than the comparison group. Close inspection of the interaction effects with male participants in the first treatment group serving as the comparison, indicates that the gender effects for participants entering the experiment with 35 points, are significant at a 10% level, whereby females have an average of 3.75 guesses less than male participants. Similarly, the results for participants who are not willing to accept a monthly wage of R500 advocates that the average treatment effects for the two groups entering with 35 and 65 points have a 10% level of significance, and both these groups have an average of 2 more guesses compared to the group that enters the experiment in close proximity to the reference in the gain domain. The interaction effects however, are insignificant across the three treatment groups under consideration.

The results of the income effects for female and male participants are presented in columns 3 and 4 respectively. The average treatment effects for female participants entering the experiment with the perception of being in the loss domain are insignificant. However, the treatment effect for the participating group entering the experiment with 65 points is significant at a 5% level, and these participants have an average of 2.03 more guesses compared to those entering the experiment with 53 points. An evaluation of the interaction effects reveals that the income effect for female participants that enter the experiment with 65 points is significant at a 10% level, and that female participants who are willing to accept a monthly wage of R500 within this treatment group have an average of 2.66 less guesses in comparison to participants who are not willing to accept a monthly remuneration of R500 for the group entering the experiment with 53 points.

***Table 11*: Effects of possible treatment heterogeneities**

	OLS	OLS	OLS	OLS
Variables	**Number of Guesses (1)**	**Number of Guesses (2)**	**Number of Guesses (3)**	**Number of Guesses (4)**
Treatment (Group 2): 47 Points	0.750 (1.772)	0.800 (0.980)	1.897 (1.382)	1.733* (0.947)
Treatment (Group3): 35 Points	3.639** (1.503)	2.000* (1.104)	1.088 (1.299)	2.933*** (1.071)
Treatment (Group 4): 65 Points	1.250 (1.664)	2.000* (1.047)	2.031** (0.862)	2.933*** (1.014)
Gender	2.607* (1.405)	-1.500 (1.551)		
Group 2 * Gender	-1.857 (2.046)	2.367 (1.999)		
Group 3 * Gender	-3.746 (1.991)	0.357 (2.008)		
Group 4 * Gender	-2.964 (2.111)	1.300 (1.719)		
Group 2 * R500 Wage			-1.917 (1.505)	-1.800 (1.212)
Group 3 * R500 Wage			-0.107 (1.642)	-0.111 (0.931)
Group 4* R500 Wage			-2.657* (1.316)	-2.500** (1.115)
Constant	4.250*** (1.344)	6.000*** (0.880)	5.769*** (0.719)	5.067*** (0.847)
Observations (N)	51	51	47	55

*** Significance at 1% level **Significance at 5%level *Significance at 10% level

Robust standard errors in parenthesis

Similarly, the average treatment effects for male participants entering the experiment with 35 and 65 points are significant at a 1% level, in which both the treatment groups have an average of 2.93 more guesses compared to the group that enters with 53 points. In addition, the average treatment effect for male participants entering the experiment with 47 points is significant at a 10% level, with an average of 1.73 more guesses than the comparison group. The interactive term demonstrates that the income effect for male participants entering the experiment with 65 points is significant at a 5% level, with a decrease in the average number of guesses for participants willing to accept a monthly wage of R500.

6 Conclusion

Behavioural decision analyses provides a valuable contribution towards the understanding of educational assessments by taking into consideration the decisions faced by students during a test or an examination. The experimental nature of this study replicates an actual test situation whereby students enter a multiple choice test with a certain aggregate based upon their average class marks, and permits the analysis of the framing effect on a student's tendency to guess under the adoption of a negative marking rule. This study eschews the mainstream approach of expected utility theory in the analysis of decision making under risk, and advocates that the framing effect of the scoring rule as a potential gain or loss relative to a reference point can be elucidated using prospect theory. In addition to evaluating whether the behaviour of students under such circumstances are consistent with the behaviour predicted by prospect theory, this study investigates whether the distance into the gain and loss domain yields significantly different results, and enables a causal relationship to be established between the different treatment groups and the tendency to guess.

Overall, the results for the full sample under consideration; for individuals who are willing and not willing to accept a monthly wage of R500; the results across both genders; as well as the findings of the gender specific treatment effects for the recipients of a lower monthly income, advocates that the lowest degree of risk aversion or the greatest degree of risk seeking occurs for participants entering the experiment with 35 points i.e. for students with the lowest aggregate marks before commencing with an examination. These results are in accordance with prospect theory from the perspective of being in the loss domain; however, they are in conflict with prospect theory from the viewpoint that the greatest degree of risk seeking occurs at the extreme end of the loss domain.

A deviation of points further into the gain domain generally results in decreasingly risk averse behaviour, with the exception of lower income participants who are willing to accept a monthly wage of R500 exhibiting an increase in risk aversion. Therefore, while higher income individuals become less risk averse in the gain domain, those in financial need become increasingly risk averse with a deviation of points further into the gain domain. This increase in risk aversion is particularly demonstrated by female participants when analysing the gender effects of these lower income participants. A potential limitation, especially for the gender specific effects of participants who are willing and not willing to receive a monthly wage of R500, is that the sample sizes are extremely small which could result in a small sample bias; therefore, caution needs to be exerted when interpreting these results. Perhaps a larger sample size for the subgroups under consideration in this study will give rise to varying results.

Lastly, the stratification of each treatment group according to gender permits the analysis of how the tendency to guess for both males and females, is affected by the way in which the scoring rule is framed. For the two treatment groups entering the experiment with 35 and 65 points respectively, the results reveal that female participants exhibit a greater degree of risk aversion in comparison to males.

7 References

Aquino, K., Steisel, V., & Kay, A. (1992). The Effects of Resource Distribution, Voice, and Decision Framing on the Provision of Public Goods. *The Journal of Conflict Resolution, 36*(4), 665-687.

Becker, W. E., & Johnston, C. (1999). The relationship between multiple choice and essay response questions in assessing economic understanding. *Economic Record, 75*(4), 348-357.

Bereby-Meyer, Y., Meyer, J., & Flascher, O. M. (2002). Prospect Theory Analysis of Guessing in Multiple Choice Tests. *Journal of Behavioral Decision Making*, 313-327.

Bliss, L. B. (1980). A Test of Lord's Assumption regarding Examinee Guessing Behaviour on Multiple-Choice Tests using Elementary School Students. *Journal of Educational Measurement, 17*(2), 147-153.

Buckles, S., & Siegfried, J. J. (2006). Using Multiple-Choice Questions to Evaluate In-Depth Learning of Economics. *The Journal of Economic Education*, 48-57.

Budescu, D., & Bar-Hillel, M. (1993). To Guess or Not to Guess: A Decision-Theoretic View of Formula Scoring. *Journal of Educational Measurement, 30*(4), 277-291.

Bush, M. (2001). A Multiple Choice Test that Rewards Partial Knowledge. *Journal of Further and Higher Education, 25*(2), 157-163.

Davis, F. B. (1967). A Note on the Correction for Chance Success. *The Journal of Experimental Education, 35*(3), 42-47.

Eckel, C. C., & Grossman, P. J. (2008). Forecasting Risk Attitudes: An Experimental Study of Actual and Forecast Risk Attitudes of Women and Men. *Journal of Economic Behaviour and Organisation, 68*, 1-17.

Espinosa, M. P., & Gardeazabal, J. (2010). Optimal Correction for Guessing in Multiple-Choice Tests. *Journal of Mathematical Psychology, 54*, 415-425.

Hartford, J., & Spearman, N. (2014). Who's Afraid of the Big Bad Wolf? Risk Aversion and Gender Discrimination in Assessment. *ERSA - Working Paper*(418).

Holt, A. (2006). An analysis of negative marking in multiple-choice assessment. *Proceedings of 19th Annual Conference of the National Advisory Committee on Computing Qualifications*, (pp. 115-118).

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Anlysis of Decisions under Risk. *Econometrica, 47*(2), 263-291.

Kniveton, B. H. (1996). A correlational analysis of multiple-choice and essay assessment measures. *Research in Education - Manchester*, 73-84.

Krawczyk, M. (2011). Framing in the Field: A simple experiment on the reflection effect. *Working paper: University of Warsaw*.

Kuhberger, A. (1998). The Influence of Framing on Risky Decisions: A Meta-Analysis. *Organizational Behavior and Human Decision Processes, 75*(1), 23-55.

Lord, F. M. (1975). Formula Scoring and Number Right Scoring. *Journal of Educational Measurement, 12*(1), 7-11.

Marin, C., & Rosa-Garcia, A. (2011). Gender bias in risk aversion: evidence from multiple choice exams. *Working paper 39987, MRPA*.

Slakter, M. J. (1969). Generality of risk taking on objective examinations. *Educational and Psuchological Measurement, 29*(1), 115-128.

Tversky, A., & Kahneman, D. (1981). The Framing of Decisions and the Psychology of Choice. *Science, New Series, 211*(4481), 453-458.

Walstad, W., & Becker, W. (1994). Achievement Differences on Multiple-Choice and Essay Tests in Economics. *American Economic Review, 84*(2), 193-196.

A Supplementary Appendix

Test Questionnaire:

Student number: _____

Gender: _____

As a student, would you be willing to accept a job that pays a monthly wage of R500?

Yes _____ No _____

Multiple Choice Questions:

1. A group of modern economists who believe that markets clear very rapidly and that expanding the money supply will always increase prices rather than employment are the:

A. Monetarists
B. New Classical School
C. Keynesians

2. What is the next term in the following sequence?

1, 1, 2, 3, 19, 34, 83, ...

A. 162
B. 115
C. 247

3. The Independent Labour and Employment Equity Action Plan was drafted by Mbazima Sithole in which year?

A. 1996
B. 1994
C. 1999

4. Who is the author of the book titled: “Random Walks and Business Cycles for Dummies”?

A. Norman Gladwell
B. Milton Savage
C. John Friedman

5. The Law of Diminishing Demand states that:

A. As more of a particular good is demanded by the economy, less of that good is demanded by an individual.
B. If good A is preferred to good B, then a higher demand for good B implies a lower demand for good A.
C. As more of a good is supplied in an economy, the less of that good is demanded by the economy.

6. The Depression of 1978 occurred as a result of:

A. Severe drought affecting subsistence agriculture and herding.
B. A banking panic which came about as a result of depositors simultaneously losing confidence in the solvency of the banks and demanding that their deposits be paid to them in cash.
C. A decline in the population growth rate.

7. A Pareto Supremum refers to the allocation of resources in which:

A. All resources are directed to a single individual and no one can be made better off.
B. It is possible to make all individuals better off.
C. A socially desirable distribution is acquired through all individuals having a higher income.

8. The principle of Malthusian Dominance states that:

A. Gains in income per person through technological advances dominates subsequent population growth.
B. An increase in the market price caused by an increase in demand dominates the higher price caused by a deficiency in supply.
C. Increased demand for subsistence consumption eliminates the non-productive elements of the economy.

9. The Population Poverty Index estimates:

A. The percentage of the population living in poor regions.
B. The number of people earning below $1 a day.
C. The average worldwide population living below the poverty line.

10. A Walrasian Balanced Growth Path refers to:

A. The act in which excess market supply counteracts excess market demand.
B. A situation in which output per worker, capital per worker and consumption per worker are growing at a constant rate.
C. An efficient allocation of goods and services in an economy, driven by seemingly separate decisions of individuals.

Solutions:

1. B
2. C
3. C
4. A
5. B
6. A
7. C
8. A
9. B
10. C

Group 1 12 August 2014

Good Day

Welcome to this decision making experiment. My name is Jesal Kika and I am currently an Economic Science Masters student at the University of the Witwatersrand, Johannesburg. Before proceeding to the test questionnaire, please take note of the experimental instructions below. At the beginning of the test, you already have 53 points to start off with and you are now being placed in a test situation in which you may gain or lose points in addition to the 53 points. These additional points may be gained or lost through a multiple choice test with the following rules:

- You are required to answer a multiple choice test consisting of 10 questions in total.
- You will receive 2 points for each correct response; lose 1 point for each incorrect response; and no points will be gained or lost for each question that you choose to omit.
- Your final amount of points will be calculated as 53 plus the number of points you obtain in the test.
- The payoff you receive will be on a rand (R1) per point basis i.e. you will receive R1 for each of your final amount of points. In addition, you will receive a bonus of R50 if your final score is above 50 points on completion of the test.
- For example: If you receive 8 points for the test, your final amount of points will be 61 (53+8). In this instance, you will receive R61 + R50 bonus since your final score is above 50 points. Therefore, your final payout will be R111.
- If however, you receive -6 points (lose 6 points) for example, your final amount of points will be 47 (53-6). In this instance, you will receive a payout of R47 (you will NOT receive a bonus of R50 because your final score is below 50 points).
- Note that your total payoff can vary between R43 and R123.
- You have 20 minutes to complete the test.

Please note that your participation in this experiment is completely voluntary, involves no risk and will not affect your academic results in any way. Your answers to these questions are completely confidential and your identity will remain anonymous in the analysis of this study. If you have any questions regarding the instructions above, please feel free to ask. Should you wish to withdraw from this experiment, you may do so at any stage. Thank you for your consideration to participate in this experiment. Should you wish to enquire about my study or access my final results, please feel free to contact me at jesalk@hotmail.com. You may now proceed to the test questionnaire.

Kind regards
Jesal C Kika
Masters Student (Economic Science)
School of Economic and Business Sciences
University of the Witwatersrand, Johannesburg

Group 2 12 August 2014

Good Day

Welcome to this decision making experiment. My name is Jesal Kika and I am currently an Economic Science Masters student at the University of the Witwatersrand, Johannesburg. Before proceeding to the test questionnaire, please take note of the experimental instructions below. At the beginning of the test, you already have 47 points to start off with and you are now being placed in a test situation in which you may gain or lose points in addition to the 47 points. These additional points may be gained or lost through a multiple choice test with the following rules:

- You are required to answer a multiple choice test consisting of 10 questions in total.
- You will receive 2 points for each correct response; lose 1 point for each incorrect response; and no points will be gained or lost for each question that you choose to omit.
- Your final amount of points will be calculated as 47 plus the number of points you obtain in the test.
- The payoff you receive will be on a rand (R1) per point basis i.e. you will receive R1 for each of your final amount of points. In addition, you will receive a bonus of R50 if your final score is above 50 points on completion of the test.
- For example: If you receive 8 points for the test, your final amount of points will be 55 (47+8). In this instance, you will receive R55 + R50 bonus since your final score is above 50 points. Therefore, your final payout will be R105.
- If however, you receive -6 points (lose 6 points) for example, your final amount of points will be 41 (47-6). In this instance, you will receive a payout of R41 (you will NOT receive a bonus of R50 because your final score is below 50 points).
- Note that your total payoff can vary between R37 and R117.
- You have 20 minutes to complete the test.

Please note that your participation in this experiment is completely voluntary, involves no risk and will not affect your academic results in any way. Your answers to these questions are completely confidential and your identity will remain anonymous in the analysis of this study. If you have any questions regarding the instructions above, please feel free to ask. Should you wish to withdraw from this experiment, you may do so at any stage. Thank you for your consideration to participate in this experiment. Should you wish to enquire about my study or access my final results, please feel free to contact me at jesalk@hotmail.com. You may now proceed to the test questionnaire.

Kind regards
Jesal C Kika
Masters Student (Economic Science)
School of Economic and Business Sciences
University of the Witwatersrand, Johannesburg

Group 3

Good Day 12 August 2014

Welcome to this decision making experiment. My name is Jesal Kika and I am currently an Economic Science Masters student at the University of the Witwatersrand, Johannesburg. Before proceeding to the test questionnaire, please take note of the experimental instructions below. At the beginning of the test, you already have 35 points to start off with and you are now being placed in a test situation in which you may gain or lose points in addition to the 35 points. These additional points may be gained or lost through a multiple choice test with the following rules:

- You are required to answer a multiple choice test consisting of 10 questions in total.
- You will receive 2 points for each correct response; lose 1 point for each incorrect response; and no points will be gained or lost for each question that you choose to omit.
- Your final amount of points will be calculated as 35 plus the number of points you obtain in the test.
- The payoff you receive will be on a rand (R1) per point basis i.e. you will receive R1 for each of your final amount of points. In addition, you will receive a bonus of R50 if your final score is above 50 points on completion of the test.
- For example: If you receive 18 points for the test, your final amount of points will be 53 (35+18). In this instance, you will receive R53 + R50 bonus since your final score is above 50 points. Therefore, your final payout will be R103.
- If however, you receive -5 points (lose 5 points) for example, your final amount of points will be 30 (35-5). In this instance, you will receive a payout of R30 (you will NOT receive a bonus of R50 because your final score is below 50 points).
- Note that your total payoff can vary between R25 and R105.
- You have 20 minutes to complete the test.

Please note that your participation in this experiment is completely voluntary, involves no risk and will not affect your academic results in any way. Your answers to these questions are completely confidential and your identity will remain anonymous in the analysis of this study. If you have any questions regarding the instructions above, please feel free to ask. Should you wish to withdraw from this experiment, you may do so at any stage. Thank you for your consideration to participate in this experiment. Should you wish to enquire about my study or access my final results, please feel free to contact me at jesalk@hotmail.com. You may now proceed to the test questionnaire.

Kind regards
Jesal C Kika
Masters Student (Economic Science)
School of Economic and Business Sciences
University of the Witwatersrand, Johannesburg

Group 4

Good Day 12 August 2014

Welcome to this decision making experiment. My name is Jesal Kika and I am currently an Economic Science Masters student at the University of the Witwatersrand, Johannesburg. Before proceeding to the test questionnaire, please take note of the experimental instructions below. At the beginning of the test, you already have 65 points to start off with and you are now being placed in a test situation in which you may gain or lose points in addition to the 65 points. These additional points may be gained or lost through a multiple choice test with the following rules:

- You are required to answer a multiple choice test consisting of 10 questions in total.
- You will receive 2 points for each correct response; lose 1 point for each incorrect response; and no points will be gained or lost for each question that you choose to omit.
- Your final amount of points will be calculated as 65 plus the number of points you obtain in the test.
- The payoff you receive will be on a rand (R1) per point basis i.e. you will receive R1 for each of your final amount of points. In addition, you will receive R50 for participating in this test.
- For example: If you receive 10 points for the test, your final amount of points will be 75 (65+10). Therefore, your final payout will be R125 (R75 + R50).
- If however, you receive -10 points (lose 10 points) for example, your final amount of points will be 55 (65-10). In this instance, you will receive a payout of R105 (R55 + R50).
- Note that your total payoff can vary between R105 and R145.
- You have 20 minutes to complete the test.

Please note that your participation in this experiment is completely voluntary, involves no risk and will not affect your academic results in any way. Your answers to these questions are completely confidential and your identity will remain anonymous in the analysis of this study. If you have any questions regarding the instructions above, please feel free to ask. Should you wish to withdraw from this experiment, you may do so at any stage. Thank you for your consideration to participate in this experiment. Should you wish to enquire about my study or access my final results, please feel free to contact me at jesalk@hotmail.com. You may now proceed to the test questionnaire.

Kind regards
Jesal C Kika
Masters Student (Economic Science)
School of Economic and Business Sciences
University of the Witwatersrand, Johannesburg