LANCASTER UNIVERSITY

DOCTORAL THESIS

Essays in Decision Making Under Risk and Uncertainty

Author:
Harrison Rolls

Supervisors:
Konstantinos Georgalos
Alexander Matros
David Rietzke

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy

Febuary 2023

Declaration of Authorship

I, Harrison Rolls, declare that this thesis titled, "Essays in Decision Making Under Risk and Uncertainty" and the work presented in it are my own. I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this or any other university. I confirm that Chapter 2 is jointly co-authored with Dr. Konstantinos Georgalos and Jessica Alam has been published in a journal.

Acknowledgements

First and foremost, the biggest thank you possible to my supervisors, Konstantinos Georgalos, David Rietzke and Alexander Matros, for their support in this, I have learnt so much about economics, formulating ideas and presenting from the three of you. A big thank you, especially to Konstantinos, for his guidance, support and patience with an endless amount of questions I promise that there will be fewer of them from here on out.

I am also grateful to everyone involved in the Lancaster University Economics Department, the lecturers, support staff, and everyone involved in my studies over the last 9 years for their support and encouragement with my projects. A special mention to Kim Kaivanto for many interesting ideas and encouragement regarding getting support with my dyslexia, which has made the remaining part of my studies much more manageable.

A big thank you/apology to my partner Nicole who has had to listen to and learn more about risky decision-making than anybody not involved in economics should ever have to. I am also grateful to everyone I consider my family and friends for their incredible support and for knowing when to ask about my progress and when not to ask.

A special thank you to my cats McVities, Clancy, Orion, Jess, Lily, Mr. Geoff and Mittens for being very cute, letting me pet them, and providing a lot of stress relief and distraction.

Contents

List of Figures

List of Tables

Introduction

This thesis contains three chapters focusing on risky decision making. This is an important field of study as most of the outcomes to decisions we make in our life have uncertain outcomes where the results are often unclear. Two of these focus on decision time, its role in the decision-making process, how a lack of it can influence a decision and how time can be used to help understand preferences. The third of these papers consider gender effects and how we can recover different risk attitudes for males and females.

In the first chapter, I introduce the Drift Diffusion Model, a common model used in psychology that can use decision time as additional information about an individual's preferences. I modify this model to consider Prospect Theory and then compare this modified model to a logit model which is commonly used within the economics literature but cannot take decision time information into account. We first estimate both models and then compare their ability to predict choices that individuals make. We find that whilst both models are good at predicting choices in a binary choice setting, the drift-diffusion model can outperform the standard model in some settings.

The second chapter re-analyses data that previously yielded inconclusive results on gender differences in risk attitudes between males and females. We use a model that allows for heterogeneity between individuals, where males and females risk attitudes are drawn from two separate distributions as opposed to one value for males and one for females. We take data from three experiments where gender effects were not found and, using this method, we find gender effects. The second chapter was co-authored with Konstantinos Georgalos and Jessica Alam and published in the Journal of Economic Behavior and Organization, 202, 2022, 168-183.

The third and final chapter focuses on the impact of time pressure on an individual's attitude to risk. We first propose and then carry out an experiment to investigate this. Unlike other papers, we consider a budget allocation experiment instead of a binary choice experiment. We then structurally recover parameters of interest and find that individuals exhibit less risk-averse behaviour in the presence of time pressure.

Chapter 1

Testing the Drift-Diffusion Model under Prospect Theory

1.1 Introduction

Decision-making is an involved and time-consuming process. We observe that decisions do not happen instantaneously in the real world, as decision-makers need time to consider options before deciding. The time taken to make decisions carries information about an individual's choices and preferences. For example, should an individual decide quickly, we can argue that there is a stronger preference for that option than the other. However, as the time to make a decision increases, we can infer that the options are of closer value and the decision is more challenging. This paper considers a model that allows us to model this relationship. We use a model that will enable us to utilise both time and choice data simultaneously, unlike more 'conventional' models, which purely consider an individual's choice. We focus on applying this framework in a risky binary choice setting and incorporate Prospect Theory (Tversky and Kahneman, 1992), which has not been considered to our knowledge.

This idea can be motivated by an example highlighting a potential limitation within risky decision-making and how considering decision time can help alleviate it. Consider a hypothetical choice task where two individuals have to choose between two 50-50 gambles, one with payoffs of £0 and £4 and the other £1 and £3. Both gambles have an identical expected yield of £2; however, the first can be considered riskier due to an increased payoff variance. If we also suppose their preferences are expressed by a utility function $U(x) = \frac{x^\rho}{\rho}$ and our two individuals' values for ρ[1] are 0.9 and 0.7. The expected utilities for the risky gamble are 1.93 and 1.89, respectively, and for the safer gamble, 2.05 and 2.26. From this, we know that both individuals would be more likely to choose the safer option with less noise, giving both individuals higher expected utilities of 2.05 and 2.26. However, if we were only to consider their choices, we would be unable to tell individuals apart, despite their different attitudes towards risk. We could argue that the more risk-averse individual ($\rho = 0.7$) is more averse to the risky gamble and values it lower than the less risk-averse individual, as shown by the larger difference in utility. As a result, we hypothesise that the more risk-averse individual would be able to make the decision easier and faster than an individual closer to risk neutrality. If we had this time data too, and the above assumption was correct, we could tell these two individuals apart, despite their identical choices. We could say that whichever individual makes the decision first is more risk averse due to the larger difference in the valuation of both options making the decision easier. The validity of this argument is supported by Clithero (2018a), which strongly suggests that decision time data can make a good substitute for extra choice data and improve prediction accuracy in choice tasks where options are more closely valued.

Decision times have been studied extensively, and a few key findings have been established[2]. Familiarity with tasks and repeatedly performing them has been shown to decrease reaction times. A larger difference in the valuation of options is associated with the higher probability of a higher-valued option being chosen in a relatively shorter period of time. Requiring more information to make a decision increases the decision time and lowers the chance of an error. This time can also be helpful if discerning if an individual has made an error, if decision time is vastly different to what you'd expect an individual to take, for example, taking 2 seconds or 3 minutes to make a decision you would expect to take 30 seconds may suggest an error has been made.

The Drift-Diffusion Model (Ratcliff, 1978) models an individual's decision-making

[1] With this utility function ρ indicates risk attitudes, if ρ is between 0 and 1, they are risk averse and above 1 they are risk seeking.

[2] See Clithero (2018b) for an in-depth review of the reaction time, where the following results can be found.

process and allows us to simultaneously consider an individual's choice and decision time. The model assumes that a noisy accumulation process drives the decision-making process. Once sufficient information has been received in favour of an option, the process stops, and the individual decides. The larger the drift rate (the rate at which an individual receives information in favour of one option compared to the other), the more information there is in favour of a given option and the faster an individual's accumulation process "drifts" towards that option. However, the smaller the difference in option values, the lower this drift rate will be and the longer it will take to make a decision[3].

Incorporating decision time has several advantages for researchers. This includes carrying information about task difficulty and how closely valued options are. Should the decision take a long time, we know that options are more closely valued, whereas a shorter reaction time carries information that an option is more heavily preferred. Combining choice and reaction time data leads to analysis being made using more data, which has several key advantages, for example, better quality estimates of an individual's preferences (Alós-Ferrer, Fehr, and Netzer, 2021). It can also reduce the number of tasks subjects need to carry out for experiments, alleviating issues such as decision fatigue while maintaining accurate estimates.

This paper considers the Drift-Diffusion model and its possible applications in risky decision-making. To do this, we must modify the existing model to include the parameters commonly associated with rank-dependent utility (Tversky and Kahneman, 1992) and risky decision-making: the curvature of a utility function and a probability weighting function. Following this, we conduct a series of increasingly complex simulation exercises to see if estimating these models is possible and if researchers can recover the parameters of interest successfully. We finally move on to actual experimental data, estimating this model and comparing its performance to a commonly used stochastic logit specification. This allows us to see how well our model handles real-world data and whether it is robust enough to use economically. We aim to provide an argument that decision time, despite being easily observable in experiments, is under-utilised, and models that take it into account should be considered when possible.

[3] This is a very brief overview of the model, which we will cover more in-depth in later sections.

1.2 Literature Review

The idea of decision-making processes has been long explored in psychology. The dual-process model (Epstein, 1994) is one of the most common theories, where an individual either makes a quick, snap decision or a more analytical decision which takes time. In economics, we assume decisions are made instantaneous and until recently, there has been little economic interest in the process by which a human makes a decision. Examples of process data include eye-tracking (Fiedler and Glöckner, 2012), neural activity (Rich, Stoll, and Rudebeck, 2018) and time (Clithero, 2018a), Fehr and Rangel (2011) provide a good summary of some of the advancements made in this field. A model that has been widely researched, however, is the Drift-Diffusion model. The model has been well established within the psychology literature for many years and continues to be used. The model was first proposed in the 1970's (Ratcliff, 1978). Since then, it has evolved into a single drifting process representation (Ratcliff, 1981) instead of a multiple racing drifts. After this, it evolved to include variability in the model's starting point. And finally, to account for non-decision time, time in the decision-making process does not result in a drift, such as understanding a particular task. Another more recent addition to the model is a diminishing bound, where the individual requires less information to decide as time progresses. Another common form of the model is the Hierarchical Bayesian Drift-Diffusion model (Wiecki, Sofer, and Frank, 2013), which can, in some cases, be easier to estimate.

The model has been used to explain the decision-making process of individuals. However, we differ from the exisiting literature by combining the model with Prospect Theory (Tversky and Kahneman, 1992). Typically experiments with this model involve binary choices, choosing between two different options based on an individual's preferences or between a correct and incorrect choice. Experiments with correct and incorrect decisions include deciding if there are more dots on a screen moving in which direction (Ratcliff and McKoon, 2008); brightness tasks, where individuals determine whether a screen gets brighter or darker (Ratcliff and Van Dongen, 2011), or memory, seeing if a word is on a given list (Ratcliff, 1978). Most research into this model to consider preferences involves food, choosing between two different foods or between a food and a reference food, (Clithero, 2018a), and can potentially explain violations in the transitivity of these choices (Baldassi et al., 2020). There has also been some interest in considering eye tracking data (Fisher, 2017), (Krajbich and Rangel, 2011), to estimate how the drift rate of the decision-making process changes as an individual focuses on a particular option. Another similar line of research involves time pressure and how the parameters of the Drift-Diffusion model change in the presence of time pressure (Milosavljevic et al., 2010).

The relationship between risky decision-making and the decision time has been shown that decisions under gains are quicker than those under losses (He et al., 2013). There is considerable literature looking at the speed-accuracy trade-off, (Wickelgren, 1977). We can see that faster reaction times can potentially lead to more errors (Edman, Schalling, and Levander, 1983), (Hale, 1968). See Heitz (2014) for a summary of this field of literature. However, Fudenberg, Strack, and Strzalecki (2018) show that shorter reaction times may lead to more accuracy if an individual's beliefs are correct. Faster reaction times can also be suggestions of biases towards a particular option, which can also be estimated by the Drift-Diffusion model (Criss, 2010). Other factors have been shown to influence decision-making times. For example, the quality of information (Zylberberg, Fetsch, and Shadlen, 2016) and age (Feenstra, Ruiter, and Kok, 2012) (adults vs adolescents) have been shown to decrease decision time and sleep deprivation (Ratcliff and Van Dongen, 2011), (Venkatraman et al., 2007) has been shown to increase the time to make a decision.

Despite the model's popularity in representing the decision-making process, there has been little progress in looking at the Drift-Diffusion model in a risky decision-making setting. The closest we have found to this comes from Webb (2019), who simulates the Holt and Laury (2002) exercises with risk neutrality to a race model. Another similar one comes from Zhao, Walasek, and Bhatia (2020), where they consider the model to explain loss aversion. However, they assume risk neutrality too. A final paper is Howlett and Paulus (2020), where they use a race model and measure risk as a portion of risky choices. We build on this previous work by using experimental data and attempting to fit both the parameters for risky decision-making and the Drift-Diffusion model simultaneously.

This paper complements the earlier papers by applying the Drift-Diffusion Model to decision tasks involving risk. Moreover, we extend the study by implementing prospect theory (Tversky and Kahneman, 1992) into the model. We then compare its performance to a more traditional model. We aim to see if this model can explain risky decision-making.

1.3 Drift-Diffusion Model

The Drift Diffusion Model belongs to the family of bounded accumulation models, which attempt to model how the decision-making process evolves throughout a decision. These models work by having the individual better understand the value of an option over time until sufficient information is received and a decision is made for a corresponding option. They have a few key predictions: the larger an option's value relative to other options, the faster a decision is made and with fewer errors. Conversely, the more information an individual requires to make a decision, the slower and less error-prone a decision will be.

Bounded accumulation models consist of several ingredients that work together to drive the process and make a decision. The first of these is an accumulation process. This process shows how quickly an individual gains information about a given option, with a faster accumulation process indicating a stronger preference for one option. The accumulation process often includes a degree of noise to account for the fact that an individual cannot always correctly perceive the valuations of the options they face. The second is a stopping rule; this shows at what point an individual has sufficiently understood their options and stops accumulating information. The final part is a choice criterion, a rule that states what is to be done once the stopping rule has been reached, for example, picking a given option[4].

The Drift Diffusion Model follows this process stated above. In this case, an individual's accumulation process drifts between an upper and lower bound. The faster the process drifts towards a given option, the more that option is preferred. The difference in the valuation of both options determines this drift. Once either of the bounds is reached (stopping rule), the accumulation process stops, and a decision is made to pick the option associated with that bound. Figure A.1 shows this example

[4]This here is a summary of this family of models, for a more in-depth description, see Webb (2019).

to help visualise how the decision-making process evolves before we go into more detail. The accumulation process drifts with noise towards the upper bound, indicating that the upper option is preferred to the lower option. When this accumulation process hits this upper bound, the decision maker stops accumulating information and chooses the upper option.

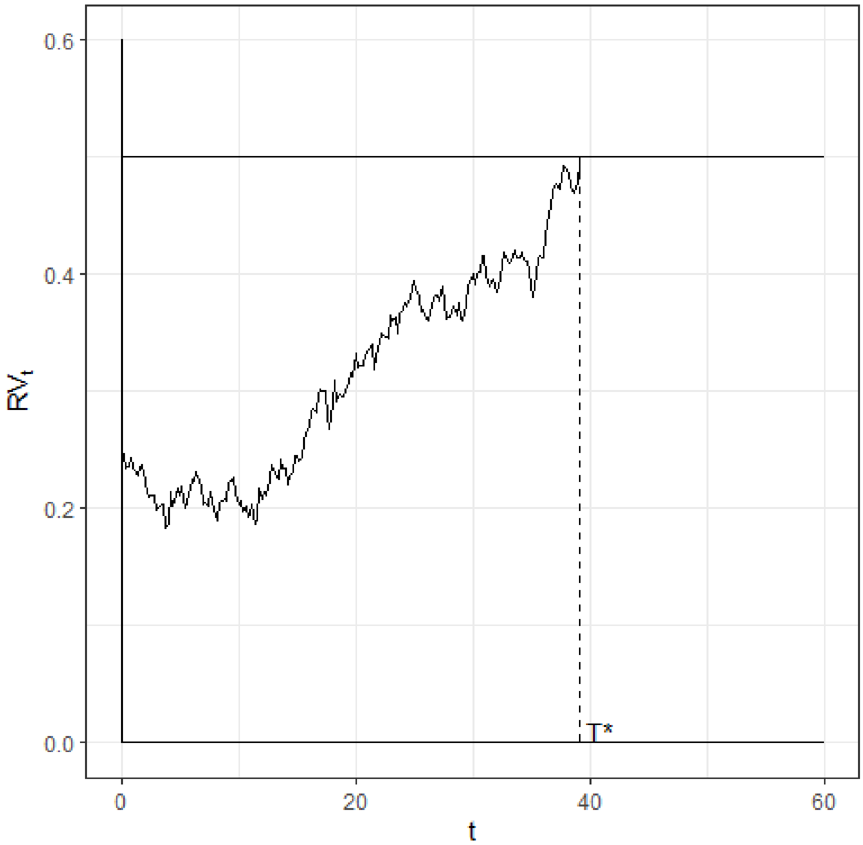


FIGURE 1.1: Visualisation of the Drift-Diffusion Model

The drift-diffusion model's accumulation process consists of both systematic and random processes. The systematic process is called the drift rate, μ and can be interpreted as the degree to which one option is preferred. As one option becomes more valued compared to the other, this drift rate becomes larger, moving farther away from 0. To calculate the drift rate and, by extension, the accumulation process, we assume that an individual has a valuation function V(.) and two choices to consider (1,2), leading to valuations V_1, V_2. Individuals perceive this difference in valuations as $\mu + \epsilon$. Where $\mu = (V_1 - V_2)$ and $\epsilon \sim N(0, \sigma^2)$. This process shows the systematic process (μ) and the stochastic process (ϵ). Following this, they update their Relative Value Signal (RV), which shows where they currently are within the decision-making process. This process evolves throughout the decision-making process as:

$$RV_t = RV_{t-1} + \mu + \epsilon_t \tag{1.1}$$

Where RV_t is the position of the Relative Value Signal at time t. The drift rate, $\mu = V_1 - V_2$, shows the preferred option. If μ is greater than 0, option 1 is preferred to option 2, with a larger number indicating a stronger preference for option 1. Likewise, if μ is negative, option 2 is preferred, and more negativity results in a stronger preference for option 2. A larger value of μ is a larger signal favouring a given option. At time 0, $RV_0 = Z$, where Z is the initial bias towards an option.

We can see that over time, this process evolves following a drifting random walk, either positively or negatively. To decide when this process stops, we implement a stopping rule in the form of an upper and lower bound. These bounds take a value of β and 0. When the RV takes a value of either of these bounds, they receive sufficient signals about which option is preferred. At this time, the decision maker chooses option 1 if they stop with $RV_{t^*} = \beta$ and 2 at $RV_{t^*} = 0$. The larger the boundary separation, the more information is needed to decide. The initial bias, Z, can be shown as a natural preference for a given option. The closer this initial bias is towards either bound, the less information is required in favour of the corresponding option. Another parameter that the model can use is the Non-Decision Time (NDT), which represents the time spent in the decision-making process not contributing towards the decision-making process, including factors such as the time spent reading and selecting an option.

The table below gives a summary of the parameters associated with the model.

TABLE 1.1: Summary of variables in Drift-Diffusion model

Parameter	Brief Description
μ	Rate at which information is accumulated in favour of an option
β	Amount of information needed to make a decision
Z	Initial bias, often assumed to be unbiased due to experiments so $\beta/2$
NDT	Time spent not making a decision

It is possible to make a mistake and pick the least preferred option. This can be done if the individual receives sufficiently large stochastic shocks in favour of the other option. However, this becomes less likely in two key ways:

1. A larger drift rate would require a more significant shock to push the signal towards the least preferred option.

2. A larger degree of boundary separation β will mean that an individual will require more information in favour of an option and thus may be able to withstand a few adverse shocks.

1.4 Estimating the Model

The model itself can prove challenging to estimate. Due to the accumulation process's stochastic nature, two identical individuals will likely take different times to make a decision as they will face a different series of shocks. We can estimate the parameters of the Drift-Diffusion model, the drift[5] and the bound, by considering the probability distribution of these decision times at each of the two bounds. From Voss, Rothermund, and Voss (2004) and Feller (1971) provide a mathematical definition of these models. The probability distributions of the RV hitting either the upper or lower bound (β or 0) at time t^*, the decision time, are given by:

$$Pr(RV_{t^*} = \beta) = g_+(t^*; \beta, \mu) = \frac{\pi}{\beta^2} \exp[(\beta - z)\mu] \sum_{n=1}^{\infty} n * sin(\frac{\pi(\beta - z)n}{\beta}) * \exp[-0.5t^* * (\mu^2 + \pi^2 n^2/\beta^2)] \tag{1.2}$$

and for the lower bound.

$$Pr(RV_{t^*} = 0) = g_-(t^*; \beta, \mu) = \frac{\pi}{\beta^2} \exp[(-z\mu] \sum_{n=1}^{\infty} n * sin(\frac{\pi z n}{\beta}) * \exp[-0.5t^* * (\mu^2 + \pi^2 n^2/\beta^2)] \tag{1.3}$$

Where t^* is the time taken to make a decision, (RT, the time from when a task starts to when a decision is made and a choice is selected). If a researcher wishes to include a non-decision time (NDT), this can be implemented by setting t equal to RT-NDT.

[5]We will modify the drift rate later to include the parameters associated with risky decision-making, the curvature of the utility function and probability weighting, in a later section.

We can integrate the above probability distribution between 0 and decision time t^*. Doing gives us the probability that a diffusion process will hit the upper bound before time t^*.

$$Pr(RV_{t \leq t^*} = \beta) = G_+(t; \beta, \mu) = \int_0^{t^*} g_+(t; \beta, \mu) dt \tag{1.4}$$

and for the lower bound

$$Pr(RV_{t \leq t^*} = 0) = G_-(t; \beta, \mu) = \int_0^{t^*} g_-(t; \beta, \mu) dt \tag{1.5}$$

These integrals and the resulting probabilities can prove to be challenging to recover due to the infinite integral with an infinite series in; however, there exist several statistical packages made to deal with this, such as Fast-dm (Voss and Voss, 2007), HDDM (Wiecki, Sofer, and Frank, 2013), ez-diffusion (Wagenmakers, Maas, and Grasman, 2007), rtdists (Singmann et al., 2020) and RWeiner (Wabersich and Vandekerckhove, 2014) We can generate a likelihood function from here, which can get tricky to estimate due to the infinite series within the equations. However, many different packages and processes can deal with this series, such as the appendix in Voss, Rothermund, and Voss (2004). Then it is a simple task of maximising the loglikelihood function

$$Loglikelihood(\theta) = \sum_{i=1}^{n} y_i log(G_{+i}) + (1 - y_i) log(G_{-i}) \tag{1.6}$$

Where θ is our variables of interest. y_i is an indicator function taking 1 if the individual chooses the "upper option" and 0 if it takes the "lower". This likelihood function can be minimised using packages such as Rsolnp (Ghalanos and Theussl, 2015), which uses a method proposed in Ye (1987).

1.4.1 Adding Risk to the model

Above, we have described the Drift-Diffusion model and how it currently stands. We now aim to apply this model to explain risky decision-making. We do this by altering how the drift rate is determined; instead of estimating a drift parameter μ, we set μ to be the difference in the expected utilities of the lotteries individuals face. We choose to model risk under Prospect Theory (Tversky and Kahneman, 1992).

1.4.2 Prospect Theory

Incorporating Prospect Theory into the model will enable us to consider individuals' attitudes towards risk and probability and how they impact decision-making. Incorporating this into the Drift-Diffusion Model can be done by altering the drift rate. If we recall, the drift rate is the difference between the valuation of both options. Therefore if we define the drift rate as:

$$\mu = U_a - U_b \tag{1.7}$$

This can be interpreted as the difference between the two utilities of options. We can see here that this behaves like a regular drift rate, an increase in U_a, showing option A becoming more appealing, will cause the drift rate to become more positive.

The utility of each option can be calculated using prospect theory. For this, we need valuations of the payoffs and how individuals perceive each payoff's probability.

For a utility function, we need a function that has a few key characteristics; it must be concave in the gains domain, convex in the loss domain, strictly monotonic and continuous. We propose using a normalised constant relative risk aversion (CRRA) utility function:

$$U(x) = \begin{cases} \frac{x^{\rho_g}}{\rho_g}, & \text{if } x \geq 0 \\ -\lambda \frac{(-x)^{\rho_l}}{\rho_l}, & \text{if } x \leq 0 \end{cases} \tag{1.8}$$

Where ρ measures the curvature of the utility function, as ρ gets closer to 1, the less risk averse an individual is, where between 0 and 1 is risk-averse and above 1 is risk seeking. We assume different values of ρ for gains and losses to allow for different attitudes under gains and losses, a result that is commonly found within the literature (Harbaugh, Krause, and Vesterlund, 2002). The parameter λ represents loss aversion. We normalise this to 1 in the case of pure losses. This is because it acts as a scalar on the drift rate and leads to estimation issues with it and the bound β.

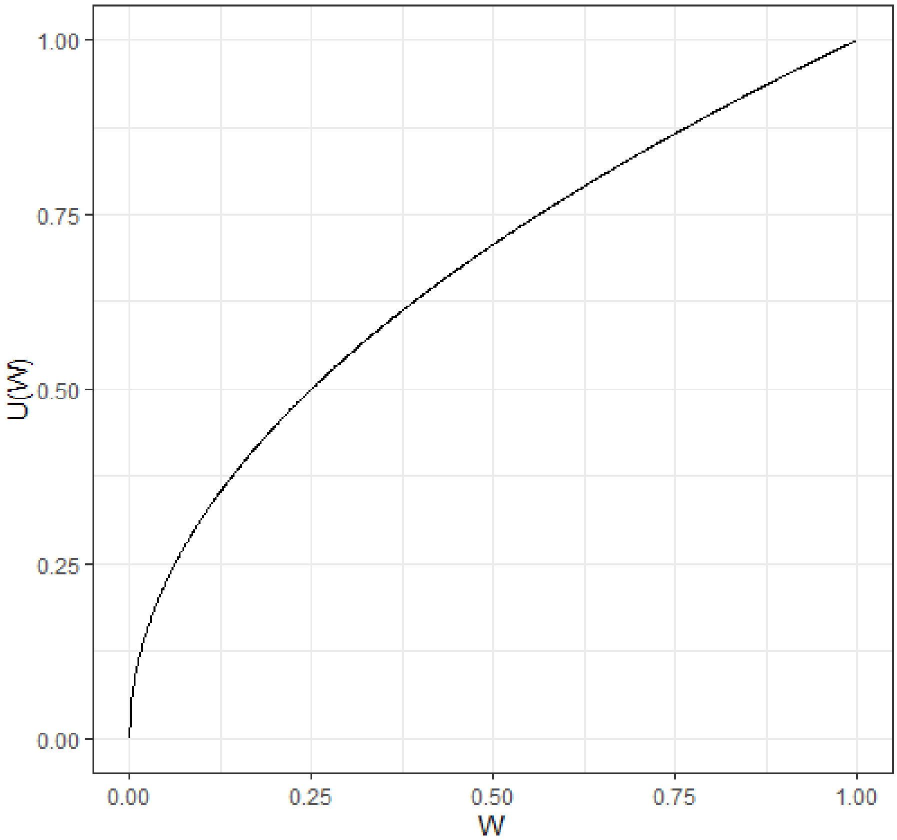


FIGURE 1.2: Visualisation of the Utility Function

The above shows an example of this function for gains; we can see that it is strictly monotonic; as wealth (W) increases, the utility increases, however, at a diminishing rate, showing diminishing marginal returns to additional wealth, a characteristic of risk aversion. As the risk parameter, ρ gets closer to 1; this graph becomes more and more linear.

The second thing we need to consider is a weighting function to allow us to model how an individual can perceive probabilities differently from what they are. These are used to alter the subjective probability to a decision weight. We consider the widely used Tversky and Kahneman (1992) function:

$$W(P) = \frac{P^{\delta}}{(P^{\delta} + (1-P)^{\delta})^{(\frac{1}{\delta})}} \tag{1.9}$$

Where δ shows the degree of probability weighting, the closer it is to 1, the less distortion and the more the perceived probability curve becomes linear. Whereas the further away δ is, the higher the degree of probability distortion. Like before, we assume different values of δ for gambles with gains and losses in our utility function.

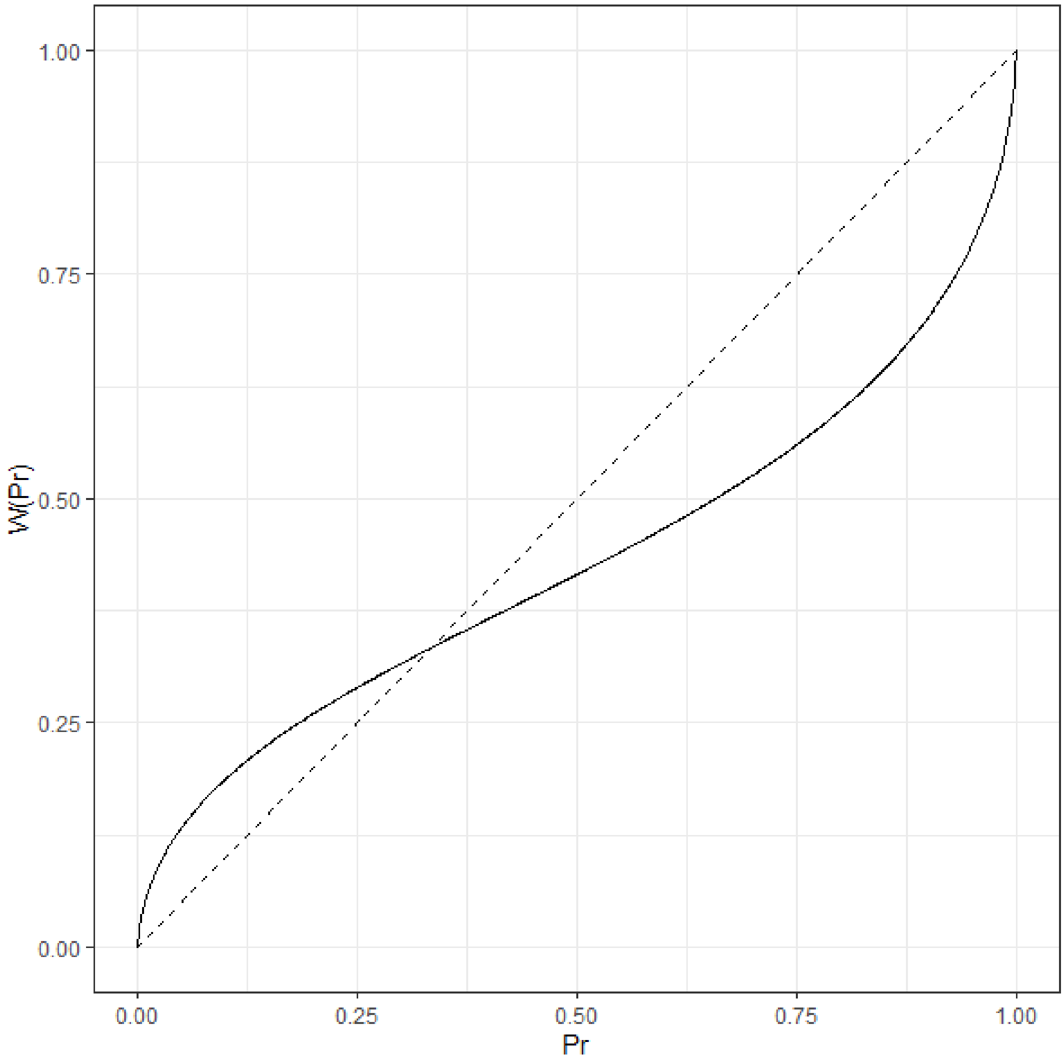


FIGURE 1.3: Visualisation of the Weighting Function

Above, we can see a plot of the weighting function in equation 1.9. From here, we can see an over-weighting of low outcomes and an under-weighing of higher probability outcomes. As δ moves closer to 1, this function becomes more linear and moves towards the dashed lines, whereas the further from 1, the larger the degree of probability distortion.

Considering Gains

The dataset we use consists of risky binary gambles with two possible outcomes. For example, a choice between lottery A with payoffs of £10 and £5 with probabilities of 0.3 and 0.7 and lottery B with payoffs of £15 and £4 with probabilities 0.2 and 0.8. Suppose we had a general lottery with payoff X with probability P and payoff Y with 1-P. Under the standard expected utility model, the expected utility of a gamble would be equal to $P * U(X) + (1 - P) * U(Y)$. The rank-dependant utility differs by including a probability weighting, replacing the higher payoff probability with decision weights of w(.) and 1-w(.) for the lower payoff. If $X > Y$ then the utility of this lottery can be defined as:

$$U_L = w(P) * U(x) + (1 - w(P)) * U(Y) \tag{1.10}$$

With the weighting function in equation 1.9 and the utility function in equation 1.8, this becomes

$$U_L = \frac{P^\delta}{(P^\delta + (1-P)^\delta)^{(\frac{1}{\delta})}} * \frac{X^{\rho_g}}{\rho_g} + (1 - \frac{P^\delta}{(P^\delta + (1-P)^\delta)^{(\frac{1}{\delta})}}) * \frac{Y^{\rho_g}}{\rho_g} \tag{1.11}$$

Similarly, if $Y > X$ then the expected utility becomes:

$$U_L = (1 - w(1 - P)) * U(X) + w(1 - P) * U(Y) \tag{1.12}$$

which can be written as:

$$U_L = (1 - \frac{(1-P)^\delta}{(P^\delta + (1-P)^\delta)^{(\frac{1}{\delta})}}) * \frac{X^{\rho_g}}{\rho_g} + \frac{(1-P)^\delta}{(P^\delta + (1-P)^\delta)^{(\frac{1}{\delta})}} * \frac{Y^{\rho_g}}{\rho_g} \tag{1.13}$$

Considering Losses

Modelling losses is very similar to how we model gains. In gains, weights were attached to the largest gain. However, in this case, they are attached to the payoff with the largest (most negative) loss. We can then carry on as we did for gains. In this case, individuals drift towards the option with the least negative payoff, which is the most preferred. For example, if X < Y (X is a larger loss), our utility function becomes

$$U_L = \frac{P^\delta}{(P^\delta + (1-P)^\delta)^{(\frac{1}{\delta})}} * -\frac{-X^{\rho_l}}{\rho_l} + (1 - \frac{P^\delta}{(P^\delta + (1-P)^\delta)^{(\frac{1}{\delta})}}) * -\frac{-Y^{\rho_l}}{\rho_l} \tag{1.14}$$

And if Y < X

$$U_L = (1 - \frac{(1-P)^\delta}{(P^\delta + (1-P)^\delta)^{(\frac{1}{\delta})}}) * -\frac{-X^{\rho_l}}{\rho_l} + \frac{(1-P)^\delta}{(P^\delta + (1-P)^\delta)^{(\frac{1}{\delta})}} * -\frac{-Y^{\rho_l}}{\rho_l} \tag{1.15}$$

Altering the Drift Rate

Now that we have defined our utility framework, we need to combine it with the Drift-Diffusion Model. We do this by defining the drift rate as the difference in utilities. For example, suppose we have the option to choose between 2 gambles, $(x_1, y_1); (p, 1-p)$ and $(x_2, y_2); (q, 1-q)$, we can calculate the rank-dependent utility of each gamble RDU_1 and RDU_2. We then define the drift rate as the difference between these utilities, $\mu = RDU_1 - RDU_2$. From here, we can substitute our drift rate into the probability distributions shown in equations 1.2 and 1.3 and follow the procedure shown above.

Due to the incorporation of risk attitudes and probability weighting, we have more parameters to estimate than the conventional Drift-Diffusion model, non-decision time, the bound, risk aversion and probability weighting.

1.5 Simulations

Before moving on to actual experimental data, we run some Monte Carlo simulation exercises to see if it is possible to recover behavioural parameters associated with the drift-diffusion model accurately. We perform several simulation exercises, starting from the simple task of Holt and Laury choice list(Holt and Laury, 2002) under standard expected utility preferences and then progressing to risky binary choices with rank-dependant utility preferences and finally adding in non-decision time.

These simulations follow the same procedure; we first set some parameters applicable for each model, such as the curvature of a utility function and the boundary separation. From there, we can calculate the expected utilities and drift rate of options with these parameters for each task. We can then simulate the process and generate a choice our simulated individual will make, as well as the time taken to make the decision. Finally, we then use this simulated choice and time data and equation 2.4 to try and recover the initial parameters we initially set. It is worth noting that due to the drift process's stochastic nature, we cannot recover parameters perfectly. The simulation procedure is repeated 100 times for each of the exercises below. This is to ensure accuracy in our simulations and not recover the parameters by chance.

1.5.1 Holt and Laury

Our first simulation exercise considers the Holt and Laury (2002) choice list. In this, subjects are presented with a list of gambles, gamble 1 being a safe gamble and gamble 2 being a more risky one. As an individual progresses down the choice list, the larger payoff of both lotteries becomes more likely, making the more risky gamble more appealing. The point where an individual switch from the safe gamble to the risky gamble carries information about an individual's risk attitudes. The further down the list that a person chooses to switch from safe to risky, the more risk-averse an individual is. The table below shows the tasks an individual faces in this experiment.

TABLE 1.2: The 10 Lotteries from Holt and Laury (2002).

	Option A				Option B			
Task	p_{A1}	x_{A1}	p_{A2}	x_{A2}	p_{B1}	x_{B1}	p_{B2}	x_{B2}
1	0.1	2.00$	0.9	1.60$	0.1	3.85$	0.9	0.10$
2	0.2	2.00$	0.8	1.60$	0.2	3.85$	0.8	0.10$
3	0.3	2.00$	0.7	1.60$	0.3	3.85$	0.7	0.10$
4	0.4	2.00$	0.6	1.60$	0.4	3.85$	0.6	0.10$
5	0.5	2.00$	0.5	1.60$	0.5	3.85$	0.5	0.10$
6	0.6	2.00$	0.4	1.60$	0.6	3.85$	0.4	0.10$
7	0.7	2.00$	0.3	1.60$	0.7	3.85$	0.3	0.10$
8	0.8	2.00$	0.2	1.60$	0.8	3.85$	0.2	0.10$
9	0.9	2.00$	0.1	1.60$	0.9	3.85$	0.1	0.10$
10	1.0	2.00$	0	1.60$	1.0	3.85$	0	0.10$

To simulate these tasks, we set a parameter commonly found within the literature and use this combined with a Standard Expected Utility model ($EU = PU(X_1) + (1-P)U(X_2)$), to calculate the expected utility for each choice lottery. We start with this simple model to minimise the parameters we recover before moving on to more complex models. From here, we calculate the drift rate from the difference in these lotteries, a positive drift rate showing lottery 1 is preferred to lottery 2 and vice versa. We then simulate the Drift-Diffusion process to generate the choice and decision time data. We then perform our estimation process to recover our behavioural parameters. Given the small number of choices an individual has to make, this process is simulated 10 times to generate a larger amount of data to work with[6]. As stated previously, this entire process is repeated 100 times to ensure robustness.

The table below shows the results of this first simulation exercise. We include the parameters used to simulate the data and the mean and variance of the parameters we recover from the 100 simulations of the data. We can see very close estimates on average with a low level of variance, confirming we can successfully identify the model and robustly recover our assumed parameters.

[6]This is similar to an example found in Webb (2019), where they repeat for 20 identical, risk-neutral subjects compared to our 10 risk averse subjects.

TABLE 1.3: Summary of Simulation for Holt and Laury choice list

	ρ	β
Parameters Used	0.75	2
Mean Parameters Recovered	0.7503	2.005
Standard Deviation	0.0066	0.0363

1.5.2 Risky Choices

We now extend this to a richer experimental design, allowing for the estimation of more complex models. We use the list of gambles from the experiment in Guo, Trueblood, and Diederich (2017). In this experiment, subjects had to choose between a certain payoff and a risky gamble. Such as a certain £39 vs a gamble of £0 and £70 with probabilities 0.4 and 0.6 respectively[7]. Subjects face 16 unique choice tasks, repeated 4 times. However, the tasks were framed in several different ways. We perform an identical process to before, calculating the drift rate for each task and then simulating the drift process. In this case, we still assume the standard expected utility model above. The table below shows the summary of these exercises. As before, we can see parameters are recovered with low levels of variance. This gives us confidence to recover parameters using the standard expected utility model and move only a more complex model.

TABLE 1.4: Summary of Simulation for Binary Choice with SEU

	ρ	β
Parameters Used	0.75	5
Mean Parameters Recovered	0.7477	4.985
Standard Deviation	0.0129	0.0399

1.5.3 Adding in Probability Weighting

Before this, we considered the standard utility model, which only considers the curvature of the utility function. We wish to extend this to consider prospect theory[8],

[7]We include the tasks individuals face in the appendix.
[8]Since we consider only gains, this reduces to the Rank Dependant Utility framework.

this considers the curvature of the utility function (ρ) and a probability weighting variable (δ). We take the same exercises as above; however, this time, when calculating drift rates, we use a rank-dependent utility specification (equations 1.10 and 1.12) instead of the standard expected utility.

The table below shows a summary of these simulations. Like before, we can consistently recover our parameters of interest with good accuracy. Different behavioural parameters are used to ensure we can recover our parameters of interest with any value.

TABLE 1.5: Summary of Simulation for Binary Choice with RDU

	ρ	δ	β
Parameters Used	0.75	0.8	5
Mean Parameters Recovered	0.7559	0.8051	4.948
Standard Deviation	0.0122	0.0119	0.0908

1.5.4 Non-Decision Time

Finally, we need to add non-decision time into the model, the time an individual spends on a task that doesn't contribute towards making a decision. This can include time for reading and understanding what is asked in a task. As before, we simulate data the same way; however, this time, we can add additional non-decision time to it. Non-Decision Time is added by adding a given amount of time onto the time it takes to make a decision. When estimating not decision time, instead of using t^*, use $t * -NDT$. As we can see from the summary, we can also recover parameters with a reasonable degree of accuracy.

TABLE 1.6: Summary of Simulation for Binary Choice with RDU and Non-Decision Time

	ρ	δ	β	NDT
Parameters Used	0.75	0.8	5	0.5
Mean Parameters Recovered	0.7336	0.7882	5.116	0.4566
Standard Deviation	0.0725	0.0561	0.8911	0.1313

1.5.5 A Simulation with Losses

We have recovered parameters of interest in the gains domain; we now wish to consider losses. Our current selection of experimental tasks does not include losses. However, we take the negative value instead to create losses. We consider losses with prospect theory and non-decision time. Below shows the results of our final simulation exercise, in which we can see that we can successfully recover the values we set for the four behavioural parameters we consider.

TABLE 1.7: Summary of Simulation for Binary Choice with PT and Non-Decision Time under Losses

	ρ	δ	β	NDT
Parameters Used	0.5	0.7	5	1
Mean Parameters Recovered	0.4973	0.695	4.916	0.9985
Standard Deviation	0.0102	0.0352	0.388	0.0167

Throughout our simulation exercises, we can see that we can successfully identify the values for the behavioural parameters we set from simple exercises to more complicated tasks and data.

1.6 Experimental Data

We now move to consider actual experimental data. The data comes from Glöckner and Pachur (2012); subjects were tasked with choosing between two risky gambles. Most of these tasks were randomly generated; for example, an individual choosing between A gamble of 400 and 320 with probabilities of 0.4 and 0.6 and a gamble of 770 and 20 with probabilities of 0.4 and 0.6. There were also a few other deliberately chosen problems, such as the Holt and Laury (Holt and Laury, 2002) tasks; however, all tasks still followed the theme of picking between two risky gambles. Subjects faced two sets of questions 1 week apart to test preference consistency. Given that they show the same preferences in both rounds, we pool the data. In total, subjects faced 140 pure gains tasks, 60 pure losses tasks, and a few tasks considering mixed gambles. For this paper, we focus on gains and losses. We also dropped several tasks where subjects mindlessly selected choices (for example deciding in 0.001 seconds). We chose to use the data from this experiment due to the high-quality, easy-to-understand decision time data, which is accurate to the microsecond (1 thousand of a second), unlike a majority of other datasets we looked at, which had poor or non-existent time data. Like our simulation exercises, we shall estimate the parameters we need using equation 1.6. We do this for both gains and losses but do not do this for the mixed gamble tasks in the experiment due to a small amount, making estimation difficult.

1.7 Fitting a Benchmark Model

To see how the Drift-Diffusion model explains an individual's actions, we compare its performance to Logit-Errors, a well-established model within the literature. Utilities can be calculated as they have for the Drift-Diffusion Model. To account for the stochastic nature of choices, we assume a Logit-Error function. Thus, the probability of choosing lottery A is given by:

$$P(A) = \frac{\exp(1/sU_A)}{\exp(1/sU_A) + \exp(1/sU_B)} \tag{1.16}$$

where s is a precision parameter to be estimated. This choice probability function can be rearranged to be:

$$P(A) = \frac{1}{1 + \exp(-\mu/s)} \tag{1.17}$$

Where $\mu = U_A - U_B$, the difference in the valuations of both gambles is also the drift rate from the Drift-Diffusion model. This can be done for gains and losses with no tweaks to the function. According to the above assumptions, the log-likelihood function is given by:

$$LL(\theta) = \sum_{i=1}^{n} y_i \ln(P_i(A)) + (1 - y_i) \ln(1 - P_{ni}(A)) \tag{1.18}$$

where n is the number of tasks, $y_i = 1(0)$ is an indicator function denoting the choice of lottery A(B) for in task i, and θ is the vector of behavioural parameters to be estimated. We also run a few simulation exercises similar to the previous specification to ensure that our model works, can recover preferences and that our coding is correct.

1.8 Comparing Both Models Under Gains

To measure the performance of the Drift-Diffusion model, we wish to compare its predictive performance to that of our logit errors to see if additional information gained from decision time can help improve predictive accuracy, as shown in Clithero (2018a). We do this by considering a k-fold cross-validation method; the dataset is split into two sets: a training set that is used to recover parameters and a test set that uses said parameters to predict choices. For both models, we split the data into 2/3 and 1/3, using the first 2/3, the training set, to estimate the parameters associated with both models. Following this, we predict the choices of the remaining 1/3, the test set, using the behavioural parameters we recover. When doing this for the Drift-Diffusion Model, we can calculate their rank-dependent utilities and drift rates and simulate this process for each of their remaining tasks. From here, we compare if the choice, predicted by their drift rate with the parameters recovered, matches the choice we observe. For our Logit-Errors model, we perform an identical process where we estimate using 2/3 of the data available on each subject. We then use the parameters recovered to estimate the probability of choosing option A using the equation 1.17. From here, we generate a random number between 0 and 1 from a uniform distribution. If the number generated is less than the probability, option 1 is selected; otherwise, option 2 is selected. We can then see if our predicted outcome is the same as an individual's choices in those particular rounds.

For both models, we repeat the process with the final third of the data 100 times for each observation. For example, simulating the drift process 100 times per dataset. We do this to ensure that our forecasting results are robust and we aren't simply getting lucky with a few randomly generated observations.

TABLE 1.8: Portion of correct predictions from the final 1/3 of data

	DDM	Logit-Error
mean	0.7589	0.7313
Median	0.7634	0.7306
Lowest	0.5445	0.5623
Highest	0.9665	0.9072
Variance	0.0063	0.0060

Table 1.8 provides a summary of the forecasting results, showing a proportion of correct predicted choices from the final 1/3 of the dataset, repeated 100 times. From both the table and 1.4 below, we can see that our Drift-Diffusion model has both a higher mean and median success at predicting this final third of chosen options. The difference between these predictive accuracies of both models is statistically significant with a Wilcoxon rank sum test with a p-value of 0.04818.

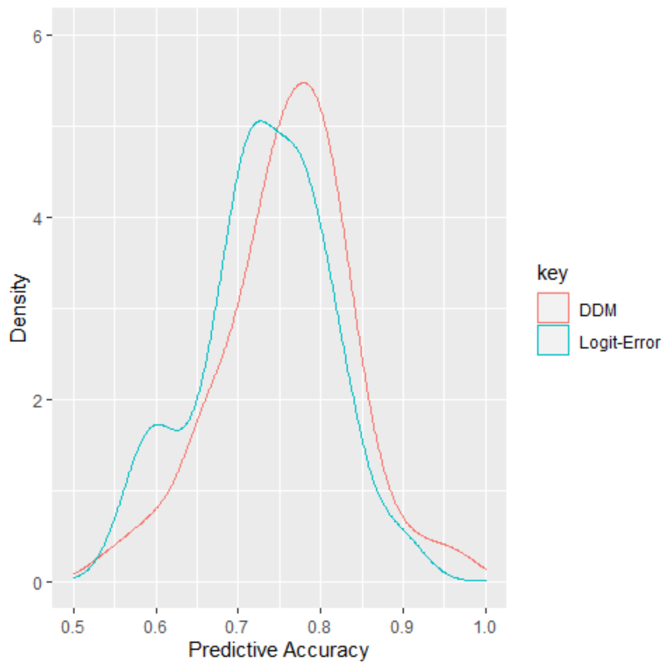


FIGURE 1.4: Density Plot for the predictive accuracy of both models

Whilst the difference in performance between both models is statistically significant, the difference is relatively small. To increase the robustness of our results, we consider altering our k-fold cross-validation technique. In this case, we split choice data into 5 groups, combine 4 of these groups and use them to be the training set to estimate individual's behavioural parameters, then use these to forecast the remaining fifth of their choice data. This process is repeated 5 times to allow each fifth to have the chance to be the test set and forecasted. This allows each part to be forecasted, allowing for a more reliable result. It prevents the possibility that the final third of the data was easier to predict than the other groups. As before, this forecasting process is repeated 100 times per individual to ensure a more accurate forecast and eliminate the chance of a lucky string of forecasted behaviour.

TABLE 1.9: Portion of correct predictions under gains

	DDM	Logit-Error
Mean	0.7520	0.7284
Median	0.7682	0.7262
Lowest	0.6188	0.6023
Highest	0.8604	0.661
Variance	0.0029	0.0039

Again we can see that from table 1.9 above and figure 1.5 below that the Drift-Diffusion model is outperforming the standard model and that this time the differences are more significant, providing more support for the in the use of the model. These differences are also significant with a Wilcoxon rank sum test with a p-value of 0.006[9]

[9]For robustness, we perform this with different utility functions and weighting functions and find that the model either outperforms the logit-errors or is equal to, depending on the specification. This can be found in Appendix A.1.

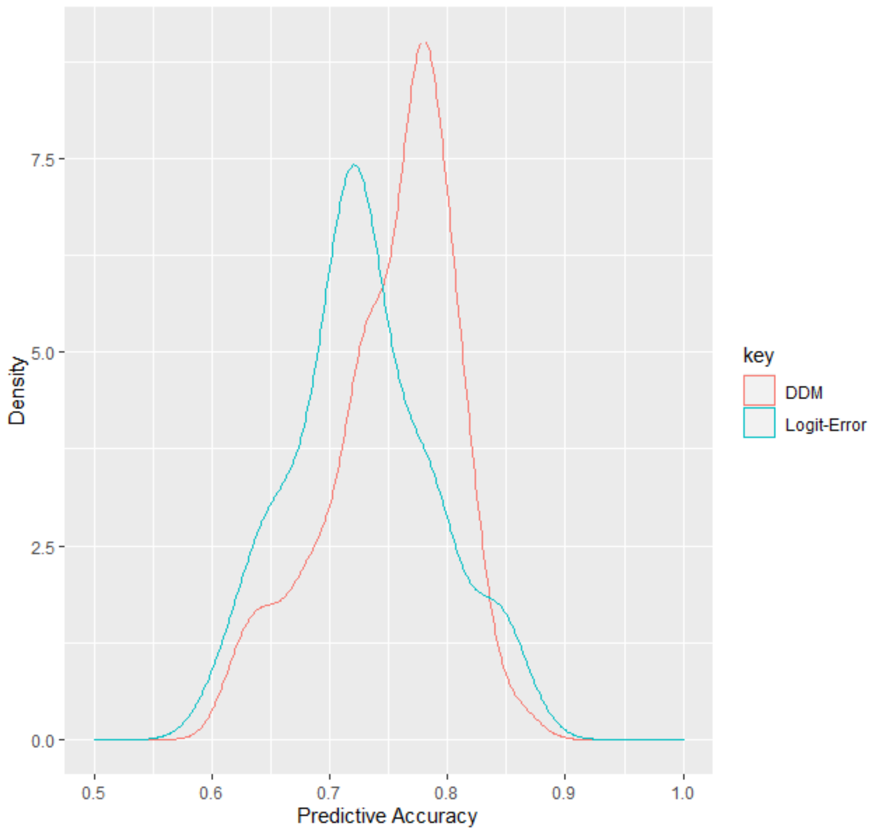


FIGURE 1.5: Density Plot for the predictive accuracy of both models

1.9 Comparing Both Models Under Losses

We now move to consider losses. The dataset contains relatively few loss tasks. Due to this, we use our second k-fold procedure to allow for more estimation accuracy. Like before, we split this into 5 groups, estimate the parameters with 4 of the groups and then forecast the final group, repeating so that every fifth can be predicted and used for estimation.

TABLE 1.10: Portion of correct predictions under losses

	DDM	Logit-Error
Mean	0.6127	0.6898
Median	0.6180	0.6784
Lowest	0.4937	0.5153
Highest	0.7326	0.9043
Variance	0.0034	0.0101

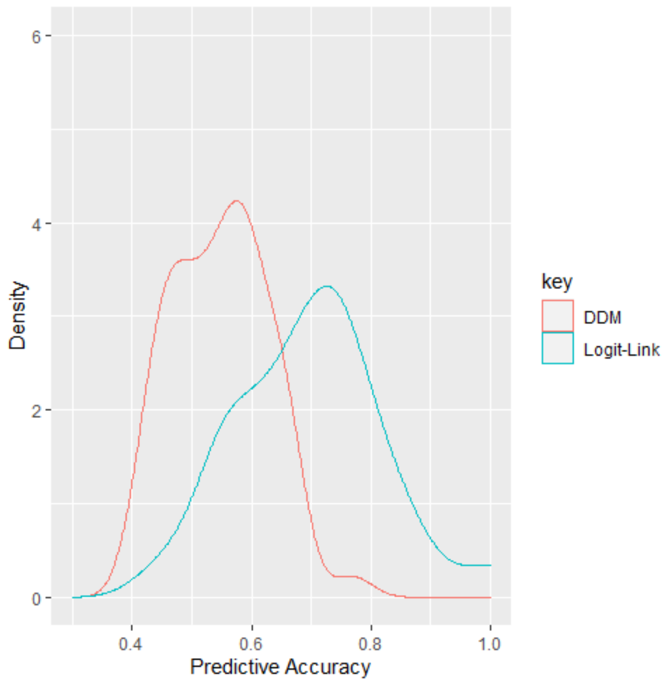


FIGURE 1.6: Density Plot for the predictive accuracy of both models under losses

We see the opposite result when considering losses: our logit error model outperforms the Drift-Diffusion model. We have a few possible explanations for this; first, we have a minimal amount of loss tasks compared to those for gains. Therefore, fitting a model with more parameters (4 in the drift model instead of 3 in the logit error model) could prove more difficult. We also note that the predictive power for both many suggests more difficulty with losses; this could lead to a different decision process being involved instead, for example, heuristics shown in Glöckner and Pachur (2012), based off Brandstätter, Gigerenzer, and Hertwig (2006) or with taking more time to make a decision[10]. Another issue could be the many tasks individuals carried out during the experiment. Some degree of decision fatigue could lead to further noise within an individual's choices. Future work could involve trying this again with a larger, more focused dataset with losses to test this domain more thoroughly.

[10]We do find individuals take longer to decide for loss tasks, a summary of this information can be found in the Appendix A.2

1.10 Conclusions

We have considered a well-established, highly researched model in the psychology literature and extended it into economics and the domain of risky decision-making. We apply it to a standard risky binary choice experiment to test how incorporating behavioural parameters associated with risky decision-making into the model can lead to better predictive powers than models considered in the economics literature.

We have taken the model and compared it to a well-established model within the economics literature. We can see that the model is suitable for explaining risky binary choices with our extensive simulation exercises. Following this, we look at the model's ability to forecast the decisions made by an individual in a given situation. We can see that the Drift-Diffusion Model outperforms the logit-error model we compared it to in the domain of gains. We perform a similar exercise with losses and find that the model can recover parameters. Still, it struggles with predicting losses; however, this could be due to the small data available. We suggest performing this exercise with a larger, more focused dataset to prevent decision fatigue and noise. The results we gain from our simulation and comparison exercises indicate that the Drift-Diffusion Model is suitable for modelling an individual's behaviour under risk and could be used instead of the logit model when the data is available. More importantly, it provides us with an argument that decision times are valuable and potentially carry a lot of information and models that can incorporate time should be considered if possible when the data is available.

While we cannot say with the data used, we can speculate that the Drift-Diffusion Model can recover behavioural characteristics in situations we previously could not. This includes experiments between a risky gamble and a certain payoff of the same amount, or in the Holt and Laury experiment shown earlier. We used this in our Holt and Laury simulation exercises to recover a parameter where previously, one could only recover a range of parameters. However, this is currently purely speculative and whilst we have unintentionally done this with simulations and will need to be investigated with experimental data.

Whilst this is undoubtedly an exciting start to incorporating the Drift-Diffusion Model into economics, there already exists a few interesting ways we can further extend this. As stated earlier, fitting mixed gambles is a natural extension; however, we need to find suitable data or run an experiment. Following this, we can follow the history of the literature and look into variance in the parameters, for example, the bounds and non-decision time varying throughout the experiment. Other examples of ways to extend this model are by adding time pressure into the mix and seeing if this causes an impact on the bounds. There is also a possibility of overcoming one of the limiting characteristics of the model, that being; it could only be used for binary choices and try to extend it further than this to multiple options. Other potential extensions can include using this framework in different environments, such as ambiguity or introducing this into game theory.

Chapter 2

Risk Preferences, Gender Effects and Bayesian Econometrics

2.1 Introduction

There is no doubt that risk preferences play a central role in every aspect of economic life. Gender differences in risk preferences is a much debated topic and it has often been argued that these differences might provide a possible explanation of the observed differences between the two genders in various aspects of economic life such as financial decision making, hold of front office roles, or entrepreneurship, to name but a few. Nevertheless, there is little agreement on whether there is a universal pattern of differences between the two genders. Early surveys from the economic literature (Eckel and Grossman, 2008; Croson and Gneezy, 2009) provide mostly supporting evidence of women being less willing to accept risks. Recently, Filippin and Crosetto (2016) conducted an extensive meta-analysis on gender differences and risk attitudes, using data from 7000 subjects and 54 replication studies of the Holt and Laury (2002) risk elicitation task. One of their main findings is that:

> "[..] gender differences appear in less than 10% of the studies and are significant but negligible in magnitude once all the data are pooled."

and they conclude that:

> "[..] the structural model seems to confirm that significant gender differences are detected in the HL task when merging all the observations. The reason is to be found in the sky-rocketing increase of the statistical power of the test, which drives fairly close to zero the likelihood of observing a false negative when data are merged."

The above statement indicates that in order to be in place to detect any potential gender effects, one needs to recruit an extremely large sample of subjects, for the standards of economic experimentation, a task which seems prohibiting given all the time, financial and practical constraints that a researcher may face. In this paper, motivated by the conclusions of Filippin and Crosetto (2016), we investigate how one can increase the extracted information from small sample datasets, and what are the implications of omitting to do so.

One of the most common approaches to explore potential differences between genders is to assume a particular preference functional, pool all the data together, and estimate a representative agent model, using demographic dummy variables to control for heterogeneity (see Harrison and Rutström, 2008, Xie, Page, and Hardy (2017), Vieider et al., 2015, Bouchouicha et al., 2019). Parameters are then obtained by using either Maximum Likelihood Estimation techniques (MLE) or Non-Linear Least Squares estimation methods, and the statistical significance of the dummies defines the existence and the size of potential differences. While the representative approach is attractive, due to its simplicity , it comes with a serious limitation. By ignoring individual heterogeneity, the estimated preferences may not be representative for any of the subjects. Consider an extreme scenario where out of 100 subjects, 50 are male and risk neutral, 25 female and risk seeking with a risk coefficient of -0.50 (assuming a power utility function as in Holt and Laury, 2002 and later in our analysis of the form $x^{1-r}/(1-r)$) and the remaining 25 subjects are females and risk averse, with a coefficient of 0.50. Pooling all the data together and fitting a representative agent model to this dataset, including a control variable to capture potential gender differences, will return an estimated risk aversion very close to zero, implying risk neutrality, and the coefficient of gender effects to be insignificant[1]. The main conclusion that a researcher could draw from a similar analysis is that the observed population has risk neutral preferences and there are no gender effects. Consider now a policy maker who aims to identify the risk seeking women in a population. By conducting a similar analysis, the policy maker will reach the conclusion that no risk seeking women exist in this sample and no action needs to be taken. While this example is extreme and perhaps improbable, it is used to highlight the impact of ignoring potential behavioural heterogeneity in identifying preferences and differences based on demographic criteria.

On the other end of the spectrum, one could estimate preference functionals at the individual subject-level (see Hey and Orme, 1994, Stott, 2006). While this approach takes into consideration the individual characteristics of each subject, a large amount of data points is required in order to obtain robust and reliable estimates. This comes at a high cost for the researcher, as larger number of decision tasks would mean longer sessions that could potentially lead to boredom and eventually to more

[1] We indeed executed a similar simulation exercise where this result was confirmed. Details are available on request.

noisy data.

In the present study we compare the representative agent modelling approach, to two more flexible and informative methods of parameter estimation that allow one to simultaneously make inferences at both the individual subject and the experimental population level. In particular, we compare the frequentist and the Bayesian methods, by analysing the data using Maximum Simulated Likelihood Estimation techniques (MSLE), as well as Hierarchical Bayesian (HB) econometric modelling. We use data from three prominent studies of decision making under risk. First, we use the original data from the Holt and Laury (2002) experiment, and assuming Expected Utility preferences, we first show how all three inference methods are able to capture gender effects. Then, we extend our analysis to non-Expected Utility preferences, and particularly to Rank Dependent Utility, since our focus in on risky choice in the gains domain. Using the dataset from Baillon, Bleichrodt, and Spinu (2020), we show that taking into consideration individual heterogeneity, improves the inference, while the MLE representative agent model fails to identify the existence of gender differences. Finally, we focus on the domain of losses, and adopting a Cumulative Prospect Theory framework, we explore the differences between the two genders, across all the components of risk preferences, namely utility curvature, probability weighting and loss aversion. We show how MLE fails to capture gender differences and we also focus on the differences between the MSLE and the HB methods in capturing these differences.

Our results can be summarised as follows. When there is a small number of parameters to estimate, any of the inference methods will be able to detect the presence of gender differences in the key behavioural parameters. As the model complexity increases, and therefore the number of parameters along with their collinearity, more flexible methods that take into consideration individual heterogeneity, provide more robust inference when the focus is on the difference between two populations. We complement our study with an extensive Monte Carlo simulation to compare the three inference methods, and we show that while all MLE, MSLE and HB methods are able to successfully recover the mean values of the simulated parameters, frequentist methods are more prone to ignore statistical significance due to overfitting, compared to Bayesian methods.

The rest of the paper is organised as follows: section 2.2 briefly introduces the idea of Hierarchical Bayesian modelling, section 2.3 focuses on the Holt and Laury (2002) risk elicitation task and presents, along with the task and the data, the econometric specification for both MLE and HB, assuming Expected Utility preferences (EU), section 2.4 relaxes the hypothesis of EU and introduces Rank Dependent Utility preferences using data that allow the estimation of such preferences, and finally, section 2.5 focuses on the domain of losses, introducing a Cumulative Prospect Theory model and loss aversion. In section 2.6 we report the results of the simulation. We then conclude.

2.2 Frequentist Vs Bayesian Parameter Estimation

The most common approach to estimate structural decision making models is by either pooling all data together and fit a representative agent model, or by assuming complete independence and fit subject-level models, using maximum likelihood estimation techniques (MLE). Fitting a representative agent model ignores much of individual behavioural heterogeneity and generates estimates which potentially, are not representative of any individual subject in the sample. A simple way to introduce heterogeneity to the representative model is to condition the parameters to a set of observable demographics and assume that subjects that belong to the same demographic group share the same behavioural parameters (see for example Harrison and Rutström, 2008, Bouchouicha et al., 2019). An alternative way to introduce heterogeneity, within the frequentist framework, is to use a *random-coefficients* model, a popular method to model unobserved heterogeneity, on top of the observed one (e.g. through demographics). In this kind of modelling, it is assumed that each behavioural parameter in the model is characterised by an underlying distribution across the population. Using MLE techniques and simulation, it is possible to combine estimates of the population distribution (mean and standard deviation) with individual choices, and make inferences at both the population and the subject level (for applications see Gaudecker, Van Soest, and Wengström, 2011; Conte, Hey, and Moffatt, 2011; Moffatt, 2016). Nevertheless, it is known that MLE is susceptible to overfitting and may generate noisy and unreliable estimates when there is a lack of a large number of observations (see Bishop, 2006, pp. 166, Nilsson, Rieskamp, and Wagenmakers, 2011). An alternative method to introduce heterogeneity and mitigate these drawbacks is to adopt Hierarchical Bayesian estimation techniques (see Balcombe and Fraser, 2015a; Ferecatu and Önçüler, 2016 and Baillon, Bleichrodt, and Spinu, 2020 for some recent applications of hierarchical models for choice models under risk and Stahl, 2014 for ambiguity models.). The key aspect of hierarchical modelling is that even though it recognises individual variation, it also assumes that there is a distribution governing this variation (individual parameter estimates originate from a group-level distribution). As Baillon, Bleichrodt, and Spinu (2020) highlight, Hierarchical Bayesian modelling is a compromise between a representative agent and subject-level type estimation. It estimates the model parameters for each subject separately, but it assumes that subjects share similarities and draw their individual parameters from a common, population level distribution. In that way, individual parameter estimates inform each other and lead to a *shrinkage* towards the group mean that reduces biases in parameter estimates. The latter leads to more efficient and reliable estimates compared to those estimated using frequentist methods. One of the most crucial aspects of Bayesian inference, is the way uncertainty is incorporated in the econometric model in the form of probability distributions. A researcher can use her subjective beliefs or objective knowledge and form a prior distribution which summarises all the available knowledge regarding a particular parameter, before observing any data. In Bayesian inference, the estimation of a parameter of interest corresponds to the calculation of the probability distribution over the parameter, given the observed data and the prior beliefs. Another aspect of the Hierarchical model is that it is applied in an hierarchical form providing both within decision unit analysis (subject level) and across unit analysis (population level). Both the way the Bayesian model incorporates uncertainty and its Hierarchical structure, allows it generate precise estimation of preferences, even when the available data are limited.

Jacquement and L'Haridon (2018, p. 247) provide a comparison between the frequentist and Bayesian methods, highlighting the most important differences, namely the way each method interprets each parameter, the nature of the point estimation, the way intervals for statistical significance are estimated, and; the way hypothesis testing can be done. For the frequentist method the parameter is an unknown constant while for the Bayesian a random variable. Similarly, the point estimation will be the value of the estimator in the former, while a posterior summary in the latter (e.g. the mode of the distribution). For statistical significance, the frequentist method requires the estimation of confidence intervals, compared to the credible intervals in the Bayesian inference. As Huber and Train (2001) point out, in the presence of small samples, the two procedures can provide numerically different results, due to the different way of treating uncertainty in the parameters of the population distribution. In what follows, we compare the three different inference methods (MLE, MSLE and HB) in their capacity to detect gender differences, focusing on three representative examples of decision making under risk.

2.3 Risk Preferences and Expected Utility

Gender differences in risky decision making has been the topic of numerous studies. Eckel and Grossman (2008) and Croson and Gneezy (2009) summarise the literature, finding that female subjects tend to be more risk averse. Charness and Gneezy (2012) and Holt and Laury (2014) discuss how the risk elicitation task affects the inference on differences, while Filippin and Crosetto (2016) challenge the early evidence by finding that the observed effects are negligible in magnitude. In this section we focus on one of perhaps the most common elicitation methods that has been used in the literature, the Holt and Laury (2002) task.

2.3.1 Decision Task and Data

For the analysis, we use the data from the original Holt and Laury (2002) study. Each subject is presented with the 10 choice tasks, as shown in Table 2.1. Each task consists of a choice between two paired lotteries A and B. The payoffs for lottery A are fixed to \$2.00 and \$1.6, while for lottery B, the payoffs are 3.85 and \$0.10. Since lottery A is characterized by less variable payoffs, one can label A as the *safe* option and B the *risky* one. In the first choice task, the probability of getting the high payoff is equal to 10% for both lotteries, and it increases as one moves down the table. At the first row, only the extremely risk seeking subjects are expected to choose lottery B. A risk neutral person is expected to choose lottery A for the first 4 tasks (since the expected value of lottery A is greater) and then switches to lottery B for the remaining tasks. Holt and Laury (2002), assuming a particular form of risky preferences, provide a mapping between then number of safe choices and the value of risk coefficient of a subject (the higher the degree of risk aversion, the higher the number of safe choices).

There are data from 212 subjects (95 females) from 4 treatments, an incentivised low-payoff treatment (LOW_1), with payoffs as those in Table 2.1, a hypothetical treatment (HYP), with the payoffs scaled up by 20, 50 or 90, an incentivised high-payoff treatment ($HIGH$), with payoffs scaled up by 20, and finally, a low-payoff treatment (LOW_2), identical to the first one. For our purposes, we use only the data from the low-payoff treatment (LOW_1).

TABLE 2.1: The 10 Lotteries from Holt and Laury (2002).

	Option A				Option B			
Task	p_{A1}	x_{A1}	p_{A2}	x_{A2}	p_{B1}	x_{B1}	p_{B2}	x_{B2}
1	0.1	2.00$	0.9	1.60$	0.1	3.85$	0.9	0.10$
2	0.2	2.00$	0.8	1.60$	0.2	3.85$	0.8	0.10$
3	0.3	2.00$	0.7	1.60$	0.3	3.85$	0.7	0.10$
4	0.4	2.00$	0.6	1.60$	0.4	3.85$	0.6	0.10$
5	0.5	2.00$	0.5	1.60$	0.5	3.85$	0.5	0.10$
6	0.6	2.00$	0.4	1.60$	0.6	3.85$	0.4	0.10$
7	0.7	2.00$	0.3	1.60$	0.7	3.85$	0.3	0.10$
8	0.8	2.00$	0.2	1.60$	0.8	3.85$	0.2	0.10$
9	0.9	2.00$	0.1	1.60$	0.9	3.85$	0.1	0.10$
10	1.0	2.00$	0	1.60$	1.0	3.85$	0	0.10$

2.3.2 Theoretical Framework and Econometric Specification

We assume that the agent holds Expected Utility preferences and receives utility from income according to a Constant Relative Risk Aversion (CRRA) utility function of the form:

$$u(x) = \frac{x^{1-r}}{1-r} \tag{2.1}$$

where x is the monetary payoff, and r is the risk coefficient with $r > 0$ indicating a concave utility for gains (risk aversion), $r < 0$ a convex utility (risk seeking) and $r = 0$ a linear utility (risk neutrality). For $r = 1$ the function collapses to the logarithmic function. A lottery is evaluated by the weighted sum of the utilities of the payoffs, therefore, the expected utility of lottery A, for a particular task, is given by

$$EU_A = p_{A1} \frac{x_{A1}{}^{1-r}}{1-r} + (1 - p_{A1}) \frac{x_{A2}{}^{1-r}}{1-r} \tag{2.2}$$

To account for the stochastic nature in choices, we assume a *logit* link function. Thus, the probability of choosing lottery A is given by:

$$P(A) = \frac{\exp(1/\xi EU_A)}{\exp(1/\xi EU_A) + \exp(1/\xi EU_B)} \tag{2.3}$$

with ξ a precision parameter to be estimated. According to the above assumptions, the log-likelihood function is given by:

$$LL(\theta) = \sum_{n=1}^{N} \sum_{i=1}^{I} y_{ni} \ln(P_{ni}(A_i)) + (1 - y_{ni}) \ln(1 - P_{ni}(A_i)) \tag{2.4}$$

where N is the total number of subjects, I is the number of tasks, $y_{ni} = 1(0)$ is an indicator function denoting the choice of lottery A(B) for subject n in task i, and θ is the vector of behavioural parameters to be estimated. Therefore, there are 2 parameters to estimate, the risk coefficient r and the precision parameter ξ. To introduce gender effects, we introduce a dummy variable y_{FEMALE} which takes the value 1 is the subject is female, otherwise it is equal to 0. For each parameter θ_n in our model, with $\theta_n \in \{r, \xi\}$ we specify

$$\theta_n = \theta_0 + \theta_{FEMALE} \times y_{FEMALE} \tag{2.5}$$

Since we consider a stochastic model which takes into consideration the errors of the decision maker, we include in the analysis the observations of all the subjects (rather than focusing only on subjects without multiple switches). There are in total 4 parameters to estimate, the risk coefficient, the precision parameter and the two parameters that capture gender effects. [2]

For the HB estimation, we follow Rouder and Lu (2005) and Nilsson, Rieskamp, and Wagenmakers (2011) set-up. Each subject n made a series of I binary choices in a given dataset and the observed choices vector is denoted by $D_n = (D_{n1} \cdots D_{nI})$. Every subject is characterised by its own parameter vector $\Theta_n = (r_n, \xi_n)$, and we assume that both the utility curvature r_n and the sensitivity parameter ξ_n are normally distributed ($\theta_n \sim N(\mu_\theta, \sigma_\theta)$), while for the hyper-parameters we assume normal priors for the mean μ_θ and uninformative priors (uniform) for σ_θ. We also follow the standard procedure and transform all the parameters to their exponential form to ensure that they lie within the appropriate bounds (see Balcombe and Fraser, 2015a). To capture gender differences, we condition the mean of all parameters to a female covariate. For each subject n, each parameter θ_n is assumed to be drawn from a normal distribution of the form: $\theta_n \sim N(\theta + \theta_{FEMALE} \times y_{FEMALE}, \sigma^2_\theta)$, with y_{FEMALE} a female dummy variable. That is, the mean between the two groups differs by θ_{FEMALE}. In what follows, we use either a normal or a log-normal distribution, depending on whether there are constraints for a parameter to be strictly positive.

The likelihood of subject's n choices is given by:

$$P(D_n|\Theta_n) = \prod_{i=1}^{I} P(D_{n,i}|\Theta_n)$$

where $P(D_{n,i}|\Theta_n)$ is given by:

$$LL(\theta) = \sum_{i=1}^{I} y_{ni} \ln(P_{ni}(A_i)) + (1 - y_{ni}) \ln(1 - P_{ni}(A_i)) \tag{2.6}$$

Combining the likelihood of the observed choices and the probability distribution of all the behavioural parameters, the posterior distribution of the parameters is given by:

$$P(\Theta|D) \propto P(D|\Theta) \times P(\Theta)$$

with $P(D|\Theta)$ being the likelihood of observed choices over all the subjects and $P(\Theta)$ the priors for all parameters in the set Θ.

[2]For the estimation we use a general nonlinear augmented Lagrange multiplier optimisation routine that allows for random initialisation of the starting parameters as well as multiple restarts of the solver, to avoid local maxima. The estimation was conducted using the *R* programming language for statistical computing (The *R* Manuals, version 3.6.1. Available at: http://www.r-project.org/).

Monte Carlo Markov Chains (MCMC) were used to estimate all the specifications. The estimation was implemented in JAGS (Plummer, 2017). The posterior distribution of the parameters is based on draws from two independent chains, with 50,000 MCMC draws each. Due to the high level of non-linearity of the models, there was a burn-in period of 25,000 draws, while to reduce autocorrelation on the parameters, the samples were thinned by 10 (every tenth draw was recorded). Convergence of the chains was confirmed by computing the $\hat{R}$ statistic (Gelman and Rubin, 1992).

Finally, for the MSLE we follow Train (2009) and Moffatt (2016) and we estimate the models with the help of simulation. As mentioned before, in this random-coefficient model, the behavioural parameters for a given subject are fixed and they vary across the experimental population according to a distribution (usually assumed Normal). Assuming that a parameter θ is drawn from a distribution with density $g(\theta)$, for a set of I choices, the likelihood of subject's n choices is given by:

$$LL(\theta) = \int \left[\prod_{i=1}^{I} P_{ni}(A_i)^{y_{ni}} \times (1 - P_{ni}(A_i)^{1-y_{ni}}) g(\theta) \right] d\theta \tag{2.7}$$

and the total log-likelihood is given by the sum of the logarithm of (2.7) across all subjects. The parameter θ is distributed over subjects according to the density function $g(\theta)$, and is known as the subject-specific random effect. The variation in θ captures the between-subject heterogeneity. When there are more than one parameters θ, the distribution $g(\theta)$ is a multivariate distribution and the integral is multidimensional. Therefore, the challenge for the estimation method is how to evaluate the integral in (2.7), since there is no analytical solution. In our analysis, we resort to simulation to approximate the integral, using Maximum Simulated Likelihood Estimation techniques. We use 100 Halton draws per subject. Following Conte, Hey, and Moffatt (2011), we assume the stochastic parameter ξ to be constant. For the Expected Utility model we therefore estimate 5 parameters, the mean and standard deviation of the risk coefficient r, the precision parameter ξ and the gender effects for both parameters.

2.3.3 Results

Table 2.2 reports the estimates from the three inference methods. The first column reports the results from the MLE, the middle from the MSLE and the last one from the HB model. For each parameter θ, we report the point estimate for the MLE, the mean of the distribution μ_θ for the MSLE, and the mode of the posterior distribution for the HB. The standard errors are reported in the Table, with the exception of the HB model where the standard deviation of the posterior distribution of each parameter is reported instead. The statistical significance is based on the respective confidence intervals (credible intervals for the HB).

TABLE 2.2: Estimates using the Holt and Laury (2002) data.

	MLE	MSLE	HB
r	0.289***	0.265***	0.292***
s.e.	0.022	0.021	0.033
r_{FEMALE}	0.103***	0.103***	0.103**
s.e.	0.034	0.021	0.050
σ_r	-	0.192	-
s.e.	-	0.000	-
ξ	0.253***	0.184***	0.086***
s.e.	0.014	0.016	0.014
ξ_{FEMALE}	0.050**	0.136	0.041
s.e.	0.023	0.193	0.333

The Table reports estimates from all three inference methods: Maximum Likelihood Estimation (MLE), Maximum Simulated Likelihood Estimation (MSLE) and Hierarchical Bayesian (HB). For each parameter θ, the Table reports the point estimate for the MLE, the mean of the distribution μ_θ for the MSLE, and the mode of the posterior distribution for the HB. Standard errors are reported (standard deviation for the HB). *p<0.1; **p<0.05; ***p<0.01

In all cases, the risk coefficient is positive and statistically significant, indicating risk averse preferences for all subjects. The coefficient of risk aversion ranges between 0.265 and 0.292 between the three inference methods, what Holt and Laury

(2002) characterise as "slightly risk averse". Focusing on the gender effects parameters, the coefficient is positive and statistically significant in all three cases, and remarkably at the same magnitude of 0.103. Finally, focusing on the precision parameter ξ, the effect of introducing more flexible inference methods to its magnitude, is apparent. The estimate of ξ using MLE is equal to 0.253 which is quite large compared to the other two methods. Since there is an inverse relationship between the size of ξ and the estimated noise (the lower the ξ the higher the precision) a larger estimate of ξ indicates issues with overfitting. As the inference methods become more flexible, the estimate of ξ becomes smaller, indicating more precise and less noisy estimates. The main conclusion from this analysis, is that by ignoring the between-subject heterogeneity, and estimating a model assuming a basic level of heterogeneity, as in the case of the MLE estimation, it is possible to detect the existence of gender differences, regardless of which estimation method is adopted. In what follows, we explore whether this result can be generalised when the complexity of the model increases. Filippin and Crosetto (2016) extend their analysis and investigate whether relaxing the expected utility assumption, has an effect to the inferred gender differences. By introducing a probability weighting function and non-expected utility preferences, they estimate a structural specification, using MLE, and show that the gender differences in the risk coefficient disappear, and they appear in the probability weighting parameter. As the original Holt and Laury (2002) task was not developed with non-expected utility preferences in mind, in the next section we repeat the same analysis as above, using data from an experiment which was particularly developed to identify risk preferences, stemming from both the curvature of the utility function and the shape of the probability weighting function.

2.4 Risk Preferences and Rank Dependent Utility

Motivated by the Allais paradox, a vast theoretical and experimental literature emerged, challenging the assumption of expected utility preferences (see Starmer, 2000 for a review of non-EU theories; Camerer, 1995 for an early discussion of the experimental work; and Hey, 2014 for a more recent review). In this section, we focus on one of the most influential alternatives to EU, the Quiggin (1982) Rank Dependent Utility model (RDU) which later led to the modification of the Original Prospect Theory model and the development of the Tversky and Kahneman (1992) Cumulative Prospect Theory model (which we explore in the next section). In the RDU model, attitudes towards risk are characterised by both the curvature of the utility function, and the shape of the probability weighting function, while there is evidence that the two components are not strongly correlated (Qui and Steiger 2011; Toubia et al. 2013). Therefore, given the extensive empirical evidence of the existence of non-EU preferences, it is crucial to take both components into consideration, when one investigates the existence of gender differences in risk preferences. We do so by using the data from Baillon, Bleichrodt, and Spinu (2020).

2.4.1 Decision Task and Data

Objective of this experiment was to identify the reference point that subjects are using when they make choices under risk. Each experimental task involved a choice between two paired lotteries again, A and B. An optimal design was employed to construct the questions of the experiment in a way that they would satisfy the following 5 criteria: (1)the questions must be diverse in terms of number of outcomes and magnitudes of probabilities involved, (2)the questions within each choice must have nonmatching maximal or minimal outcomes, (3) the questions must be diverse in terms of relative positioning in the outcome space, (4) they must have similar expected value to avoid trivial or statistically noninformative choice situations, and; (5) they must be "orthogonal" in some sense to maximise statistical efficiency. The number of the outcomes within each lottery varied between tasks, from 2 to 4 outcomes, all in the gains domain (strictly positive). An example of a task is provided below:

$$A = \begin{cases} 135, & \text{with probability } 0.55 \\ 290, & \text{with probability } 0.35 \\ 329, & \text{with probability } 0.10 \end{cases} \quad B = \begin{cases} 159, & \text{with probability } 0.05 \\ 259, & \text{with probability } 0.55 \\ 359, & \text{with probability } 0.10 \\ 409, & \text{with probability } 0.30 \end{cases}$$

The order of the tasks was randomised, and there was a total of 70 tasks per subject, with varying payoff and probability levels, generating a rich dataset for structural estimations. There are in total data from 139 subjects (49 females).[3] The experimental population consisted of students in Moldova, and the payoffs were expressed in the local currency. To incentivise the experiment, each subject had a one-third chance to be selected among all the subjects, to play out one of their choices for real. The experiment involved high stakes with payoffs up to a week's salary.

2.4.2 Theoretical Framework and Econometric Specification

As mentioned before, the RDU model consists of two components, the utility function and the probability weighting function, which transform every objective probability p to the decision weight $w(p)$ in the interval $[0, 1]$. We again assume a CRRA utility function, while for the probability weighting function, we assume the widely used Tversky and Kahneman (1992) function of the form:

$$w(p) = \frac{p^{\gamma}}{(p^{\gamma} + (1-p)^{\gamma})^{1/\gamma}} \tag{2.8}$$

where γ is the probability weighting parameter. The form of the function is inverse-S shaped for $\gamma < 1$, indicating overweighting of low probabilities and underweighting of moderate and high probabilities. To evaluate the RDU of a lottery, we first need to rank the outcomes of the lottery from the best to the worst, such that $x_1 \geq x_2, \cdots, \geq x_n$. The decision weight associated with each outcome is given by:

$$\begin{aligned}
\pi(x_1) &= w(p_1) \\
\pi(x_2) &= w(p_1 + p_2) - w(p_1) \\
&\dots \\
\pi(x_n) &= 1 - w(p_1 + p_2 + \cdots + p_n)
\end{aligned}$$

[3]For our analysis we use the data from 136 subjects as there were missing data on the gender of 2 subjects, and 1 subject had missing data.

The RDU of lottery A is then given by

$$RDU(A) = \sum_{n=1}^{N} \pi(x_n) \frac{{x_n}^{1-r}}{1-r} \tag{2.9}$$

We assume the same stochastic function as in Equation 2.3, by replacing the expected utility with the corresponding Rank Dependent Utility, and we form the log-likelihood function as in Equation 2.4 for the MLE estimation. As there are no multiple treatments, we control only for gender differences by introducing a gender dummy for all the parameters (r, γ, ξ, giving in total 6 parameters to estimate).

For the HB model, on top of the specifications for r and ξ, which are exactly the same as in the EU case, we need an additional specification for the γ parameter. This parameter must be positive, with a lower bound equal to 0.279 to ensure the monotonicity of the function. For the MSLE estimation, we need to estimate the parameters of the two distributions for r and γ, namely the means μ_r and μ_γ and their standard deviations σ_μ and σ_γ[4].

2.4.3 Results

Table 2.3 reports the results from all the three inference methods. The results are quite similar to what is usually observed in this literature. The estimated risk coefficient r is between 0.360 and 0.480, indicating moderate risk averse preferences, while the estimate for the probability weighting function is equal to 0.586 and 0.621, indicating an inverse-S shape of the function. While these results are quite uniform and the estimates look quite close in terms of magnitude and statistical significance, there are contradictory results regarding the presence of gender effects. Assuming heterogeneity only at the gender level (MLE) fails to capture any kind of effects for any of the parameters, while a same pattern is observed when MSLE is used to estimate the model. Nevertheless, when HB is used, one can infer that there is a significant difference between males and females in the way objective probabilities are transformed. With an estimate of γ equal to 0.705 (compared to 0.621 for men), it

[4]In the framework of MSLE, the coefficient vector θ is assumed to be normally distributed, across the population, with mean equal to a vector b and covariance matrix W. To maintain a manageable number of parameters, we assume that the off-diagonal elements of W are equal to zero and estimate the variance of each distribution. Allowing for correlation between the parameters led to worse performance of the model.

seems that women tend to exhibit lower probability distortion. Again, the effect of the different estimation methods on the precision parameter ξ is similar as in the Expected Utility case (the noise in the estimates decreases when more flexible inference methods are introduced).

TABLE 2.3: Estimates using the Baillon, Bleichrodt, and Spinu (2020) data.

	MLE	MSLE	HB
r	0.479***	0.424***	0.360**
s.e.	0.032	0.038	0.029
r_{FEMALE}	-0.025	-0.030	-0.054
s.e.	0.111	0.203	0.099
σ_r	-	0.291**	-
s.e.	-	0.122	-
γ	0.598***	0.586**	0.621***
s.e.	0.016	0.251	0.029
γ_{FEMALE}	0.050	-0.060	0.084*
s.e.	0.032	0.244	0.049
σ_γ	-	0.943	-
s.e.	-	0.821	-
ξ	0.121***	0.076***	0.083***
s.e.	0.001	0.007	0.010
ξ_{FEMALE}	0.041	-0.003	0.026***
s.e.	0.044	0.013	0.005

The Table reports estimates from all three inference methods: Maximum Likelihood Estimation (MLE), Maximum Simulated Likelihood Estimation (MSLE) and Hierarchical Bayesian (HB). For each parameter θ, the Table reports the point estimate for the MLE, the mean of the distribution μ_θ for the MSLE, and the mode of the posterior distribution for the HB. Standard errors are reported (standard deviation for the HB). *p<0.1; **p<0.05; ***p<0.01

The analysis above provides an example of the implications of ignoring heterogeneity between (as well as within) participants. While a basic MLE estimation provides no evidence of any kind of gender differences, allowing for a more informative approach reveals the existence of such differences. In the next section, we extend our analysis to one of the most important domains of decision theory under risk, that of loss aversion.

2.5 Risk Preferences and Loss Aversion

In this section we focus on the three components that characterise risk preferences in the losses domain, as these are articulated in the Tversky and Kahneman (1992) Cumulative Prospect Theory (CPT) model. The CPT model adopts a similar approach to the RDU model, on the way it handles monetary payoffs and probabilities, with the additional feature of *loss aversion*, the concept that "losses loom larger than gains". The results from the literature are mixed. Some studies find that women are more less loss averse (see Schmidt and Traub, 2002, Brooks and Zank, 2005), others that males are more loss averse (Booij, Van praag, and Kuillen, 2009), others that there is no difference (Harrison and Rutström, 2008), and others with a mixed result (Bouchouicha et al., 2019). As Bouchouicha et al. (2019) argue, currently, there is no consensus of what is the appropriate definition of loss aversion in the literature[5]. Nevertheless, for the sake of the example, we will focus on the CPT definition of loss aversion, while our approach can be extended to alternative definitions.

2.5.1 Decision Task and Data

To estimate a CPT specification when losses are present, we use the data from Bouchouicha et al. (2019) which is a subset from the data used in Vieider et al. (2015). There are in total observations of almost 3000 subjects, from 30 countries, on decision making under risk and ambiguity, in both the gains and losses domain. As our focus is on small samples, we use only the USA data. This set includes the choices of 95 subjects (47 females) in 12 choice tasks (6 in the gains domain, 5 in the losses domain, and 1 in the mixed domain to identify the loss aversion parameter). While there are available data on a larger set of risky tasks (28 tasks), we follow Bouchouicha et al. (2019) and use only the smaller subset for two reasons: (1) this set

[5] See Schmidt and Zank (2005) for the various definitions of loss aversion.

of tasks includes only 50:50 gambles, which allows the estimation of a functional-free probability weighting function, and; (2) estimating a structural model from a small set of observations per participant is one of the strengths of the Hierarchical approach, and this dataset allows to test the limits of this approach.

All tasks are in the form (x, y), representing the prospect of getting the monetary payoff x with probability 50% or y with the residual probability, with x and y being positive, negative or zero, depending on the task (Table 2.4 lists the 12 tasks). The subject had to express her *certainty equivalent* for each of the tasks. For the mixed domain prospect, the amount l was elicited, that would make the subject indifferent between a 50:50 gamble of (20,l) and the status quo of zero. The experiment was incentivised and an endowment equal to the largest possible loss was provides to the subject, to cover for potential losses.

TABLE 2.4: The tasks from Bouchouicha et al. (2019)

Gains	Losses	Mixed
(5,0)	(-5,0)	(20,-l)
(10,0)	(-10,0)	
(20,0)	(-20,0)	
(30,0)	(-20,-5)	
(30,10)	(-20,-10)	
(30,20)		

2.5.2 Theoretical Framework and Econometric Specification

We model preferences assuming a CPT decision maker. We employ a power utility function as before, of the form:

$$u(x) = \begin{cases} \frac{x^{1-r}}{1-r}, & \text{if } x \geq 0 \\ -\lambda \frac{(-x)^{1-r}}{1-r}, & \text{if } x < 0 \end{cases}$$

with r the risk coefficient, and λ the parameter of loss aversion. The status quo of zero is assumed as a reference point. We assume a common parameter for r for gains and losses, for two reasons: (1) there is extensive empirical evidence of no difference between the two domains (see Fox and Poldrack 2009), and; (2) to avoid any potential identification issues of the loss aversion parameter (see Wakker, 2010). As mentioned before, since only 50:50 gambles are used in the analysis, there is no need to specify a functional form for the probability weighting function. Therefore, we introduce two parameters to estimate, w_g and w_l, which represent the probability weighting for gains and losses respectively. Summarising, a prospect $L = (x, y)$ can be evaluated as:

$$U(L) = w_s u(x) + (1 - w_s)u(y)$$

with $s \in \{g, l\}$, while for the mixed prospect $L = (x, l)$, the prospect is evaluated as:

$$U(L) = w_g u(x) + w_l u(l)$$

The certainty equivalent $\hat{ce}$ for a prospect L is then given by:

$$\hat{ce} = u^{-1}[w_s u(x) + (1 - w_s)u(y)]$$

To form the likelihood function we need a different approach to the one used in the previous sections. In particular we assume that a decision maker states her certainty equivalent with some noise. The observed certainty equivalent of a subject in a task i is equal to $ce_i = \hat{ce} + \varepsilon_i$, where $\hat{ce}$ is the theoretical optimal certainty equivalent, for a set of behavioural parameters, and $\varepsilon \sim \mathcal{N}(0, \xi^2)$ with ξ being the standard deviation of the Fechner error (see Hey and Orme, 1994). We assume that this error is domain-specific (for mixed gambles we use the error for losses) and we also take into consideration a *contextual* error Wilcox 2011 by making the parameter ξ to be dependent on the difference between the best and the worst outcome of each prospect. That is, $\xi_i = \xi|x_i - y_i|$. The loglikelihood function for N subjects and I tasks is then given by:

$$LL(\theta) = \sum_{n=1}^{N} \sum_{i=1}^{I} \ln[\psi(\theta_n, L_i)] \tag{2.10}$$

with θ a vector of behavioural parameters to be estimated, L_i a task i and ψ the contribution to the likelihood function given by:

$$\psi(\theta_n, L_i) = \phi\left(\frac{\hat{ce}_{ni} - ce_{ni}}{\xi_{nis}}\right)$$

where ϕ is the standard normal density function. For the MLE estimation, we follow Bouchouicha et al. (2019) and we assume heterogeneity of the parameters at the gender level, and the domain level for the decision weights and the precision parameters. We need to estimate 12 parameters in total ($r, \lambda, w_g, w_l, \xi_g, \xi_l$ along with the controls for gender).

For the HB model, the specification of the likelihood function remains the same as in the MLE case. We specify distributions for the six parameters as above, with the decision weights constrained to the interval $[0, 1]$ and the loss aversion parameter to the interval $[0, 10]$, while for the MSLE, we estimate the parameters of the distributions for the risk attitude, the loss aversion and the probability weighting for gains and losses.

2.5.3 Results

Table 2.5 reports the estimates from the three inference methods. Three points are worth to mention: (1) there is significant loss aversion in this sample with a λ parameter statistically significant ranging between 1.596 and 1.672, (2) the risk coefficient is not statistically different than zero for the MLE and the HB cases, indicating a linear utility function, (3) the probabilities in the gains domain are distorted more that the probabilities in the losses domain (for instance, the decision weight of 0.5 is estimated to be 0.426 for gains and 0.478 for losses in the MLE case), and; (4) the control coefficient for gender differences is insignificant for all the major parameters of interest, in the MLE case with the exception of the noise parameter. Once again, using MLE techniques, one can conclude that there are no gender differences in the way females and males perceive monetary outcomes, transform probabilities to decision weights or perceive losses. Focusing on the more flexible methods of MSLE and HB, two points are interesting. First, the estimates of the mean, for all the parameters, are remarkably close between the two methods reinforcing the result of Huber and

Train (2001). Nevertheless, when gender effects are considered, while both methods find differences in the loss aversion parameters between the two groups, the MSLE methods fails to detect any gender effects in the key parameter or risk attitude. A potential explanation for this result could be the larger estimate of the precision parameter (a lower value indicates more precise estimates).

In this additional example, we provide further evidence that as the model complexity increases, by ignoring heterogeneity at the subject level, it may lead to incorrect inference regarding the difference between different demographic groups. Both methods that allow for this kind of heterogeneity (MSLE and HB) managed to detect the existence of such effects. Nevertheless, the results are not uniform. To identify which method is the most appropriate to use, in the next section we report the results of an extensive simulation exercise were we compare the performance of each of the methods.

2.6 Exploring the Advantages of HB Modelling

In the previous sections, we have shown that the identification on gender effects largely depends on the adopted inference method. We have provided a rigorous comparison of the representative agent model against two alternative methods that allow for extensive behavioural heterogeneity, even when the available sample size is small. Given that all three methods result is quantitatively different estimates, it raises the question of which method should one adopt. In this section we aim to provide an answer to this question, by means of an extensive Monte Carlo simulation exercise. Several studies have focused on the comparison between classical and Bayesian estimates, providing support on the latter (see for example Nilsson, Rieskamp, and Wagenmakers, 2011 or Gao, Harrison, and Tchernis, 2020). Here we repeat a similar exercise, suitably adapted to our objective of identifying gender differences in the elicited behaviour.

The main goal of this simulation study is two-fold. First, we want to confirm whether all estimation procedures are able to accurately recover the true parameter values from simulated data. Secondly, we test whether the inference methods under

TABLE 2.5: Estimates using the Bouchouicha et al. (2019) data.

	MLE	MSLE	HB
λ	1.596***	1.672***	1.615***
s.e.	0.099	0.112	0.113
λ_{FEMALE}	0.321	0.430***	0.398***
s.e.	0.211	0.141	0.136
σ_λ	-	0.467***	-
s.e.	-	0.044	-
r	-0.133	0.146**	-0.026
s.e.	0.07	0.061	0.019
r_{FEMALE}	-0.011	-0.061	-0.082^{*}
s.e.	0.115	0.083	0.041
σ_r	-	0.000	-
s.e.	-	0.053	-
w_g	0.426***	0.433***	0.444***
s.e.	0.019	0.023	0.017
$w_{gFEMALE}$	-0.009	-0.032	-0.037
s.e.	0.031	0.031	0.056
σ_{w_g}	-	0.38***	-
s.e.	-	0.051	-
w_l	0.478***	0.467***	0.501***
s.e.	0.019	0.022	0.010
$w_{lFEMALE}$	0.007	-0.010	-0.011
s.e.	0.031	0.031	0.037
σ_{w_l}	-	0.436***	-
s.e.	-	0.048	-
ξ	0.173***	0.178***	0.097***
s.e.	0.007	0.012	0.012
ξ_{FEMALE}	0.028***	0.000	0.051
s.e.	0.012	0.000	0.261
ξ_l	0.150***	0.112	0.063***
s.e.	0.006	0.008	0.007
$\xi_{lFEMALE}$	0.033***	0.000	0.055
s.e.	0.011	0.000	0.311

The Table reports estimates from all three inference methods: Maximum Likelihood Estimation (MLE), Maximum Simulated Likelihood Estimation (MSLE) and Hierarchical Bayesian (HB). For each parameter θ, the Table reports the point estimate for the MLE, the mean of the distribution μ_θ for the MSLE, and the mode of the posterior distribution for the HB. Standard errors are reported (standard deviation for the HB). *p<0.1; **p<0.05; ***p<0.01

consideration, are equally efficient in detecting gender effects. To make the simulation as general as possible, we focus on the Bouchouicha et al. (2019) design and the CPT model, which satisfies the conditions for which researchers usually resort to pool their data (a relatively large number of parameters to estimate using a relatively low number of data points per subject). For our exercise, we simulate data of 100 subjects which we then estimate using each of the three inference methods: MLE, MSLE and HB.

We assume that gender differences exist only in two of the model's parameters, the coefficient of loss aversion and the risk coefficient[6]. The parameters used in the simulation, are normally distributed across the experimental population with mean θ_n and standard deviation σ_θ. In the simulation we set the gender difference in the risk coefficient to be small but significant (mean of 0.500 for males and 0.600 for females) with a standard deviation equal to 0.05[7]. The loss aversion is set to 1.648 for males and 2.013 for females[8] with a standard deviation of 0.100. The probability weighting coefficient for gains w_g is set equal to 0.540 while the probability coefficient for losses w_l is set equal to 0.510. We assume no heterogeneity by setting the standard deviation equal to 0 for the weighting parameters and we also assume a common Fechnerian error for gains and losses. We conducted the simulation for three different levels of noise by setting the value of the error term equal to 0.130 (low noise), 0.150 (medium noise) and 0.200 (high noise). We report the results of the medium noise specification as they are the most representative[9]. For each simulation, we generate the data of the 100 artificial subjects by drawing parameters from the relevant distributions that were described above. This dataset was then estimated using each of the methods. Table 2.6 reports the results of 100 simulations. In particular, we report the mean and the standard deviation of the point estimates, in the case of MLE, the mean of the distribution means in the case of MSLE, and the

[6]We make this assumption in order to keep the simulation as simple as possible. Of course this analysis can be extended to any of the parameters of the model (i.e. probability weighting function, noise coefficient) since empirically, gender effects are observed in all components of preferences.

[7]We confirmed that the statistical significance of the two distributions is indeed significant based on a two-sided t-test ($p<0.000$).

[8]Since we transform the parameters to be drawn from a log-normal distribution, the values of loss aversion correspond to $\exp(0.500)$ for men, and $\exp(0.700)$ for women.

[9]Bouchouicha et al., 2019 using this dataset, estimate the noise parameter to be equal to 0.170. For our simulations, we are using a noise parameter of 0.150, which is in the middle of the interval between the low noise parameter (0.130) and the empirically observed parameter (0.170). We delegate the estimates from the low and high noise simulations to the online Appendix (see Tables B.1 and B.2).

mean of the posterior means of the distributions in the case of HB.

TABLE 2.6: Mean and standard deviations of the parameters.

Parameter	True value	MLE	MSLE	HB
λ	1.648	1.575	1.637	1.657
s.e.	-	0.056	0.062	0.067
λ_{FEMALE}	0.365	0.698	0.374	0.389
s.e.	-	0.132	0.118	0.129
σ_λ	0.100	-	0.081	-
s.e.	-	-	0.046	-
r	0.500	0.538	0.500	0.493
s.e.	-	0.019	0.031	0.025
r_{FEMALE}	0.100	0.064	0.105	0.108
s.e.	-	0.034	0.031	0.032
σ_r	0.050	-	0.046	-
s.e.	-	-	0.014	-
w_g	0.540	0.559	0.543	0.536
s.e.	-	0.013	0.014	0.014
w_l	0.510	0.528	0.510	0.510
s.e.	-	0.014	0.013	0.014
ξ	0.150	0.153	0.150	0.148
s.e.		0.007	0.005	0.006

The Table reports estimates from the simulation exercise on the three inference methods : Maximum Likelihood Estimation (MLE), Maximum Simulated Likelihood Estimation (MSLE) and Hierarchical Bayesian (HB), for the medium level of noise. For each parameter θ, the Table reports the mean of the point estimates, in the case of MLE, the mean of the distributions in the case of MSLE, and of the posterior mean of the distributions in the case of HB. Standard deviations in parentheses.

We first focus on the parameter recovery performance of each of the methods. The first column of the Table reports the true values of the coefficients that were used in the simulation. Compared to the true value, it is apparent that the MLE estimates have the worst performance in terms of precision. First, most of the parameters deviate significantly from the true value, compared to the other two methods. Then, in terms of gender effects, there is significant overestimation of the difference in loss aversion where the parameter is estimated to be almost twice the true value (0.698 compared to the true value of 0.365) while there is underestimation of the difference in the risk coefficient (0.064 compared to the true value of 0.100). As far as the MLE and MSLE estimates are concerned, both are remarkably close to each other and both have recovered the true parameters with quite high precision. The first conclusion from this simulation exercise is that if one is interested in the mean values of the parameters of different groups, then both MSLE and HB are equally good in recovering unbiased parameter values compared to the MLE.

We now turn to the identification of gender effects. For each of the simulations, we generate the 95% confidence interval (credible interval in the case of HB) to test the statistical significance of the estimate. When we focus on the gender effect for the risk coefficient, the MLE estimate is statistically significant for 55% of the simulations, the MSLE for 66% while the HB for 96%. Similarly, when we focus on the loss aversion parameter the MLE estimate is statistically significant for 53% of the simulations, the MSLE for 67% while the HB for 89%. Table 2.7reports the frequency with which statistically significant gender effects were detected, for each of the three inference methods, and for each of the three levels of noise (low, medium and high). The Table confirms the pattern that higher levels of noise lead to lower detection levels of gender effects, with MLE having the worst performance, HB the best, and MSLE in the between.

TABLE 2.7: Identification of gender effects.

	r_{FEMALE}		
	$\xi = 0.130$	$\xi = 0.150$	$\xi = 0.200$
MLE	62%	55%	36%
MSLE	82%	66%	59%
HB	98%	96%	76%
	λ_{FEMALE}		
	$\xi = 0.130$	$\xi = 0.150$	$\xi = 0.200$
MLE	63%	53%	37%
MSLE	81%	67%	50%
HB	97%	89%	66%

The Table reports the rate of success of each inference method to identify gender effects for each of the three levels of noise, for the gender specific parameter for risk attitude (r_{FEMALE}) and loss aversion (λ_{FEMALE}) are statistically significant, at the 5% level.

Our results mirror the conclusions of Huber and Train (2001). In this study the authors compare classical and Bayesian estimates by providing a comparison between MSLE and HB. They show that both methods result in virtually equivalent conditional estimates of the parameters. Then, they provide a list of differences between the two methods including (1) the difficulty of MSLE to locate the maximum of the likelihood function; (2) the computational burden that the variance-covariance matrix poses to the estimation of the MSLE parameters, and; (3) the identification issues that the classical approach faces compared to the Bayesian estimation. Our simulation shows that when the identification of differences between different populations is the objective, then HB is the clear winner as the most appropriate inference method. This result can be attributed to the way each of the methods handles uncertainty in the estimates and the fact that the estimate of the unobserved heterogeneity in the MSLE estimates is much noisier (larger standard errors) compared to the HB ones.

To investigate the role of the sample size in the detection of gender effects, we ran some additional simulations for the MSLE methods, varying the sample size.

Assuming a fixed level of noise ($\xi = 0.150$), we repeated the simulation exercise for $N = 200$ and $N = 500$ and again we report the rate of success to identify gender effects for the risk attitude (r_{FEMALE}) and the loss aversion (λ_{FEMALE}) gender specific parameters[10]. When the sample size is equal to 200, the risk (loss aversion) coefficient is significant for 89% (90%) of the simulations, while when the sample size increases to 500, the risk (loss aversion) coefficient is significant for 93% (94%) of the simulations. This analysis further highlights the advantages of the HB modelling since this inference method needs only half of the sample that MSLE needs in order to achieve the same detection rate of success in the case of loss aversion, while it needs only one fifth of the sample that MSLE needs, to reach the same success rate, in the case of the risk coefficient.

2.7 Concluding remarks

In this study, we focus on gender differences and compare the inference made by three econometric methods, Maximum Likelihood Estimation, Maximum Simulated Likelihood Estimation and Hierarchical Bayesian modelling, on three representative domains of risk preferences. We show that when all the data are assumed to come from a representative agent, and assume heterogeneity (gender differences or any other demographic differences) at a very basic level (e.g. all black females have the same level of loss aversion), valuable information might be ignored, and therefore, distorted conclusions may be drawn. Nevertheless, opting for a more flexible approach, and taking into consideration both the individual variation and the population-level characteristics, the inference about individual risk preferences is massively improved, and significant differences are captured.

In particular, we compare the representative agent modelling approach, to two more flexible and informative methods of parameter estimation that allow one to simultaneously make inferences at both the individual subject and the experimental population level. We compare the frequentist and the Bayesian methods, by analysing the data using Maximum Simulated Likelihood Estimation techniques (MSLE), as well as Hierarchical Bayesian (HB) econometric modelling. We use data

[10]The estimates are delegated to the online Appendix (see Table B3). There, it can be seen that as the sample size increases, the standard errors decrease, which allows for better identification of the effects.

from three representative studies on decision making under risk and we study Expected Utility preferences, for a simple analysis of risk attitudes, Rank Dependent Utility preferences, to incorporate probability weighting, and Cumulative Prospect Theory, to investigate loss averse behaviour. We show that by ignoring heterogeneity at the subject level, it may lead to incorrect inference regarding the difference between distinct demographic groups.

Recent research on Hierarchical Bayesian modelling has shown that MLE estimates are both susceptible to overfitting and dominated by outliers (Nilsson, Rieskamp, and Wagenmakers, 2011, Murphy and Brincke, 2018), while Bayesian modelling improves the robustness of the estimation, by shrinking the parameters towards the group's mean. This method allows the robust estimation of preferences, and it is particularly useful, especially when one has a limited number of data points from each subject, as is often the case with field studies, or when additional tasks are used, along with the main experimental design, to control for particular preferences. With the aid of an extensive simulation exercise, we show that Bayesian methods are better placed to capture differences between groups, and this result can be attributed to the way that each of the methods handles uncertainty in the estimates.

In this study, we do not argue in favour of any particular preference functional or model, nor we claim that there is a uniform pattern of gender differences. In our analysis, we opted for the models and the preference functionals that are often assumed in this literature. These models acted as "vehicles" to illustrate the machinery behind both estimation techniques, and this approach could be extended to any alternative model. Our main objective is to warn researchers on the dangers of small sample datasets and ignoring heterogeneity of the subjects. Of course this method could be extended to other important fields of decision making such as ambiguity preferences, time preferences or social preferences. Even more, as Gao, Harrison, and Tchernis (2020) highlight, HB methods are particularly useful when one is interested in joint estimation of perhaps non-correlated preferences (e.g. joint estimation of risk and time preferences) where the need of robust estimates is important at the individual level.

Chapter 3

Disentangling the Effect of Time-Pressure on Risky-Decision Making

3.1 Introduction

Pressure is a naturally occurring part of most people's everyday life, appearing in places such as time limits for exams, coursework or work project deadlines or competing as part of a sports competition. As a result of having to perform an activity under pressure, an individual may become flustered and not perform as well as they would if they were carrying the same task without the pressure existing (De Paola and Gioia, 2016). However, there could be situations where this pressure allows individuals to thrive and may perform better, such as in some athletes in competitions or students cramming before an exam.

Risk attitudes also play an essential part in our decision-making; for example, those opposed to risk will require a much larger incentive than others to partake in any form of gambling or risky activity. However, a person's attitude to risk may change in the presence of time pressure. They may become stressed and may make decisions they would not make otherwise, for example, acting in a more risk-seeking or reckless way. Several papers in the literature consider the effect of time pressure on risky decision-making. Most focus on binary choices between two lotteries, such as a safe and risky gamble. We believe that these kinds of tasks have the limitation of preventing an individual from fully expressing their preferences and, by extension, their attitudes towards risk as we lack information about how closely valued alternatives are. We work around this limitation by considering a budget allocation task in which individuals can choose their preferred gamble and, by extension, control the level of risk they are exposed to, allowing us to observe their preferences better.

In this paper, we wish to investigate time pressure's effect on an individual's attitude towards risk. We do this by carrying out an experiment where subjects face multiple budget allocation tasks with and without time pressure. We then structurally recover an individual's behavioural parameters associated with risky decision-making, such as the curvature of a utility function and probability distortion and compare how they differ with and without time pressure to determine how their attitudes towards risk change. We do this by considering a non-linear least squares method and see that, on average, an individual is less risk-averse under time pressure.

3.2 Literature Review

Time pressure has long been understood to impact an individual's behaviour under risk. Under time pressure, an indiviudals attitude towards risk may "flip", becoming more risk-seeking for gains and risk-averse for losses (Saqib and Chan, 2015). We have seen individuals exhibit less risk-averse behaviour under time pressure (Huber and Kunz, 2007), (Hu et al., 2015), (Madan, Spetch, and Ludvig, 2015), (Nursimulu and Bossaerts, 2014), (Chandler and Pronin, 2012), (Olschewski and Rieskamp, 2021), (Hu et al., 2015). We also note there have been some papers with increased levels of risk-averse behaviour in the presence of time pressure (Ben Zur and Breznitz, 1981), (Haji et al., 2019). However, different levels of risk may explain this difference in risk attitude under time pressure (Dror, Basola, and Busemeyer, 1999), where people have been observed to act more conservatively with low levels of risk and more risk-seeking with higher risk levels. Other results from the literature on time pressure on risky decision-making include time pressure resulting in a higher degree of probability distortion (Young et al., 2012), and increased levels of loss aversion (Kocher, Pahlke, and Trautmann, 2013). Time pressure has also had an impact on behaviour in environments such as ambiguity (Baillon et al., 2018), auctions (Haji et al., 2019) and game theory (Lindner and Sutter, 2013). We have also seen time pressure impact the real world, such as crossing roads in more dangerous situations (Cœugnet, Cahour, and Kraiem, 2019).

Outside of risk, several factors also impact an individual's attitudes towards risk. These include include age (Mather et al., 2012), religion (Bartke and Schwarze, 2008) and education (Rosen, Tsai, and Downs, 2003). Other things shown to alter risk attitudes include experiencing traumatic events (Moya, 2018) and the recent pandemic (Heo, Rabbani, and Grable, 2021).

We also note that there is the notion of a speed-accuracy trade-off, where an individual can make a decision quickly or accurately. The trade-off relates to time pressure, where individuals have less time to decide or carry out a task. We can see this in papers such as Hobbs and Williamson (2003), showing more mistakes in maintenance under time pressure, in medical diagnostics (Hellyer, 2019),(Betsch et al., 2004), where we see time pressure, leads to more errors but error rate falls as time pressure becomes less severe. We can also see that people resort to using heuristics as time pressure increases (Lewis et al., 2014), (Rieskamp and Hoffrage, 2008).

The paper by Young et al. (2012) considers rank-dependant utility and certainty equivalence under time pressure, which is a similar premise to what we investigate in this paper. However, we differ in this paper in several ways to contribute to the literature differently. For example, our subjects faced both time-pressure and non-time-pressure tasks, which allowed us to see how time pressure impacted different individuals' attitudes towards risk compared to having two groups of subjects. We also use certainty equivalence in separate ways; we calculate certainty equivalence from an individual's behavioural parameters, whereas Young et al. used an individual's self-reported certainty equivalence to recover behavioural parameters (the reverse of what we are doing). We also do not scale down payoffs; subjects had their payoffs scaled down to 1/5 of the gambles, which can potentially cause differences given that subjects will behave differently in high and low-stakes situations (Dror, Basola, and Busemeyer, 1999). Finally, we also use fewer tasks; they used a large number of problems in their experiment, which can lead to decision fatigue and what is called the "Ramsey Trifle Problem", where a slight increase in the increased average payoff from acting in a more computationally intensive way is not worth the increased payoffs[1].

We differ from the broader literature by moving away from binary choice exercises and considering a budget allocation choice. We do this as we feel it can allow individuals to choose their gamble and the level of risk they would be willing to take and better express their risk attitudes. Compared to binary choices, where we do not know how strongly preferred the chosen gamble is. We can disentangle the effect of time pressure on an individual's behavioural parameters to see which parameters change, for example, the curvature of a utility function of the degree of probability distortion.

3.3 Experimental Design

Our experiment is by the experiments proposed in Choi et al. (2007) and Choi et al. (2014), in which an individual faced a budget allocation problem. They had to invest tokens into two risky accounts and then receive a payoff based on the number of tokens invested and an account-specific exchange rate for the chosen account, randomly selected with a 50-50 chance. We extend this design in several key ways:

1. We have two different sets of tasks for individuals to face, one with time pressure and one without time pressure.
2. We use a wide range of probabilities to cover the whole probability spectrum and estimate a probability weighting function.
3. A different interface and way to represent the lotteries.

3.3.1 Experimental Tasks

In our experiment, subjects were given 100 tokens to allocate between two accounts. Each account had an associated price to convert tokens into real money. For example, a price of 6 would require investing 6 tokens to receive £1, or the exchange rate between tokens and real money is the inverse of this price. The two accounts have a probability of being chosen of P and 1-P and, when selected, will pay out money

[1]See Abdellaoui et al. (2020) for more information on this.

based on the number of tokens invested and the exchange rate. For example, accounts A and B have probabilities of being chosen of 0.7 and 0.3 and prices of 5 and 10 (exchange rates of 0.2 and 0.1). If an individual invests 80 tokens into A and 20 into B, they will receive £16 with a probability of 0.7 and £2 otherwise.

The tasks subjects faced in the experiment were randomly generated ahead of time. The probabilities of accounts being chosen for each round were drawn randomly from a list of probabilities[2]. Prices were also randomly generated by generating an integer between 5 and 15. We generated a total of 40 tasks, 20 to be used under time pressure and 20 without[3]

Subjects faced 2 rounds of 20 tasks, one of these rounds were performed under time pressure in which subjects had 15 seconds to make and submit a decision. The second round of tasks was carried out without time pressure, where an individual had up to a minute to submit their decision but could not submit for the first 15 seconds to ensure subjects had a longer decision time than under time pressure. There was a small break between the sets of tasks where subjects were made aware of the change in decision time. Should an individual not submit within the time, the currently selected allocation bundle was submitted.

3.3.2 Interface

The interface subjects faced consisted of a few key elements, summarised in the table below:

TABLE 3.1: Summary of Interface Elements

Element	Use
Timer	Shows the amount of time remaining in the current task
Slider	Selecting how many tokens to allocate to each account
Pie Chart	Visualise the probability of each account being selected
Bar Chart	Shows the payoffs of each account, and the width represents the probability
Submit	Submit the choices and move on to the next task
Text	Shows information on rounds, tokens and prices

[2]The list of probabilities were drawn randomly from 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.25, 0.75, 1/3 and 2/3.

[3]The set of tasks faced by subjects can be found in the Appendix C.1.

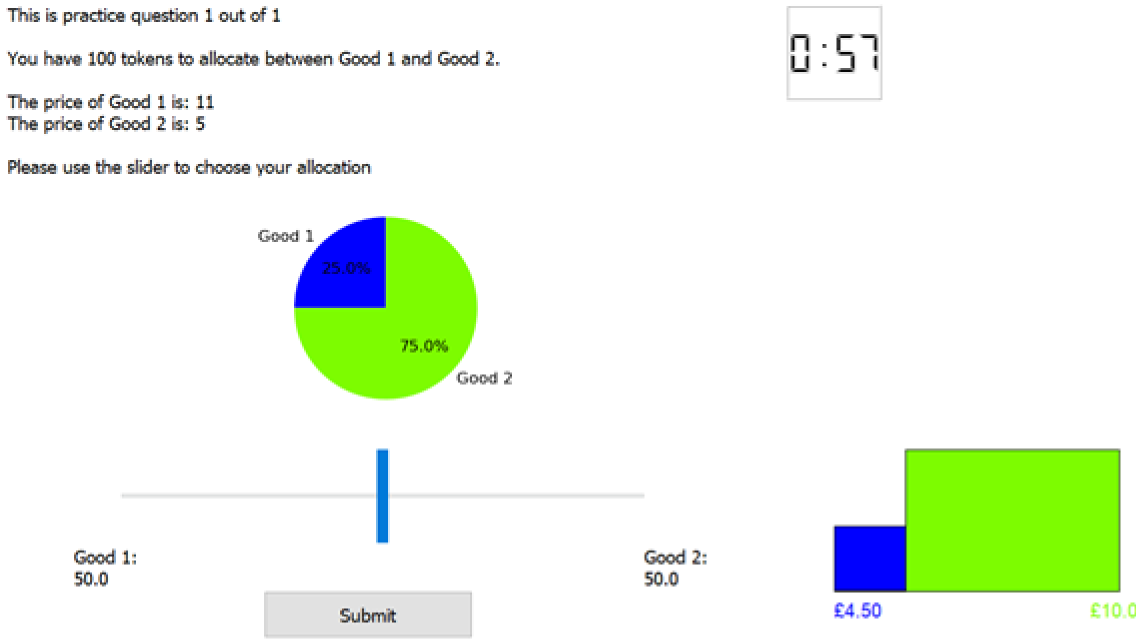


FIGURE 3.1: Example of the experiments interface

Figure 3.1 above shows a snapshot of what the subjects faced. They can drag the slider around to control the number of tokens in each account. The slider completely to the left will allocate all 100 of their tokens into Account 1. Whereas the slider fully to the right allocated all tokens to Account B. Should the slider be at a point between, it will result in X into Account 1 and 100-X into Account 2 to ensure all 100 tokens are used. At the selected allocation, an individual would receive $\frac{X}{p_1}$ if account 1 is selected $\frac{100-X}{p_2}$ if account 2 is selected. Tokens can also be divided into units of 0.1 to allow for better freedom to express preferences. As the slider moves, the bar charts change in height and show the payoffs, with the width of the bar chart also representing the probability (the larger the probability, the wider the bar will be to scale). Once an individual is happy with their selection, they click the submit button and move on to the next task. In the example above, Accounts 1 and 2 have prices of 11 and 5 and probabilities of 0.25 and 0.75 of being selected, respectively. The current allocation, 50 tokens in each, yields payoffs of £4.50 with a probability of 0.25 and £10 with 0.75.

3.3.3 Procedure, Incentives and Payments

The experiment was carried out at the LExEL laboratory at Lancaster University in December 2019. Subjects were predominantly students of the university, drawn randomly from the pool of subjects that the laboratory has. 51 subjects (mean age 22 and 20 of which were female) were split into 4 sessions, two of these sessions were randomly selected to have the time pressure tasks carried out first, and the others had time pressure second. Subjects received a set of written instructions[4] which explained what to do and how to use the interface. They could then carry out 3 practice rounds to ensure they were familiar with the interface, which had no time limit. Following this, subjects then moved on to the actual experimental tasks. Subjects were paid a show-up fee (£3), plus the outcome of one task being randomly selected and played out for its actual value. Once the experiment was finished, subjects were taken to a separate to sort out payments. For this, they picked out numbers from different hats and followed this procedure:

1. Either 1 or 2 to represent either the first or second round of tasks.
2. A second from 1-20 to represent which task is used in the selected round of tasks.
3. A final number from 1-100 to represent probability.
4. If the probability number drawn is less than the probability of option A in the selected task, then receive the payoff from Account A OR B otherwise

For example, if a task was selected in the experiment with payoffs of £10 with a probability of 0.8 and £5 with a probability of 0.2, if they drew the number 73, A was selected, and they received £10. The maximum amount a subject could win was £20 plus a show-up fee of £3.

3.4 Theoretical Model

We assume Prospect Theory (Kahneman and Tversky, 1979) to model an individual's preferences. This differs from the standard expected utility model by allowing

[4] Available in Appendix C.1.

for both the curvature of the utility function and probability distortion. Suppose an individual is faced with a lottery of X with a probability of P and Y with a probability of 1-P. Under the standard expected utility model, the utility of this gamble is given by $P * U(X) + (1 - P) * U(Y)$, where U(.) is the utility function. Under prospect theory, these probabilities are transformed by a weighting function $w(.)$ to yield decision weights. to calculate the decision weights and the Rank Dependant Utilities (RDU), we order a given gamble's payoffs such as $x_1 \geq x_2, \cdots, \geq x_n$. Then the decision weights can be given by:

$$\begin{aligned}
\pi_{x_1} &= w(p_1) \\
\pi_{x_2} &= w(p_1 + p_2) - w(p_1) \\
\cdots & \\
\pi_{x_n} &= 1 - w(p_1 + p_2 + \cdots + p_n)
\end{aligned}$$

The RDU of lottery A is then given by

$$RDU(A) = \sum_{n=1}^{N} \pi(x_n) U(x_n) \tag{3.1}$$

In our experiment, we only have two possible payoffs, which can simplify this to:

If $X > Y$, then this rank dependant utility is given by:

$$RDU = w(P) * U(X) + (1 - w(P)) * U(Y)$$

Similarly, if $Y > X$, then the rank-dependent utility becomes:

$$RDU = (1 - w(1 - P)) * U(X) + w(1 - P) * U(Y)$$

The utility function we consider is the Constant Relative Risk Aversion (CRRA) function:

$$U(W) = \frac{W^{1-\rho}}{1-\rho} \tag{3.2}$$

In this, the parameter ρ represents risk attitude. When $\rho = 0$, an individual is risk neutral, and the utility function is linear; when $\rho > 0$, an individual is increasingly risk averse with an increasingly concave utility function. The function can be defined for $rho < 0$, however, we chose not to consider this as it implies risk-seeking behaviour. This paper does not consider risk-seeking behaviour as our experimental design does not allow for this. It would require investing over 100 tokens into one account and a negative amount (making a loss) into the other. We chose to use this function for the previous reason above and power functions being shown to represent preferences effectively (Balcombe and Fraser, 2015b).

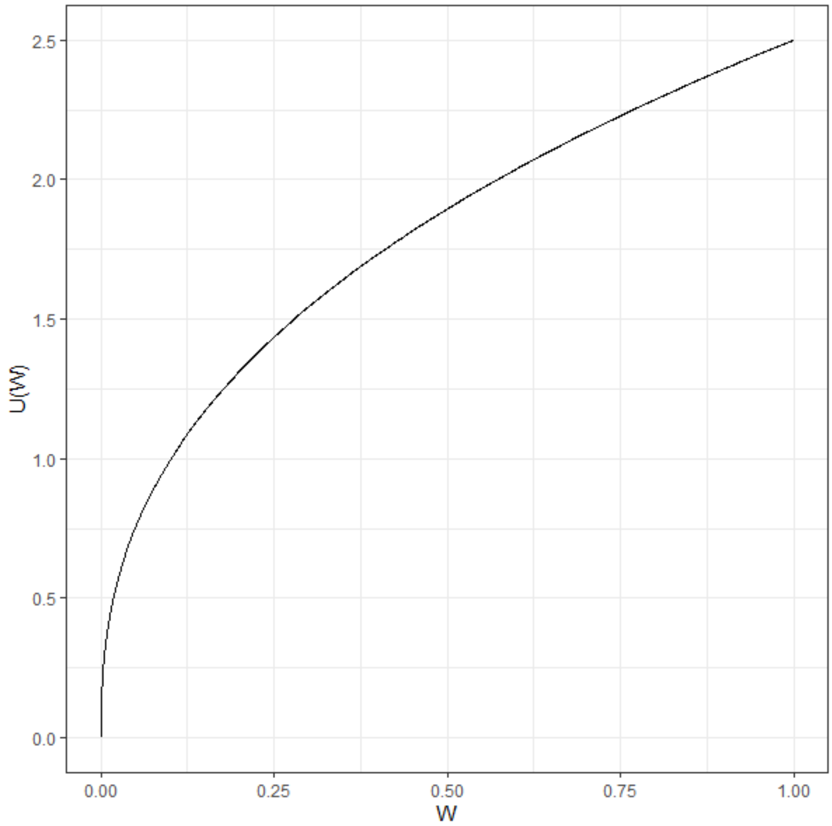


FIGURE 3.2: Visualisation of the utility function with $\rho = 0.6$

For the probability weighting function, we use a two-parameter Prelec weighting function (Prelec, 1998)[5].

$$w(P) = \exp\left(-((-\gamma \ln(P))^{\delta})\right) \tag{3.3}$$

Unlike many weighting functions, this has two parameters. The δ represents the size of probability distortions. The closer to zero, the more pronounced the "inverse-s-shape" the function becomes, which larger overweighting of low probabilities and underweighting of larger probabilities, the closer to one, the more linear. The second

[5]For robustness, we also consider some other weighting functions proposed in the literature. These results can be found in Appendix C.4.

parameter, γ, can be interpreted as a degree of optimism. The smaller γ, the larger the set of probabilities are over-weighted and the fewer that are under-weighted

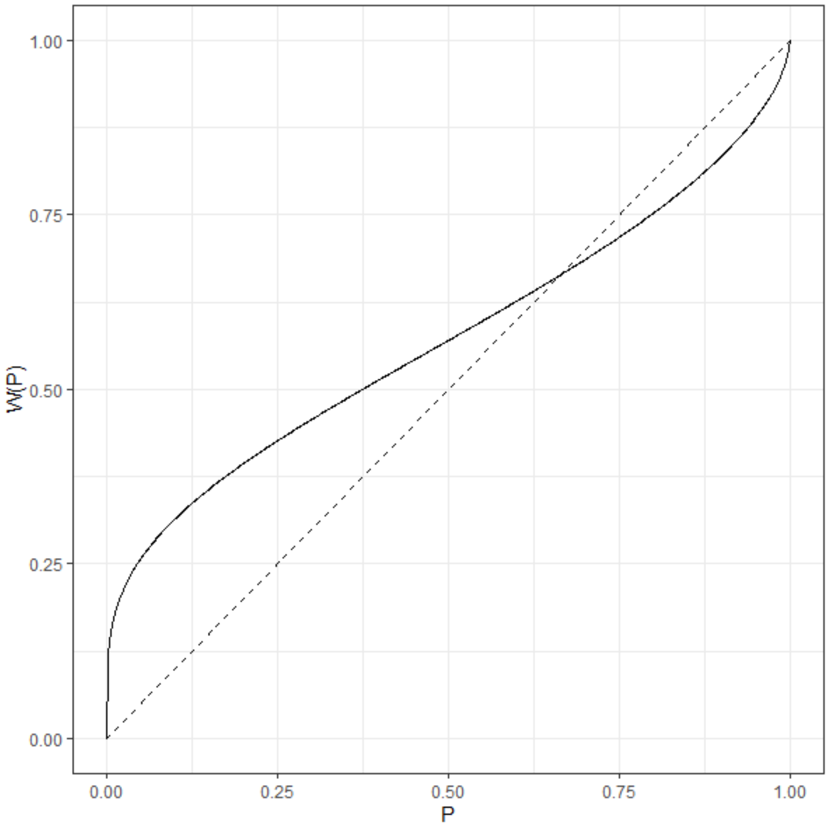


FIGURE 3.3: Visualisation of the weighting function with $\delta = 0.6$ and $\gamma = 0.7$

The diagram above shows a visualisation of the weighting function, as δ increases, the closer to 1, the more linear this is, indicating less distortion in probability. As γ increases, the lower the intersection of the weighting function and the 45-degree line, showing that more probabilities are underweight, and fewer are over-weighted.

If $X > Y$, we combine the utility and the weighting functions, and we can calculate the rank-dependent utility to be given as:

$$RDU = exp(-((-\gamma \ln(P))^{\delta})) * \frac{X^{1-\rho}}{1-\rho} + (1 - \exp(-((-\gamma \ln(P))^{\delta}))) * \frac{Y^{1-\rho}}{1-\rho} \quad (3.4)$$

and a similar equation if $Y > X$:

$$RDU = (1 - \exp(-((-\gamma \ln(1-P))^{\delta}))) * \frac{X^{1-\rho}}{1-\rho} + \exp(-((-\gamma \ln(1-P))^{\delta})) * \frac{Y^{1-\rho}}{1-\rho} \tag{3.5}$$

It is also important to note that we cannot use the drift-diffusion model illustrated in chapter one for this analysis. This is due to the Drift-Diffusion model is currently only suitable for binary choices [6] or some cases of trinary choices, which is not suitable due to the vast amount of ways that an individual can allocate their tokens. The second reason for this is for the tasks without time pressure, subjects were unable to submit their choices for the first 15 seconds of decision tasks. This would make it difficult to determine how long a decision took if they made it in less than 15 seconds as they had to wait to submit.

3.4.1 Optimal Solutions

We assume that individuals are utility maximisers, aiming to maximise their utility given any restrictions. In this case, subjects face a budget restriction, where they have a maximum of 100 tokens to invest between both accounts. This gives a feasibility constraint of $X + Y = 100$, where if they invest all 100 tokens into one good, they receive 100 divided by the price of that account. When maximising their utility, the bundle chosen will be at the point tangential to the budget constraint and the indifference curve representing all bundles with the same expected utility.

[6] We do make sure to save the decision time data for individuals to use for future work where this is possible potentially.

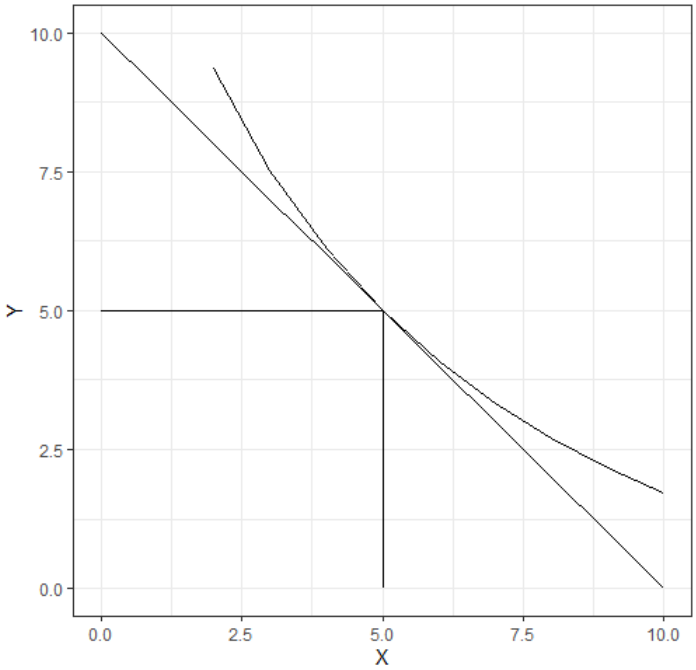


FIGURE 3.4: Visualisation of Utility Maximisation

Figure 3.4 above illustrates this tangency point. If a point on the budget curve will provide less utility for the same level of expenditure and points above and to the right of this point give a higher level of utility but are unobtainable with the current budget.

For our experiment and framework, subjects had 100 tokens to invest into the two accounts, X and Y, with probabilities P and 1-P. The payoff of an account is determined by dividing the number of tokens by the account price. This leads to payoffs of $X * e_x$ and $Y * e_y$ for each account, where e_x and e_y are $\frac{1}{p_x}$ and $\frac{1}{p_y}$ respectively, the exchange rate between tokens invested into a given account and the associated payoffs. The maximisation problem subjects face can then be written as:

$$\max \pi_x * \frac{(x*e_x)^{1-\rho}}{1-\rho} + \pi_y * \frac{(y*e_y)^{1-\rho}}{1-\rho}$$
$$\text{s.t } x + y = 100$$

where $\pi_x = w(P)$ if $X * e_x > Y * e_y$ and $\pi_x = 1 - w(1-P)$ if $Y * e_y > X * e_x$. Suppose we assume the weighting function and utility function used above. In that case, we can maximise this and yield demand equations for accounts X and Y[7]

[7]Derivations of these equations can be found in the Appendix C.2.

$$X(\rho,\delta,\gamma) = \frac{100 * (\pi_x * e_x^{1-\rho})^{\frac{1}{\rho}}}{(\pi_x * e_x^{1-\rho})^{\frac{1}{\rho}} + (\pi_y * e_y^{1-\rho})^{\frac{1}{\rho}}} \tag{3.6}$$

$$Y(\rho,\delta,\gamma) = \frac{100 * (\pi_y * e_y^{1-\rho})^{\frac{1}{\rho}}}{(\pi_x * e_x^{1-\rho})^{\frac{1}{\rho}} + (\pi_y * e_y^{1-\rho})^{\frac{1}{\rho}}} \tag{3.7}$$

3.4.2 Considering Time Pressure

In the previous subsection, we have illustrated the model to represent risky decision-making. We wish to change this to consider how this risk attitude changed in the presence of time pressure. We can do this by tweaking the behavioural parameters above to include a dummy variable which allows the behavioural parameters of the model to take different values on exercises with time pressure and without it. The specification for the behavioural parameters we consider can be shown below:

$$\theta = \theta_{NTP} + \alpha_{TP}\theta_{TP} \tag{3.8}$$

Where θ is our parameter of interest (ρ, δ and γ). α_{TP} is a dummy variable that takes a value of 1 if a subject faces time pressure in a particular exercise and 0 on takes without time pressure. θ_{TP} then represents the difference between an individual's attitude with and without time pressure. To ensure that an individual exhibits a different behaviour with and without time pressure, we require that at least one of the time pressure behavioural parameters ($\rho_{NTP}, \delta_{NTP}, \gamma_{NTP}$) to be statistically different from zero.

3.5 Estimating the Model

We use a non-linear least squares (NLLS) method to recover the parameters of interest. We fit parameters of interest that minimise the sum of the differences between individuals observed chosen bundle and the allocation predicted by demand equations 3.6 and 3.7.

$$NLLS(\theta) = \sum_{i=1}^{N}(X_i^* - X_i^C)^2 + (Y_i^* - Y_i^c)^2 \tag{3.9}$$

Where X_i^C and Y_i^C represent the observed chosen token allocation an individual selects in round i and X_i^* and Y_i^* represent the predicted optimal token allocation from the demand equations, for a given set of behavioural parameters ρ, γ and δ. N is the amount of decision tasks subjects faced; in this case, N=40. We estimate these parameters in R and use the package Rsolnp(Ghalanos and Theussl, 2015), which uses multiple restarts and starting values in its optimisation strategy. The advantage of this is that the multiple restarts allow for us to be sure that we have converged to the global optimal of our equation 3.9 and not a local minimum.

3.6 Experimental Results

3.6.1 Parameter Estimates

Despite a majority of the behavioural parameters for time pressure being statistically different from zero(shown in the table below), we find that the actual distribution of the behavioural parameters with and without time pressure (θ_{NTP} and $\theta_{NTP} + \theta_{TP}$) to be similar. This is shown by the P-values of Wilcoxon tests of 0.28658, 0.1121 and 0.8845 for ρ, δ and γ respectively[8]. A possible explanation is that time pressure impacts people differently, with some becoming more and some becoming less risk-averse. These individual changes cancel each other out at the aggregate level. There will also exist a sizable portion of individuals for each group with multiple time pressure parameters being statistically zero, which can cause these distributions to be similar. As a result of this, we consider the aggregate distributions to be an unsuitable way to determine if people are more or less risk-averse under time pressure and move to consider how individuals' parameters change.

[8] Plots of the parameter distributions can be found in Appendix C.3.

TABLE 3.2: Summary of behavioural parameters

Mean ρ_{NTP}	1.37500507
Mean ρ_{TP}	0.10864338
Mean δ_{NTP}	1.02516864
Mean δ_{TP}	0.21754409
Mean γ_{NTP}	1.19760454
Mean γ_{TP}	-0.01214881
Standard Deviation ρ_{NTP}	1.3074645
Standard Deviation ρ_{TP}	0.6348932
Standard Deviation δ_{NTP}	0.6288316
Standard Deviation δ_{TP}	0.7745950
Standard Deviation γ_{NTP}	0.4718599
Standard Deviation γ_{TP}	0.5957599
Statistically Significant ρ_{TP}	29
Statistically Significant δ_{TP}	35
Statistically Significant γ_{TP}	27
No Statistical Significance	8

Table 3.2 above shows a summary of the parameters we recover and whether or not their time pressure parameter is statistically significant. We can see that most individuals have at least one of their time pressure parameters statistically different from zero[9]. We find 8 subjects in which all the time pressure behavioural parameters are statistically 0, indicating the same behaviour with and without time pressure.

Considering how these behavioural parameters move under time pressure, it is difficult to deduce whether an individual has become more or less risk averse when they experience time pressure. This is because their behavioural parameters might move in different ways. For example, their utility function becomes more concave ($\rho_{TP} > 0$), indicating an increased level of risk aversion, and their weighting function becomes more linear ($\delta_{TP} > 0$), suggesting less risk aversion. To tell if an individual has become more or less risk averse under time pressure, we consider Certainty

[9]This is shown by a wald test, significant at the 5% level.

Equivalence as a measurement of risk attitude and how it moves to determine how their risk attitudes change with and without time pressure.

3.6.2 Considering Certainty Equivalence

Since the behavioural parameters can move in different ways under time pressure, it becomes difficult to determine if an individual has become more or less risk averse. We can solve this problem by proposing a theoretical gamble, then using the recovered behavioural parameters, calculate the certainty equivalence of said gamble with and without time pressure. Certainty equivalence is the minimum money required to make an individual indifferent between taking the gamble and the certain payoff. A lower level of certainty equivalence would indicate an increased risk aversion as they would prefer to take a lower level of certain money than partake in a gamble. For this exercise, we propose a 50-50 gamble between £10 and £5. We can then calculate the certainty equivalence of this gamble by finding the amount of money such that the utility of said money is equal to the expected utility, with and without pressure, by solving the following equation:

$$U(CE) = w(0.5)U(10) + (1 - w(0.5))U(5) \tag{3.10}$$

Substituting our parameters into equation 3.10 and rearranging yields an equation for certainty equivalence.

$$CE = [w(0.5)10^{1-\rho} + (1 - w(0.5))5^{1-\rho}]^{\frac{1}{1-\rho}} \tag{3.11}$$

We generate this value for individuals with and without pressure using their recovered behavioural parameters ($\theta_i = \theta_{i,NTP}$ without time pressure and $\theta_i = \theta_{i,NTP} + \theta_{i,TP}$ with time pressure). Once these values have been calculated, we can compare how they differ with and without time pressure to determine when they are more or less risk-averse. If, under time pressure, the certainty equivalence value we generate is lower than without time pressure, then an individual requires less money to be indifferent and is, therefore, more risk averse under time pressure. We consider an individual to have a different value for certainty equivalence and, therefore, different risk attitudes if at least one of their behavioural time pressure parameters is statistically different from 0.

TABLE 3.3: Summary of Certainty Equivalence Results

Mean Certainty Equivalence without Time Pressure	6.8002
Mean Certainty Equivalence under Time Pressure	7.0211
More Risk Averse without Time Pressure	24
More Risk Averse under Time Pressure	13
No Change in Risk Attitudes	8

We can see from Table 3.3 that, on average, a higher level of certainty equivalence under time pressure suggests that individuals exhibit less risk-averse behaviour under time pressure. When considering individual parameter changes, we can see that most individuals (24) are more risk-averse without time pressure than when time pressure is faced. The 8 subjects with no change and in risk attitudes do not have a time pressure parameter statistically different from zero.

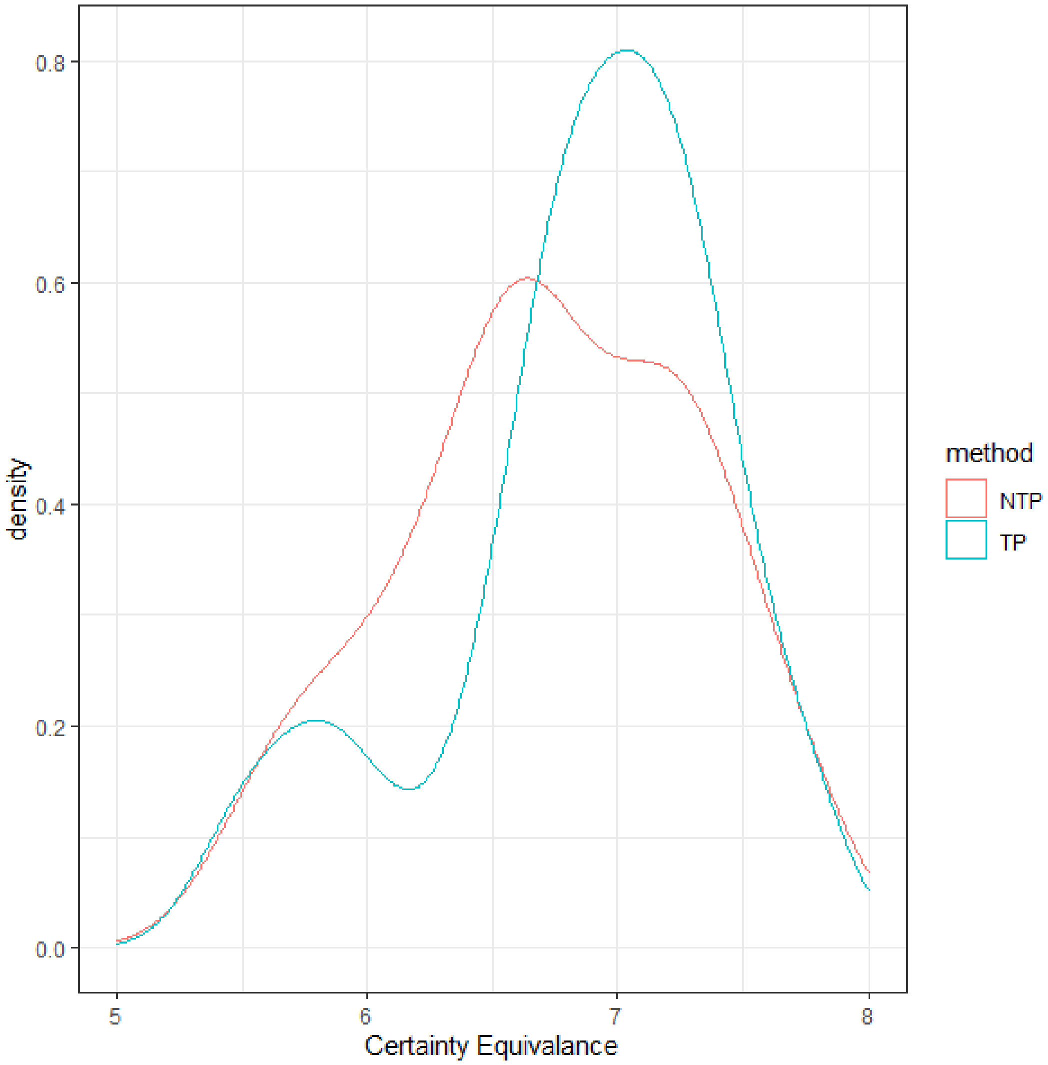


FIGURE 3.5: Density Plot of Certainty Equivalence

Figure 3.5 above shows the distributions of the certainty equivalence with and without time pressure. We can see a large peak near 6.5 without time pressure and closer to 7 when time pressure is faced, suggesting that the distribution is lower without time pressure. A Wilcoxon test supports this claim, which confirms this with a p-value of 0.0259.

Overall when considering how risk attitudes change under time pressure, we suggest from our results that individuals exhibit less risk-averse behaviour under time pressure than when they have more time to make a decision.

3.7 Concluding Remarks

We have carried out an experiment considering the effect of time pressure on an individual's risk attitudes under a budget allocation task. We do this by recovering behavioural parameters of interest with and without time pressure. Then, we propose a theoretical gamble and calculate an individual's certainty equivalence to this gamble with and without time pressure using the parameters we previously recovered. We compare this calculated certainty equivalence with and without time pressure to determine how individual's risk attitudes have changed. We find that our results are in line with the literature, which until this point has solely considered binary choices, in which individuals act in a less risk-averse manner (for gains) under time pressure. There are several ways to extend this for future work. For example, having a constant probability instead of different probabilities in each round would allow us to consider revealed preferences and if an individual's choices remain consistent under time pressure. Other ideas include extending this to losses and ambiguity or having varying degrees of time pressure.

Conclusion

We have looked at 3 different papers focusing on decision-making under uncertainty. We have shown that time can play a role in decision-making. It can be used as a substitute for choice data in increasing subject observations. The availability of time has also been shown to have an impact on individual preferences, with less time available resulting in decreased levels of risk aversion. We have also shown that there exist issues with using representative agent models with small datasets in that you may not be able to recover gender effects and propose a way in which gender effects can be observed.

There exist several paths for an extension to these papers. For Chapter 1, it would be ideal to consider an experiment with mixed gambles, utilise other domains such as ambiguity and time preferences, or alter the model to be able to consider more than binary choices. For chapter 2, considering other domains such as ambiguity or time preferences would be natural extensions or applying the framework to the time pressure experiment in Chapter 3. Finally, for Chapter 3, we aim to perform a similar experiment with consistent probabilities to rest revealed preferences and move into other domains, such as ambiguity.

Overall, what ties these papers and the key takeaway is that sometimes using different methods can lead to higher-quality results. For example, using time as a substitute for additional data, using a more flexible model to find gender effects and certainty equivalence to show how risk attitudes change under time pressure.

Appendix A

Appendix for Chapter 1

A.1 Robustness Checks

For robustness, we consider other specifications, we first consider a second probability weighting function proposed in Prelec, 1998 $w(P) = \exp\left(-((-\ln(P))^{\delta})\right)$. δ in this represents the degree of probability distortion. The closer to 1, the less distortion, and the closer to 0, the more probability distortion. We also consider a second utility function $U(W) = \frac{W^{1-\rho}}{1-\rho}$. Where ρ represents risk aversion, the higher ρ represents an increase in risk aversion and risk neutrality at $\rho = 0$

We perform the same cross-fold validation technique proposed in section 1.8, changing either the utility function, the weighting function or both.

TABLE A.1: Robustness Checks

	1	2	3	4
Utility Function	POWER	POWER	CARA	CARA
Weighting Function	TK	PRL	TK	PRL
Mean DDM Accuracy	0.7520415	0.7192152	0.7519902	0.7191441
Mean Logit Accuracy	0.7283764	0.729377	0.7286108	0.730232
Significance	0.006366	0.403	0.01028	0.349
Median DDM Accuracy	0.7685954	0.7207354	0.7686338	0.7207354
Median Logit Accuracy	0.726203	0.7261634	0.7235012	0.7238919
Variance DDM Accuracy	0.002855772	0.002409153	0.00286247	0.002394783
Variance Logit Accuracy	0.003682566	0.00360727	0.003826338	0.003498794

Where POWER, CARA, TK and PRL represent different utility and weighting functions being used. A utility function of power denotes $U(W) = \frac{W^{\rho}}{\rho}$. A Utility function of CARA represents $U(W) = \frac{W^{1-\rho}}{1-\rho}$. A weighting function of TK represents $w(P) = \frac{P\delta}{(P^{\delta}+(1-P)^{\delta}}) ^{\frac{1}{\delta}}$. A weighting function of PRL represents $w(P) = \exp\left(-((-\ln(P))^{\delta})\right)$. Significance represents the p-value of a Wilcoxon rank sum test, testing if the distributions of predicted choices are different.

From A.1, we can see that the Drift Diffusion Model performs better at predicting choices compared to the logit-errors model when considering a standard weighting function, regardless of the utility function. However, when we use the Prelec weighting function, we find that the predictive power of both models is roughly identical, as shown by the similar means, medians and the Wilcoxon rank sum test result. This suggests that the Drift-Diffusion Model has the potential to be an improvement or at least equal to that of the logit-errors model.

A.2 Information on Decision Time Data

Here, we include some information on the decision time data of subjects in the experiment

TABLE A.2: Summary of Decision Time data under Gains and Losses

	Gains	Losses
Mean Times	7711.446	10665.89
Median	6502	9018.5
Standard Deviation	5474.487	6992.715

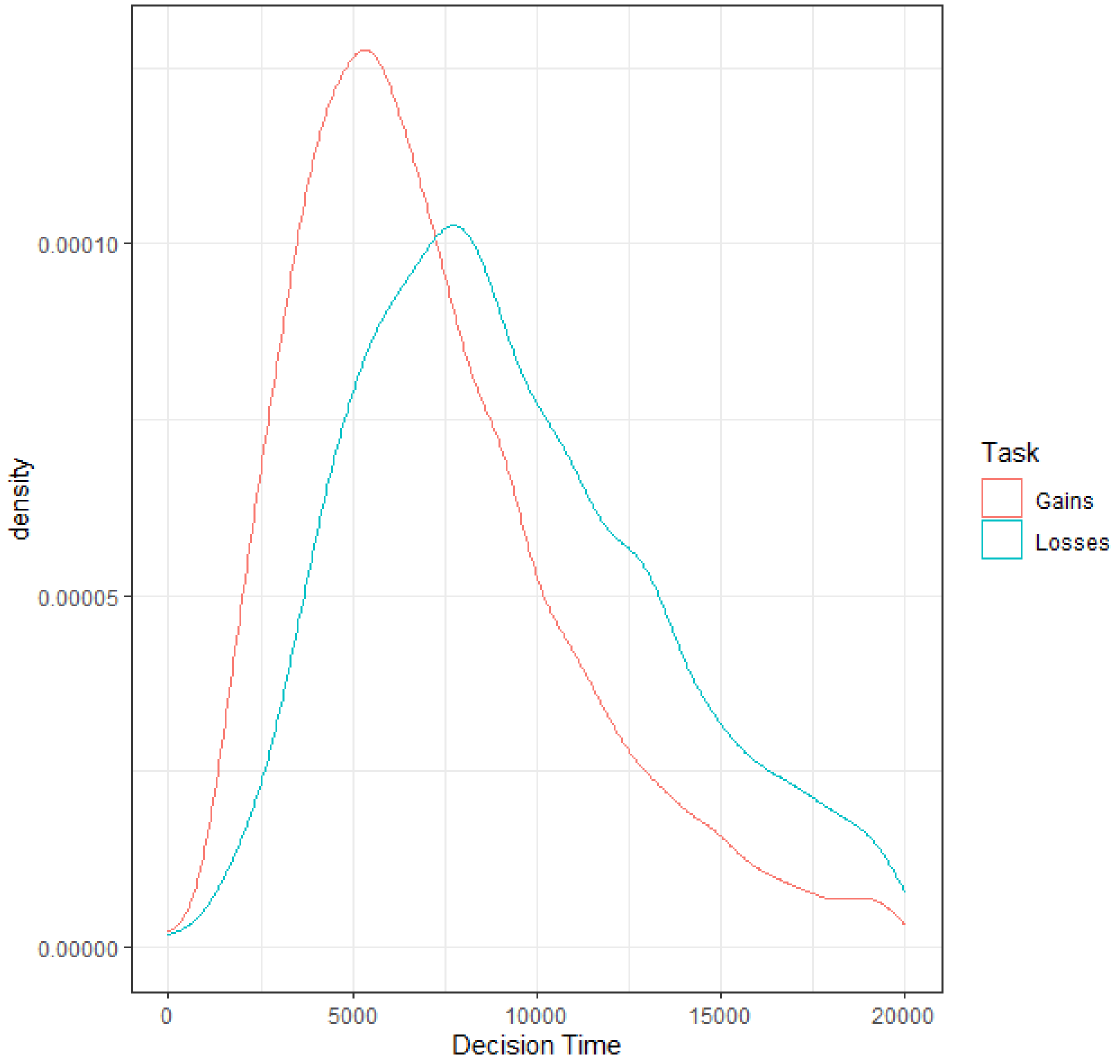


FIGURE A.1: Density Plots for decision times under gains and losses

We can see from the table and diagram that subjects took, on average, a significantly longer time to decide on losses than tasks considering gains. This could suggest some difficulty in making these decisions or require more information to make these decisions.

Appendix B

Appendix For Chapter 2

B.1 Monte Carlo Simulation

This Appendix presents further results of the simulation in Section 6 when the noise is low ($\xi = 0.130$) in Table B.1, and when it is high ($\xi = 0.200$) in Table B.2.

TABLE B.1: Mean and standard deviations of the parameters.

Parameter	True value	MLE	MSLE	HB
λ	1.648	1.576	1.640	1.648
s.e.	-	0.050	0.035	0.035
λ_{FEMALE}	0.365	0.696	0.368	0.379
s.e.	-	0.117	0.057	0.055
σ_{λ}	0.100	-	0.088	-
s.e.	-	-	0.036	-
r	0.500	0.537	0.500	0.497
s.e.	-	0.017	0.023	0.021
r_{FEMALE}	0.100	0.065	0.105	0.107
s.e.	-	0.030	0.028	0.027
σ_r	0.050	-	0.047	-
s.e.	-	-	0.012	-
w_g	0.540	0.559	0.540	0.536
s.e.	-	0.012	0.050	0.022
w_l	0.510	0.527	0.509	0.506
s.e.	-	0.012	0.048	0.023
ξ	0.130	0.134	0.130	0.128
s.e.	-	0.005	0.005	0.043

The Table reports estimates from the simulation exercise on the three inference methods : Maximum Likelihood Estimation (MLE), Maximum Simulated Likelihood Estimation (MSLE) and Hierarchical Bayesian (HB), for the low level of noise (0.13). For each parameter θ, the Table reports the mean of the point estimates, in the case of MLE, the mean of the distributions in the case of MSLE, and of the posterior mean of the distributions in the case of HB. Standard deviations in parentheses.

TABLE B.2: Mean and standard deviations of the parameters.

Parameter	True value	MLE	MSLE	HB
λ	1.648	1.572	1.640	1.670
s.e.	-	0.072	0.051	0.054
λ_{FEMALE}	0.365	0.705	0.371	0.398
s.e.	-	0.173	0.081	0.089
σ_λ	0.100	-	0.072	-
s.e.	-	-	0.057	-
r	0.500	0.539	0.500	0.489
s.e.	-	0.024	0.031	0.032
r_{FEMALE}	0.100	0.062	0.106	0.111
s.e.	-	0.045	0.040	0.041
σ_r	0.050	-	0.043	-
s.e.	-	-	0.022	-
w_g	0.540	0.560	0.540	0.534
s.e.	-	0.018	0.077	0.036
w_l	0.510	0.528	0.509	0.502
s.e.	-	0.018	0.075	0.038
ξ	0.200	0.202	0.199	0.198
s.e.		0.008	0.006	0.042

The Table reports estimates from the simulation exercise on the three inference methods : Maximum Likelihood Estimation (MLE), Maximum Simulated Likelihood Estimation (MSLE) and Hierarchical Bayesian (HB), for the high level of noise (0.20). For each parameter θ, the Table reports the mean of the point estimates, in the case of MLE, the mean of the distributions in the case of MSLE, and of the posterior mean of the distributions in the case of HB. Standard deviations in parentheses.

B.2 Sample size

Table B3 reports the results of the simulation exercise when the size sample of 100 increases by a factor of 2 (N=200) and 5 (N=200). All the parameter values

TABLE B3: Mean and standard deviations of the parameters.

	True value	N=100	N=200	N=500
λ	1.648	1.637	1.640	1.648
s.e.	-	0.062	0.029	0.021
λ_{FEMALE}	0.365	0.374	0.370	0.368
s.e.	-	0.118	0.046	0.031
σ_λ	0.100	0.081	0.088	0.093
s.e.	-	0.046	0.029	0.016
r	0.500	0.500	0.503	0.504
s.e.	-	0.031	0.017	0.015
r_{FEMALE}	0.100	0.105	0.099	0.096
s.e.	-	0.031	0.022	0.017
σ_r	0.050	0.046	0.048	0.050
s.e.	-	0.014	0.008	0.005
w_g	0.540	0.543	0.542	0.542
s.e.	-	0.014	0.041	0.029
w_l	0.510	0.510	0.511	0.511
s.e.	-	0.013	0.043	0.029
ξ	0.150	0.150	0.150	0.150
s.e.	-	0.005	0.003	0.002

The Table reports estimates from the simulation exercise using Maximum Simulated Likelihood Estimation (MSLE) for three levels of sample size (N) namely 100, 200 and 500. Standard deviations in parentheses.

Appendix C

Appendix For Chapter 3

C.1 Experimental Tasks

These are the 40 tasks subjects faced in the experiment, 20 with time pressure and 20 without, the ordering of the tasks were randomised for each subject so they faced tasks in a different order. However, each member of an experimental session faced time pressure or non time pressure tasks first and the other second.

TABLE B1: Experimental Tasks

	Time Pressure				Non-Time Pressure			
Round	p_x	p_y	$Prob_x$	$Prob_y$	p_x	p_y	$Prob_x$	$Prob_y$
1	12	7	0.75	0.25	12	12	0.25	0.75
2	6	10	0.3	0.7	9	12	0.4	0.6
3	9	14	0.75	0.25	13	9	0.9	0.1
4	7	15	0.75	0.25	5	9	0.1	0.9
5	13	15	0.5	0.5	12	7	0.3	0.7
6	5	8	0.2	0.8	6	13	0.5	0.5
7	11	11	0.8	0.2	12	10	0.3	0.7
8	11	14	0.8	0.2	10	9	0.5	0.5
9	11	5	0.25	0.75	6	12	0.33	0.67
10	7	9	0.5	0.5	9	15	0.25	0.75
11	14	5	0.75	0.25	15	9	0.67	0.33
12	6	5	0.8	0.2	12	12	0.8	0.2
13	8	6	0.4	0.6	13	8	0.4	0.6
14	12	12	0.67	0.33	10	7	0.75	0.25
15	8	9	0.1	0.9	11	12	0.67	0.33
16	14	7	0.9	0.1	6	9	0.8	0.2
17	10	12	0.9	0.1	5	7	0.33	0.67
18	15	9	0.67	0.33	15	14	0.5	0.5
19	10	5	0.7	0.3	8	7	0.3	0.7
20	5	13	0.9	0.1	8	10	0.6	0.4

We have two sets of instructions for the experiment. One for those who did time pressure tasks first and ones for two had those tasks second. They are both similar but both have been included for completion.

Instructions for Subjects with Time Pressure Tasks First

Instructions

Welcome

Welcome to this experiment. These instructions are to help you understand what you are being asked to do during the experiment and how you will be paid. The experiment is simple and gives you the chance to earn a considerable amount of money, which will be paid to you in cash after you have completed the experiment. The payment described below is *in addition* to a participation fee of £3.00 that you will be paid independently of your answers.

The experiment

This experiment consists of two parts. You will receive instructions on the second part once you have completed the first one.

In the first part of this experiment, you will be presented with a total of 20 questions. Each question is of the same form. For each question, you will have 100 tokens to allocate between two different goods, Good 1 and Good 2. The questions will vary in two ways: first the price of each good will be different in every question; second, the chances of obtaining either of the goods will differ. After you have responded on all questions, in both parts, one of them will be chosen at random and played out. 'Playing out' a question means that one of the two goods will occur (in a way we describe below) with chances as specified in that question, and you will be paid the money value of the number of tokens that you allocated to that Good.

Example of a question

Imagine the question as shown in the Figure below. The chances of the two goods are:

Good 1: 70%

Good 2: 30%.

The prices of the two goods are:

Good 1: 10

Good 2: 5

Suppose that you make the following allocation:

Good 1: 80 tokens

Good 2 : 20 tokens

In this case you will receive 80/10=£8 if Good 1 will occur, or 20/5=£4 if Good 2 will occur.

Specifying the Chances

Each question will specify the chances for Good 1 and Good2. These may differ from question to question but they will always be something out of 100. In the previous example, the chance of Good 1 occurring is 70 out of 100, and of Good 2 occurring 30 out of 100. If this question were to be played out at the end of the experiment, you will be provided with an opaque bag filled with 100 numbered tickets. The bag would then be shaken and you would be asked to draw a ticket at random (and without looking into the bag) out of the bag. If the ticket has a number from 1 to 70 (70 out of 100), you will receive whatever you have allocated to Good 1 (80 tokens) divided by the price of Good 1 (10), that is £8. If the ticket has a number from 71 to 100 (30 out of 100), you will receive whatever you have allocated to Good 2 (20 tokens) divided by the price of Good 2 (5), that is £4.

For example if you were to draw ticket number 45, then Good 1 would occur and your payment would be £11=£8 (your payment from Good 1)+£3 (the show-up fee).

The Experimental Interface

The computer will tell you in each question the prices of the two goods. The chances of either good occurring are shown with the aid of a pie chart. The larger the proportion for one of the goods, the higher the chances for this good to occur. You can slide the slider to denote how many tokens you would like to allocate to each good. You can see in the screen how much you have allocated to each of the goods. A picture at the bottom right indicates the amounts implied by your allocations in money terms, and are represented by the heights of the coloured bars. The widths of the coloured bars represent the chance of the goods being realised in that question. The questions are independent of each other, meaning you start with a fresh 100 tokens in the start of each question. When you click on ‘Confirm’ or hit the ‘Space’ key you will move onto the next question.

Timing

You will be given a maximum of 12 seconds to decide; if you have not confirmed your decision by the time 12 seconds have elapsed, it will be assumed that the last position of the slider reflects your preferred allocation.

Payment

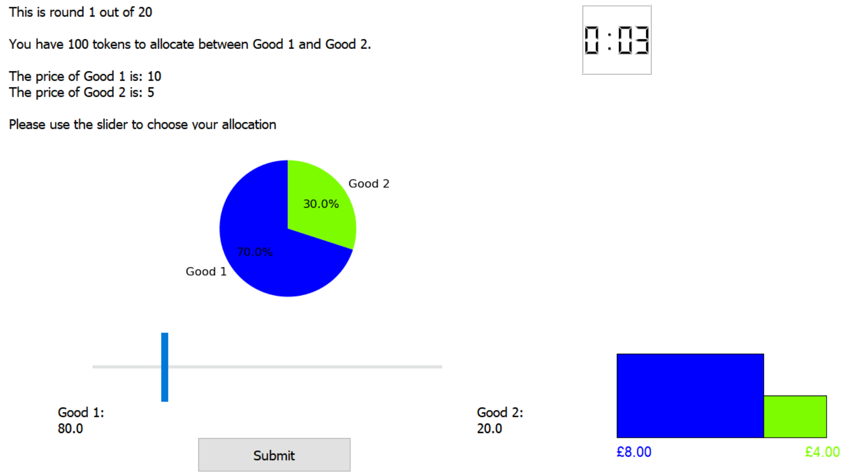
At the end of the experiment one of the two parts will be randomly chosen for payment. One of the 20 questions from this part will be randomly selected by drawing a ticket from an opaque bag which contains 20 numbered tickets. The computer will recover your choices and the details of this particular question. You will then draw a ticket from the bag with the 100 numbered tickets to determine which good will occur.

Practice Rounds

To familiarise yourself with the experimental interface and the task, there will be 3 practice questions which will not count towards your final payment. Feel free to ask any questions you may have during the practice questions.

Questionnaires

At the end of parts 1 and 2, there will be 3 short questionnaires collecting various demographics and personality data.



Thank you for participating

Harry Rolls

Konstantinos Georgalos

December 2019

Instructions for Subjects with Time Pressure Tasks Second

Instructions

Welcome

Welcome to this experiment. These instructions are to help you understand what you are being asked to do during the experiment and how you will be paid. The experiment is simple and gives you the chance to earn a considerable amount of money, which will be paid to you in cash after you have completed the experiment. The payment described below is *in addition* to a participation fee of £3.00 that you will be paid independently of your answers.

The experiment

This experiment consists of two parts. You will receive instructions on the second part once you have completed the first one.

In the first part of this experiment, you will be presented with a total of 20 questions. Each question is of the same form. For each question, you will have 100 tokens to allocate between two different goods, Good 1 and Good 2. The questions will vary in two ways: first the price of each good will be different in every question; second, the chances of obtaining either of the goods will differ. After you have responded on all questions, in both parts, one of them will be chosen at random and played out. 'Playing out' a question means that one of the two goods will occur (in a way we describe below) with chances as specified in that question, and you will be paid the money value of the number of tokens that you allocated to that Good.

Example of a question

Imagine the question as shown in the Figure in the last page. The chances of the two goods are:

Good 1: 70%

Good 2: 30%.

The prices of the two goods are:

Good 1: 10

Good 2: 5

Suppose that you make the following allocation:

Good 1: 80 tokens

Good 2: 20 tokens

In this case you will receive 80/10=£8 if Good 1 will occur, or 20/5=£4 if Good 2 will occur.

Specifying the Chances

Each question will specify the chances for Good 1 and Good2. These may differ from question to question but they will always be something out of 100. In the previous example, the chance of Good 1 occurring is 70 out of 100, and of Good 2 occurring 30 out of 100. If this question were to be played out at the end of the experiment, you will be provided with an opaque bag filled with 100 numbered tickets. The bag would then be shaken and you would be asked to draw a ticket at random (and without looking into the bag) out of the bag. If the ticket has a number from 1 to 70 (70 out of 100), you will receive whatever you have allocated to Good 1 (80 tokens) divided by the price of Good 1 (10), that is £8. If the ticket has a number from 71 to 100 (30 out of 100), you will receive whatever you have allocated to Good 2 (20 tokens) divided by the price of Good 2 (5), that is £4.

For example if you were to draw ticket number 45, then Good 1 would occur and your payment would be £11=£8 (your payment from Good 1)+£3 (the show-up fee).

The Experimental Interface

The computer will tell you in each question the prices of the two goods. The chances of either good occurring are shown with the aid of a pie chart. The larger the proportion for one of the goods, the higher the chances for this good to occur. You can slide the slider to denote how many tokens you would like to allocate to each good. You can see in the screen how much you have allocated to each of the goods. A picture at the bottom right indicates the amounts implied by your allocations in money terms, and are represented by the heights of the coloured bars. The widths of the coloured bars represent the chance of the goods being realised in that question. Again, the larger the width for one of the goods, the higher the chances for this good to occur. The questions are independent of each other, meaning you start with a fresh 100 tokens in the start of each question. When you click on 'Confirm' or hit the 'Space' key you will move onto the next question.

Timing

You will not be able to confirm your decision on any question until at least 15 seconds have elapsed. You will be given a maximum of 45 seconds to decide; if you have not confirmed your decision by the time 45 seconds have elapsed, it will be assumed that the last position of the slider reflects your preferred allocation.

Payment

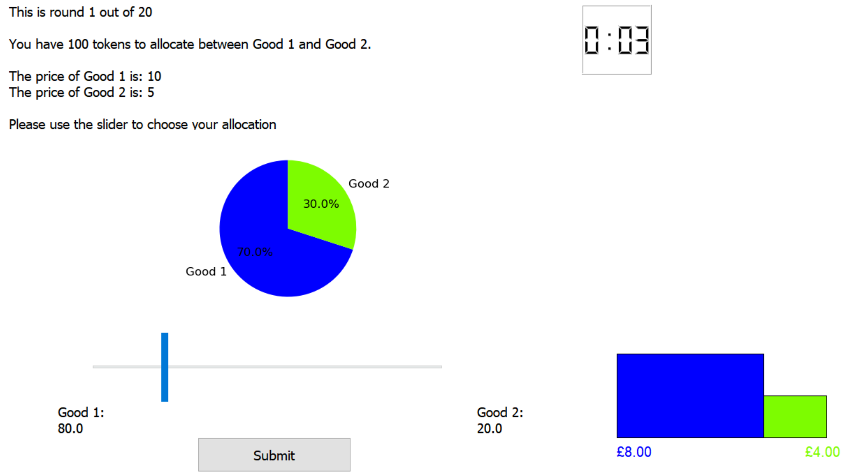
At the end of the experiment one of the two parts will be randomly chosen for payment. One of the 20 questions from this part will be randomly selected by drawing a ticket from an opaque bag which contains 20 numbered tickets. The computer will recover your choices and the details of this particular question. You will then draw a ticket from the bag with the 100 numbered tickets to determine which good will occur.

Practice Rounds

To familiarise yourself with the experimental interface and the task, there will be 3 practice questions which will not count towards your final payment. Feel free to ask any questions you may have during the practice questions.

Questionnaires

At the end of parts 1 and 2, there will be 3 short questionnaires collecting various demographics and personality data.



Thank you for participating

Harry Rolls

Konstantinos Georgalos

December 2019

C.2 Derivation of Demand Equations

Here we shall show the derivation of the optimal allocation

$$\max \pi_x * \frac{(x*e_x)^{1-\rho}}{1-\rho} + \pi_y * \frac{(y*e_y)^{1-\rho}}{1-\rho}$$
$$\text{s.t } x + y = 100$$

First Order Conditions:

$$\pi_x * e_x * (x * e_x)^{-\rho} = \lambda$$
$$\pi_y * e_x * (y * e_y)^{-\rho} = \lambda$$
$$x + y = 100$$

Rearrange

$$\frac{\pi_x e_x}{\pi_y e_y}\left(\frac{xe_x}{ye_y}\right)^{-\rho} = 1$$
$$\left(\frac{ye_y}{xe_x}\right)^{\rho} = \frac{\pi_y e_y}{\pi_x e_x}$$
$$y = \frac{xe_x}{e_y}\left(\frac{\pi_y e_y}{\pi_x e_x}\right)^{\frac{1}{\rho}}$$

Put this into the budget constraint:

$$x + \frac{xe_x}{e_y}\left(\frac{\pi_y e_y}{\pi_x e_x}\right)^{\frac{1}{\rho}} = 100$$
$$X^* = \frac{100}{1+\frac{e_x}{e_y}\left(\frac{\pi_y e_y}{\pi_x e_x}\right)^{\frac{1}{\rho}}}$$

Solve for y:

$$x + y = 100$$
$$Y^* = 100 - \frac{100}{1+\frac{e_x}{e_y}\left(\frac{\pi_y e_y}{\pi_x e_x}\right)^{\frac{1}{\rho}}}$$

Can be expressed more tidily as:

$$X^* = \frac{100*(\pi_x * e_x^{1-\rho})^{\frac{1}{\rho}}}{(\pi_x * e_x^{1-\rho})^{\frac{1}{\rho}} + (\pi_y * e_y^{1-\rho})^{\frac{1}{\rho}}}$$
$$Y^* = \frac{100*(\pi_y * e_y^{1-\rho})^{\frac{1}{\rho}}}{(\pi_x * e_x^{1-\rho})^{\frac{1}{\rho}} + (\pi_y * e_y^{1-\rho})^{\frac{1}{\rho}}}$$

C.3 Plots for distributions of behavioural parameters

Here, we include the density plots for individuals behavioural parameters. For this we compare the distribution of behavioural parameter θ_{NTP} and $\theta_{NTP} + \theta_{TP}$ for comparing the parameters used for decision making with and without time pressure.

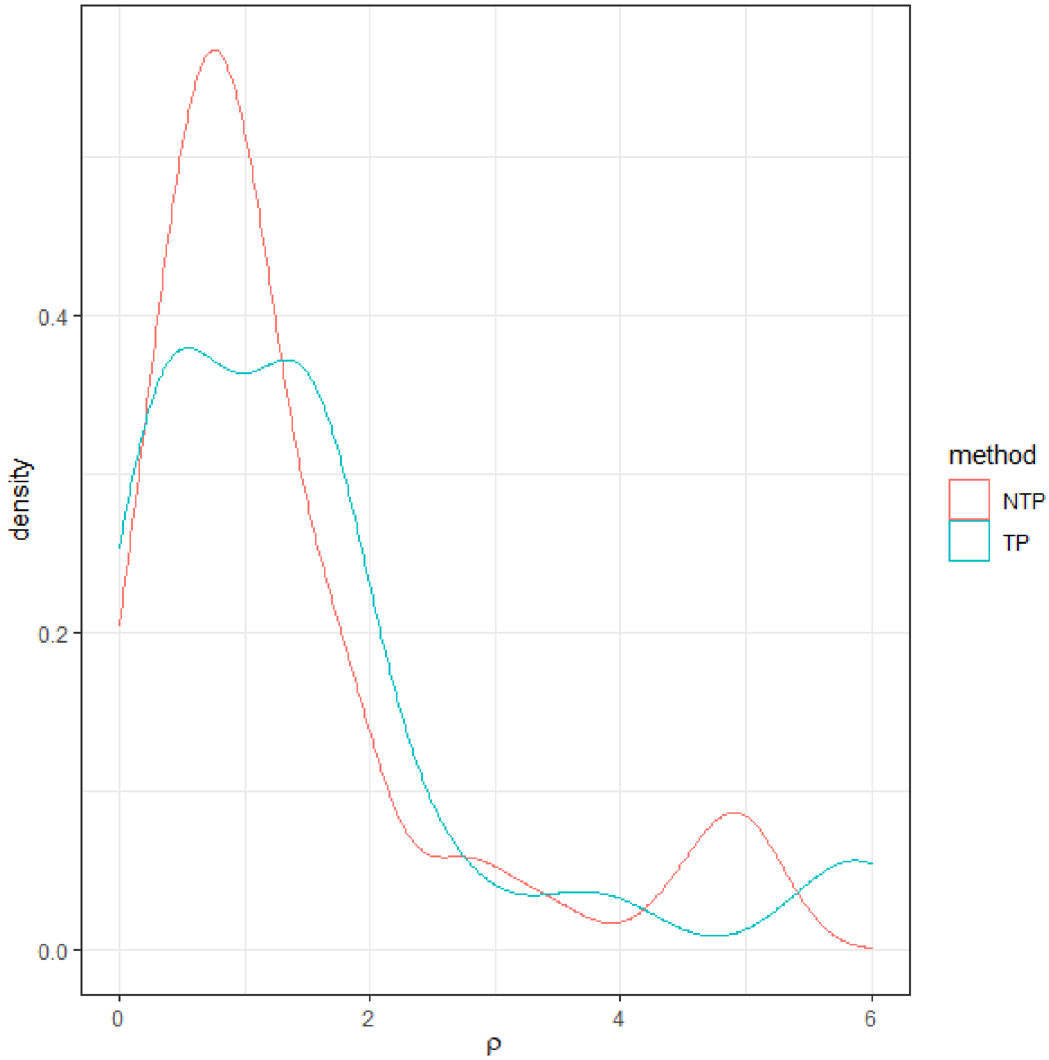


FIGURE C.1: Density Plot of ρ, the curvature of utility function

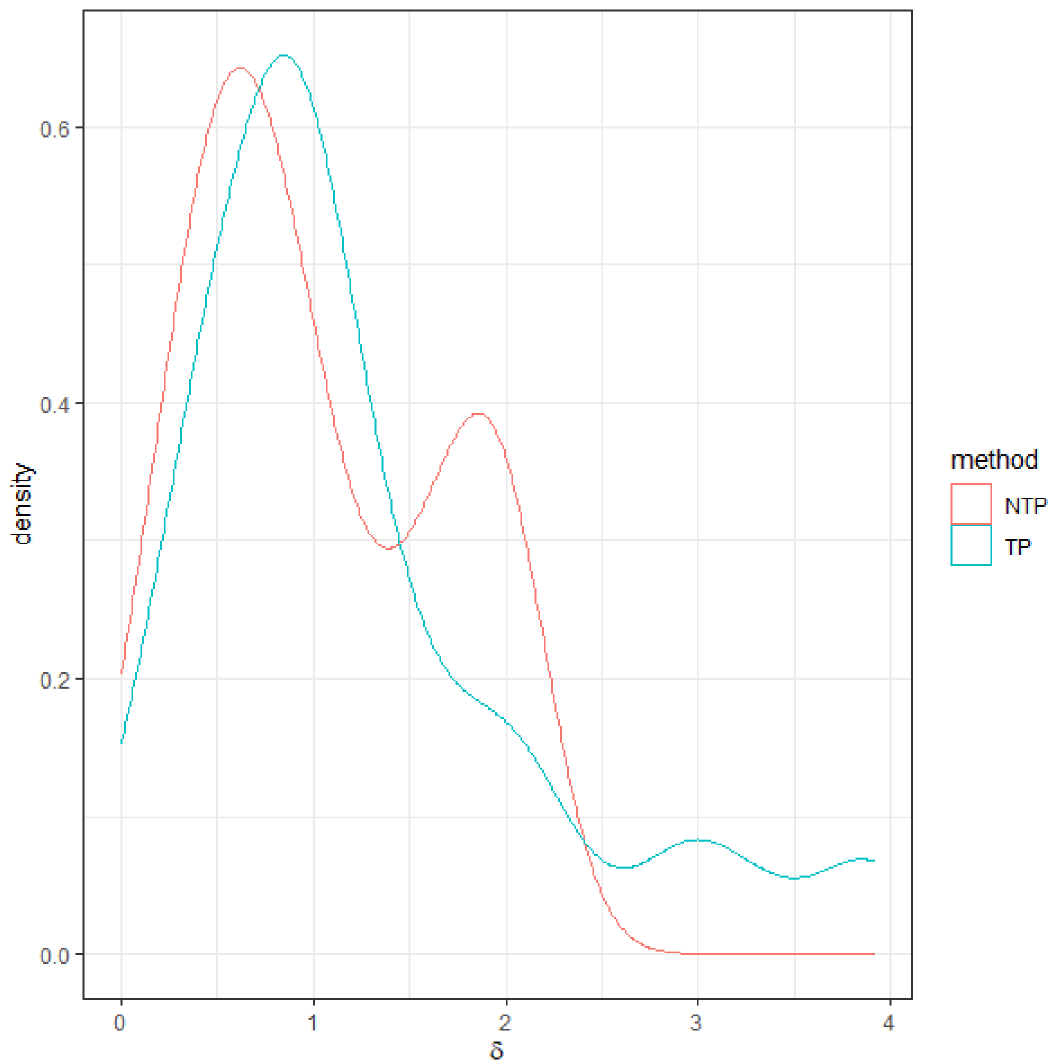


FIGURE C.2: Density Plot of δ, the degree of probability distortion

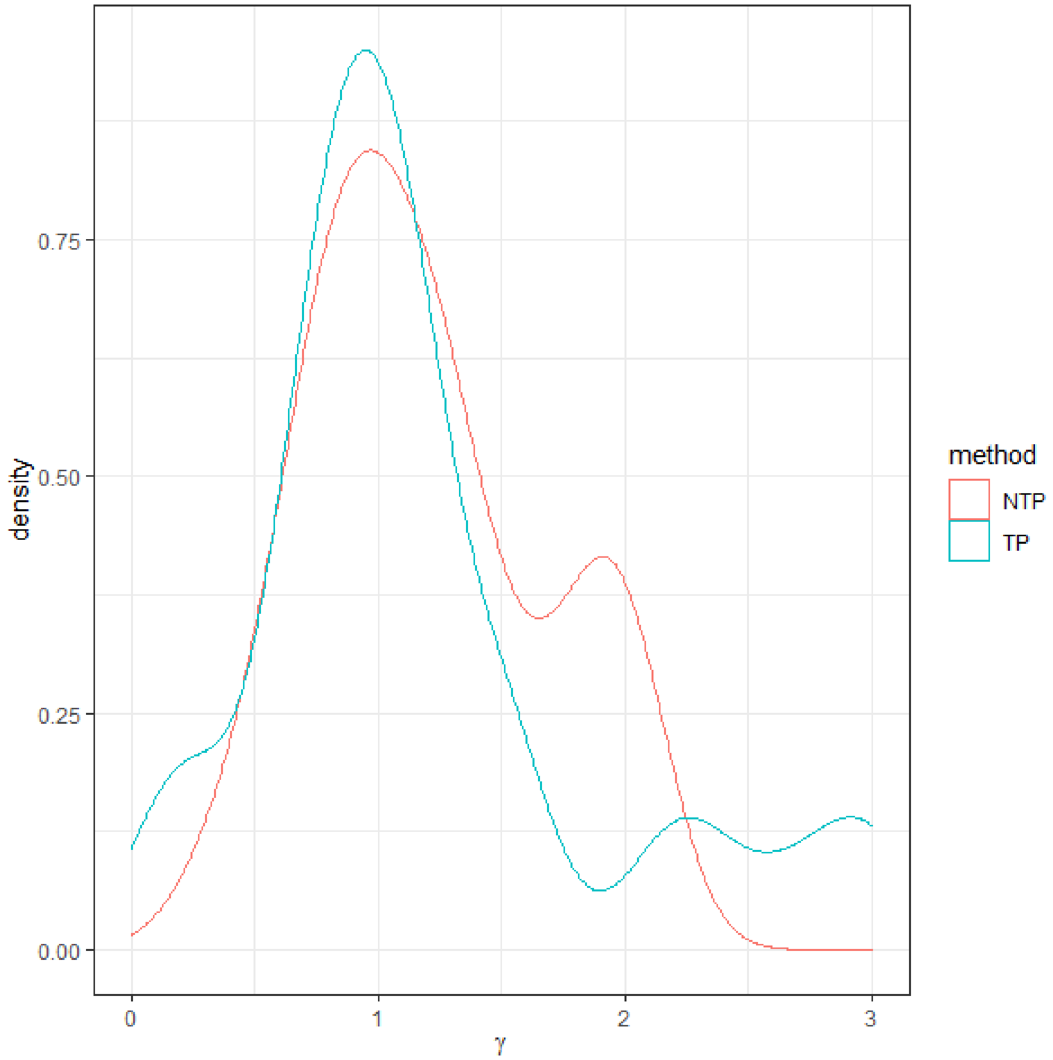


FIGURE C.3: Density Plot of γ, degree of probability over or under weighting

C.4 Considering other Weighting Functions

In order to see if our results proposed above are robust, we compare our Two Parameter Prelec function (PRL2) with several other functions proposed within the literature in order to see is we get the same result that a majority of individuals are more risk averse without time pressure.

TABLE B2: Other weighting functions to consider

Name	Specification
Goldstein-Einhorn (GE)	$\frac{\gamma P^{\delta}}{\gamma P^{\delta}+(1-P)^{\delta}}$
1 Parameter Prelec (PRL1)	$\exp\left(-((-\ln(P))^{\delta})\right)$
Tversky-Kahneman (TK)	$\frac{P^{\delta}}{(P^{\delta}+(1-P)^{\delta})^{\frac{1}{\delta}}}$

We then perform the same procedure as in the main paper, recovering the behavioural parameters of interest with and without time pressure, determining what is significant and finally calculating and comparing certainty equivalence in order to see how risk attitudes change with and without time pressure. Below we show a summary of the results for the other weighting functions we consider as well as the 2 Parameter Prelec function for comparison.

TABLE B3: Summary Results for all weighting functions

	PRL2	GE	PRL1	TK
Mean ρ_{NTP}	1.3750	1.3761	1.3532	1.1057
Mean ρ_{TP}	0.1086	0.0799	0.0824	-0.1010
Mean δ_{NTP}	1.0252	1.1325	0.9121	1.0613
Mean δ_{TP}	0.2175	0.0916	0.2027	-0.0087
Mean γ_{NTP}	1.1976	0.9890		
Mean γ_{TP}	-0.01215	0.1667		
Statistically Different ρ_{TP}	29	24	40	43
Statistically Different δ_{TP}	35	33	41	42
Statistically Different γ_{TP}	27	33		
No statistical changes under time pressure	8	7	5	1
Mean Certainty Equivalence without Time Pressure	7.0211	6.9929	6.9798	6.8714
Mean Certainty Equivalence with Time Pressure	6.8002	6.8148	6.9046	6.7913
More risk averse under time pressure	13	12	18	17
More risk averse without time pressure	24	24	25	29

We can see that for all weighting functions, we have the same results, with a majority of subjects having at least one of their behavioural parameters for time pressure statistically different from zero. When considering their certainty equivalences with each of these specifications, we can see the same results with a majority of subjects being more risk averse without time pressure, as shown by a lower level of certainty equivalence.

Bibliography

Abdellaoui, Mohammed et al. (2020). “A defense of prospect theory in Bernheim & Sprenger’s experiment”. In: *unpublished paper*.

Alós-Ferrer, Carlos, Ernst Fehr, and Nick Netzer (2021). “Time Will Tell: Recovering Preferences When Choices Are Noisy”. eng. In: *The Journal of political economy* 129.6, pp. 1828–1877. ISSN: 0022-3808.

Baillon, A., H. Bleichrodt, and V. Spinu (2020). “Searching for the Reference Point”. In: *Management Science* 66.1, pp. 93–112.

Baillon, Aurélien et al. (2018). “Measuring Ambiguity Attitudes for All (Natural) Events”. In: *Econometrica* 86.5, pp. 1839–1858. DOI: https://doi.org/10.3982/ECTA14370. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA14370. URL: https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA14370.

Balcombe, K. and I. Fraser (2015a). “Parametric Preference Functionals under Risk in the Gain Domain: A Bayesian Analysis”. In: *Journal of Risk and Uncertainty* 50.2, pp. 161–187. ISSN: 1573-0476.

Balcombe, Kelvin and Iain Fraser (2015b). “Parametric preference functionals under risk in the gain domain: A Bayesian analysis”. eng. In: *Journal of risk and uncertainty* 50.2, pp. 161–187. ISSN: 0895-5646.

Baldassi, Carlo et al. (2020). “A Behavioral Characterization of the Drift Diffusion Model and Its Multialternative Extension for Choice Under Time Pressure”. eng. In: *Management science* 66.11, pp. 5075–5093. ISSN: 0025-1909.

Bartke, Stephan and Reimund Schwarze (2008). “Risk-averse by nation or by religion? Some insights on the determinants of individual risk attitudes”. In: *Some Insights on the Determinants of Individual Risk Attitudes (October 2008). SOEPpaper* 131.

Ben Zur, Hasida and Shlomo J Breznitz (1981). “The effect of time pressure on risky choice behavior”. eng. In: *Acta psychologica* 47.2, pp. 89–104. ISSN: 0001-6918.

Betsch, Tilmann et al. (2004). “Oops, I did it again—relapse errors in routinized decision making”. In: *Organizational Behavior and Human Decision Processes* 93.1, pp. 62–74. ISSN: 0749-5978. DOI: https://doi.org/10.1016/j.obhdp.2003.09.002. URL: https://www.sciencedirect.com/science/article/pii/S0749597803001043.

Bishop, C. (2006). *Pattern Recognition and Machine Learning*. Springer-Verlag New York.

Booij, A., B. Van praag, and G. Kuillen (2009). “A Parametric Analysis of Prospect Theory’s Functionals for the General Population”. In: *Theory and Decision* 68, pp. 115–148.

Bouchouicha, R. et al. (2019). “Gender effects for loss aversion: Yes, no, maybe?” In: *Journal of Risk and Uncertainty* 59, pp. 171–184.

Brandstätter, Eduard, Gerd Gigerenzer, and Ralph Hertwig (2006). “The Priority Heuristic: Making Choices Without Trade-Offs”. eng. In: *Psychological review* 113.2, pp. 409–432.

Brooks, P. and H. Zank (2005). “Loss Averse Behavior”. In: *Journal of Risk and Uncertainty* 31, pp. 301–325.

Camerer, C. (1995). “Individual Decision Making”. In: *The Handbook of Experimental Economics*. Ed. by J. Kagel and A. Roth. Princeton University Press.

Chandler, Jesse J. and Emily Pronin (2012). “Fast Thought Speed Induces Risk Taking”. eng. In: *Psychological science* 23.4, pp. 370–374. ISSN: 0956-7976.

Charness, G and Y. Gneezy (2012). “Strong Evidence for Gender Difference in Risk Taking”. In: *Journal of Economic Behavior & Organization* 83.1, pp. 50–58.

Choi, Syngjoo et al. (2007). “Consistency and heterogeneity of individual behavior under uncertainty”. In: *American economic review* 97.5, pp. 1921–1938.

Choi, Syngjoo et al. (2014). “Who Is (More) Rational?” In: *American Economic Review* 104.6, pp. 1518–50. DOI: 10.1257/aer.104.6.1518. URL: https://www.aeaweb.org/articles?id=10.1257/aer.104.6.1518.

Clithero, John A (2018a). "Improving out-of-sample predictions using response times and a model of the decision process". eng. In: *Journal of economic behavior & organization* 148, pp. 344–375. ISSN: 0167-2681.

— (2018b). "Response times in economics: Looking through the lens of sequential sampling models". eng. In: *Journal of economic psychology* 69, pp. 61–86. ISSN: 0167-4870.

Conte, Anna, John D. Hey, and Peter G. Moffatt (2011). "Mixture models of choice under risk". In: *Journal of Econometrics* 162.1, pp. 79 –88. ISSN: 0304-4076.

Criss, Amy H (2010). "Differentiation and Response Bias in Episodic Memory: Evidence From Reaction Time Distributions". eng. In: *Journal of experimental psychology. Learning, memory, and cognition* 36.2, pp. 484–499. ISSN: 0278-7393.

Croson, R. and Y. Gneezy (2009). "Gender Differences in Preferences". In: *Journal of Economic Literature* 47.2, pp. 448–474.

Cœugnet, Stéphanie, Béatrice Cahour, and Sami Kraiem (2019). "Risk-taking, emotions and socio-cognitive dynamics of pedestrian street-crossing decision-making in the city". In: *Transportation Research Part F: Traffic Psychology and Behaviour* 65, pp. 141–157. ISSN: 1369-8478. DOI: https://doi.org/10.1016/j.trf.2019.07.011. URL: https://www.sciencedirect.com/science/article/pii/S1369847818307472.

De Paola, Maria and Francesca Gioia (2016). "Who performs better under time pressure? Results from a field experiment". eng. In: *Journal of economic psychology* 53, pp. 37–53. ISSN: 0167-4870.

Dror, Itiel E, Beth Basola, and Jerome R Busemeyer (1999). "Decision making under time pressure: An independent test of sequential sampling models". eng. In: *Memory cognition* 27.4, pp. 713–725. ISSN: 0090-502X.

Eckel, Catherine C. and Philip J. Grossman (2008). "Men, Women and Risk Aversion: Experimental Evidence". In: ed. by Charles R. Plott and Vernon L. Smith. Vol. 1. Handbook of Experimental Economics Results. Elsevier, pp. 1061 –1073. DOI: https://doi.org/10.1016/S1574-0722(07)00113-8. URL: http://www.sciencedirect.com/science/article/pii/S1574072207001138.

Edman, G., D. Schalling, and S.E. Levander (1983). "Impulsivity and speed and errors in a reaction time task: A contribution to the construct validity of the concept of impulsivity”. In: *Acta Psychologica* 53.1, pp. 1–8. ISSN: 0001-6918. DOI: https://doi.org/10.1016/0001-6918(83)90012-4. URL: https://www.sciencedirect.com/science/article/pii/0001691883900124.

Epstein, Seymour (1994). “Integration of the Cognitive and the Psychodynamic Unconscious”. eng. In: *The American psychologist* 49.8, pp. 709–724. ISSN: 0003-066X.

Feenstra, H, R.A.C Ruiter, and G Kok (2012). “Go fast! Reaction time differences between adults and adolescents in evaluating risky traffic situations”. eng. In: *Journal of health psychology* 17.3, pp. 343–349. ISSN: 1359-1053.

Fehr, Ernst and Antonio Rangel (2011). “Neuroeconomic Foundations of Economic Choice—Recent Advances”. eng. In: *The Journal of economic perspectives* 25.4, pp. 3–30. ISSN: 0895-3309.

Feller, William (1971). *An introduction to probability theory and its applications. Vol. II.* Second edition. New York: John Wiley & Sons Inc., pp. xxiv+669.

Ferecatu, A. and A. Önçüler (2016). “Heterogeneous Risk and Time Preferences”. In: *Journal of Risk and Uncertainty* 53.1, pp. 1–28.

Fiedler, Susann and Andreas Glöckner (2012). “The Dynamics of Decision Making in Risky Choice: An Eye-Tracking Analysis”. In: *Frontiers in Psychology* 3. ISSN: 1664-1078. DOI: 10.3389/fpsyg.2012.00335. URL: https://www.frontiersin.org/article/10.3389/fpsyg.2012.00335.

Filippin, A. and P. Crosetto (2016). “A Reconsideration of Gender Differences in Risk Attitudes”. In: *Management Science* 62.11, pp. 3138–3160.

Fisher, Geoffrey (2017). “An attentional drift diffusion model over binary-attribute choice”. In: *Cognition* 168, pp. 34–45. ISSN: 0010-0277. DOI: https://doi.org/10.1016/j.cognition.2017.06.007. URL: https://www.sciencedirect.com/science/article/pii/S0010027717301695.

Fox, Craig R. and Russell A. Poldrack (2009). “Chapter 11 - Prospect Theory and the Brain”. In: *Neuroeconomics*. Ed. by Paul W. Glimcher et al. London: Academic Press, pp. 145 –173. ISBN: 978-0-12-374176-9. DOI: https://doi.org/10.1016/B978-0-12-374176-9.00011-7. URL: http://www.sciencedirect.com/science/article/pii/B9780123741769000117.

Fudenberg, Drew, Philipp Strack, and Tomasz Strzalecki (2018). "Speed, Accuracy, and the Optimal Timing of Choices". eng. In: *The American economic review* 108.12, pp. 3651–3684. ISSN: 0002-8282.

Gao, X., G. Harrison, and R. Tchernis (2020). "Estimating Risk Preferences for Individuals: A Bayesian Approach". CEAR Working Paper 2020-15.

Gaudecker, H. von, A. Van Soest, and E. Wengström (2011). "Heterogeneity in Risky Choice Behaviour in a Broad Population". In: *American Economic Review* 101.2, pp. 664–694.

Gelman, A. and D. Rubin (1992). "Inference from Iterative Simulation Using Multiple Sequences". In: *Statistical Science* 7.4, pp. 457–472. ISSN: 08834237.

Ghalanos, Alexios and Stefan Theussl (2015). *Rsolnp: General Non-linear Optimization Using Augmented Lagrange Multiplier Method*. R package version 1.16.

Glöckner, Andreas and Thorsten Pachur (2012). "Cognitive models of risky choice: Parameter stability and predictive accuracy of prospect theory". eng. In: *Cognition* 123.1, pp. 21–32. ISSN: 0010-0277.

Guo, Lisa, Jennifer S Trueblood, and Adele Diederich (2017). "Thinking Fast Increases Framing Effects in Risky Decision Making". eng. In: *Psychological science* 28.4, pp. 530–543. ISSN: 0956-7976.

Haji, Anouar El et al. (2019). "Time pressure and risk taking in auctions: A field experiment". eng. In: *Journal of behavioral and experimental economics* 78, pp. 68–79. ISSN: 2214-8043.

Hale, David (1968). "The relation of correct and error responses in a serial choice reaction task". In: *Psychonomic Science* 13.6, pp. 299–300. ISSN: 2197-9952. DOI: 10.3758/BF03342595. URL: https://doi.org/10.3758/BF03342595.

Harbaugh, William T, Kate Krause, and Lise Vesterlund (2002). "Risk Attitudes of Children and Adults: Choices Over Small and Large Probability Gains and Losses". eng. In: *Experimental economics : a journal of the Economic Science Association*. Experimental Economics 5.1, pp. 53–84. ISSN: 1386-4157.

Harrison, G. and E. Rutström (2008). "Expected Utility Theory and Prospect Theory: one Wedding and a Decent Funeral". In: *Experimental Economics* 12.2, p. 133. ISSN: 1573-6938.

He, Fei et al. (2013). "Neural Mechanisms of Risky Decision Making in Monetary Gain and Loss Situations". eng. In: *Social behavior and personality* 41.10, pp. 1725–1736. ISSN: 0301-2212.

Heitz, Richard P (2014). "The speed-accuracy tradeoff: History, physiology, methodology, and behavior". eng. In: 8.8, pp. 150–150. ISSN: 1662-4548.

Hellyer, Paul (2019). "Less time, more mistakes?" eng. In: *British dental journal* 226.11, pp. 857–857. ISSN: 0007-0610.

Heo, Wookjae, Abed Rabbani, and John E Grable (2021). "An evaluation of the effect of the COVID-19 pandemic on the risk tolerance of financial decision makers". In: *Finance Research Letters* 41, p. 101842.

Hey, J. (2014). "Choice Under Uncertainty: Empirical Methods and Experimental Results". In: *Hanbook of the Economics of Risk and Uncertainty*. Ed. by M. Machina and K. Viscusi W. Eslevier B.V., pp. 809–850.

Hey, J. and C. Orme (1994). "Investigating Generalizations of Expected Utility Theory Using Experimental Data". In: *Econometrica* 62.6, pp. 1291–1326.

Hobbs, Alan and Ann Williamson (2003). "Associations between Errors and Contributing Factors in Aircraft Maintenance". In: *Human Factors* 45.2. PMID: 14529193, pp. 186–201. DOI: 10.1518/hfes.45.2.186.27244. eprint: https://doi.org/10.1518/hfes.45.2.186.27244. URL: https://doi.org/10.1518/hfes.45.2.186.27244.

Holt, Charles A and Susan K Laury (2002). "Risk Aversion and Incentive Effects". eng. In: *The American economic review* 92.5, pp. 1644–1655. ISSN: 0002-8282.

Holt, Charles A. and Susan K. Laury (2014). "Chapter 4 - Assessment and Estimation of Risk Preferences". In: *Handbook of the Economics of Risk and Uncertainty*. Ed. by Mark Machina and Kip Viscusi. Vol. 1. Handbook of the Economics of Risk and Uncertainty. North-Holland, pp. 135 –201. DOI: https://doi.org/10.1016/B978-0-444-53685-3.00004-0. URL: http://www.sciencedirect.com/science/article/pii/B9780444536853000040.

Howlett, Jonathon R and Martin P Paulus (2020). "A parallel accumulator model accounts for decision randomness when deciding on risky prospects with different expected value". eng. In: *PloS one* 15.7, e0233761–e0233761. ISSN: 1932-6203.

Hu, Yixin et al. (2015). "The effect of emotion and time pressure on risk decision-making". In: *Journal of Risk Research* 18.5, pp. 637–650. DOI: 10.1080/13669877.2014.910688. eprint: https://doi.org/10.1080/13669877.2014.910688. URL: https://doi.org/10.1080/13669877.2014.910688.

Huber, J. and K. Train (2001). "On the Similarity if Classical and Bayesian Estimates of Individual Mean Partworths". In: *Marketing Letters* 12.3, pp. 259–269.

Huber, Oswald and Urs Kunz (2007). "Time pressure in risky decision-making: effect on risk defusing". In: *Psychology Science* 49.4, p. 415.

Jacquement, N. and O. L'Haridon (2018). *Experimental Economics: Methods and Applications*. Cambridge University Press.

Kahneman, Daniel and Amos Tversky (1979). "Prospect Theory: An Analysis of Decision under Risk". In: *Econometrica* 47.2, pp. 263–91. URL: https://EconPapers.repec.org/RePEc:ecm:emetrp:v:47:y:1979:i:2:p:263-91.

Kocher, Martin G., Julius Pahlke, and Stefan T. Trautmann (2013). "Tempus Fugit: Time Pressure in Risky Decisions". eng. In: *Management science* 59.10, pp. 2380–2391. ISSN: 0025-1909.

Krajbich, Ian and Antonio Rangel (2011). "Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions". eng. In: *Proceedings of the National Academy of Sciences - PNAS* 108.33, pp. 13852–13857. ISSN: 0027-8424.

Lewis, Penny J et al. (2014). "Exploring the causes of junior doctors' prescribing mistakes: a qualitative study". In: *British journal of clinical pharmacology* 78.2, pp. 310–319.

Lindner, Florian and Matthias Sutter (2013). "Level-k reasoning and time pressure in the 11–20 money request game". In: *Economics Letters* 120.3, pp. 542–545. ISSN: 0165-1765. DOI: https://doi.org/10.1016/j.econlet.2013.06.005. URL: https://www.sciencedirect.com/science/article/pii/S0165176513002875.

Madan, Christopher R, Marcia L Spetch, and Elliot A Ludvig (2015). "Rapid makes risky: Time pressure increases risk seeking in decisions from experience". eng. In: *Journal of cognitive psychology (Hove, England)* 27.8, pp. 921–928. ISSN: 2044-5911.

Mather, Mara et al. (2012). "Risk preferences and aging: The "certainty effect" in older adults' decision making." In: *Psychology and aging* 27.4, p. 801.

Milosavljevic, Milica et al. (2010). "The Drift Diffusion Model Can Account for the Accuracy and Reaction Time of value-based Choices under High and Low Time Pressure". eng. In: *Judgment and decision making*. Judgment and Decision Making 5.6, pp. 437–449. ISSN: 1930-2975.

Moffatt, P. (2016). *Experimetrics*. Macmillan Palgrave.

Moya, Andrés (2018). "Violence, psychological trauma, and risk attitudes: Evidence from victims of violence in Colombia". In: *Journal of Development Economics* 131, pp. 15–27. ISSN: 0304-3878. DOI: https://doi.org/10.1016/j.jdeveco.2017.11.001. URL: https://www.sciencedirect.com/science/article/pii/S0304387817300949.

Murphy, R. and R. ten Brincke (2018). "Hierarchical Maximum Likelihood Parameter Estimation for Cumulative Prospect Theory: Improving the Reliability of Individual Risk Parameter Estimates". In: *Management Science* 64, pp. 308–326.

Nilsson, H., J. Rieskamp, and E-J. Wagenmakers (2011). "Hierarchical Bayesian Parameter Estimation for Cumulative Prospect Theory". In: *Journal of Mathematical Psychology* 55, pp. 84–93.

Nursimulu, Anjali D. and Peter Bossaerts (2014). "Risk and reward preferences under time pressure". eng. In: *REVIEW OF FINANCE* 18.3, pp. 999–1022. ISSN: 1572-3097.

Olschewski, Sebastian and Jörg Rieskamp (2021). "Distinguishing three effects of time pressure on risk taking: Choice consistency, risk preference, and strategy selection". In: *Journal of Behavioral Decision Making* 34.4, pp. 541–554. DOI: https://doi.org/10.1002/bdm.2228. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/bdm.2228. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/bdm.2228.

Plummer, M. (2017). *JAGS Version 4.3.0 User Manual*. Tech. rep.

Prelec, Drazen (1998). "The Probability Weighting Function". In: *Econometrica* 66.3, pp. 497–527. ISSN: 00129682, 14680262. URL: http://www.jstor.org/stable/2998573 (visited on 06/30/2022).

Qui, J. and E. Steiger (2011). "Understanding the two Components of Risk Attitudes: an Experimental Analysis". In: *Management Science* 57, pp. 193–199.

Quiggin, J. (1982). "A Theory of Anticipated Utility". In: *Journal of Economic Behavior and Organization* 3.4, pp. 323–343.

Ratcliff, Roger (1978). "A theory of memory retrieval". eng. In: *Psychological review* 85.2, pp. 59–108. ISSN: 0033-295X.

— (1981). "A theory of order relations in perceptual matching". In: *Psychological review* 88.6, pp. 552–572. ISSN: 0033-295X.

Ratcliff, Roger and Gail McKoon (2008). "The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks". eng. In: *Neural computation* 20.4, pp. 873–922.

Ratcliff, Roger and Hans P A Van Dongen (2011). "Diffusion model for one-choice reaction-time tasks and the cognitive effects of sleep deprivation". eng. In: *Proceedings of the National Academy of Sciences - PNAS* 108.27, pp. 11285–11290. ISSN: 0027-8424.

Rich, Erin L, Frederic M Stoll, and Peter H Rudebeck (2018). "Linking dynamic patterns of neural activity in orbitofrontal cortex with decision making". In: *Current Opinion in Neurobiology* 49. Neurobiology of Behavior, pp. 24–32. ISSN: 0959-4388. DOI: https://doi.org/10.1016/j.conb.2017.11.002. URL: https://www.sciencedirect.com/science/article/pii/S0959438817301885.

Rieskamp, Jörg and Ulrich Hoffrage (2008). "Inferences under time pressure: How opportunity costs affect strategy selection". In: *Acta Psychologica* 127.2, pp. 258–276. ISSN: 0001-6918. DOI: https://doi.org/10.1016/j.actpsy.2007.05.004. URL: https://www.sciencedirect.com/science/article/pii/S0001691807000492.

Rosen, Allison B., Jerry S. Tsai, and Stephen M. Downs (2003). "Variations in Risk Attitude across Race, Gender, and Education". In: *Medical Decision Making* 23.6. PMID: 14672111, pp. 511–517. DOI: 10.1177/0272989X03258431. eprint: https://doi.org/10.1177/0272989X03258431. URL: https://doi.org/10.1177/0272989X03258431.

Rouder, J. and J. Lu (2005). "An Introduction to Bayesian Hierarchical Models with an Application in the Theory of Signal Detection". In: *Psychonomic Bulletin & Review* 55, pp. 84–93.

Saqib, Najam U. and Eugene Y. Chan (2015). "Time pressure reverses risk preferences". In: *Organizational Behavior and Human Decision Processes* 130, pp. 58–68. ISSN: 0749-5978. DOI: https://doi.org/10.1016/j.obhdp.2015.06.004. URL: https://www.sciencedirect.com/science/article/pii/S0749597815000722.

Schmidt, U. and S. Traub (2002). "An Experimental Test of Loss Aversion". In: *Journal of Risk and Uncertainty* 25.3, pp. 233–249.

Schmidt, U and H. Zank (2005). "What is Loss Aversion". In: *Journal of Risk &Uncertainty* 30.2, pp. 157–167.

Singmann, Henrik et al. (2020). *rtdists: Response Time Distributions*. R package version 0.11-2. URL: https://CRAN.R-project.org/package=rtdists.

Stahl, D. (2014). "Heterogeneity of Ambiguity Preferences". In: *The Review of Economics and Statistics* 96.5, pp. 609–617.

Starmer, C. (2000). "Developments in Non-expected Utility Theory: The Hunt for a Descriptive Theory of Choice under Risk". In: *Journal of Economic Literature* 38.2, pp. 332–382. DOI: 10.1257/jel.38.2.332.

Stott, H. (2006). "Cumulative Prospect Theory's Functional Menagerie". English. In: *Journal of Risk and Uncertainty* 32.2, pp. 101–130. ISSN: 0895-5646. DOI: 10.1007/s11166-006-8289-6.

Toubia, O. et al. (2013). "Dynamic Experiments for Estimating Preferences: an Adaptive Method of Eliciting Time and Risk Parameters". In: *Management Science* 59, pp. 613–640.

Train, K. (2009). *Discrete Choice Methods with Simulation*. 2nd. Cambridge University Press.

Tversky, A. and D. Kahneman (1992). " Advances in Prospect Theory: Cumulative Representation of Uncertainty". In: *Journal of Risk and Uncertainty* 5.4, pp. 297–323.

Venkatraman, Vinod et al. (2007). "Sleep deprivation elevates expectation of gains and attenuates response to losses following risky decisions". eng. In: *Sleep (New York, N.Y.)* 30.5, pp. 603–609. ISSN: 0161-8105.

Vieider, Ferdinand M. et al. (2015). "Common Componetns of Risk and Uncertainty Attitudes Across Contexts and Domains: Evidence from 30 Countries". In: *Journal of the European Economic Association* 13.3, pp. 421–452. DOI: https://doi.org/

10.1111/jeea.12102. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/jeea.12102. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/jeea.12102.

Voss, Andreas, Klaus Rothermund, and Jochen Voss (2004). "Interpreting the parameters of the diffusion model: An empirical validation". eng. In: *Memory & cognition* 32.7, pp. 1206–1220. ISSN: 0090-502X.

Voss, Andreas and Jochen Voss (2007). "Fast-dm: A free program for efficient diffusion model analysis". eng. In: *Behavior research methods* 39.4, pp. 767–775. ISSN: 1554-351X.

Wabersich, Dominik and Joachim Vandekerckhove (2014). "The RWiener package: An R package providing distribution functions for the Wiener diffusion model". In: *The R Journal* 6.1. R package version 1.3-3, pp. 49–56.

Wagenmakers, E.J, H.L.J van der Maas, and R.P.P.P Grasman (2007). "An EZ-diffusion model for response time and accuracy". eng. In: *Psychonomic bulletin I& review* 14.1, pp. 3–22. ISSN: 1069-9384.

Wakker, P. (2010). *Prospect Theory*. Cambridge University Press.

Webb, Ryan (2019). "The (Neural) Dynamics of Stochastic Choice". eng. In: *Management science* 65.1, pp. 230–255. ISSN: 0025-1909.

Wickelgren, Wayne A. (1977). "Speed-accuracy tradeoff and information processing dynamics". In: *Acta Psychologica* 41.1, pp. 67–85. ISSN: 0001-6918. DOI: https://doi.org/10.1016/0001-6918(77)90012-9. URL: https://www.sciencedirect.com/science/article/pii/0001691877900129.

Wiecki, Thomas V, Imri Sofer, and Michael J Frank (2013). "HDDM: Hierarchical bayesian estimation of the drift-diffusion model in Python". eng. In: 7.2013, pp. 14–14. ISSN: 1662-5196.

Wilcox, N. (2011). "Stochastically more Risk Averse: A Contextual Theory of Stochastic Discrete Choice under Risk". In: *Journal of Econometrics* 162.1, pp. 89 –104.

Xie, Zhixin, L. Page, and B. Hardy (2017). "Investigating Gender Differences under Time Pressure in Financial Risk Taking". In: *Frontiers in Behavioural Neuroscience* 11.246.

Ye, Yinyu (1987). "Interior Algorithms for Linear, Quadratic, and Linearly Constrained Non-Linear Programming". PhD thesis. Department of ESS, Stanford University.

Young, Diana L et al. (2012). “Decision making under time pressure, modeled in a prospect theory framework”. In: *Organizational behavior and human decision processes* 118.2, pp. 179–188.

Zhao, Wenjia Joyce, Lukasz Walasek, and Sudeep Bhatia (2020). “Psychological mechanisms of loss aversion: A drift-diffusion decomposition”. eng. In: *Cognitive psychology* 123, pp. 101331–101331. ISSN: 0010-0285.

Zylberberg, Ariel, Christopher R Fetsch, and Michael N Shadlen (2016). “The influence of evidence volatility on choice, reaction time and confidence in a perceptual decision”. eng. In: *eLife* 5.2016. ISSN: 2050-084X.