The Probability Discounting Model of Choice Under Risk:

A Critique

by

Andre Hofmeyr[†]

July 2020

ABSTRACT

The probability discounting model has become a popular framework for investigating choice under risk in psychology. But it is not clear how the model relates to standard theories of choice under risk, such as expected utility theory and prospect theory. I critically review the theoretical development of the model and argue that it rests on an outdated conception of reinforcement learning. I also show that it is formally isomorphic to the dual theory of choice under risk, but that it is limited to simple prospects with a specific parametric probability weighting function. I discuss the methodological and statistical approaches typically used in studies of probability discounting and develop a structural statistical framework to test the efficacy of the model's implied probability weighting function in applied work. Using data from a widely-cited study, I show that the probability discounting model is needlessly restrictive since it does not allow for the simultaneous overweighting and underweighting of probabilities, which is a common pattern identified in the experimental literature on choice under risk.

Keywords: probability discounting, choice under risk, risk aversion, probability weighting

[†] School of Economics, University of Cape Town (South Africa)

I. INTRODUCTION

Probability discounting (PD) is a popular model in psychology for investigating the instantaneous or atemporal attitudes toward risk of individuals in experimental settings. It is particularly common in studies of addiction where delay discounting data are also often obtained.[1] This model was developed by Rachlin, Logue, Gibbon and Frankel (RLGF) [1986], Rachlin, Castrogiovanni and Cross (RCC) [1987], and Rachlin, Raineri and Cross (RRC) [1991].[2] As the name suggests, the PD model draws its inspiration from models of temporal or delay discounting, where the delay to a reward is replaced with the odds against receiving a reward (see Green and Myerson [2004] and Madden and Bickel [2010] for reviews). I explain the model in language familiar to computational learning theorists, statisticians, and economists; show how the model relates to standard theories of choice under risk; highlight the shortcomings of the PD approach; and provide empirical support for the claim that the model suffers from needless limitations.

This critique of the PD model is a case study of the long-term risks associated with methodological silos in cognate disciplines. The technical literature on choice under risk has a long, impressive lineage, and theoretical, methodological, empirical, and statistical advances keep being made (see Harrison and Rutström [2008] for a review). By contrast, the PD model inadequately characterises choice under risk because it is based on a theoretical reformulation of the probability concept that is outdated and suffers from a number of methodological and statistical restrictions.

It is critical that researchers investigating choice under risk cross this methodological divide, because the substantive topics to which the PD model has been applied, such as addiction, remain of paramount importance across several disciplines. Researchers in psychology would be better served by adopting incentive-compatible elicitation tasks; theories that incorporate multiple pathways to explain the risk premium, and not just probability weighting as in the PD model; and statistical tools that are

---

[1] See, for example, Mitchell [1999], Richards, Zhang, Mitchell and de Wit [1999] and Reynolds, Richards, Horn and Karraker [2004]. For a recent meta-analysis of the relationship between probability discounting and gambling behaviour see Kyonka and Schutte [2018].

[2] According to Google Scholar (22 July 2020), these articles have been collectively cited 1,869 times.

appropriate for estimating models of choice under risk to ensure that valid inferences can be drawn.

Section II outlines some canonical models of choice under risk. Section III discusses the derivation of the PD model and focusses on some theoretical, methodological, and statistical limitations of this approach. Section IV shows that the model is formally isomorphic to the dual theory of choice under risk due to Yaari [1987]. Section V develops a structural statistical framework to estimate the PD model's implied probability weighting function, along with other functions that are commonly used in the literature. Section VI re-analyses data from a widely-cited study to test the relevance of the PD model in applied work, and Section VII concludes.

II. THEORIES OF CHOICE UNDER RISK WITH SIMPLE PROSPECTS

RLGF set themselves the task of translating prospect theory (PT), developed by Kahneman and Tversky (KT) [1979], into a behavioural model of choice inspired by the work of Herrnstein [1961] on the "matching law,"[3] which relates behavioural outputs to environmental inputs. PT was developed to account for a number of purported anomalies in choice among lotteries[4] that the canonical model of choice under risk, expected utility (EU) theory, allegedly could not explain. PT proposes that two phases take place during the choice process: an initial editing phase which typically yields a simpler representation of the gambles, and a subsequent evaluation phase where the edited gambles are evaluated and the gamble with the highest value is chosen. KT provide a descriptive explanation of the editing phase and develop a formal model of the evaluation phase. The model of the evaluation phase will be outlined briefly, along with other models of choice under risk, to show how the PD model is related to them.

[3] Other early contributions to the literature on the matching law are Davenport [1962], Logan [1965], Chung [1965], Chung and Herrnstein [1967], and Herrnstein [1970].
[4] The terms "lottery," "gamble," and "prospect" refer to a probability distribution over outcomes and will be used interchangeably.

Using the notation in KT, $(x, p; y, q)$ is a prospect yielding outcome x with probability p, outcome y with probability q, and 0 with probability $1 - p - q$, where $p + q \leq 1$. KT [p. 276] define a "regular" prospect as one where $p + q < 1$, or $x \geq 0 \geq y$, or $x \leq 0 \leq y$. During the evaluation phase of PT, regular prospects are evaluated by the function:

$$V(x, p; y, q) = \pi(p)v(x) + \pi(q)v(y), \tag{1}$$

where $v(0) = 0$, $\pi(0) = 0$, $\pi(1) = 1$, $\pi(\cdot)$ is an increasing function of p, and $\pi(p) \in [0, 1]$.

According to KT, individuals respond differently to gains and losses. This matters in (1) because the outcomes x and y are evaluated relative to some reference point r which determines whether the outcomes are perceived as gains or losses. KT were agnostic about the determination of r. Their subsequent model in Tversky and Kahneman [1992], called cumulative prospect theory, assumes that the reference point coincides with zero income (i.e., $r = 0$). I adopt the latter assumption because it implies that when $x, y \geq 0$, the lottery is in the gain frame, and I need not complicate the exposition by representing the possibility that the outcomes are perceived as losses.

For the present discussion I will be concerned with binary prospects of the form $(x, p; 0, q)$: prospects that pay x with probability p and 0 with probability q, where $q = 1 - p$. These simple prospects are used because the theory of probability discounting developed by Rachlin et al. [1986, 1987, 1991] is limited only to gambles involving the outcome 0 and one positive or negative outcome x. In the language of PT, I will only consider lotteries in the gain frame, where $x > 0$, to keep the discussion focussed on essentials.

Since $v(0) = 0$, and we are concerned with simple prospects, $(x, p; 0, q)$, (1) reduces to:

$$V(x, p; 0, q) = \pi(p)v(x). \tag{2}$$

In words, the prospective utility V of lottery $(x, p; 0, q)$ is determined solely by the product of some function $\pi(\cdot)$ of the probability p assigned to outcome x and some utility function $v(\cdot)$ over the outcome x. The function $\pi(\cdot)$ is referred to as a probability weighting function (PWF) in the literature because it takes probabilities as its argument and returns decision weights. Consequently, the function $\pi(\cdot)$ incorporates the possibility of subjective distortions of objective probabilities: the objective probability 0.1 may be subjectively perceived as greater or less than 0.1. Under the PT formulation, risk preferences in the gain frame are determined both by the shape of the utility function $v(\cdot)$ and the shape of the PWF $\pi(\cdot)$.

By contrast, EU theory assumes $\pi(p) = p$ and defines the expected utility V of lottery $(x, p; 0, q)$ as:

$$V(x, p; 0, q) = pv(x). \tag{3}$$

EU is not defined relative to a reference point, so it does not treat gains and losses differently, and probabilities are interpreted objectively. Thus, risk preferences are determined solely by the shape of the utility function $v(\cdot)$ because $V(\cdot)$ is linear in p.[5]

Finally, the dual theory of choice under risk developed by Yaari [1987], sometimes referred to as a rank-dependent expected value (RDEV) model (see Harrison and Rutström [2008]), defines the utility V of a simple prospect as:

$$V(x, p; 0, q) = \pi(p)x. \tag{4}$$

Like EU, RDEV is not defined relative to a reference point and does not treat gains and losses differently. As V is linear in outcome x in (4), risk preferences in this model are determined solely by the shape of the PWF $\pi(p)$.

It will be shown that risk preferences in the PD model are determined by the shape of the PWF, and it is therefore equivalent to RDEV, except that the PD model is limited to simple prospects and employs a specific parametric PWF. To establish the context for this result, it is useful to first discuss the theoretical and empirical development of the PD model.

[5] $V(\cdot)$ is unique up to a positive affine transformation so $U(\cdot) = a + bV(\cdot)$, for $b > 0$, represents the same preferences as (3).

III. THE PD MODEL: A BEHAVIOURAL THEORY OF CHOICE UNDER RISK

RLGF argue that choice under risk can be tied to a temporal framework by interpreting the probability associated with a reward as the delay to, or rate of reinforcement of, this reward. To understand this reformulation of the probability concept, consider the following thought experiment presented in RLGF [p. 36]. A person is presented with a gamble (x, ⅓; 0, ⅔) that pays x with probability p = ⅓ and 0 with probability $1 - p$ = ⅔. A physical randomisation device is used to determine whether the subject wins x and the gamble is played out repeatedly with the subject observing the result of each gamble.[6] Suppose that the physical randomisation device takes c seconds to deliver a result and the intertrial interval between gambles is t seconds. Then, timed from the start of the first gamble, the average or expected delay (D) to the person's first win is given by a waiting-time function:

$$D = [(t + c) / p] - t. \tag{5}$$

In this function t is subtracted under the assumption that there is no delay to the first gamble. Expression (5) shows how, at least in the repeated-gambles case, probability and delay are linked. As probability increases the delay to reward falls, and *vice versa*. In addition, probability affects the rate of reinforcement of a reward in a specific way. Using the parameters from the above formalisation (i.e., p = ⅓) over a long series of repeated gambles the rate at which the person would receive money is $x / [3(t + c)]$ per second. Now suppose that the person is in one room and the physical randomisation device is in another room, so that the subject is not aware of it. If the person wins x then it is delivered through a trapdoor. Suppose further that there are two randomisation devices for two different lotteries that are placed behind two trapdoors, and that the person must choose between them. According to RLGF, the choice can be viewed as one between gambles or as one between rates of reinforcement.[7]

[6] Since the gamble is repeated and the person wins x on successful realisations of the gamble, payoffs in this thought experiment are cumulative. RLGF implicitly assume away the effect that cumulative payoffs have on choice behaviour. In other words, they assume that the person has an additively-separable intertemporal utility function which is not affected by changes in income.

[7] Appendix A explains the behavioural model of choice developed by RLGF in more detail and critically discusses an experiment that was conducted to test the implications of the model.

RLGF [p. 36] state that the physical randomisation devices in this thought experiment are "… unseen by and unknown to the subject …," implying that he or she merely chooses a trapdoor and then waits for it to pay out. In other words, the person only observes the outcome of the gamble when it is favourable rather than choosing a gamble and observing when it does and does not pay out. These procedural differences likely have a large impact on how the experiment is perceived by the subject. Choosing a gamble and then waiting until it pays out, without observing each trial, frames the task as one involving amounts and delays. Watching a gamble pay out or not, across repeated trials, frames the experiment as one involving risk and allows the subject to estimate the probability of receiving x, along with the complementary probability of receiving 0. In the former case, interpreting probability as delay to, or rate of reinforcement of, reward is valid because that is how the subject experiences the task. In the latter case, which is certainly empirically relevant, the link between probability and delay is tenuous.

In addition, these procedural differences might affect the way people learn about the reward contingencies in this experiment. Choosing between trapdoors and waiting for them to pay out is suited to Rescorla-Wagner-style associative strength learning where prizes and probabilities are bundled together as rewards (see Rescorla and Wagner [1972]). As Glimcher [2011, p. 401] explains, "… when you learn the value of an action in this way, the probability is bundled directly into your estimate of the value of that action. Your estimate of the value of the action includes the probability, but you cannot extract the probability from this estimate." By contrast, if a subject is presented with a gamble which conveys the prize-probability information directly, as in the first stage of the RLGF thought experiment, and then gets to observe each realisation of this gamble, the person is more likely to encode probability and prize information separately, and can estimate what Glimcher [2011, p. 402-405] refers to as "expected subjective value."[8]

Another issue with this thought experiment is that it relies on the assumption that individuals are exposed to repeated gambles. How then do people behave when faced with a one-shot gamble? RLGF [p. 38] argue that, "… the behavioral model must

---

[8] For a review of the different ways in which people learn and encode value see Clark [2016].

infer the existence of past external events (events that had been paired with the stated probability).” Thus, the underlying theoretical assumption of RLGF appears to be that Rescorla-Wagner-style conditioned reinforcement provides the basis for all individual responses to uncertainty in choice. That is, it is assumed that people interpret the probability of winning a reward as the rate of reinforcement of that reward, which is linked to unknown past events. This is problematic because in a one-shot setting, probability-prize information may be available to subjects to inform choice, which arguably should be expected to influence the way they encode value. It would require a very strong, indeed metaphysical, prior assumption to license simply ignoring the possible effects of this information. Furthermore, RLGF must maintain the assumption that probability and rate of reinforcement are linked even though in a one-shot gamble with a prize of 0 and a prize of $x > 0$, a simple prospect, the probability of receiving x represents the likelihood that one receives x and the complementary probability represents the likelihood of receiving nothing. Thus, if one does not receive x then it will not be received, regardless of the length of time that one waits.

Notwithstanding these issues, RLGF and RCC[9] laid the foundations for the PD model, but it is the method of RRC that has been replicated numerous times[10] and defines the model as it is currently employed. RRC drew on the work of Mazur [1984], who ran a series of delay discounting experiments with pigeons, to further develop the hypothesised link between probability and delay. Mazur [1984, p. 427] argued that the pigeons’ delay discounting data was best explained by a hyperbolic discounting function:

$$V = x / (1 + \delta D), \tag{6}$$

where V is the present or discounted value of the delayed reward x, D is the delay to the reward, and δ is a parameter that captures the extent to which future values are discounted. As δ increases in (6) so the present value V of a delayed reward declines.

[9] RCC conducted an experiment with real rewards to bolster the claim of RLGF that probabilistic choice can be tied to a temporal framework. The experiment was based on Rachlin and Green [1972], who studied the behaviour of pigeons in a delay-commitment paradigm. The major difference with this experiment, other than that RCC used human subjects, was that delays were replaced with probabilities: long delays with low probabilities and short delays with high probabilities. RCC is not directly relevant to the development of the PD model so it is discussed in Appendix B.

[10] See, for example, Ostaszewski [1997], Mitchell [1999], Richards, Zhang, Mitchell and de Wit [1999], Reynolds, Richards, Horn and Karraker [2004], Ohmura, Takahashi and Kitamura [2005], Reynolds [2006] and Reynolds et al. [2007].

RRC argue, once again, that probability is best interpreted as delay to, or rate of reinforcement of, reward and then use the waiting-time function (5) to derive a specific result. On the assumption that c, which is the time it takes for a physical randomisation device to deliver a result, is small relative to the intertrial interval t, the waiting-time function (5) can be re-written as:

$$D = (t / p) - t = t[(1 - p) / p] = t\Theta, \tag{7}$$

where $\Theta = (1 - p) / p$ represents the odds against receiving a reward. Thus, if the probability of receiving reward x under some gamble is 0.2, the odds against receiving x are 4:1. As RRC [p. 235] note, in the context of repeated gambles, "odds against" is the average number of losses expected before a win.

Adopting this logic, RRC develop a function, which they argue is analogous to the delay discounting function (6), to describe how people value or "discount" lotteries of the form (x, p; 0, q):

$$V = x / (1 + \gamma\Theta), \tag{8}$$

where γ performs the same role as δ in (6), and captures the extent to which the probabilistic reward x is "discounted" as a function of the odds against receiving it. Expression (8) defines the PD model as it is typically employed.

This derivation is sensible if one accepts the premises, and (8) certainly "looks like" (6), with γ taking the place of δ and Θ taking the place of D. However, presumably the correct substitution for D in (6) is $t\Theta$, which is the result derived in (7) under the assumptions that $c = 0$ and that there is no intertrial interval prior to the outcome of the first gamble in a set of repeated gambles. This latter assumption is why t is subtracted in the waiting-time function (5) and why $\Theta = (1 - p) / p$ represents the odds against receiving a reward. If t is not subtracted in (5) then $\Theta = 1 / p$ which is clearly not the same as odds against winning.

Ignoring this issue, RRC set out to test the PD model (8) and the delay discounting model (6) by recruiting 80 undergraduate students to take part in experiments with hypothetical rewards: 40 subjects were used to obtain a PD function and 40 were used to obtain a delay discounting function. In the PD experiment subjects made binary choices between \$1,000 available with different probabilities (0.95, 0.9, 0.7, 0.5, 0.3, 0.1, and 0.05) and an amount of money to be received with certainty.[11] For each probability, half of the subjects in the PD experiment were presented with these certain amounts of money in decreasing and then increasing order, while the other subjects were presented with these amounts in increasing and then decreasing order. This elicitation procedure was used to find certainty equivalents (CEs) for the $1,000 lotteries.[12]

CEs were obtained by averaging the amounts of money before and after a switch was made. For example, if someone chose the lottery paying $1,000 with a probability of 0.9 over $850 with certainty, and then chose $900 with certainty over the lottery paying $1,000 with a probability 0.9, then the CE was calculated as $875 for that lottery.[13] The procedure in the delay discounting experiment was identical except that probabilities were replaced with delays: 1 month, 6 months, 1 year, 5 years, 10 years, 25 years, and 50 years.

RRC fit the hyperbolic delay discounting function (6) and hyperbolic PD function (8) to the median indifference points and CEs in the sample, respectively. They argue that these hyperbolic functions provide a better fit to these data than exponential functions, but do not conduct formal statistical tests of this assertion. Nevertheless, they conclude that, "The corresponding form of Equations [(6) and (8)] implies that odds against in probabilistic discounting acts like delay in delay discounting and tends to confirm the speculation of Rachlin et al. [1986] that stated probability and stated delay have corresponding effects on behaviour" (RRC [p. 239-240]). To what extent is this claim justified?

[11] The full set of amounts was: $1,000, $990, $980, $960, $940, $920, $900, $850, $800, $750, $700, $650, $600, $550, $500, $450, $400, $350, $300, $250, $200, $150, $100, $80, $60, $40, $20, $10, $5 or $1.

[12] Loosely, a lottery's certainty equivalent (CE) is the amount of money received with certainty such that a decision maker is indifferent between playing out the lottery and receiving the amount CE. I provide a formal definition in the next section.

[13] Taking the average of these amounts is arbitrary and discards information on the uncertainty of the estimate, as explained in detail below. The correct approach for analysing interval data, such as where a person's CE lies in the interval ($850, $900), is interval regression (see Harrison and Rutström [2008, p. 62-69]) but RRC do not adopt this approach. Alternatively, one can employ a complementary full information maximum likelihood framework, which directly incorporates the uncertainty of the estimates and models it statistically, so that one can draw robust inferences about the ability of different PWFs to characterise PD data.

The link between probability and delay has been investigated in delay discounting experiments where researchers have varied the probability that subjects receive payment for one of their choices on the task. Keren and Roelofsma [1995] conducted an experiment, using hypothetical rewards, where subjects were offered a choice between 100 Dutch Guilders now or 110 Guilders in 4 weeks. When the probability of payment was 1, 82% of 60 subjects chose the smaller, sooner (SS) reward. When the payment probability was reduced to 0.9, only 54% of 70 subjects chose the SS reward, and when the payment probability declined to 0.5, only 39% of 100 subjects chose the SS reward. These results suggest that probability and delay are behaviourally, empirically linked. However, follow-up studies by Weber and Chapman [2005], using hypothetical rewards, and Andersen, Harrison, Lau and Rutström [2014], using real rewards, failed to replicate this result.

Methodologically, the elicitation method of RRC suffers from a serious drawback. Suppose that someone's CE is $850 for a lottery that pays $1,000 with a probability of 0.9 and $0 with a probability of 0.1. This implies that certain amounts greater than $850 will be preferred to the gamble but certain amounts less than $850 will not. Given that people cannot state indifference in these experiments, and that the amounts used in the procedure were typically in increments of $50, the elicitation procedure will always over-estimate or under-estimate this true CE.[14] More generally, this elicitation method will always over-estimate or under-estimate a subject's CE unless it lies exactly midway between two of the certain values used in the task. This problem is magnified in experiments with hypothetical rewards because subjects have no incentive to represent their preferences truthfully (see Smith [1982]). In addition, if people make mistakes in their decisions[15] by, for example, selecting a less preferred option to a more preferred option then it becomes impossible to recover their "true," latent CEs.

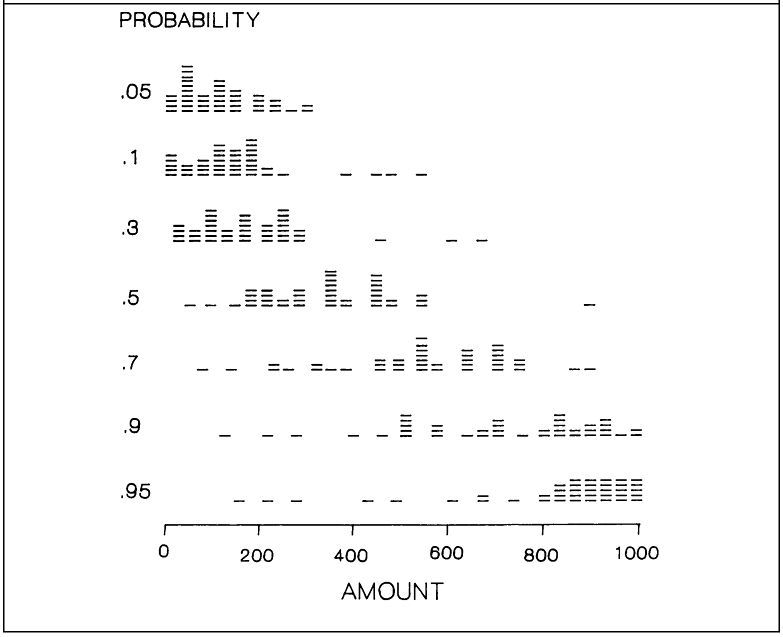
---

[14] For example, suppose a subject with this CE is offered the choice between $1,000 with a probability of 0.9 and $850 for sure. Given that the subject is indifferent between these options she may select either one. If the subject selects $850 for this pair but then switches to the gamble when offered $800 with certainty, as her preferences dictate, her assumed CE will be $825. If, on the other hand, the subject selects $1,000 with probability 0.9 over $850 with certainty, having selected the certain $900 previously, then her assumed CE will be $875. Thus, we derive two different and incorrect CEs for this subject given the algorithm that is used in the elicitation task.

[15] See Wilcox [2008] for a detailed survey of stochastic models of choice under risk that have been developed to account for this possibility.

This methodological issue raises an important statistical issue. Given that the estimate of a CE contains some error, one should be cognisant of the uncertainty of the estimate and model it statistically. In other words, every point estimate has a standard error and to ignore this sampling distribution is to assume it is degenerate. As the example above shows, this assumption cannot be the case, by design. Furthermore, to use some measure of central tendency, such as the median in RRC, to select the sample's CE for each probability, is to ignore the distribution of CEs in the sample. As Figure I shows, for every probability in the experimental task there is a distribution of CEs, and these distributions are often skewed, bimodal, or trimodal. Statistical techniques that treat each choice by every subject in the elicitation procedure as a datum should be used so that important information is not discarded and the uncertainty surrounding elicited CEs is modelled appropriately.

Figure I: Distributions of Certainty Equivalents in RRC



Source: RRC [p. 237, Figure 3].

Finally, even though the hyperbolic PD function may have better characterised the data than the exponential PD function, despite the lack of a formal statistical test showing this, there are a rich variety of PWFs that may better characterise these data and were not ruled out. For all of these reasons, it seems premature for RRC to claim that stated probability and stated delay have corresponding effects on behaviour.

IV. THE PD MODEL AND OTHER THEORIES OF CHOICE UNDER RISK

To show how the PD model relates to other theories of choice under risk it is important to recognise the lotteries to which the theory applies. The PD model is limited to a particular class of lotteries: simple prospects that take the form (x, p; 0, q). While the class of simple prospects is undoubtedly interesting, the PD model does not address gambles involving two non-zero outcomes (i.e., gambles with rewards $x > 0$ and $y > 0$), nor gambles with mixed domains (i.e., gambles with rewards $x > 0$ and $y < 0$), nor gambles involving more than two outcomes (i.e., gambles of the form (x, p; y, q; z, $1 - p - q$). The last characteristic requires that one move to a "rank dependent" specification, as proposed by Quiggin [1982] and incorporated by Tversky and Kahneman [1992] into cumulative prospect theory.

An issue that has been neglected in the literature on the PD model is that it employs the implicit assumption that the utility of a lottery is linear in outcomes. To see this, note that under the PD approach, researchers use an elicitation procedure to find the CE for the simple gamble (x, p; 0, q). Formally, the CE of a gamble is defined as the outcome or amount of money Z received with certainty that provides the same *utility* as the prospective utility of the gamble. In the case of a simple gamble:

$$v(Z) = \pi(p)v(x). \tag{9}$$

Compare equation (9) to equation (8) for the PD function: $V = x / (1 + \gamma\Theta) = \pi(\Theta)x = \pi(p)x$ since $\Theta = (1 - p) / p$. What this comparison shows is that PD researchers use a task to elicit $V = v(Z)$ but when they estimate γ they assume that the utility assigned to outcome x, $v(x)$ in (9), can be replaced with the nominal amount x in (8). In effect, they assume that the utility of a lottery is linear in outcome x such that $v(Z) = \pi(p)x$. A large body of empirical research[16] suggests that v is typically concave in outcomes. To assume that v is linear in x implies that risk preferences in the PD model are determined solely by the function $\pi(p)$ where $\pi(p) = 1 / (1 + \gamma\Theta)$ and $\Theta = (1 - p) / p$.

Thus, apart from the specific functional form for $\pi(p)$, the PD model is therefore exactly the dual theory of choice under risk due to Yaari [1987], the RDEV model,

[16] See, for example, Harless and Camerer [1994], Hey and Orme [1994], Holt and Laury [2002], Andersen, Harrison, Lau and Rutström [2008] and Harrison and Rutström [2008].

limited to simple prospects. Recall that RDEV incorporates the potential for non-linear transformations of probabilities while assuming that utility is linear in outcomes. This model is more general than the PD approach to risk preferences, however, because it admits different PWFs, lotteries with more than two prizes[17], and lotteries that incorporate both positive and negative prizes. Thus, the RDEV model is arguably the preferable theory if one insists on assuming linear utility.

When one recognises that the PD model is just the dual theory of choice under risk with a specific functional form for the PWF, the question of interest changes to whether this function is useful in applied research. Writing the PWF $\pi(p)$ in terms of probabilities p rather than odds against Θ, this function takes the following form:

$$\pi(p) = p \,/\, [p + \gamma(1 - p)] \tag{10}$$

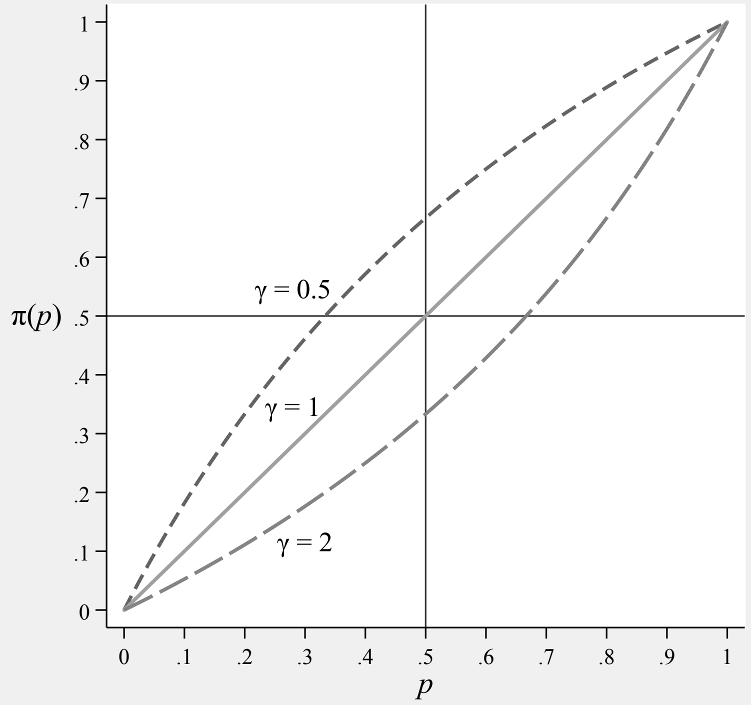
Figure II plots this PWF for different values of γ. When $\gamma = 1$, $\pi(p)$ is linear. When $\gamma < 1$, $\pi(p)$ is concave, which represents probability optimism and risk seeking, and when $\gamma > 1$, $\pi(p)$ is convex, which represents probability pessimism and risk aversion. This function is linear, concave, or convex throughout its range, so it is very similar to the power PWF used in some economic contexts: $\pi(p) = p^{\gamma}$.

KT [p. 280-284] argue on the basis of empirical evidence that people tend to overweight low probabilities ($\pi(p) > p$ for small p) and underweight moderate to high probabilities ($\pi(p) < p$ for moderate to high probabilities). This overweighting and underweighting yields an "inverse S-shaped" PWF $\pi(\cdot)$ with the following properties: subcertainty, subproportionality, and subadditivity.[18] These properties do not *define* the PT model, but are simply *common* assumptions within the model. To the extent that these assumptions are valid, the PWF of the PD model might be too restrictive because it cannot incorporate this inverse S-shape. To test this hypothesis, one can

[17] In these cases, the RDEV model applies rank-dependent non-linear transformations of probabilities so that first-order stochastic dominance is not violated.
[18] Subcertainty (see KT [p. 281-282]) means that the sum of decision weights, for $0 < p < 1$, is less than 1: $\pi(p) + \pi(1 - p) < 1$. Subproportionality (see KT [p. 282]) means, for a fixed ratio of probabilities, that the ratio of the corresponding decision weights is closer to 1 with low probabilities as opposed to high probabilities. Finally, subadditivity (see KT [p. 282]) arises if subproportionality holds and low probabilities are overweighted. Specifically, if $\pi(p) > p$, $\pi(\cdot)$ is monotone over (0, 1), and subproportionality holds, then $\pi(\alpha p) > \alpha\pi(p)$, when $0 < \alpha < 1$.

compare the PWF of the PD model to two functions commonly used in the economics literature.[19]

Figure II: PD Model Probability Weighting Function



Tversky and Kahneman (TK) [1992] popularised the PWF:

$$\pi(p) = p^{\gamma} / [p^{\gamma} + (1 - p)^{\gamma}]^{1/\gamma}, \tag{11}$$

which is defined for $1 > p > 0$. This function permits linear, inverse S-shaped, and S-shaped forms. Gonzalez and Wu [1999] review some early empirical evidence on this function and find that $1 > \gamma > 0$ in the bulk of the studies, most of which, unfortunately, used hypothetical rewards.[20] This range of γ gives the function[21] an inverse S-shape with overweighting of low probabilities up to a crossover point where

---

[19] Stott [2006] reviews the "menagerie" of PWFs that have been developed for models that incorporate subjective distortions of probabilities.

[20] By contract, Wilcox [2019], using real rewards and three different experiments, finds that probability optimism, i.e., overweighting of all interior probabilities, is the most common pattern of probability weighting exhibited by his experimental subjects. Three experiments employing different prizes and probabilities were used to test the sensitivity of his results to the lottery pairs presented to subjects, and evidence of the inverse S-shape was most prominent in his third experiment (see Wilcox [2019, p. 30-35]). This leads to Wilcox [2019, p. 36] to conclude that, "Asked what the probability weighting function looks like, the reply of a worldly experimenter might resemble that of the famously broad-minded corporate accountant: What do you want it to look like?"

[21] Ingersoll [2008] shows that this function is not monotonic at *very* small values of γ. However, these low values of γ are not empirically relevant.

$\pi(p) = p$, and then underweighting of moderate to high probabilities, thereby supporting the empirical hypothesis of KT.

Prelec [1998] axiomatically derived a flexible two-parameter PWF:

$$\pi(p) = \exp[-\eta(-\ln p)^{\gamma}], \qquad (12)$$

which is defined for $1 > p > 0$, $\eta > 0$, and $\gamma > 0$.[22] This function admits linear, inverse S-shaped, and S-shaped forms when $\eta = 1$, and incorporates objective weighting, underweighting, or overweighting *of all probabilitie*s when $\gamma = 1$. Thus, the Prelec [1998] function incorporates the qualitative properties of both the PD and TK functions for different parameter values.

The data from Richards, Zhang, Mitchell and de Wit (RZMW)[23] [1999] will be used to test the relevance of the PWF of the PD model in applied research in comparison to those provided by TK and Prelec [1998].

V. AN EMPIRICAL ANALYSIS OF THE PWF OF THE PD MODEL

RZMW recruited 24 subjects to take part in a within-subject experimental study, using real as opposed to hypothetical rewards, of the acute effects of moderate doses of alcohol on delay and probability discounting; I only discuss the probability discounting task here. A titration procedure was used to elicit CEs for simple prospects of the form (\$10, p; 0, q) where p took on the values: 1, 0.9, 0.75, 0.5, and 0.25. Participants attended four experimental sessions or treatments (a pre-placebo session, a post-placebo session, a pre-ethanol session, and a post-ethanol session) so each person provides 20 CEs for analysis. Appendix C of RZMW [p. 140] includes all of these elicited CEs, which can be used to determine the empirical validity of the PWF of the PD model.[24] These data are used because they are readily available, the

---

[22] Prelec [1998, proposition 1, part C, p. 503] provides these parameter restrictions. Prelec [1998, proposition 1, part B, p. 503] constrains $1 > \gamma > 0$, but this constraint can be quite restrictive in practice because it restricts the PWF to be inverse S-shaped.

[23] According to Google Scholar (22 July 2020), this article has been cited 794 times.

[24] Unfortunately RZMW do not provide the choice data that was used to derive the CEs and one is therefore obliged to analyse the CEs rather than the choice data itself. Nevertheless, the statistical approach adopted here is the appropriate method for analysing these data.

experiment was incentivised, and the experimental methodology of RZMW has been replicated numerous times.[25]

RZMW convert the five probabilities listed above into odds against winning and then use non-linear least squares (NLLS) estimation to find the best fitting hyperbolic function (8) for the CE data. RZMW [Appendix A, p. 138] estimate the value of γ in (8) for each subject in each experimental session. They also average the CEs in the pre-placebo and pre-ethanol sessions, and then estimate the value of γ for each subject (see RZMW [Table 1, p. 131]). Finally, they use the median CE of the sample for each probability to estimate a grand value of γ over all subjects (see RZMW [Figure 4, p. 132]).[26,27]

I adopt a different approach to data analysis that uses all of the data, elicited CEs and not statistics such as the mean or median of elicited CEs, provided by all of the subjects across all of the sessions to estimate the parameters of PWFs *at the level of the sample*. In addition, I formally incorporate the fact that each subject made multiple choices in the task and across the sessions by clustering the standard errors of the estimates. By using all of the information that the data provide while accounting for the lack of independence of observations, I am able to draw valid statistical inferences from these data. Note that NLLS and maximum likelihood (ML) estimators are consistent as the sample size n tends toward infinity. All of the RZMW estimates rely on 5 observations, so it is highly questionable whether any of the asymptotic properties of the estimators can be invoked to support the inferences they draw.

To compare the PWF of the PD model to those provided by TK and Prelec [1998], probabilities rather than odds against are used for estimation. Expressions (8) and (10) can be used derive the non-linear equation for estimating γ when odds against has been transformed back into probability:

---

[25] For example Mitchell [1999], Reynolds, Karraker, Horn and Richards [2003] and Reynolds, Richards, Horn and Karraker [2004] used this approach in incentivised studies, and Ohmura, Takahashi and Kitamura [2005] and Sheffer et al. [2013] used this approach in non-incentivised studies.

[26] As noted earlier, using the average and/or the median of elicited CEs discards information on the distribution of these data and does not allow one, therefore, to draw valid statistical inferences from these data.

[27] I have replicated these results but do not present them here because they are available in RZMW.

$$V_i/x = (p_i / [p_i + \gamma(1 - p_i)]) + \varepsilon_i, \qquad (13)$$

where the subscript i denotes each observation, V is the elicited CE that has been normalised by the reward x = \$10 for a probability p, and ε is the regression error term assumed to be a Normal distribution with mean zero and variance σ^2.

To estimate this model in a ML framework one must explicitly identify the log-likelihood by expanding (13):

$$\ln L_i(\gamma, \sigma; y, X) = \ln \Phi \ \{[V_i/x - (p_i / [p_i + \gamma(1 - p_i)])] / \sigma_i\} - \ln \sigma_i, \qquad (14)$$

where Φ is the standard Normal density with mean 0 and variance 1, y represents the data that are used to estimate γ and σ (i.e., V, x, and p), and X is a vector of individual characteristics and task parameters such as age, gender, and experimental treatment.

A ML framework is attractive because it is straightforward to allow for multiple responses by the same subject (i.e., clustering), perform non-nested model selection tests which rely on comparisons of the log-likelihoods of each observation in each model, make the parameter of interest, γ in (14), a linear function of observable characteristics, and estimate a mixture model of the different PWFs.[28] It is also a simple matter to adjust (14) to incorporate the TK and Prelec [1998] PWFs.

VI. RESULTS

Table I presents homogenous preference[29] estimates of the PD, TK and Prelec [1998] PWFs, which account for multiple responses by the same subject. The estimate of γ = 1.303 for the PD function implies underweighting of all probabilities, but γ is not statistically significantly greater than 1 (p = 0.094) so the function is effectively linear. By contrast, the estimate of γ = 0.754 for the TK PWF, which is significantly less than 1 (p < 0.001), implies overweighting of low probabilities and underweighting of moderate to high probabilities. Similarly, the estimates of η = 1.039 and γ = 0.793 for the Prelec [1998] PWF, where the former is not significantly different to 1 (p = 0.637) but the latter is significantly different to 1 (p < 0.001), also

[28] With NLLS it is also straightforward to incorporate clustering and to make the parameter of interest a linear function of observable characteristics. Thus, one benefits differentially from using a ML approach if one wants to conduct non-nested model selection tests or estimate mixture models.

[29] Homogenous preferences means that the models are estimated at the level of the sample without admitting heterogeneity in the coefficient estimates by incorporating observable characteristics such as age, gender, etc.

yields an inverse S-shaped function. Figure III plots the PWFs for the estimates in Table I.[30]

TABLE I: PROBABILITY WEIGHTING FUNCTION ML ESTIMATES
HOMOGENOUS PREFERENCES

	Model 1	**Model 2**	**Model 3**
	PD	TK	Prelec [1998]
PWF parameter γ	1.303***	0.754***	0.739***
	(0.181)	(0.053)	(0.062)
PWF parameter η			1.039***
			(0.082)
Sigma (σ)	0.141***	0.136***	0.136***
	(0.016)	(0.014)	(0.014)
N	480	480	480
log-likelihood	257.853	274.844	278.315

Results account for clustering at the individual level
Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

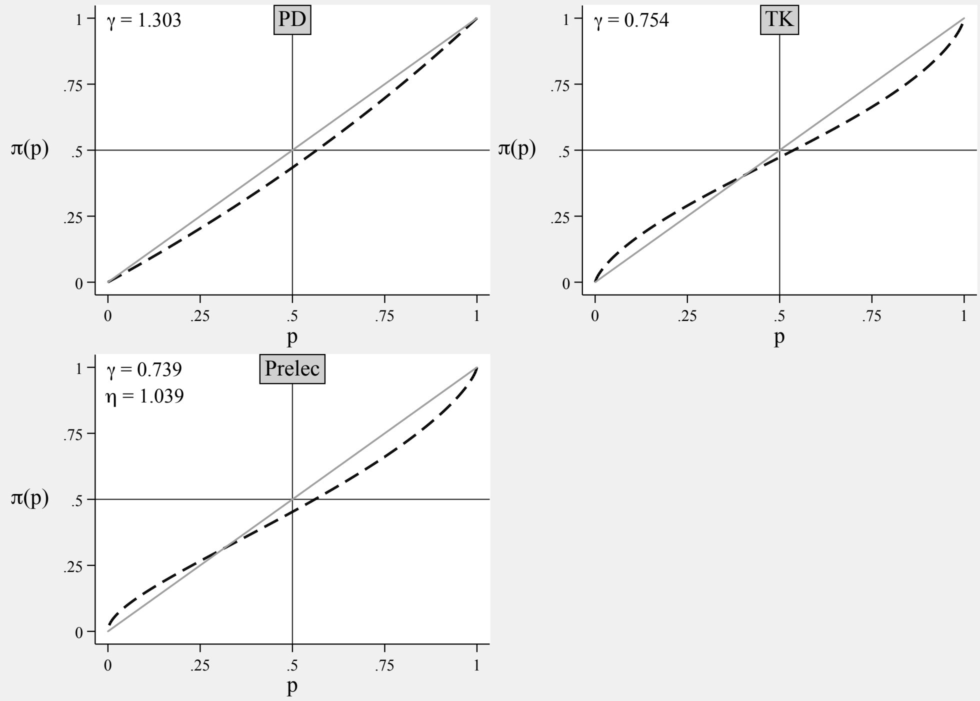
The log-likelihoods for the TK and Prelec [1998] functions exceed the log-likelihood for the PD function, suggesting that the TK and Prelec [1998] functions better characterise the data. This hypothesis can be tested formally using Vuong [1989] and Clarke [2007] non-nested model selection tests.[31] The Clarke test is more appropriate for these data (see Appendix D for more details) and leads to the following transitive ranking of PWFs: the Prelec [1998] function finds the most support in the data, followed by the TK function, and then the PD function. That the PD function finds the least support in the data is not surprising given the presence of inverse S-shaped probability weighting: in effect, the PD function is "confused" because it has to be linear, concave, or convex throughout its range.

Clarke [2007] tests are based on the implicit assumption that the observations are produced by only one data generating process (DGP), e.g., the PD, TK, or Prelec [1998] PWFs, when more may be present in the data. In other words, the PD function

[30] In Appendix C heterogenous preference models are estimated by making the parameters of interest a linear function of observable characteristics and task parameters.
[31] RZMW fit hyperbolic and exponential PD functions to each participant's elicited CEs, after probabilities are transformed into odds against winning. R^2 values are saved for each function and for each participant, and then used as data to construct tests of whether the hyperbolic or exponential functions provide better fits to the risk preference data. Using the point estimate of a statistic such as R^2 as a datum ignores the statistical imprecision of this point estimate and, thus, does not produce a valid test of one function's ability to better explain subject choices. Ignoring this issue, RZMW use Wilcoxon matched-pairs signed-rank tests on these *estimated* R^2 values and "find" that the hyperbolic PD functions explain the subject's choices significantly better than the exponential PD functions. These statistical procedures are, to be blunt, incoherent.

may explain some choices in the data better than the TK function whereas the TK function may explain other choices in the data better than the PD function. The assumption that only one DGP characterises all of the data clearly precludes such a possibility.

Figure III: Estimated Probability Weighting Functions



Mixture models[32] allow two or more DGPs to account for the data and also provide a measure of the proportion of the data that is explained by each process. In the current context one can estimate a mixture model of, say, the PD and TK PWFs and then ask the data to determine how much support each function has. To do so, one specifies a "grand likelihood" function which is just a probability-weighted average of the likelihoods of the two models.

Letting π^{PD} represent the probability that the PWF of the PD model is correct, and $\pi^{TK} = (1 - \pi^{PD})$ the probability that the TK function is correct, the grand likelihood is the probability-weighted average of the two conditional likelihoods L^{PD} and L^{TK} for the

[32] For detailed discussions of mixture models see McLachlan and Peel [2000], and for economic applications see Harrison and Rutström [2009] and Conte, Hey and Moffatt [2011].

PD and TK models, respectively. Thus, the likelihood for the mixture model is given by:

$$\ln L_i(\gamma^{PD}, \gamma^{TK}, \sigma, \kappa; y, X) = \Sigma_i \ln [(\pi^{PD} \times L^{PD}) + (\pi^{TK} \times L^{TK})], \qquad (15)$$

where κ is a parameter that defines the log odds of the probability of the PD model: $\pi^{PD} = 1 / (1 + \exp(\kappa))$. Note that this transformation allows the parameter κ to take on any value during the maximisation process but constrains the probabilities π^{PD} and π^{TK} to lie within the unit interval. The grand likelihood in (15) is maximised to estimate the parameters of each model and the weight accorded to each model in the data.

TABLE II: MIXTURE MODEL ML ESTIMATES
PD AND TK FUNCTIONS

	Estimate	Std error	*p*-value	95% Confidence interval	
	PD probability weighting function				
PWF parameter (γ^{PD})	4.802*	2.163	0.026	0.563	9.042
Mixture probability (π^{PD})	0.130	0.089	0.146	-0.045	0.304
	TK probability weighting function				
PWF parameter (γ^{TK})	0.855***	0.052	0.000	0.752	0.957
Mixture probability (π^{TK})	0.870***	0.089	0.000	0.696	1.045
	Sigma				
Constant (σ)	0.110***	0.016	0.000	0.079	0.141
N	480				
log-likelihood	298.681				

H_0: $\pi^{TK} = 1$, p-value = 0.146

Results account for clustering at the individual level
* p<0.05, ** p<0.01, *** p<0.001

Table II presents estimates of the mixture model of the PD and TK PWFs. The estimate of $\gamma^{PD} = 4.802$ is large and implies extreme underweighting of probabilities, but the 95% confidence interval shows that it is estimated very imprecisely. The mixture probability $\pi^{PD} = 0.130$ implies that approximately 13% of the data are best characterised by the PWF of the PD model, but this estimate is not significantly different to zero (p = 0.146).[33] The estimate of γ^{TK} = 0.855, which is significantly less than 1 (p = 0.005) implies overweighting of low probabilities and underweighting of moderate to high probabilities. Finally, the estimate of π^{TK} = 0.870 implies that approximately 87% of the data are best characterised by the TK PWF, but I cannot reject the hypothesis that this estimate is equal to 1 (p = 0.146). Thus, the PWF of the PD model finds almost no support in the data, even when it is allowed to account for only a fraction of the choices in the experiment.

The preceding results suggest that the PD model is too restrictive when used to classify risk preferences in the RZMW data. According to Clarke [2007] tests, the TK and Prelec [1998] PWFs find more support in the data than the PD function. A mixture model of the PD and TK PWFs confirms this result and shows that there is almost no support for the PD model.

VII. DISCUSSION AND CONCLUSIONS

The theoretical justification for the PD model is that choice under risk can be tied to a temporal framework by reinterpreting the probability of a reward as the delay to, or rate of reinforcement of, that reward. This reformulation of the probability concept is based on a behaviourist foundation of learning and conditioned reinforcement that the latest research in cognitive science, artificial intelligence, and neuroscience has called into question. For example, Daw et al. [2011, p. 1204] argue that, "… it has long been known that the reinforcement principle offers at best an incomplete account of learned action." Thus, Daw et al. [2011] discuss two different strategies for encoding value: a model-free strategy, which employs a temporal difference learning mechanism associated with the midbrain dopamine system and *reward* prediction errors, that is linked to habitual responses; and a model-based strategy, which is more closely associated with cortical regions and *state* prediction errors, that drives goal-directed actions. However, the latest research in this literature[34] suggests that this theoretical

[33] The lower bound of the 95% confidence interval for the mixture probability π^{PD} is less than 0 and the upper bound of the 95% confidence interval for the mixture probability π^{TK} is greater than 1, even though the log odds transformation constrains these probabilities to lie within the unit interval. These values lie outside the unit interval due to the use of the delta method (see Oehlert [1992]), which is an approximation, much like a Taylor series, to transform κ into the mixture probabilities.
[34] See Daw et al. [2011], Gershman and Daw [2012], Decker et al. [2016], Fleming and Daw [2017], Momennejad et al. [2017], and Russek et al. [2017].

dichotomy of two strategies is fast giving way to a more integrated view that, according to Clark [2016, p. 254], "combine[s] dense enabling webs of habit with sporadic bursts of genuine prospection."

The upshot is that the brain does appear to encode probability information directly under certain circumstances (see the review in Bach and Dolan [2012]) so reducing probability to delay to, or rate of reinforcement of, reward can be unnecessarily restrictive. Glimcher [2011] argues that this reduction is only justified in some contexts and the literature on the description-experience gap in choice under risk (see Hertwig et al. [2004] and Hertwig and Erev [2009]) is relevant in this regard.

The description-experience gap refers to the finding that people make different choices when presented with lotteries symbolically (i.e., ones that are described in terms of probabilities and associated prizes) as opposed to lotteries where subjects have to learn the probabilities of the prizes through experience (i.e., by choosing and playing out the lotteries repeatedly). In the former case, which is the approach that RRC and RZMW used, people sometimes exhibit inverse S-shaped PWFs and thereby overweight low probabilities, while in the latter case, which is closest to the thought experiment in RLGF, people apparently tend to underweight low probabilities instead.

Glimcher [2011] attributes these divergent results to the ways in which the brain represents probability information. In the experiential task, the brain employs a temporal difference learning mechanism that encodes probability and prize information in a single variable. It therefore embeds probability information but in a non-invertible way through reward prediction errors in the midbrain dopamine system, the learning rate of which leads to the underweighting of low probabilities. In the symbolic task, by contrast, the probability information is encoded directly, and in an invertible way, in parts of the prefrontal cortex and ventral striatum, leading to the overweighting of low probabilities.

To the extent that the argument of Glimcher [2011] is valid[35] it seems misguided to reinterpret probability as delay to, or rate of reinforcement of, reward, particularly in

[35] See the review in Clark [2016] that provides further evidence to support it.

experiments that represent lotteries symbolically. Given these theoretical issues with the PD model, should it still be used in applied research?

The PD model is formally isomorphic to the dual theory of choice under risk due to Yaari [1987], but limited to a certain class of lotteries and with a specific functional form for the PWF, implying that it is an unnecessarily restrictive theory for use in empirical work. In addition, there are separate issues with the ways in which PD data are typically elicited and analysed. With regard to elicitation, experimental procedures used in this literature have the unfortunate feature that they will always over-estimate or under-estimate a subject's true CE, unless it lies exactly halfway between two of the certain values used in the task. In addition, the data derived from these procedures (i.e., midpoints of an interval) are inherently noisy and this has to be, but seldom is, taken into account when drawing inferences from these data. Thus, statistically, researchers need to be cognisant of the DGP and build this into their framework for analysis.

This is the approach adopted in re-analyses of the RZMW data: all of the data provided by all the subjects are used to estimate ML models at the level of the sample, while clustering the standard errors of the estimates to incorporate the fact that participants made multiple choices in the task. This allows one to conduct formal Clarke [2007] tests of the ability of the different PWFs to characterise the RZMW data. The PWF of the PD model found the least support in the data because it could not account for the simultaneous overweighting and underweighting of probabilities which is a feature of these data. Furthermore, in a complementary mixture model analysis, the PWF of the PD model found virtually no support, even when it was allowed to account only for a portion of the data.

The preceding discussion suggests that the PD model inadequately characterises choice under risk. It is based on a theoretical reformulation of the probability concept that is outdated and suffers from a host of methodological and statistical limitations. If one insists on assuming linear utility then the model of Yaari [1987], which admits different PWFs, lotteries with more than two prizes, and lotteries that incorporate both positive and negative prizes, is clearly the preferable theory. But researchers interested in choice under risk would arguably be better served by adopting incentive-compatible elicitation tasks; theories that incorporate multiple pathways to explain the risk premium, not just probability weighting; and statistical tools that respect the DGP and incorporate the possibility of multiple DGPs.

REFERENCES

ANDERSEN, S., G. W. HARRISON, M. I. LAU, AND E. E. RUTSTRÖM (2008): "Eliciting Risk and Time Preferences," *Econometrica*, 76, 583-618.

ANDERSEN, S., G. W. HARRISON, M. I. LAU, AND E. E. RUTSTRÖM (2014): "Discounting Behavior: A Reconsideration," *European Economic Review*, 71, 15-33.

BACH, D. R., AND R. J. DOLAN (2012): "Knowing How Much You Don't Know: A Neural Organization of Uncertainty Estimates," *Nature Reviews Neuroscience*, 13, 572-586.

CHUNG, S. H. (1965): "Effects of Delayed Reinforcement in a Concurrent Situation," *Journal of the Experimental Analysis of Behavior*, 8, 439-44.

CHUNG, S. H., AND R. J. HERRNSTEIN (1967): "Choice and Delay of Reinforcement," *Journal of the Experimental Analysis of Behavior*, 10, 67-74.

CLARK, A. (2016): *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*. Oxford: Oxford University Press.

CLARKE, K. A. (2007): "A Simple Distribution-Free Test for Nonnested Model Selection," *Political Analysis*, 15, 347-363.

CONTE, A., J. D. HEY, AND P. G. MOFFATT (2011): "Mixture Models of Choice under Risk," *Journal of Econometrics*, 162, 79-88.

DAVENPORT, J. W. (1962): "The Interaction of Magnitude and Delay of Reinforcement in Spatial Discrimination," *Journal of Comparative and Physiological Psychology*, 55, 267-73.

DAW, N. D., S. J. GERSHMAN, B. SEYMOUR, P. DAYAN, AND R. J. DOLAN (2011): "Model-Based Influences on Humans' Choices and Striatal Prediction Errors," *Neuron*, 69, 1204-1215.

DECKER, J. H., A. R. OTTO, N. D. DAW, AND C. A. HARTLEY (2016): "From Creatures of Habit to Goal-Directed Learners: Tracking the Developmental Emergence of Model-Based Reinforcement Learning," *Psychological Science*, 27, 848-58.

FLEMING, S. M., AND N. D. DAW (2017): "Self-Evaluation of Decision-Making: A General Bayesian Framework for Metacognitive Computation," *Psychological Review*, 124, 91-114.

GERSHMAN, S. J., AND N. D. DAW (2012): "Percetion, Action and Utility: The Tangled Skein," in *Principles of Brain Dynamics: Global State Interactions*, ed. by M. Rabinovich, M. K. Friston, and P. Varona. Cambridge, MA: MIT Press, 293-312.

GLIMCHER, P. W. (2011): *Foundations of Neuroeconomic Analysis*. Oxford: Oxford University Press.

GONZALEZ, R., AND G. WU (1999): "On the Shape of the Probability Weighting Function," *Cognitive Psychology*, 38, 129-166.

GREEN, L., AND J. MYERSON (2004): "A Discounting Framework for Choice with Delayed and Probabilistic Rewards," *Psychological Bulletin*, 130, 769-92.

HARLESS, D. W., AND C. F. CAMERER (1994): "The Predictive Utility of Generalized Expected Utility Theories," *Econometrica*, 62, 1251-1289.

HARRISON, G. W., AND E. E. RUTSTRÖM (2008): "Risk Aversion in the Laboratory," in *Research in Experimental Economics: Volume 12. Risk Aversion in Experiments*, ed. by J. C. Cox, and G. W. Harrison. Bingley: Emerald, 41-196.

HARRISON, G. W., AND E. E. RUTSTRÖM (2009): "Expected Utility Theory *and* Prospect Theory: One Wedding and a Decent Funeral," *Experimental Economics*, 12, 133-158.

HERRNSTEIN, R. J. (1961): "Relative and Absolute Strength of Response as a Function of Frequency of Reinforcement," *Journal of the Experimental Analysis of Behavior*, 4, 267-72.

HERRNSTEIN, R. J. (1970): "On the Law of Effect," *Journal of the Experimental Analysis of Behavior*, 13, 243-66.

HERTWIG, R., G. BARRON, E. U. WEBER, AND I. EREV (2004): "Decisions from Experience and the Effect of Rare Events in Risky Choice," *Psychological Science*, 15, 534-9.

HERTWIG, R., AND I. EREV (2009): "The Description-Experience Gap in Risky Choice," *Trends in Cognitive Sciences*, 13, 517-23.

HEY, J. D., AND C. ORME (1994): "Investigating Generalizations of Expected Utility Theory Using Experimental Data," *Econometrica*, 62, 1291-1326.

HOLT, C. A., AND S. K. LAURY (2002): "Risk Aversion and Incentive Effects," *The American Economic Review*, 92, 1644-1655.

INGERSOLL, J. (2008): "Non-Monotonicity of the Tversky-Kahneman Probability-Weighting Function: A Cautionary Note," *European Financial Management*, 14, 385-390.

KAHNEMAN, D., AND A. TVERSKY (1979): "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47, 263-291.

KEREN, G., AND P. ROELOFSMA (1995): "Immediacy and Certainty in Intertemporal Choice," *Organizational Behavior and Human Decision Processes*, 63, 287-297.

KYONKA, E. G. E., AND N. S. SCHUTTE (2018): "Probability Discounting and Gambling: A Meta-Analysis," *Addiction*, 113, 2173-2181.

LOGAN, F. A. (1965): "Decision Making by Rats: Delay Versus Amount of Reward," *Journal of Comparative and Physiological Psychology*, 59, 1-12.

MADDEN, G. J., AND W. K. BICKEL (2010): *Impulsivity: The Behavioral and Neurological Science of Discounting*. Washington, DC: American Psychological Association.

MAZUR, J. E. (1984): "Tests of an Equivalence Rule for Fixed and Variable Reinforcer Delays," *Journal of Experimental Psychology: Animal Behavior Processes*, 10, 426-437.

MCLACHLAN, G., AND D. PEEL (2000): *Finite Mixture Models*. New York, NY: Wiley.

MITCHELL, S. H. (1999): "Measures of Impulsivity in Cigarette Smokers and Non-Smokers," *Psychopharmacology (Berl)*, 146, 455-64.

MOMENNEJAD, I., E. M. RUSSEK, J. H. CHEONG, M. M. BOTVINICK, N. D. DAW, AND S. J. GERSHMAN (2017): "The Successor Representation in Human Reinforcement Learning," *Nature Human Behaviour*, 1, 680-692.

OEHLERT, G. W. (1992): "A Note on the Delta Method," *The American Statistician*, 46, 27-29.

OHMURA, Y., T. TAKAHASHI, AND N. KITAMURA (2005): "Discounting Delayed and Probabilistic Monetary Gains and Losses by Smokers of Cigarettes," *Psychopharmacology*, 182, 508-15.

OSTASZEWSKI, P. (1997): "Temperament and the Discounting of Delayed and Probabilistic Rewards: Conjoining European and American Psychological Traditions," *European Psychologist*, 2, 35-43.

PRELEC, D. (1998): "The Probability Weighting Function," *Econometrica*, 66, 497-527.

QUIGGIN, J. (1982): "A Theory of Anticipated Utility," *Journal of Economic Behavior & Organization*, 3, 323-343.

RACHLIN, H., A. CASTROGIOVANNI, AND D. CROSS (1987): "Probability and Delay in Commitment," *Journal of the Experimental Analysis of Behavior*, 48, 347-53.

RACHLIN, H., AND L. GREEN (1972): "Commitment, Choice and Self-Control," *Journal of the Experimental Analysis of Behavior*, 17, 15-22.

RACHLIN, H., A. W. LOGUE, J. GIBBON, AND M. FRANKEL (1986): "Cognition and Behavior in Studies of Choice," *Psychological Review*, 93, 33-45.

RACHLIN, H., A. RAINERI, AND D. CROSS (1991): "Subjective Probability and Delay," *Journal of the Experimental Analysis of Behavior*, 55, 233-44.

RESCORLA, R. A., AND A. R. WAGNER (1972): "A Theory of Pavlovian Conditioning: Variations in the Effectiveness of Reinforcement and Nonreinforcement," in *Classical Conditioning II: Current Research and Theory*, ed. by A. H. Black, and W. F. Pofasky. New York, NY: Appleton-Century-Crofts, 64-99.

REYNOLDS, B. (2006): "The Experiential Discounting Task Is Sensitive to Cigarette-Smoking Status and Correlates with a Measure of Delay Discounting," *Behavioural Pharmacology*, 17, 133-42.

REYNOLDS, B., K. KARRAKER, K. HORN, AND J. B. RICHARDS (2003): "Delay and Probability Discounting as Related to Different Stages of Adolescent Smoking and Non-Smoking," *Behavioural Processes*, 64, 333-344.

REYNOLDS, B., M. PATAK, P. SHROFF, R. B. PENFOLD, S. MELANKO, AND A. M. DUHIG (2007): "Laboratory and Self-Report Assessments of Impulsive Behavior in Adolescent Daily Smokers and Nonsmokers," *Experimental and Clinical Psychopharmacology*, 15, 264-71.

REYNOLDS, B., J. B. RICHARDS, K. HORN, AND K. KARRAKER (2004): "Delay Discounting and Probability Discounting as Related to Cigarette Smoking Status in Adults," *Behavioural Processes*, 65, 35-42.

RICHARDS, J. B., L. ZHANG, S. H. MITCHELL, AND H. DE WIT (1999): "Delay or Probability Discounting in a Model of Impulsive Behavior: Effect of Alcohol," *Journal of the Experimental Analysis of Behavior*, 71, 121-43.

RUSSEK, E. M., I. MOMENNEJAD, M. M. BOTVINICK, S. J. GERSHMAN, AND N. D. DAW (2017): "Predictive Representations Can Link Model-Based Reinforcement Learning to Model-Free Mechanisms," *PLoS Computational Biology*, 13, e1005768.

SHEFFER, C. E., M. MENNEMEIER, R. D. LANDES, W. K. BICKEL, S. BRACKMAN, J. DORNHOFFER, T. KIMBRELL, AND G. BROWN (2013): "Neuromodulation of Delay Discounting, the Reflection Effect, and Cigarette Consumption," *Journal of Substance Abuse Treatment*, 45, 206-14.

SMITH, V. L. (1982): "Microeconomic Systems as an Experimental Science," *American Economic Review*, 72, 923-955.

STOTT, H. P. (2006): "Cumulative Prospect Theory's Functional Menagerie," *Journal of Risk and Uncertainty*, 32, 101-130.

TVERSKY, A., AND D. KAHNEMAN (1992): "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, 5, 297-323.

VUONG, Q. H. (1989): "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses," *Econometrica*, 57, 307-333.

WEBER, B., AND G. B. CHAPMAN (2005): "The Combined Effects of Risk and Time on Choice: Does Uncertainty Eliminate the Immediacy Effect? Does Delay Eliminate the Certainty Effect?," *Organizational Behavior and Human Decision Processes*, 96, 104-118.

WILCOX, N. T. (2008): “Stochastic Models for Binary Discrete Choice under Risk: A Critical Primer and Econometric Comparison,” in *Research in Experimental Economics: Volume 12. Risk Aversion in Experiments*, ed. by J. C. Cox, and G. W. Harrison. Bingley: Emerald, 197-292.

WILCOX, N. T. (2019): “Unusual Estimates of Probability Weighting Functions,” Working Paper WP 2019-05, Center for the Economic Analysis of Risk (CEAR), Georgia State University

YAARI, M. E. (1987): “The Dual Theory of Choice under Risk,” *Econometrica*, 55, 95-115.

APPENDIX A
[ONLINE WORKING PAPER]

RLGF [p. 35] explain that a behavioural model of choice is one where an animal's behaviour in a specific environment is described by two sets of rules. The first consists of reinforcement schedules that are independent of the animal's behaviour while the second describes an animal's behaviour as a function of exposure to environmental stimuli. The following behavioural model of choice has been used to explain a large proportion of animal choice data:

$$B_1/B_2 = (A_1/A_2)^a(R_1/R_2)^r(D_2/D_1)^d, \quad (1)$$

where B_1 and B_2 are rates of responding on two levers, buttons etc., during a particular time interval, A_1 and A_2 are the respective reinforcement amounts that are delivered over that interval, R_1 and R_2 are the respective rates of reinforcement over that interval, D_1 and D_2 are the respective delays to the rewards, and a, r and d represent an animal's sensitivity to amount, rate and delay, respectively.

This function occupies a venerable space in the experimental literature on animals' choice behaviour between different schedules of reinforcement. It is a generalised version of Herrnstein's (1961) matching law and it has been used to explain the commonly found tendency of animals to sharply discount delayed rewards (see RLGF [p. 36]).[1]

Although this function was derived to provide an account of animals' delay discounting behaviour, RLGF employ it to explain how people value probabilistic rewards. As discussed in the main text, the crucial step in this reformulation is to interpret the probability associated with a reward as the delay to, or the rate of reinforcement of, this reward.

In experiments where people choose between lotteries, the two effects of probability (i.e., on the rate of reinforcement of a reward and the delay to a reward) are confounded. Following RLGF [p. 39], to see how this affects the generalised matching law equation (1), let $R_1 = 1/D_1$

[1] Kagel, Battalio and Green [1995] critique the matching law and its implications for the commodity-choice behaviour [p. 51-71], labour-supply behaviour [p. 110-128], and time discounting [p. 178-180] of animals, predominantly rats and pigeons, under experimental conditions. They argue in favour of an economic account of animal choice behaviour that relies on maximising, i.e., the comparison of marginal rates of return, rather than matching, i.e., the comparison of average rates of return.

and $R_2 = 1/D_2$ so that rate and delay are perfectly confounded. This yields the following equation:

$$B_1/B_2 = (A_1/A_2)^a(D_2/D_1)^{r+d} \tag{2}$$

Using (5) in the main text and assuming that $c_1 = c_2 = c$ and $t_1 = t_2 = t$, (2) becomes:

$$B_1/B_2 = (A_1/A_2)^a([p_1(t(1 - p_2) + c)] / [p_2(t(1 - p_1) + c)])^{r+d} \tag{3}$$

Expression (3) represents the matching law for probabilistic rewards when the rate of reinforcement of a reward is perfectly confounded with the delay to that reward.

RLGF derive this expression to provide a behavioural account of choice among gambles: preference for lottery 1 over lottery 2, as measured by rate of responding (B_1/B_2) over a time interval, is a function of the rewards in these lotteries (A_1 and A_2), the probabilities associated with these rewards (p_1 and p_2), the reward delivery period c, the ITI length t, and the sensitivity exponents a, r, and d. RLGF use (3) and set $p_2 = 1$, $t = 1$, $c = 0.2$, and $A_1 = A_2$, in an attempt to replicate KT's assertion that people overweight low probabilities and underweight moderate to high probabilities. Under these assumptions, (3) becomes:

$$B_1/B_2 = [(0.2p_1) / (1.2 - p_1)]^{r+d} \tag{4}$$

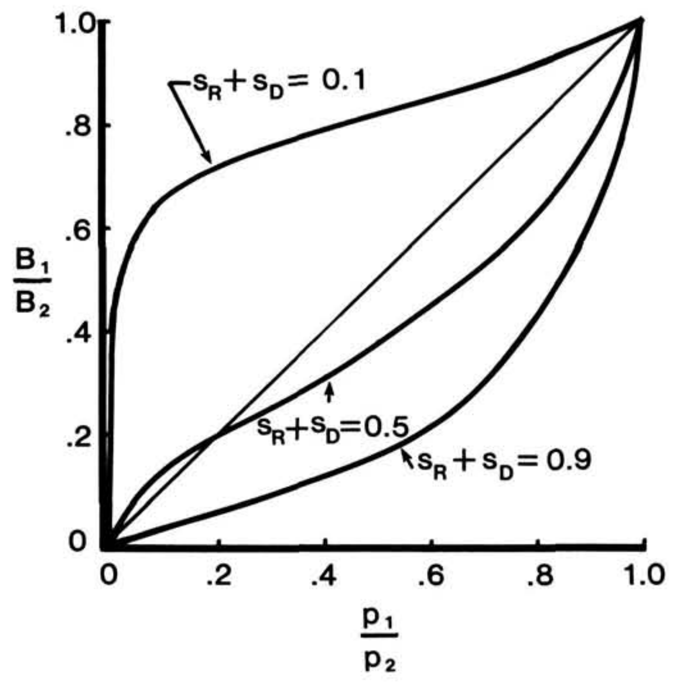
Plotting B_1/B_2 as a function of p_1/p_2 (where $p_2 = 1$) for different values of $r + d$, yields Figure A:I. In the figure, $s_R = r$ and $s_D = d$ according to the notation that I have used. RLGF [p. 39] argue that when $r + d = 0.5$, the plotted function is their model's counterpart to the inverse S-shaped PWF of KT which is based on the overweighting of low probabilities and the underweighting of moderate to high probabilities.[2] Note that a PWF relates stated probabilities (i.e., p_1 and p_2) to subjective decision weights: $\pi(p_1)$ and $\pi(p_2)$. Figure A:I, by contrast, plots relative rate of responding (B_1/B_2) as a function of relative probabilities (p_1/p_2). While the shape of the curve in Figure A:I when $r + d = 0.5$ resembles an inverse S-shaped PWF, it does not relate probabilities to decision weights and hence it is inaccurate for RLGF to claim that they have translated KT's model into a behavioural model of choice.

Furthermore, Figure A:I was plotted on the assumptions that $A_1 = A_2$ and $p_2 = 1$. In other words, the rewards under both lotteries are the same but reward A_2 is received with certainty

[2] As stated previously under expression (1) in the main text, $\pi(0) = 0$ and $\pi(1) = 1$, which, when coupled with the properties of subadditivity, subcertainty, and subproportionality, means the KT PWF has jump discontinuities at $p = 0$ and $p = 1$. KT [p. 282-283] suggest that these discontinuities may capture the distinction between certainty and uncertainty and conclude that their function is not well-behaved near the end points of the probability interval [0, 1]. Figure A:I does not incorporate these jump discontinuities.

under lottery 2. Over almost its entire range, Figure A:I has $p_1 < 1$ but the figure suggests that lottery 1 will still be chosen some of the time (i.e., $B_1/B_2 \neq 0$). This means that a lottery which pays, say, \$100 with a probability less than one and \$0 otherwise will be chosen over a degenerate lottery which pays \$100 for sure. While KT [p. 284] recognise that their model can lead to violations of stochastic dominance, and subsequently developed cumulative prospect theory to account for this, they argued that such violations were unlikely because dominated prospects would be removed during the initial editing phase. Clearly this is not the case in the model of RLGF because Figure A:I implies that the stochastically dominated lottery will be chosen at least some of the time.

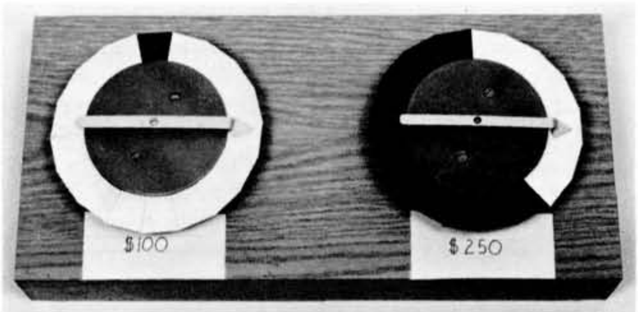
Figure A:I



Source: RLGF [p. 40, Figure 4].

A testable implication of (3) is whether a longer ITI t will affect choice among probabilistic rewards. RLGF conducted an experiment to test this hypothesis, one which was replicated by Silberberg, Murray, Christensen and Asano (SMCA) [1988]. Both studies used hypothetical rewards so their results should be treated with caution. In the experiment, subjects chose between two spinners, each of which rotated over a circle made up of 18 pie-shaped wedges (see Figure A:II). The wedges were black on one side and white on the other. If the spinner landed on a white wedge, the subject was told that she had won, if it landed on a black wedge, the subject was told that she had lost. One spinner, referred to as the "sure thing," had 17 white wedges and 1 black wedge whereas the other spinner, referred to as the "risky gamble," had 7 white wedges and 11 black wedges. The hypothetical payoff for the sure thing was \$100 and the hypothetical payoff for the risky gamble was \$250.

Figure A:II



Source: RLGF [p. 40, Figure 4].

The number of white wedges on the sure thing spinner remained constant throughout the experiment. However, the number of white wedges on the risky gamble spinner was adjusted depending on the subject's previous choice: a choice of the sure thing resulted in one more white wedge on the risky gamble (i.e., an increase in the odds of winning on the risky gamble) whereas a choice of the risky gamble led to one less white wedge on the risky gamble (i.e., a decrease in the odds of winning on that spinner). This titration procedure[3] was used so that, at equilibrium, the subject would be indifferent between the sure thing and the risky gamble (i.e., B_1 and B_2 in (3) would be equal at equilibrium). Subjects made 10 choices in total in the experiment.[4]

To see the effect of a longer ITI on choice between the two gambles, note that $A_1 = 100$, $A_2 = 250$, p_1 is approximately equal to 1[5], $c_1 = c_2 = 5s$, $t_1 = t_2 = t$ and, for simplicity, $a = r + d = 1$.

[3] Titration procedures are susceptible to being "gamed" by subjects and do not, therefore, promote truthful revelation of preferences, i.e., they lack incentive compatibility. For example, subjects may disproportionately choose the sure thing spinner on the first few trials so as to increase the odds of winning on the risky gamble in the final trials of the experiment. This point is moot for studies involving hypothetical rewards as these lack incentive compatibility to begin with, but should be taken into account for studies with titration procedures and real rewards.

[4] Whether 10 choices are enough to reach equilibrium is an open question and one which is not taken up by RLGF.

[5] KT [p. 265] argue that people respond differently to certain outcomes as opposed to near-certain outcomes. RLGF ignore this point by assuming that when $p = 17/18$ this is the same as $p = 1$. In other words, they treat an inherently risky prospect as one involving no risk.

Substituting these values into (3), remembering that at equilibrium $B_1 = B_2$, and solving for p_2:

$$p_2 = (t + 5) / (t + 12.5) \tag{5}$$

When the ITI $t = 0$, $p_2 = 0.4$. In other words, a person choosing between the spinners would be indifferent between the sure thing and the risky gamble when the probability of receiving $250 under the risky gamble is 0.4. However, as t increases so $p_2 \rightarrow 1$. Thus, with a long ITI, a subject is only indifferent between the sure thing and the risky gamble when the likelihood of receiving $250 under the risky gamble is approximately 1. This is precisely the implication that RLGF and SMCA set out to test.[6]

RLGF found evidence in support of this hypothesis in a between-subject experimental design involving 30 subjects, where one group's ($n_1 = 15$) ITI was 30*s* and the other group's ($n_2 = 15$) ITI was 90*s*.[7] Specifically, the group with the longer ITI selected the sure thing spinner more often, which means that the odds of winning on the risky gamble was higher, than the group with the shorter ITI. This comparison used the number of risky gamble choices of each subject over the course of the experiment as the data on which to conduct a *t*-test. Subjects in the 30*s* ITI group selected the risky gamble an average of 5.87 times, whereas the subjects in the 90*s* ITI group selected the risky gamble an average of 3.93 times, over the course of the experiment; this difference was statistically significant, $t = 4.65$, df = 28, $p < 0.01$.

SMCA replicated RLGF's experiment, albeit with three procedural differences: SMCA used a computer, rather than an experimenter, to present and record subjects' choices; SMCA added an additional experimental treatment: some of the subjects began the experiment with a choice trial while others[8] started the experiment with an ITI; and one group had an ITI of 25*s* as compared to the RLGF study where the comparable group had an ITI of 30*s*.

[6] This implication is at odds with the literature on the matching law applied to choice among delayed rewards. This literature, cited previously, suggests that as the delay to all rewards increases, the likelihood of selecting the larger, more delayed reward, rather than the smaller, more immediate reward, increases. If people understand probability as delay then a low probability is equivalent to a long delay. So as the ITI increases (i.e., as the delay to both rewards increases), subjects should be more likely to select the larger, more uncertain reward (viz., the larger, more delayed reward). This works against the hypothesis that RLGF sought to test.
[7] With different ITIs, temporal discounting behaviour could drive the results in the experiment. RLGF implicitly assume that their subjects did not discount delayed rewards.
[8] Unfortunately SMCA do not provide the exact number of subjects in each experimental treatment.

SMCA recruited 101 subjects to take part in the experiment. They compared the number of white wedges on the risky gamble spinner in the final round of the experiment across the groups (choice trial vs. ITI × 25*s* vs. 90*s*), using a Kruskal-Wallis test, and found a significant between-group difference: $H(3) = 15.8$, $p < 0.05$. Using pairwise post hoc contrasts, they found no statistically significant differences[9] between subjects who began the experiment with a choice trial and subjects who began the experiment with an ITI. However, subjects in the short ITI group, regardless of whether they started the experiment with a choice trial or ITI, selected the risky gamble spinner significantly more often than subjects in the long ITI group. Thus, SMCA replicated the result of RLGF.

In a follow-up experiment, SMCA decided to test the robustness of this result by telling subjects how many choices (i.e., 10) they would make in the experiment; this information was not provided in SMCA's original experiment nor the experiment of RLGF. Using a sample of 20 students, SMCA found that there was no significant difference in the number of risky gamble choices between the group with the short ITI and the group with the long ITI, although they did not provide test statistics or *p*-values for their comparisons.[10]

Finally, SMCA conducted another experiment with 40 students where, in addition to the different ITIs, they told one group of subjects that they had been endowed with \$10 of hypothetical money and the other group was told that they had been endowed with \$10,000 of hypothetical money. This information was given to subjects at the start of the experiment and they were also told that the experiment consisted of 10 choice trials. They found differences between the group endowed with \$10 and the group endowed with \$10,000 but no differences between the short ITI and the long ITI groups. In sum, SMCA replicated the result in RLGF but found that the ITI effect disappeared as soon as subjects were told how many choices they would have to make in the experiment. Thus, the results reported by RLGF appear to be very sensitive to the information provided to subjects.

Ignoring the sensitivity of these results for the moment, what does the RLGF model imply when these different ITIs are used? Using (5), when $t = 30s$, $p_2 \approx 0.82 \approx 15/18$ white wedges on the risky gamble. When $t = 90s$, $p_2 \approx 0.84 \approx 15/18$ white wedges on the risky gamble. So,

[9] SMCA do not report the test statistics nor *p*-values for the pairwise post hoc contrasts.
[10] SMCA dropped the treatment where one group of subjects started the experiment with a choice trial while the other group started with an ITI.

for these parameters, the RLGF model implies very little difference in the choice behaviour of the two groups.[11] RLGF [p. 41] found that, at the end of the experiment, the group with the longer ITI had an average of 9 white wedges on the risky gamble whereas the group with the shorter ITI had an average of 6 white wedges on the risky gamble. Thus, the model of RLGF not only predicts very little difference between the groups it also vastly overestimates the number of white wedges on the risky gamble spinner.

Finally, note what the RLGF model implies for the relationship between risk aversion and the length of the ITI. A person is risk averse if she prefers the certain outcome x to a gamble with an expected value of x. From (5) it is clear that as the ITI increases, p_2 increases, which means that as the ITI increases, risk aversion increases too. Pushing this logic to its natural conclusion, a long enough ITI would generate an aversion to risk that would make a small, certain reward preferable to a far larger, near-certain reward. While this may apply to some agents in some circumstances, empirical evidence supporting the general validity of this prediction is not provided by RLGF.

ADDITIONAL REFERENCES

BARDSLEY, N., R. CUBITT, G. LOOMES, P. MOFFATT, C. STARMER, AND R. SUGDEN (2010): *Experimental Economics: Rethinking the Rules*. Princeton, NJ: Princeton University Press.

HARRISON, G. W. (1989): "Theory and Misbehavior of First-Price Auctions," *American Economic Review*, 79, 749-762.

HARRISON, G. W. (1992): "Theory and Misbehavior of First-Price Auctions: Reply," *American Economic Review*, 82, 1426-1443.

HARRISON, G. W. (1994): "Expected Utility Theory and the Experimentalists," *Empirical Economics*, 19, 223-253.

[11] Harrison [1989, 1992, 1994] provides a detailed critique of inferences drawn from experimental studies when the studies fail to satisfy Smith's [1982] precepts for controlled microeconomic experiments. Harrison [1989, 1992, 1994] shows that deviations from "optimal" behaviour in many experiments often entail such a low cost in terms of forgone expected income that experimental subjects have almost no incentive to find the theoretical optimum nor reveal their preferences with any real precision. In other words, some experimental tasks do not adequately incentivise subjects to reveal their "true" preferences because the expected increase in earnings from doing so does not offset the required cognitive effort. Thus, many experiments fail to satisfy Smith's [1982] precept of "dominance," which requires that the experimental reward medium dominates the subjective costs of decision making in the experiment; Harrison refers to this as the "payoff-dominance critique" whereas Bardsley et al. [2010] refer to this as the "flat maximum critique." Morgan and Tustin [1992, p. 1142-1143] make a similar point in the context of labour supply choices by pigeons: to locate the optimum in some decision problems requires the comparison of margins which are so small that only a fully informed, optimising agent could conceivably evaluate. In these cases, the decision problems do not provide a valid test of the theory. The experiment of RLGF is a case in point.

KAGEL, J. H., R. C. BATTALIO, AND L. GREEN (1995): *Economic Choice Theory: An Experimental Analysis of Animal Behavior*. New York, NY: Cambridge University Press.

MORGAN, P. B., AND D. TUSTIN (1992): "The Perception and Efficiency of Labour Supply Choices by Pigeons," *The Economic Journal*, 102, 1134-1148.

SILBERBERG, A., P. MURRAY, J. CHRISTENSEN, AND T. ASANO (1988): "Choice in the Repeated-Gambles Experiment," *Journal of the Experimental Analysis of Behavior*, 50, 187-195.

SMITH, V. L. (1982): "Microeconomic Systems as an Experimental Science," *American Economic Review*, 72, 923-955.

APPENDIX B
[ONLINE WORKING PAPER]

RCC (1987): Probability and Delay in Commitment

RCC conducted an experiment with real rewards to bolster RLGF's claim that probabilistic choice can be tied to a temporal framework. The experiment was based on Rachlin and Green [1972] who studied the behaviour of pigeons in a delay-commitment paradigm. The major difference with this experiment, other than that RCC used human subjects, was that delays were replaced with probabilities: long delays with low probabilities and short delays with high probabilities.

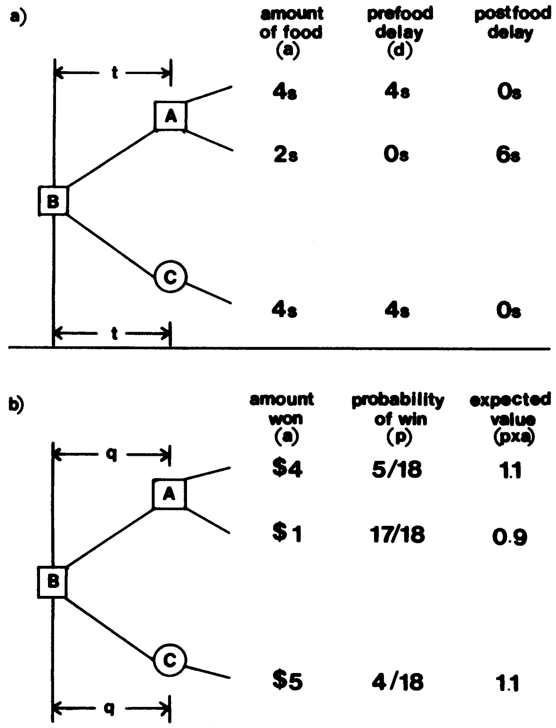
To understand the RCC experiment, it helps to discuss the original experiment of Rachlin and Green [1972]. Figure B:Ia shows the experimental design of Rachlin and Green [1972] whereas Figure B:Ib shows the experimental design of the RCC study. Rachlin and Green [1972] gave pigeons a choice, at point B, between the path leading to point A or the path leading to point C; pigeons pecked at illuminated keys to make this choice. If pigeons chose the path leading to point A, they made a subsequent choice between a smaller, sooner (SS) reward (a = 2s, d = 0s) and a larger, later (LL) reward (a = 4s, d = 4s).[12] Thus, choice of the path leading to point A gave pigeons *flexibility* in their choice in the second stage. If pigeons chose the path leading to point C, there was no subsequent choice and the pigeons automatically received the LL reward (a = 4s, d = 4s). Thus, choice of the path leading to point C *committed* pigeons to the LL reward.

Rachlin and Green [1972] manipulated the delay t to points A and C from a choice at B to determine whether this delay affected choice behaviour at point B. They found that when t was short (t < 4s), pigeons at point B predominantly chose the path leading to point A and then subsequently chose the SS reward (a = 2s, d = 0s) over the LL reward (a = 4s, d = 4s) on almost every trial (i.e., more than 90% of the time). By contrast, when t was relatively long (t > 4s), pigeons at point B predominantly chose the path leading to point C where only the LL reward was available (a = 4s, d = 4s). This is a delay-commitment paradigm because it tests whether commitment (i.e., choice of the path leading to point C over the path leading to point

[12] Note that the reward a is measured in seconds because this is the amount of time that pigeons were given access to food.

A) increases as the delay to all rewards increases. Rachlin and Green [1972] found that longer delays lead to a preference for commitment.

B:I



Source: RCC [Figure 1, p. 348].

As mentioned previously, RCC replaced delays with probabilities, and pigeons with humans. Their experiment played out across two stages. In the first stage, subjects had to choose whether to allocate a blue chip (probability q = 3/18 of advancing to stage 2) or red chip (probability q = 15/18 of advancing to stage 2) to one of two cards: X or Y. Card X corresponds to point A and card Y corresponds to point C in Figure B:Ib. After allocating a red or blue chip to card X or Y, a spinner, programmed with the relevant probability (i.e., q = 3/18 for the blue chip or q = 15/18 for the red chip), was used to determine whether the subject proceeded to stage 2.

If a subject allocated her chip to card X and was successful, she moved on to stage 2 (point A) where she had to choose whether to play a low reward, high probability gamble (i.e., $1 reward with probability 17/18) or a high reward, low probability gamble (i.e., $4 reward with probability 5/18). Thus, if the subject allocated a chip to card X and was successful, she had *flexibility* in her choice at stage 2. If, by contrast, she placed her chip on card Y during stage

1 and was successful, she then played a high reward, low probability gamble (i.e., $5 reward with probability 4/18) in stage 2 (point C). Thus, by allocating a chip to card Y, the subject was *committed* to the high reward, low probability gamble in stage 2, if it was reached. As mentioned previously, the rewards used in this study were real, rather than hypothetical, and subjects were paid their winnings at the end of the experiment.

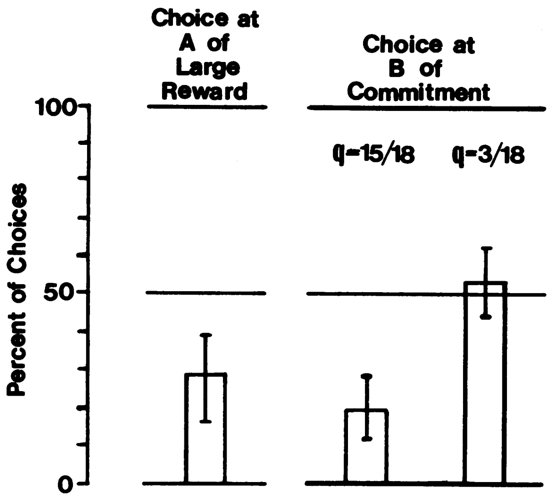
At the start of the experiment each subject was given 10 blue chips and 10 red chips which they could allocate, in any order, to card X or Y across 20 trials. This differs from the design in Rachlin and Green (1972) because pigeons could not choose the delays themselves; the pigeons were exposed to the different delay treatments. RCC provide a strange justification for this difference:

> "This method of having the subjects themselves select trial order was chosen because pilot experiments of ours as well as published accounts of human laboratory analogs to animal experiments … indicate that corresponding results are more likely when people's tasks are made more complicated and varied than corresponding animal tasks. For similar reasons, the two larger rewards ($5 and $4) were not identical (although their expected values were identical)." (see RCC [p. 349]).

RCC argue that their experiment is the probabilistic choice analogue of the experiment in Rachlin and Green [1972]. Rather than manipulate the delay t to points A and C from B, RCC manipulated the probability q of reaching points A and C from B. Choice of the path leading to point A (i.e., allocating a token to card X) gives a subject flexibility in her choice in the second stage of the experiment, *if it is reached*. By contrast, choice of the path leading to point C (i.e., allocating a token to card Y) commits the subject to a high reward, low probability gamble in the second stage of the experiment, *if it is reached*. By varying q, RCC could test whether a low continuation probability, which RCC argue is analogous to a long delay, is associated with more commitment choices than a high continuation probability.

RCC found that at point A (i.e., after allocating a token to card X and successfully proceeding to stage 2), 28% of choices were for the high reward, low probability gamble, as represented by the bar on the far left of Figure B:II. Note that this fraction of choices was significantly less (t = 3.28, df = 10, one-tailed test) than 50%. In other words, if point A was reached, there was a preference for the low reward, high probability gamble. Note that this result does not line up perfectly with Rachlin and Green [1972], who found an almost exclusive preference for the smaller, sooner reward at point A.

Figure B:II



Source: RCC [Figure 3, p. 351].

In stage 1, approximately 18% of the high continuation probability chips (i.e., red chips where q = 15/18) and 53% of the low continuation probability chips (i.e., blue chips where q = 3/18) were placed on card Y, as represented by the bars on the right of Figure B:II.[13] Note that the allocation of red chips to card Y was significantly less than the allocation of blue chips to card Y (t = 5.42, df = 10, two-tailed test). RCC interpret the preceding set of results as evidence that a preference for the low reward, high probability gamble changed to indifference when the continuation probability q fell from 15/18 to 3/18.

The conclusion which RCC reached relies on a strange and dubious comparison: the allocation, approximately 53%, of low continuation probability blue chips to card Y during stage 1, as shown by the bar on the far right of Figure B:II, and the choice of the \$4 low probability gamble during stage 2 of card X (approximately 28%), as shown by the bar on the far left of Figure B:II. In other words, the comparison is between the choice of gambles after the resolution of stage 1 uncertainty, and the initial stage 1 choice between cards prior to the resolution of uncertainty.

[13] RCC do not provide standard deviations for these estimates but these are represented by the whiskers in Figure B:II.

A more appropriate comparison would be with the allocation of red and blue chips to cards X and Y in stage 1. Subjects allocated significantly more red chips than blue chips to card X ($t = 5.42$, df = 10, two-tailed test). This means that subjects had a preference for flexibility in stage 2, rather than commitment in stage 2, when using the high continuation probability red chips.[14] This does not imply the converse though: that subjects had a preference for commitment over flexibility when using low continuation probability blue chips. Subjects were practically and statistically indifferent between flexibility and commitment when allocating blue chips. RCC [p. 350] simply state that the 53% allocation of blue chips to card Y is not significantly different to 50% without providing test statistics, although this can be seen to be true by looking at the whiskers of the box on the far right of Figure B:II.

Thus, RCC replicated the Rachlin and Green [1972] result of a preference for flexibility at short delays (viz., high probabilities), but failed to replicate the result of a *preference* for commitment at long delays (viz., low probabilities). While the fraction of commitment choices was greater with low probabilities than with high probabilities, this fraction was not significantly different to 0.5. Thus, it is not valid for RCC [p. 350] to claim that, "these results parallel those obtained with pigeons choosing among rewards of various amounts and delays."

A major experimental design issue of the RCC study was the sequential allocation of tokens: subjects had to allocate a token, observe the result of the ensuing gamble, and then allocate another token, until all of their red and blue tokens were finished. Consequently, each person's idiosyncratic payoff history may have influenced her subsequent choices. In other words, this design is not immune to order or wealth effects across trials.[15] This is a point which RCC [p. 350] acknowledge, but their approach to the problem is not satisfactory.

RCC focussed on the last four trials of the experiment to see whether choices during these trials were markedly different to the choices made in previous trials. At the level of the sample as a whole, there was an equal number of red and blue tokens left for allocation over

[14] Recall that allocating a chip to card X gives one freedom of choice (viz., flexibility) in stage 2, if it is reached. By contrast, allocating a chip to card Y commits one to playing the high reward, low probability gamble in stage 2, if it is reached.

[15] A cleaner experimental design would be to ask subjects to allocate their red and blue tokens across the cards and across the gambles at stage 2 at the outset of the experiment, as if they were constructing a portfolio of risky assets, and then play out all of the gambles.

the last four trials. Thus, the last four trials, at least in terms of the proportion of red and blue tokens in the sample, were comparable to the trials at the start of the experiment. RCC found that the allocation of tokens to cards X and Y was very similar in the last four trials as in the experiment as a whole.[16] In addition, they found that choices in stage 2 of card X (i.e., between the high reward, low probability gamble and the low reward, high probability gamble) were very similar in the last four trials as in the experiment as a whole.

While these results suggest that order and wealth effects were unlikely to be driving RCC's findings, one should heed the warning of Harrison [2007] that appropriate statistical techniques need to be used to draw inferences from experimental data when there is the potential for correlation of responses at the level of the individual and over time. RCC ignore these possibilities by treating the *K* choices of each subject as independent and by not taking into account the time path of choices in the experiment.

In addition to the issues outlined above, only 11 subjects took part in the experiment so there was minimal power for the statistical tests that were conducted. Furthermore, RCC specifically asked their subjects not to make any mathematical calculations even though they were presented with options that had real financial consequences. Thus, the experiment of RCC does little to support their contention that probability is best interpreted as delay.

ADDITIONAL REFERENCES

HARRISON, G. W. (2007): "House Money Effects in Public Good Experiments: Comment," *Experimental Economics*, 10, 429-437.

[16] RCC compared the last four trials to the full twenty trials. Ideally they should have compared the last four trials to the first sixteen trials of the experiment.

APPENDIX C
[ONLINE WORKING PAPER]

As mentioned in the main text, it is straightforward to make the parameter(s) of interest in our models a linear function of observable characteristics and thereby admit heterogeneity in the PWF estimates. RZMW collected each participant's gender and Appendix C (see RZMW [p. 140]) groups the elicited certainty equivalents by experimental session and divides participants by whether they received a low or high dose of ethanol. Table C:I presents estimates of the three models that incorporate these variables.

TABLE C:I: PROBABILITY WEIGHTING FUNCTION ML ESTIMATES
HETEROGENOUS PREFERENCES

	Model 1	**Model 2**	**Model 3**
	PD	TK	Prelec
PWF parameter (γ)			
Male	0.035	-0.011	0.007
	(0.341)	(0.096)	(0.130)
Ethanol - high dose	0.45	-0.08	0.023
	(0.378)	(0.107)	(0.130)
Post-placebo session	-0.006	-0.002	-0.002
	(0.048)	(0.018)	(0.027)
Pre-ethanol session	0.223**	-0.045	0.005
	(0.108)	(0.031)	(0.045)
Post-ethanol session	0.062	0.005	0.044
	(0.099)	(0.036)	(0.042)
Constant	1.003***	0.811***	0.715***
	(0.184)	(0.073)	(0.118)
PWF parameter (η)			
Male			0.027
			(0.172)
Ethanol - high dose			0.215
			(0.172)
Post-placebo session			0.001
			(0.021)
Pre-ethanol session			0.111**
			(0.045)
Post-ethanol session			0.056
			(0.052)
Constant			0.880***
			(0.116)
Sigma (σ)			
Constant	0.138***	0.135***	0.132***
	(0.016)	(0.014)	(0.014)
N	480	480	480
log-likelihood	269.434	278.331	289.727

Results account for clustering at the individual level
Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Gender and ethanol dose are not statistically significant in any of the models and none of the experimental treatment variables are statistically significant in the TK model. However, the estimate of γ in the PD model is significantly higher in the pre-ethanol session than in the pre-placebo session (the omitted base category). Wald tests show that the estimate in the pre-ethanol session is also significantly greater than the estimates in the post-placebo ($p = 0.038$) and the post-ethanol ($p = 0.034$) sessions. This result is contrary to the hypothesis that ethanol increases probabilistic discounting, which the researchers set out to test, and differs to RZMW who found no statistically significant difference between the estimates of γ in the pre-ethanol and post-ethanol sessions.[17] Note that RZMW used *estimates* of γ as *data* to conduct *t*-tests of potential differences across the pre- and post-ethanol sessions. The valid approach to analysis that I have adopted uses all of the information that a dataset imparts to estimate the parameters of a model and conduct hypothesis tests on these estimates. These differences in analysis likely explain the contradictory findings.

The estimates of γ for the Prelec PWF do not differ significantly according to observable characteristics and task parameters but the estimate of η in the pre-ethanol session is significantly higher than in the pre-placebo session (the omitted base category). In addition, Wald tests show that the estimate of η in the pre-ethanol session is significantly higher than estimates in the post-placebo ($p = 0.017$) and post-ethanol ($p < 0.038$) sessions. These results mirror those for γ in the PD model.

[17] RZMW did not compare the estimate of γ in the pre-ethanol session to the estimates of γ in the pre-placebo and post-placebo sessions.

APPENDIX D
[ONLINE WORKING PAPER]

As discussed in the main text, the log-likelihoods for the TK and Prelec (1998) functions exceed the log-likelihood for the PD function, suggesting that the TK and Prelec (1998) functions better characterise the data. This hypothesis can be tested formally using Vuong [1989] and Clarke [2007] non-nested model selection tests. The Clarke test is asymptotically more efficient and has greater power in discriminating between models than the Vuong test when the distribution of the models' individual log-ratios is highly peaked. Thus, when the distribution of these log-ratios is leptokurtic[18], the Clarke test is superior, from both statistical efficiency and power perspectives, to the Vuong test.

Figure D:I: Distribution of Individual Log-ratios

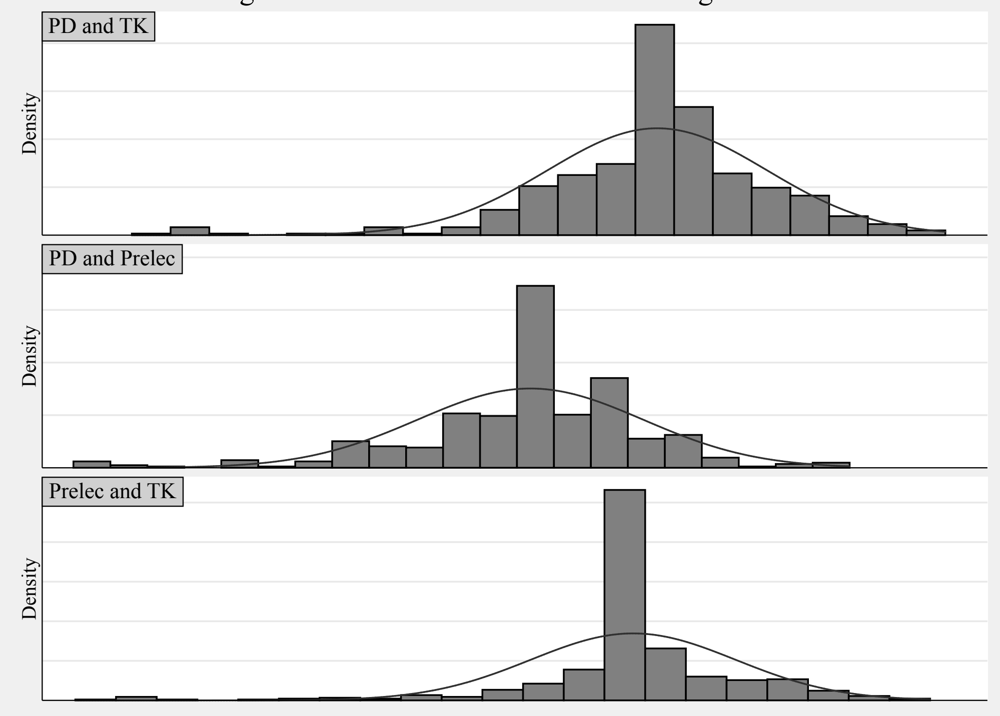


Figure D:I plots the distribution of the individual log-ratios, with a normal density overlay, for the three PWF comparisons. The distribution of these log-ratios is leptokurtic (i.e., highly peaked) which suggests that the Clarke test is more appropriate for these data. The Clarke test yields a test statistic based on the binomial distribution which must be compared to a critical value to determine which model, in a pairwise comparison, receives the most support in the

[18] The normal distribution is the quintessential mesokurtic distribution. A distribution which has positive excess kurtosis (i.e., a highly peaked distribution) is leptokurtic.

data. A Clarke test comparing the PD and TK functions yields a test statistic of 197, which is below the critical value of 240, implying that the TK function better characterises the data ($p < 0.001$).

A Clarke test comparing the PD PWF and the Prelec (1998) PWF finds in favour of the Prelec (1998) function ($p < 0.001$). Finally, a Clarke test of the TK and Prelec (1998) PWFs suggests that the Prelec (1998) function finds more support in the data ($p < 0.001$). Based on these tests the following transitive ranking of PWFs emerges: the Prelec (1998) function finds the most support in the data, followed by the TK function, and then the PD function.