Flexible estimation of parametric prospect models using hierarchical Bayesian methods.

Experimental Economics (2025), 1–25
doi:10.1017/eec.2025.10012

ORIGINAL PAPER

Flexible estimation of parametric prospect models using hierarchical bayesian methods

Kelvin Balcombe[1] and Iain Fraser[2]

[1]Department of Agri-Food Economics and Marketing, School of Agriculture, Policy and Development, University of Reading, Earley, Reading, Berkshire, UK
[2]School of Economics, Politics and International Relations and Durrell Institute of Conservation and Ecology (DICE), University of Kent, Canterbury, Kent, CT, United Kingdom
Corresponding author: Iain Fraser; Email: i.m.fraser@kent.ac.uk

(Received 8 January 2024; revised 28 April 2025; accepted 2 May 2025)

Abstract:
In this paper, we present a flexible approach to estimating parametric cumulative Prospect Theory using Hierarchical Bayesian methods. Bayesian methods allow us to include prior knowledge in estimation and heterogeneity in individual responses. The model employs a generalised parametric specification of the value function allowing each individual to be risk-seeking in low-stakes mixed prospects. In addition, it includes parameters accounting for varying levels of model noise across domains (gain, loss, and mixed) and several aspects of lottery design that can influence respondent behaviour. Our results indicate that enhancing value function flexibility leads to improved model performance. Our analysis reveals that choices within the gain domain tend to be more predictable. This implies that respondents find tasks in the gain domain cognitively less challenging in comparison to making choices within the loss and mixed domains.

Keywords: Cumulative prospect theory; complexity; Hierarchical Bayesian methods; generalised parametric specification

JEL Codes: C11; C52; D81

1. Introduction

In this paper, we employ Hierarchical Bayesian methods (HBMs) to estimate model preference parameters for a generalized form of cumulative Prospect Theory (PT) (Tversky & Kahneman, 1992). In estimating PT models, researchers face significant challenges. One key challenge is to strike a balance between building models that are sufficiently general to capture behaviour patterns while avoiding excessive generality that leads to overfitting in-sample data and poor out-of-sample predictions. Additionally, researchers must build upon existing knowledge and theories, aiming to integrate new insights without merely confirming prior findings. Bayesian approaches such as the HBM offer a transparent approach for researchers to navigate these conflicting demands effectively.

These aforementioned challenges are interconnected. The Hierarchical Bayesian approach serves as a framework for reconciling generality and parsimony on two fronts. Firstly, it permits heterogeneity in preferences while imposing restrictions on the extent of this heterogeneity. Secondly, because parsimony relates not only to the number of parameters in a model but to the freedom they are allowed, HBMs enable the integration of conditions that not only align with the theoretical foundations of the models but also draw from prior empirical investigations involving these models, but in a way that will still allow new data to modify or contradict previous work.

At first sight PT models may not seem particularly complex or daunting to estimate. However, given two prospects, a specific parameterizations of a PT model may have a wide range of parameter values that would assign a reasonably high probability to choosing either, and where people have heterogeneous noisy preferences, the identification of parameters becomes an increasingly difficult task. As one considers increasingly flexible aspects either in terms of value functions or probability warping, any model can rapidly become weakly or even non-identified in the sense of having a singular set of parameters that will maximize the probability of the choices we are able to see. Empirical implementations of PT commonly employ restrictions that are made either without strong justification, or in many cases, no justification at all, probably in order to make the models empirically tractable. Yet, at the very same time, one can argue that even the most general parametric structures that are commonly used are highly restrictive.

The advantages of employing HBMs for estimating PT models have been previously highlighted (e.g., Nilsson et al., 2011, Balcombe & Fraser, 2015, Balcombe et al., 2019, Alam et al., 2022, Gao et al., 2023). As emphasized by Gao et al. (2023), it is important to acknowledge that Bayesian methods are not the sole approach to achieving what some refer to as "shrinkage" (Murphy & ten Brincke, 2018). Moreover, the benefits of Bayesian estimation go beyond the mere utilization of priors. Equally, it is essential to consider the extent that prior knowledge, whether theoretical or empirical, can aid in identifying parameters in more general PT models, irrespective of whether this is done within an explicitly Bayesian framework or not. It is crucial not to limit the use of more flexible PT models due to a methodological stance that parsimony is permissible only if it comes in the form of "tight restrictions".

Our specific use of HBMs enables us to estimate a generalized value function that provides greater flexibility around small payoffs in the set of prospects, whereby we allow each individual to be risk seeking in low-stakes mixed prospects. The need for this extension has previously been noted in the literature. For example, Neilson and Stowe (2002) identified limitations with employing the standard power function (the constant relative risk aversion (CRRA)) specification and noted the need for parameter estimation for lotteries over likely gains and unlikely gains. In addition, our model also includes parameters allowing model noise to vary across the gain, loss and mixed domains. The reason for introducing this extension into the model specification is that there is no a priori reason to assume that the noise we observe in the different domains will be the same. Recent evidence on the need to incorporate this type of flexibility is presented by (Chapman et al., 2024). Finally, we have also taken account of several aspects of lottery design in our model specification that can influence individual respondent behaviour.

HBMs can accommodate a variety of individual preference distributions where the "posterior distributions" for each individual can be estimated. This means that our HBMs allow us to capture heterogeneity in individual responses effectively bridging the gap between "representative agent" methods and models that allow for individual effects. This is achieved in HBMs by assuming that different individuals are drawn from a common distribution. The assumption that there is a common distribution allows information about parameters to be "shared" or "borrowed" across individuals (Gao et al., 2023). Another benefit from employing HBMs is that we can use priors in model estimation. We recognize that researchers should avoid constraining models so rigidly that estimates merely mirror their priors. But, they should feel confident in employing clear and transparent parametric priors that align with consensus regions and theoretical boundaries. Embracing an "Informative Bayesian" approach provides a transparent method for incorporating these considerations.

The specific benefits of employing HBMs, for estimating individual respondent risk preferences, has previously been considered by Alam et al. (2022), Balcombe and Fraser (2015), Balcombe et al. (2019), Nilsson et al. (2011) and Gao et al. (2023), drawing on a well-established set of HBMs (e.g.,Rossi et al., 2005, Allenby & Rossi, 2006). Within the risk literature, Balcombe et al. (2019) examined issues regarding loss aversion when model specific parametric assumptions are relaxed

https://doi.org/10.1017/eec.2025.10012 Published online by Cambridge University Press

for individual level choices. More recently, Gao et al. (2023) used HBMs to estimate individual risk preferences and insurance purchase decisions.

In this paper, we add to the existing HBMs applications in the risk literature by introducing greater flexibility into the estimation of the value function. When it comes to the choice of value function to populate a PT model, there is arguably no clearly superior choice but the impact on the estimates can be substantial (Stott, 2006). The range of value functionals that have been employed within the literature is extensive (see Kobberling & Wakker, 2005, Fehr-Duda et al., 2010, Bouchouicha & Vieider, 2017) and there continues to be ongoing debate regarding what sort of value functional to employ. In an extensive examination of functional forms employed within the PT literature, Balcombe and Fraser 2015 examined six value functions frequently encountered in the literature (i.e., power function, quadratic function, and one and two parameter exponential and logarithm functionals) in addition to several probability weighting functions including linear and Prelec I and II. In keeping with Stott (2006) they found that the power utility specification was the preferred value function. Balcombe et al. (2019) subsequently examined CRRA and constant absolute risk aversion (CARA) utility specifications at the individual level finding in favour of CRRA. In related research presenting a meta-analysis of loss aversion, Balcombe and Fraser (2024) provide an excellent illustration of the scope of approaches to the estimation of PT models. In Table 3 of Brown et al. (2024) the most common value function employed in the literature is the piecewise power utility specification that yields CRRA. The popularity of this value function is probably in no small part due to the seminal paper by Tversky and Kahneman (1992), with almost 60% of the 522 papers examined producing estimates of loss aversion employing a CRRA functional form. Papers were also differentiated in that they used different levels of aggregation (e.g., representative agent vs individual estimates) and different reference points. Finally, Zrill (2024) examines the predictive ability of simple functional forms (e.g., CRRA) in risk settings and reports that they are useful modelling choices despite being restrictive.

The use of HBMs enables researchers to balance the competing demands of model complexity and parsimony in a way that contributes and extends the long-standing tradition of estimating preference parameters and examining specific facets of PT using experimental data (e.g., Buschena & Atwood, 2011, Rieger et al., 2017, Murphy & ten Brincke, 2018, l'Haridon & Vieider, 2019). We also extend parametric estimation of PT models at the individual level taking explicit account of frequently encountered aspects of lottery design. Although many specific features of lottery design and implementation have been extensively researched (Holzmeister & Stefan, 2021) there is little research examining these influences on parameter estimation.

The set of lottery tasks used in this paper are described in detail by (Balcombe & Fraser, 2024). The experimental design employed PT model parameter priors that reflect the high-density regions of the consensus distributions in the PT literature (Tversky & Kahneman, 1992). The approach to experimental design aimed to minimize the impact of prior information in the experimental design on the resulting model parameter estimates. In addition, the set of tasks always includes prospects where there is a sure-thing option, that is an outcome of a lottery that if selected guarantees payment of the amount stated ex ante, can effect model performance. Many researchers reduce task complexity by employing a sure-thing option (e.g., Bruhin et al., 2010, Falk et al., 2018, l'Haridon & Vieider, 2019) but this can induce certainty effect bias, where respondents chose options that have payoffs that are certain but cannot be explained by expected utility theory (EUT). Kahneman and Tversky (1986) view the certainty effect as a framing bias, but it can potentially be understood as a phenomenon that arises because respondents choose options that they can more easily understand (albeit because of framing). Interest in the certainty effect is ongoing, see Zilker and Pachur (2022) and Frydman and Jin (2022).

The set of lottery tasks also include prospects with either two or three payoffs. This design feature is justified because it provides greater insights into the nature of probability weightings. This, in turn, enhances model estimation and identification. We, explicitly take account of this feature in our experimental design.

https://doi.org/10.1017/eec.2025.10012 Published online by Cambridge University Press

Both of the lottery design features considered can in principle change the degree of task complexity. Task complexity, that is the cognitive difficulty associated with undertaking a task, has been examined in various research areas, including stated preference discrete choice experiments (e.g., Pfeiffer et al., 2014, Johnson et al., 2017, Regier et al., 2014) and the risk literature in terms of similarity of prospect pairs (e.g. Buschena & Atwood, 2011, Buschena & Zilberman, 1999) and with regard to preference reversals (Loomes & Pogrebna, 2017). Amador-Hidalgo et al. (2021) and Andersson et al. (2016) report that as task complexity increases, that the number of inconsistent choices increases. Inconsistent choices also increase as the realization of the probabilities of the payoffs get closer because the computation required by a respondent to identify the preferred option becomes harder. The impact of task complexity has also been discussed in relation to the legitimacy of rank dependency (see Bernheim & Sprenger (2020) and Bernheim et al. (2022)).

The importance of task complexity more generally in decision making in relation to lotteries has also been examined by Enke and Shubatt (2023). This is part of a wider research agenda (Oprea, 2022, Oprea, 2024) that is assesses how complexity can in fact generate empirical results that are associated with PT that is, loss aversion and probability weighting, even when there is no explicit risk involved in the decision. These findings suggest that designing risk tasks that are cognitively less complex is required if risk preferences are to be identified. This would also indicate that taking account of task complexity in an experimental design is also important if key features of PT are to be identified.

As we will show, the least flexible model specification generated model parameter estimates for curvature in the loss domain that are above unity. This is broadly inconsistent with the proposed view within PT that people are generally "risk-seeking" in the loss domain (Tversky & Kahneman, 1992, Wakker, 2010). However, we find that by allowing for a more generalized value function that also takes account of uncertainty effects, which is enabled by our use of HBMs, improves statistical model performance and yields estimates that are consistent with being below or at one. Our findings remain stable across alternative specifications of the value function. Importantly, we find evidence that both probability warping and certainty effects seem more prominent in the loss domain, and in a sense less consistent with what we would loosely characterise as consensus values. Probability weightings in the gain domain are directionally the same as in Tversky and Kahneman (1992) but to a lesser degree. However, for the probability weightings, we find in the loss domain that they are strongly opposite to that proposed in Tversky and Kahneman (1992).

The structure of the paper is as follows. We begin in Section 2 by outlining our generalized PT model. Section 3 presents our econometric specification which extends the standard PT model and in Section 4 we briefly describe the experimental data employed in our analysis. In Section 5, we report our results and in Section 6 we conclude.

2. Econometric specification

2.1. The value function

Define an ordered prospect $\mathcal{Z} = ((x, ..., x_{K,}), (p_1, ..., p_K))$ where p_k is the probability of the k^{th} payoff x_k where $x_1 \leq x_2 \leq ... \leq x_K$. A prospect is said to be a gain (G) (loss, (L)) prospect if the payoffs (with positive probability) are non-negative (negative). A prospect is said to be mixed (M) if it contains both positive and negative payoffs. Let $j = 1, ..., J$ denote the j^{th} individual who has to complete $t = 1, ..., T$ tasks which involves choosing between two prospects. The prospects are not necessarily ordered from smallest to largest payoff, but it is assumed that the respondent is able to order them. In outlining the functions below, we initially suppress the j, and t subscripts to simplify notation.

There is an array of potential value functions that can be employed. Six forms over the positive domain are given in Stott (2006), though a number of these nest or approximate others as special cases. Some value functions are only defined over the positive range and thus need to be generalized to be used in the context of PT which requires assigning utility to both negative as well as positive payoffs. However, these utility functions can be employed in a "piecewise" fashion. That is, two

https://doi.org/10.1017/eec.2025.10012 Published online by Cambridge University Press

distinct utility/value functions $u^{+}(x), u^{-}(x)$ are defined where $u^{+}(x)$ is the utility for a positive payoff ($x > 0$) and $u^{-}(x)$ is the utility for a negative payoff ($x < 0$). These two are "joined" at zero and they employ an additional parameter λ such that piecewise utility becomes

$$u(x) = I_{(x \geq 0)} u^{+}(x) + \lambda I_{(x<0)} u^{-}(x) \tag{1}$$

This transformation can be used to give additional flexibility or to define utility over negative as well as positive regions.

Bouchouicha and Vieider (2017) provide a good discussion of four piecewise functions. These include the two most popular parametric functions employed in the context of risk, the CARA and the CRRA, but also a restricted version of the EXPO (Saha, 1993) of the CARA function along with the normalized logarithmic (NLL) utility function.

Since it was used by Tversky and Kahneman (1992) the CRRA form has been the predominant one implemented within PT, with the linear version second and CARA third (see Brown et al., 2024, Table 3). There has been limited work comparing alternative models, but Stott (2006) recommends the power CRRA form with models estimated at the individual level in the gain domain with Balcombe and Fraser (2015) obtaining a similar result on the same data. By contrast Bouchouicha and Vieider (2017) find that the both the CARA and NLL version to be superior using a model estimated at the representative agent level (and no mixed domains). Alternatively, Balcombe et al. (2019) find the CRRA to outperform the CARA employing a HBM framework most similar to the one employed here.

We believe that there should be ongoing research examining the performance functional forms in full PT models. However, this task is complicated by the interactive nature of the different functionals, representation of noise and treatment of exogeneity. Here, we follow the most trodden path and employ the CRRA in a "power parameterisation"

$$v(x, \theta) = I_{(x \geq 0)} x^{\alpha} - I_{(x<0)} \lambda |x|^{\beta} \tag{2}$$

where $I_{(.)}$ denotes an indicator function which is equal to one where the condition in the brackets is satisfied. We define the set of model parameters as $\theta = (\alpha, \lambda, \beta)$ where $\alpha > 0, \lambda > 0$, and $\beta > 0$. However, we generalize this to allow for a "wobble" in small payoff regions as explained below.

2.1.1. Flexibility of the value function near zero

A possible feature of risky behaviours is that they may differ for smaller payoffs relative to larger ones in a way that requires the curvature of a value function that is not reflected by CRRA or CARA. With this in mind we introduce an alteration of the value function that creates a "wobble" in the value function close to zero. We present the general case below, but in order to motivate the "wobble", let us first examine the simplified piecewise power-utility function below:

$$v(x, \alpha) = I_{(x \geq 0)} x^{\alpha} - I_{(x<0)} |x|^{\alpha} \tag{3}$$

The function [3] has a constant relative risk aversion coefficient (CRRA) of $\pm(1 - \alpha)$ that is positive in the gain domain and negative in the loss domain if $\alpha < 1$. Thus, any risky version prospect with the same expected value as a safer one will have positive utility difference in favour of the safer option in the gain domain for any value for $0 < \alpha < 1$, but a negative utility difference for the safer option in the loss domain.

Under [3], if we take two prospects that have the same expected value but simply scale them (by multiplying all the payoffs by some number), then the scaled up version will have a larger utility difference (in absolute terms) relative to the smaller option, but the sign of that utility difference would always be the same. What we wish to consider, however, is that people might not be consistent with this behaviour in the sense that a scaled version of two prospects might not have the same signed utility difference in the upper ranges of α. For example, given an even chance of zero and x or a sure

https://doi.org/10.1017/eec.2025.10012 Published online by Cambridge University Press

$\frac{x}{2}$, one might expect that the sure-thing will be chosen for some value of x, but when scaled down such that the choice is between zero and sx and $\frac{sx}{2}$ for some values of $s < 1$ and $\alpha < 1$. That is, they are risk averse for the larger prospect but risk seeking for the smaller.

In order to reflect this possible behaviour, we modify the value function so as to allow flexibility around small payoffs. Specifically, we consider a logistic component (the "wobble") with the parameters ω_1 and ω_2 that will determine the curvature of the value function:

$$\varphi(x, \omega_1, \omega_2) = \frac{1}{1 + \omega_1 e^{-\omega_2 x}} \quad \text{where } \omega_1, \omega_2 > 0 \tag{4}$$

In order to allow a difference between the loss (L) and gain (G) domains, we multiply the power components by this logistic component that is potentially different across the domains:

$$v(x, \theta, \omega) = I_{(x \geq 0)}\, x^{\alpha} \varphi(x, \omega_G) - I_{(x<0)} \lambda |x|^{\beta} \varphi(x, \omega_L) \quad \text{where } \omega = (\omega_{G,1}, \omega_{G,2}, \omega_{L,1}, \omega_{L,2}) \tag{5}$$

It should be evident that for very large payoffs the utility will be approximately equal to the standard power utility as the payoff x becomes large. Equally, it should be evident that as ω_1 becomes small then utility will be arbitrarily close to power utility. However, the behaviour of the utility for prospects involving the smaller amount may be quite different.

This parameterization allows for the modification of risk seeking and risk aversion behaviour depending on the size of the gambles. Importantly, however, in our empirical application we do not assume that adjustments to utility are the same in both domains and we allow the parameters ω_1, ω_2 to be domain specific.

The behaviour above essentially extends into comparing choices for small fair mixed prospects with large prospects from any domain. The utility differences are approximately invariant to the transformed utility providing any one of the payoffs is large (positive or negative). Small mixed gambles will, however, have the potential for reversals.

2.1.2. A note on the "loss aversion" parameter λ

Within the existing literature $\alpha = \beta$ is commonly imposed for model estimation. In this case the parameter λ has commonly been interpreted as a measure of loss aversion providing $\lambda > 1$. However, unless $\alpha = \beta$, the value of λ is dependent on the denomination (e.g., pounds, pence, dollars or cents) of the payoffs such that alternative values for λ can be obtained with identical data simply by redenominating the payoffs.

Brown et al. (2024) observe that 221 out of 302 studies employing the CRRA formulation impose $\alpha = \beta$ but do not comment on whether this restriction was imposed without pre-testing. Tversky and Kahneman (1992) themselves shied away from promoting the CRRA form used in their paper, characterizing it only as a way to parsimoniously describe their data. Their imposition of $\alpha = \beta$ is made without discussion, perhaps because without this restriction the role and meaning of λ becomes muddied (see Balcombe et al., 2019). However, given the array of possible functional form specifications, it is unsurprising that many researchers have followed the lead of Tversky and Kahneman (1992). Our perspective is that there is no strong theoretical nor empirical reason to impose $\alpha = \beta$ other than as one which clarifies the interpretation of λ.

2.2. The probability weighting function

In PT, respondents are assumed to modify the decumulative and cumulative distributions of the products in the G and L domains respectively to obtain "decision weights". Given this, we present the probability weighting function in a form that is consistent with how PT requires respondents to

https://doi.org/10.1017/eec.2025.10012 Published online by Cambridge University Press

determine their preferred choice. Formally, denote $D(p)$ and $C(p)$ as the decumulative and cumulative functions over the probability simplex p, (where p has been ordered from smallest to largest x) and they are defined as

$$C(p) = \left(p_1, \sum_{i=1}^{2} p_i, \sum_{i=1}^{3} p_i, ..., 1 \right) \tag{6}$$

$$D(p) = \left(1, 1 - p_1, 1 - \sum_{i=1}^{2} p_i, ..., p_k \right) \tag{7}$$

Also, denote the inverse operators such that $C^{-1}(C(p)) = p$ and $D^{-1}(D(p))$ for any simplex p . In common with much of the literature (Wakker, 2010, Gao et al., 2023), we employ the Prelec II probability weighting function, whereby for some vector $\gamma = (\gamma_1, \gamma_2) > 0$

$$q(p, \gamma) = exp(-\gamma_2(-\ln(p))^{\gamma_1})) \tag{8}$$

The domain specific decision weights (w) are

$$\begin{aligned} w^+(p, \gamma) &= D^{-1}(q(D(p), \gamma)) \\ w^-(p, \delta) &= C^{-1}(q(C(p), \delta)) \end{aligned} \tag{9}$$

where $\delta = (\delta_1, \delta_2) > 0$, and across domains can be combined to give

$$w(x, \gamma, \delta) = I_{(x<0)} w^-(p, \delta) + I_{(x \geq 0)} w^+(p, \gamma) \tag{10}$$

The Prelec II function is linear at $\gamma_1 = \gamma = 1$ and $\delta_1 = \delta = 1$ and can generate a wide range of transformations, including the "inverse S" shapes favoured by PT (Wakker, 2010). While the domain specific decision weights sum to unity, mixed domain weights do not in general have this property.

2.3. Systematic utility

PT is not a stochastic theory in the sense that it attributes a probability to the decisions that people make. For this reason, we define the utility given by PT as systematic utility which must then be embedded in a probabilistic model of choice. Systematic utility of a prospect is, under PT (using the definitions [3] and [10]):

$$\begin{aligned} V(\mathcal{Z}, \Omega_0) &= \sum_k w(x_k, \gamma, \delta)\, v(x_k, \theta, \omega) \\ \text{where } \Omega_0 &= (\theta, \omega, \gamma, \delta) \end{aligned} \tag{11}$$

2.3.1. Allowing for complexity bias

In our model, we generalize this relationship by allowing the respondents to prefer (i.e., be biased towards) the sure-thing option that is used in the experimental design in an effort to reduce task complexity. Take a prospect pair $\mathcal{P} = (\mathcal{Z}_a, \mathcal{Z}_b)$ where $\mathcal{Z}_b = (x_b, p_b)$. In our lottery data $\mathcal{Z}_b$ will always be the sure-thing where any zero probability payoffs are zero, whereas $\mathcal{Z}_a$ can involve either one or two non-zero payoffs (with non-zero probability). Denote the number of non-zero payoffs in $\mathcal{Z}_a$ as n_0. It then follows that respondents are assumed to have the respective utilities:

$$\begin{aligned} V_a(\mathcal{Z}_a, \Omega_0, \eta) &= (1 + \eta) I_{(n_0=1)} V(\mathcal{Z}_a, \Omega_0) + I_{(n_0=2)} V(\mathcal{Z}_a, \Omega_0) \\ V_b(\mathcal{Z}_b, \Omega_0, \pi) &= \pi_L I_{(x_b<0)} V(\mathcal{Z}_b, \Omega_0) + \pi_G I_{(x_b \geq 0)} V(\mathcal{Z}_b, \Omega_0) \end{aligned} \tag{12}$$

$$\text{where } \pi = (\pi_G, \pi_L) \tag{13}$$

The parameter η will deviate from zero if the simpler prospect is more or less preferred than would otherwise be suggested by the systematic utility in [11]. In this case, we assume that the simpler

https://doi.org/10.1017/eec.2025.10012 Published online by Cambridge University Press

prospect requires $\mathcal{Z}_a$ to have only one non-zero payoff (with non-zero probability). This means that our estimate of η assesses the preferences of respondents in regard to this aspect of choice complexity.

Turning to π, these are assumed to be constant parameters, where π_L will deviate from one if there is bias for or against the sure loss, and π_G will deviate from one if there is bias for or against a sure gain.

2.4. Stochastic linkage

We assume that respondents make decisions based ostensibly on utility differences. However, it is possible that not only is the systematic utility of a domain-specific, but also the "noise" components are as well. Thus, in our flexible model specification, we allow for the possibility that given choices in different domains may not be as predictable even when the utility differences are the same.

To do this, the stochastic choices are modeled by allowing for different levels of noise across the three domains of choice:

$$\begin{aligned} \rho\left(\mathcal{Z}_a, \mathcal{Z}_b, \rho\right) &= \rho_G I\left(\mathcal{P} \in \mathcal{G}\right) + \rho_L I\left(\mathcal{P} \in \mathcal{L}\right) + \rho_M I\left(\mathcal{P} \in \mathcal{M}\right) \\ &\text{where } \rho = \left(\rho_G, \rho_L, \rho_M\right) \end{aligned} \tag{14}$$

where $\mathcal{G}$ denotes the set of all gain domain pairs, $\mathcal{L}$ denotes the set of all loss domain pairs and $\mathcal{M}$ denotes the set of all mixed domain pairs.

2.4.1. Model noise specification

The model we present employs what has become the conventional approach to incorporating noise into a PT model. We view using a PT model within a HBM Logit as an approximation of a Random Utility Model (RUM). While the PT is deterministic, its use within the HBM Logit, assumes that people use PT stochastically. Thus, we would see the model as one where people are able to formulate a random utility about a prospect, but where the deterministic assumptions about preferences are replaced with stochastic versions. That is, respondents are allowed to make what would be considered preference reversals within a deterministic framework and are stochastically transitive rather than transitive in a deterministic sense. That said, the claim that one is implementing an approximation to a RUM, is not the same as saying that one has given an explanation as to how such stochastic behaviour comes about.

Importantly, the warping of probabilities within PT was largely introduced as way of rationalizing empirical behaviours such as the "Allais paradox". While replacing the independence axiom with comonotonic independence is an elegant generalisation, the PT literature is arguably vague on why comonotonic preferences might occur. Additionally, complexity of prospects have no explicit role within PT, and while PT is an explicitly domain sensitive theory, it does not suggest whether levels of noise in one domain should be the same or different.

In the light of this, it is worth considering how and why complexity effects and warping might occur. Vieider (2024) introduces the Bayesian Inference Model (BIM) where warping, inter alia, may occur because people add "cognitive noise" to signals. The BIM also treats the payoffs as subject to "cognitive noise" through the act of "mental decoding". The BIM model is Bayesian in the sense that once it is assumed that there is noise in the probabilities and payoffs, there is Bayesian updating of the quantities using a prior, and the parameters that determine the extent of the updating determine the probability of choice. A survey respondent would use priors to adapt such perceived noisy quantities when making decisions seems inherently plausible. However, the BIM would also work if people believed that there was noise in the quantities they are given (i.e., not due to cognitive processes). Vieider (2024) demonstrates that the parameters of the noise and prior distributions combine to give a rationalization as to why noisy choice would occur. He also shows that the form of this equation would be similar to probability warping of the Goldstein & Einhorn (1987) type. If one adopts this type of

https://doi.org/10.1017/eec.2025.10012 Published online by Cambridge University Press

approach, then either differential noise or domain-specific priors would suggest different behaviours across domains.

In Vieider (2024) noise is ascribed to logged ratios of the prospect attributes, where one of the prospects is a sure thing, and the other is a two-payoff prospect, but states that the approach could be extended to multi attribute "wagers". The BIM approach, as implemented in Vieider (2024), requires comparison across states in the construction of log-ratios and the attachment of priors to these ratios. As such, it is not clearly amenable to a RUM framework that requires a random utility to be assigned to each prospect. However, the essential idea that the outcomes and probabilities are received or used by respondents as being noisy and are updated in a Bayesian way opens the door to a multitude of parameterizations to which randomness is assigned, some of which might be compatible with a RUM. However, we also believe that the role of complexity within these approaches requires further thought and investigation, as we would contend that much noise will arise from the calculation of parameters rather than the underlying quantities (probabilities and outcomes) themselves, with greater complexity leading to potentially greater noise.

Importantly, we think there is a distinction between the existence of uncertainty and its explanation. We view that estimating risk models assuming a HBM Logit is consistent with the approximation of a RUM. Random utility does not strictly assume deterministic preferences although there is a non-stochastic component which happens to be non-linear in the case of PT or EUT models.

2.5. Model summary

We now have the full set of model parameters

$$\Omega = (\alpha, \lambda, \beta, \omega, \gamma, \delta, \rho, \pi, \eta)$$

It then follows that the stochastic utilities are

$$\begin{aligned} U(\mathcal{Z}_a, \Omega) &= \rho(\mathcal{Z}_a, \mathcal{Z}_b, \rho)\, V_a(\mathcal{Z}_a, \Omega_0, \eta) + e_a \\ U(\mathcal{Z}_b, \Omega) &= \rho(\mathcal{Z}_a, \mathcal{Z}_b, \rho)\, V_b(\mathcal{Z}_b, \Omega_0, \pi) + e_b \end{aligned} \tag{15}$$

where e_a and e_b are Gumbel distributed. With this specification, it follows that the probability of $\mathcal{Z}_a$ being preferred is

$$\Pr ob\left(U(\mathcal{Z}_a, \Omega) > U(\mathcal{Z}_b, \Omega)\right) = \frac{\exp\left(U(\mathcal{Z}_a, \Omega)\right)}{\exp\left(U(\mathcal{Z}_a, \Omega)\right) + \exp\left(U(\mathcal{Z}_b, \Omega)\right)} \tag{16}$$

With this model specification, the elements of Ω are:

- α, λ, β are the parameters that determine the conventional power value functions;
- ω are the parameters that allow additional flexibility for the power value functions close to zero;
- γ, δ are the parameters that capture the probability warping of the probability weighting function;
- ρ are the parameters that determine the noise (across all the domains); and
- π, η are the parameters that capture the certainty effect and task complexity.

All parameters with the exception of ω will have "representative agent" versions but are also estimated at the individual level.

3. Model estimation

As indicated in the Introduction, we employ a HBM to estimate a logit model specification, using Hamiltonian Monte Carlo Markov Chain (HMCMC) simulation using the software PyStan[1]. This

[1]For documentation about Pystand visit https://pystan.readthedocs.io/en/latest/. All Pystan code used in estimation is available on GitHub along with the data herein and copies of survey instruments.

https://doi.org/10.1017/eec.2025.10012 Published online by Cambridge University Press

form of estimation requires the likelihood function to be specified along with each of the priors and we outline them below. PyStan (or more generally Stan) uses a version of HMCMC that is referred to as the No-U-Turn Sampler (NUTS, seeHoffman & Gelman, 2014). As with all MCMC algorithms this requires a phase at the beginning which allows the sampler to find a high-density region as well as to finding "tuning parameters" that help the sampler to operate efficiently. For all results here in we set this phase to have 2,000 iterations, followed by a further 1,250 iterations using eight independent chains which provide 10,000 sample values to summarize the posterior distribution.

Convergence of MCMC has a slightly different meaning in the context of Bayesian estimation compared to Maximum-Likelihood, and is generally inferred by statistics such as the R-hat statistic (Vehtari et al., 2019), along with "trace-plots" for the sample statistics. For all of the models presented here the R-hats were within what is considered to be tolerance (< 1.05). We present these in the Appendix (Figures A2 and A3).

3.1. The likelihood

Equation [16] gives the probability of choosing option A for a given set of parameters. The HBM logit allows for each respondent to have their own preference parameters, such that we can define $y_{j,t} = 1$ if the j^{th} person chooses the non-sure option A in the t^{th} task. Therefore, we can write:

$$\Pr ob\left(y_{jt} = 1|\Omega_j\right) = \frac{\exp\left(U_j\left(\mathcal{Z}_{a,t}, \Omega_j\right)\right)}{\exp\left(U_j\left(\mathcal{Z}_{a,t}, \Omega_j\right)\right) + \exp\left(U\left(\mathcal{Z}_{b,t}, \Omega_j\right)\right)} \tag{17}$$

The parameters for the j^{th} individual determine this probability. These are:

$$\Omega_j = \left(\alpha, \lambda, \beta, \omega_{1,L}, \omega_{2,L}, \omega_{1,G}, \omega_{2,G}, \gamma_1, \gamma_2, \delta_1, \delta_2, \rho_G, \rho_L, \rho_M, \pi_L, \pi_G, \eta\right)_j$$

where all parameters except $\{\omega_{1,L}, \omega_{2,L}, \omega_{1,G}, \omega_{2,G}, \}$ are respondent specific.

The ω parameters will only shift the mean value for all respondents who complete the set of prospect pairs. It is worth re-iterating that this model specification is not only able to estimate a flexible value function but it is doing so for each individual respondent. This is an important advantage of employing the HBMs compared to Classical estimation.

Given that $\mathbb{P} = (\mathcal{P}_1, \mathcal{P}_2, ..., \mathcal{P}_T)$ prospect pairs are given to each individual, and with associated outcomes (choices) $y_j = \left(y_{1,j}, y_{2,j}, ..., y_{T,j}\right)$, the data can be defined as $Y = \left\{y_j\right\}_{j=1}^{J}$ and $\Omega = \left\{\Omega_j\right\}_{j=1}^{J}$ and the log likelihood of the data is as follows:

$$\ln f\left(Y|\Omega\right) = \sum_{j=1}^{J}\sum_{t=1}^{T}\left(\ln \Pr ob\left(y_{jt} = 1|\Omega_j\right) y_{jt} + \ln \Pr ob\left(y_{jt} = 0|\Omega_j\right)\left(1 - y_{jt}\right)\right) \tag{18}$$

3.2. Model priors

For a HBM Logit the priors are set at two levels. First for what one might think of as the "representative agent" level, and second at the level of the individual. The hierarchical nature of the HBM means that the representative agent level parameters form the means around which the individual estimates are distributed. Full details of the priors are given in Appendix A. Here we give a more partial and simplified outline.

Priors can be used to set regions over the values that the parameter can take and place higher prior likelihood for some values over others. We do both, though in a way that will allow the data to largely determine the parameter values within the bounds we set. These priors are derived from theory and past empirical work. First, we set a baseline which gives high prior density to the point

https://doi.org/10.1017/eec.2025.10012 Published online by Cambridge University Press

$\alpha = \beta = \gamma_1 = \gamma_2 = \delta_1 = \delta_2 = \lambda = 1$ (at both the representative agent and individual level) which is tantamount to a model in which expected values determine choices both within and across domains. Where additional parameters are added, these are done so in a way that their modes give high prior density to the "standard" parameterisation of Tversky and Kahneman (1992). Second, we allow the priors to be wide enough and diffuse enough to reflect what we would characterize as the "consensus region" for these parameters.

Across the different models the priors were the same for those parameters occurring in different models. For example, for the power parameter in the gain domain α, the prior for the mean of $\bar{\alpha}$ is considered normal with the bounds $[\bar{l}_\alpha, \bar{u}_\alpha]$. The parameters for each individual are then hierarchically structured such that for each individual j, such that α_j is normally distributed around $\bar{\alpha}$ within some bounds $[l_\alpha, u_\alpha]$ which is wider than those for $\bar{\alpha}$. How far the individual values α_j diverge is governed by a standard deviation σ_α which is estimated within the model which also requires a prior which in this case is σ_α^{-2} and is given a prior Gamma distribution.

The parameters that fully determine the nature of the priors are referred to as hyper parameters. The hyper parameters for α are $\mu_\alpha = 1$, where for $\sigma_\alpha^{-2} \sim Gamma(h_{\alpha,1}, h_{\alpha,2})$, $h^2_{\alpha,0} = 1$, $[\bar{l}_\alpha, \bar{u}_\alpha] = [0.05, 2.5]$, $[l_\alpha, u_\alpha] = [0.01, 4]$, $h_{\alpha,1} = 1$, and $h_{\alpha,2} = 0.25$ (i.e., the power parameter had a prior mean of 1 but was allowed to vary between $[0.05, 2.5]$ whereas the parameters pertaining to individuals had a slightly wider boundary between $[0.01, 3]$). The same structure but with some different means, variances and boundaries was then used for $\beta, \gamma_1, \gamma_2, \delta_1, \delta_2, \pi_1, \pi_2$ and η. The parameters $\omega_{1,G}$ and $\omega_{1,L<text></text>}$were half normal, and $\omega_{2,G}$ and $\omega_{2,L<text></text>}$ were uniform, but were not hierachical (that is they do not have values for individuals). Given the potential long-tail in the loss aversion parameter, we specified this to be log normal with 50% of the mass below unity and 50% above, again some boundary conditions that contain the vast majority of values (1/10, 10) for the mean and (1/12, 12) for individuals. The unitary priors for the warping parameter reflect a prior weighting towards no probability warping and together give high prior mass to the case where individuals act according to expected value maximization within and across domains. Again, readers are referred to Appendix A for full details.

As already noted, these priors were set to cover what we would see as the "consensus region". Consensus regions are difficult to establish because even within papers using Power-Utility, estimates have often been obtained using different estimation methods, weighting functions (or no probability weighting at all) and/or with parametric restrictions that are "symmetrical" in either the value space or probability weighting space or both. However, while some papers use the weighting function of Tversky and Kahneman (1992) both directionality and in terms of magnitude the probability weightings broadly correspond to what will happen with the same parameter used within the Prelec I function, and they are thus they comparable. Therefore, a reading of literature in terms of Abdellaoui (2000), Alam et al. 2022, Andersen et al. 2010, (Balcombe et al., 2019), Baillon et al. (2020), Brown et al. (2024), Bouchouicha and Vieider (2017), (Chapman et al., 2024), Gao et al. (2023), Gonzalez and Wu (1999), Murphy and ten Brincke (2018), Nilsson et al. (2011), Tversky and Kahneman (1992) and Walasek et al. (2024) broadly support the bounds and priors that we have set, where there tendency to find $\alpha, \beta, \gamma_1, \rho_1 < 1$ and $\lambda > 1$. As remarked by Wakker & Zank (2002) among those studies for which $\alpha = \beta$ has not been imposed there is evidence for $\beta > \alpha$, though this is not universal (e.g., Bouchouicha & Vieider, 2017).

3.3. Model versions

Given the flexible model specification presented, we have four models that are nested as follows:

- Model 0: (*Mdl*0) $\omega_{1,G} = \omega_{1,L} = \omega_{2,G} = \omega_{2,L} = 0$, $\eta_j = \eta = 0$ and $\pi_L = \pi_G = \pi_{L,j} = \pi_{G,j} = 1$ for all j
- Model 1: (*Mdl*1) $\eta_j = \eta = 0$ and $\pi_L = \pi_G = \pi_{L,j} = \pi_{G,j} = 1$ for all j

https://doi.org/10.1017/eec.2025.10012 Published online by Cambridge University Press

- Model 2: ($Mdl2$) $\eta_j = \eta = 0$ for all j
- Model 3: ($Mdl3$) Unrestricted

Our base model ($Mdl0$) is a standard power form without transforming the value function or allowing for flexibility. The next model ($Mdl1$) uses a generalized value function to allow greater flexibility around small payoffs in a way that allows respondents to be risk seeking in low-stakes mixed prospects. $Mdl2$ further generalizes the specification to allow for the certainty effect only and $Mdl3$ is the most general model specification allowing for the certainty effect as well as the selection of prospects based on whether they have more or less-zero payoffs.

4. Data

As explained by Balcombe and Fraser (2024), the set of lotteries have been generated by employing a statistical procedure based on two design principles. First, it is assumed that each lottery task should be "informative". A task is "non-informative" if the same option would be chosen regardless of preferences. That is, if two individuals with very different preferences are likely to make the same choice, then that task is not informative about the preferences of those individuals. By contrast, an informative task is likely to reveal different choices by individuals with different preferences. Second, any task should not be rendered redundant by any other task (pairwise redundancy). The most obvious example is that tasks should not be repeated. While repeated tasks do provide some level of additional information, there are usually differentiated tasks that are more informative. However, more generally, the answer to one task should not be able to predict the answer to any other task across the range of possible preferences.

These principles have been formalized by Balcombe and Fraser (2024) using Bayesian inference that seeks to estimate a "posterior" distribution for the parameters in question. The posterior is the distribution given the choices of respondents and is constructed from the data along with a prior distribution. The more informative this posterior distribution is, the better we can make inferences about the parameters of interest. The set of tasks chosen yield a posterior distribution with low entropy, or equivalently, high Kullback-Leibler divergence from a uniform distribution. As such, this approach to experimental design focuses on the informational content of the lotteries generated to enable effective recovery of key PT parameter estimates. With this approach it can be shown, for example, that the lottery set used by Harrison and Swarthout (2016) can be reduced given that several of lottery pairs (up to 40%) are "redundant" in terms of estimating preference parameters.

The data was collected in 2019 across a series of experimental sessions. Each session began with the distribution of detailed instructions along with examples of questions and an explanation of the payment process. Next, we distributed the survey instrument containing the choice tasks plus a small set of socio-economic questions. We required all survey participants to answer the 100 lottery tasks constructed, each composed of two options, with one option always being a sure-thing. The experimental design was composed of 21 tasks in the gain domain, 26 in the loss domain, and 53 in the mixed domain. All we required participants to do was to indicate which option they preferred. Our sample contained 143 respondents (both undergraduate and postgraduate students) meaning our data set is composed of 14,300 responses.

The payoffs for the tasks were generated to lie within the bounds -100 to 100 in the first instance. We then scaled down the lotteries by a factor of five to give a low payoff treatment and we multiplied this by two and half times to derive a high payoff treatment. In total 74 respondents completed the low payment treatment, and 69 the high payment treatment. To make the tasks easier to understand, we presented the tasks in the form of pie charts. The lottery rewards were presented as segments (pieces of pie) where the size of each segment was proportional to the chance of winning the reward. Additionally, we included labels that specified the exact reward and the corresponding chance of winning. An example of a lottery task is shown in Figure 1.

https://doi.org/10.1017/eec.2025.10012 Published online by Cambridge University Press

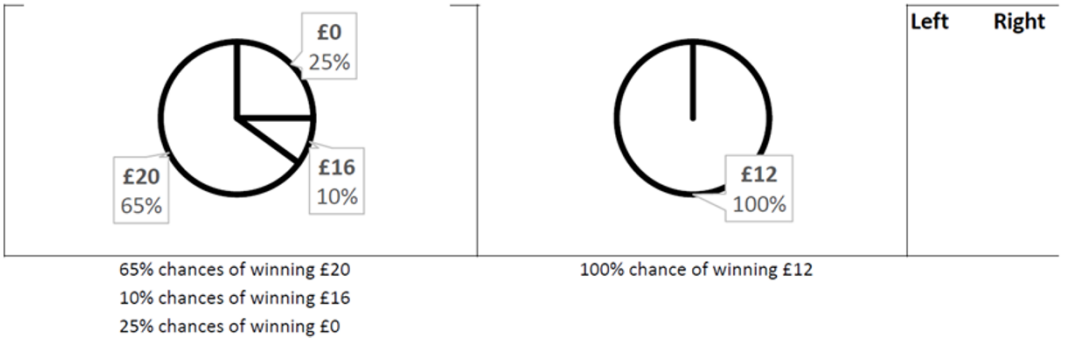

Figure 1. Example lottery task

In the example shown in Figure 1, the left option consists of the rewards £ 0, £ 16, and £ 20 with 25%, 10%, and 65% chances of winning. The right option consists of a single, certain, reward of £ 12. In the instructions handed out to participants, the information presented in Figure 1 was also summarized in words. Once all participants had completed the 100 tasks, all answer sheets were collected. Full anonymity was ensured throughout the experiment, and at no point was any participant required to divulge their identity.

On agreeing to participate in the experimental sessions, all participants knew that they would receive a £ 15 participation fee. In this experiment, we also employed real monetary incentives whereby we randomly selected 20 % of participants in each experimental session to play for one of their 100 choice tasks for real money. This meant that participants had the chance to win additional money, up to a maximum of £ 50 once all data collection was completed. This approach to implementing real incentives frequently selects 10 % of participants e.g. Amador-Hidalgo et al. (2021) and it has been shown not to negatively affect a participant's motivation or the validity of the random lottery incentive scheme. For lotteries involving negative outcomes, if selected, a fixed endowment was employed that equalled the largest possible loss that could be realized during the experiment. Employing an endowment that covers losses is the standard procedure for dealing with negative outcomes in experiments (see Fehr-Duda et al., 2010, Etchart-Vincent & L'Haridon, 2011, Vieider et al., 2015, Bouchouicha & Vieider, 2017, Amador-Hidalgo et al., 2021).

5. Results

Table 1 presents the estimates for the mean parameters and standard deviations (Sd) (in brackets) of the four models described. Model selection is undertaken by employing the Watanabe Akaike Information Criteria (WAIC) (for which a lower score indicates a preferred model, see Watanabe (2013)). The WAIC is a fully Bayesian information criterion, which awards "fit" but penalizes additional model parameters. The WAIC scores and standard errors (SEs) (in brackets) are presented at the bottom of Table 1.

In Table 1, we see that there is a monotonic decrease in the WAIC from *Mdl*0 to *Mdl*3. The decrease occurs because of the addition of the extra model flexibility of the value function relative to the "standard model" *Mdl*0. Comparing *Mdl*0 *vs Mdl*1, we see that this yields a decrease in the WAIC that is approximately twice the SE suggesting robust evidence that increase in model flexibility improves performance. The inclusion of the sure-thing via *Mdl*2, is, however, far more definitive, with a larger reduction in the WAIC, which is approximately fourfold the SE. Finally, there is a much smaller reduction for in the WAIC in moving from *Mdl*2 *vs Mdl*3 and the reduction is less than the associated SE. While this result does support model *Mdl*3 that includes the payoff dimension, the statistical evidence in support of this additional model flexibility is much weaker than for the change for *Mdl*1 to *Mdl*2.

https://doi.org/10.1017/eec.2025.10012 Published online by Cambridge University Press

Table 1. Parameter estimates - mean and standard deviation

	Mdl0 **Mean (Sd)**	Mdl1 **Mean (Sd)**	Mdl2 **Mean (Sd)**	Mdl3 **Mean (Sd)**
α	0.835(0.027)	0.722(0.040)	0.722 (0.041)	0.742(0.049)
β	1.130(0.033)	1.084(0.040)	0.985(0.048)	0.982(0.049)
λ	0.484(0.062)	0.393(0.065)	0.547(0.103)	0.575(0.124)
ρ_G	0.656(0.097)	0.934(0.154)	0.832(0.141)	0.785(0.146)
ρ_L	0.337(0.036)	0.492(0.069)	0.541(0.066)	0.484(0.083)
ρ_M	0.436(0.040)	0.594(0.076)	0.618(0.082)	0.599(0.095)
δ_1	1.674(0.065)	1.705(0.067)	1.601(0.076)	1.583(0.076)
δ_2	1.133(0.051)	1.113(0.053)	1.221(0.066)	1.233(0.064)
γ_1	0.927(0.039)	0.978(0.044)	0.913(0.052)	0.890(0.051)
γ_2	1.103(0.038)	1.045(0.040)	1.101(0.054)	1.133(0.055)
ω_{1G}		1.412(0.620)	1.596(0.593)	1.635(0.668)
ω_{2G}		0.362(0.108)	0.341(0.090)	0.340(0.105)
ω_{1L}		4.650(1.737)	4.921(1.659)	4.988(1.640)
ω_{2L}		0.640(0.110)	0.520(0.090)	0.516(0.086)
π_G			0.957(0.022)	0.936(0.023)
π_L			0.878(0.026)	0.854(0.024)
η				-0.041(0.014)
WAIC	15222.1(115.8)	15190.9(115.5)	15114.15(115.2)	15103.06(114.8)
Model Comparison*				
Mdl0vsMdl1		-31.67 (13.6)		
Mdl1vsMdl2			-76.83 (18.2)	
Mdl2vsMdl3				-11.089 (15.5)

*Notes:**Difference WAIC (SE Difference WAIC) A negative WAIC⇒ meaning that the model on the right is preferred; Sd - Standard deviation; SE - Standard Error

Comparisons for the mean parameter estimates, between the four models are made using the results in Table 1 and Figure 2, with only the parameters common to all models presented. Comparative kernel density plots are also presented in the Appendix (Figure A1).

First, we note that in some respects the key PT parameters are similar across all of the models, with a couple of notable exceptions. While the power parameter in the gain domain (α) is sub-unity (reflecting concavity), the power parameter in the loss domain (β) is above unity (also reflecting concavity) for models *Mdl*0 and *Mdl*1 (1.13, and 1.084 respectively) which is not in keeping with the expected finding of convexity in the loss domain. However, a similar finding has been reported by Kpegli et al. (2023) who employ a novel semi-parametric method. Furthermore, even though Kahneman and Tversky (1979) and Tversky and Kahneman (1992) discuss convex functions for losses, such functions have proved the exception rather than the rule in empirical estimations (Abdellaoui, 2000, Bruhin et al., 2010). There is also the possibility that the highly non-linear nature of the PT model, such as the value function estimates combining with a highly optimistic parameter of the Prelec-II function can imply that there might be noise in the model that makes parameter identification more problematic (see e.g. Zeisberger et al., 2012, for a discussion).

Another possible reason for the concavity could be that the functional form was not sufficiently flexible in the sense that for small payoffs respondents were risk seeking. It is for this reason, that we examined the model generalization allowing for non-uniform curvatures in the loss and gain domains

https://doi.org/10.1017/eec.2025.10012 Published online by Cambridge University Press

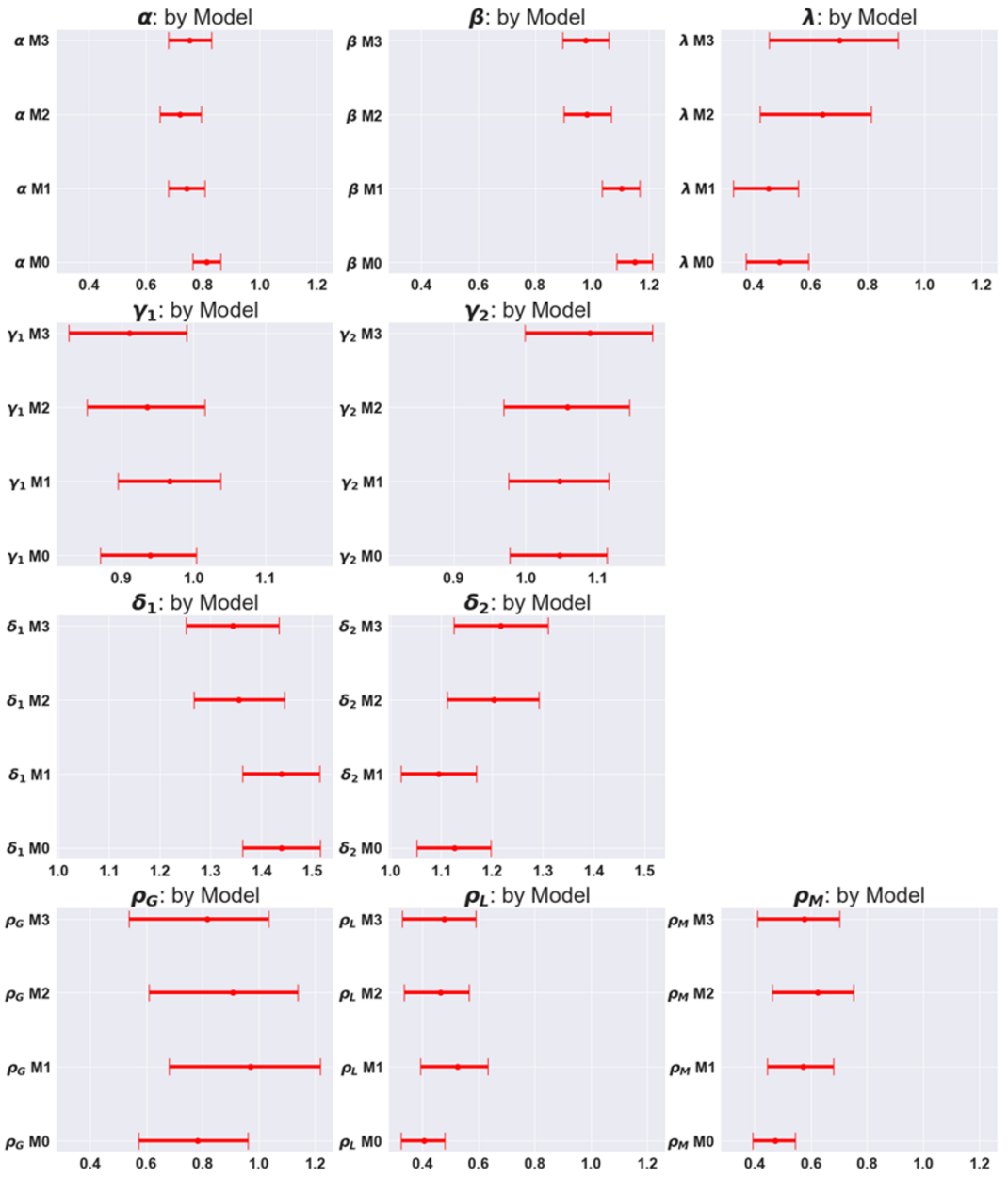

Figure 2. Comparison of PT parameters for all models

respectively. The introduction of the logistic transformation with *Mdl*1 that allowed for this, while reducing the mean value for β to a small extent, did not induce a sub-unity mean. How the value function is transformed is illustrated in Figure 3.

The left panel gives the value function for payoffs less than £ 10 in absolute terms and for payoffs under £ 50 in absolute terms on the right. These figures illustrate the relatively small departures from the power form for the lower values, but they leave the value function largely unchanged for absolute values higher than £|10|.

https://doi.org/10.1017/eec.2025.10012 Published online by Cambridge University Press

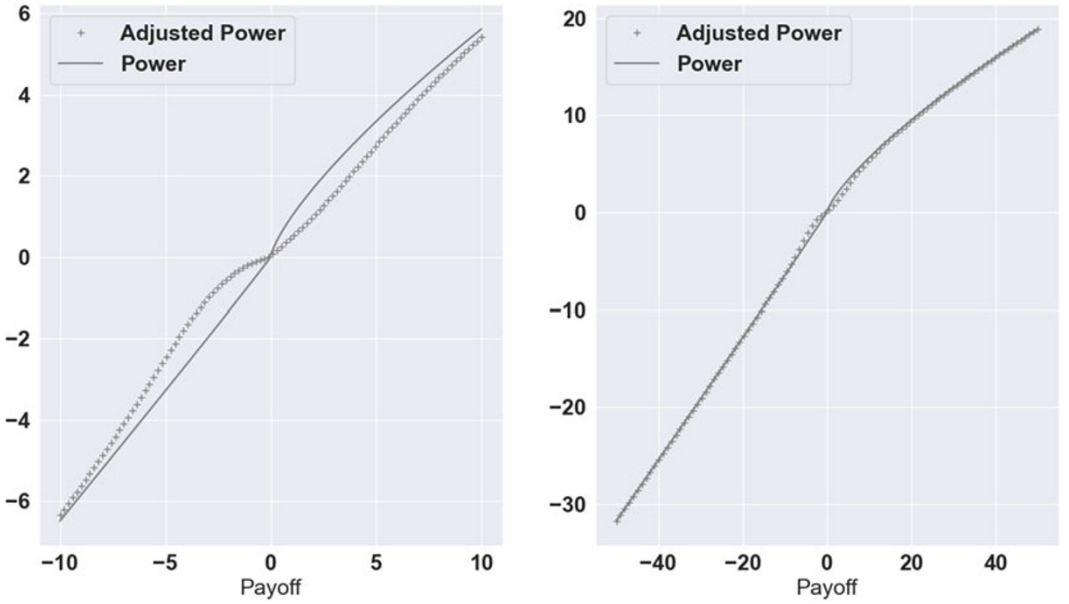

Figure 3. The mean value function (for $\mathbf{M}dl1$)

A more substantive change in the mean value of β occurs when the certainty effect correction is introduced in models *Mdl*2 and *Mdl*3. In both models, the values of β are reduced to become sub-unity, but as can be seen from Figure 2, a confidence interval would include unity. We also observe that the value of λ correspondingly shifts upwards for models that allow for the certainty effect (*Mdl*2 and *Mdl*3).

Readers are reminded that providing $\beta \neq \alpha$ the value of λ being above or below unity is largely determined by the denomination of the payoffs and does not reflect a lack of "loss aversion". The greater pivot downwards of the value function in the loss domain the greater the aversion to risky prospects in the mixed domain. However, if $\beta > \alpha$ there will be this tendency provided the payoffs are large enough. In Figure 4 we give an illustration of how people would react to a series of 50:50 prospects with a positive and negative payoff of the same size be calculating the certainty equivalent at each payoff size using the parameters for *Mdl*3, with and without the impact of the "wobble" in the value function.[2] We also look at just the pure value effect as opposed to factoring in the probability warping components. We can see that for very small sized prospects people would accept the risky prospect, but that as the size grows they eventually become risk averse. The role of the "wobble" exacerbates this tendency with the probability weighting also increasing people's "tenancy"? to take the riskier option. [3] Using our approach, β has a reduced value because of the bias towards a negative sure-thing in the loss domain (as reflected in Table 1 for $\pi_L = 0.878$ and 0.854 for *Mdl*2 and *Mdl*3 respectively). The sub-unity value of this parameter means that the negative utility associated with a negative sure-thing is less negative than that suggested by the value function alone. Thus, in models *Mdl*0 and *Mdl*1, which did not allow for the certainty effect, the fact that respondents were choosing a

[2]Note that the model with the "wobble" does not have an analytical certainty equivalent but can be approximated.

[3]We also estimated our model using only data in the gain and loss domains to assess if the utility wobble result was robust (See Appendix B). We found that when using the partial data set, there is much less of a wobble in the gain domain. However, the wobble for the loss domain parameters remained remarkably similar and it was in this domain that the larger "wobble" was observed for the full data set. Given these results, we conclude that evidence for the smaller wobble in the gain domain was only evident with the inclusion of mixed prospects. However, the larger departure from the standard power form which was in the loss domain is exhibited with or without the inclusion of mixed prospects.

https://doi.org/10.1017/eec.2025.10012 Published online by Cambridge University Press

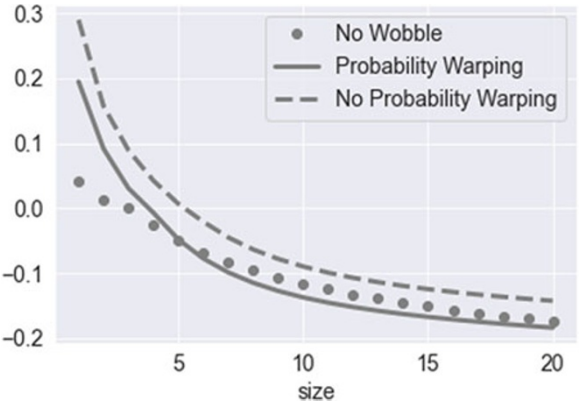

Figure 4. Certainty equivalent for a 50:50, EV=0 prospect (for $\mathbf{M}dl3$)

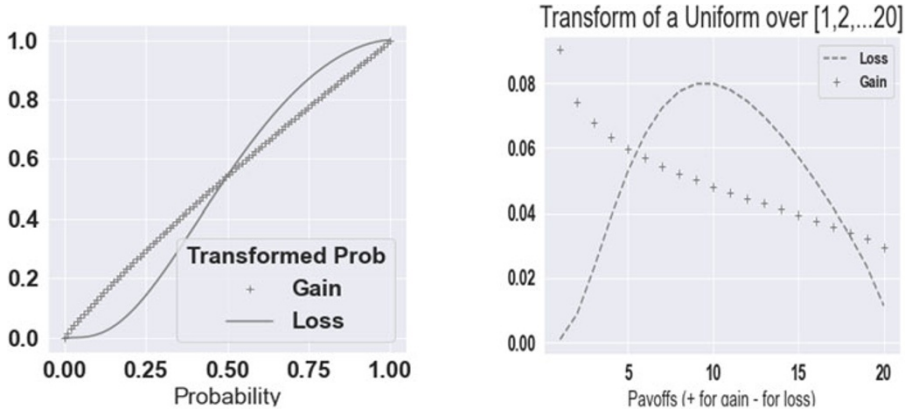

Figure 5. Mean probability weightings (for $\mathbf{M}dl3$)

negative sure-thing resulted in convexity of the value function in the loss domain. There also appears to be weaker evidence for a smaller certainty effect for positive sure-things (as reflected in $\pi_G = 0.957$ and 0.936 for *Mdl*2 and *Mdl*3 respectively), which would reflect a slight tendency to avoid the sure-thing in the gain domain, but not in a way that dramatically alters the curvature in the gain domain. Finally, the parameter $\eta = -0.041$ indicates the utility of those with only two payoffs (with one being non-zero) having down-weighted utility relative to the three-payoff prospects (with two being non-zero). As noted earlier this model addition (*Mdl*2 *vs Mdl*3) only has weak support from the WAIC but has a posterior that has more of its mass below zero.

Turning to the probability weightings, these vary to some extent over the four models but do not change the essential shape of the probability weighting functions, which are illustrated in Figure 5 for the "representative agent". What is evident for these parameters is that while γ_1 is just sub-unity (in conformance with previous literature) the value for ρ_1 (1.583 for *Mdl*3) which much larger than unity, which is counter to what we would regard as consensus though this is also reported by Kpegli et al. (2023).

On the left-hand side of Figure 5 are the transformations of a linear cumulative function. These can be understood as the how the probability of the larger payoff (in absolute terms) is transformed for a two-payoff prospect. On the right-hand side are the transformations of uniform probability mass of a 20-payoff ordered prospect in the gain and loss domains (only to illustrate the shape). These show that

https://doi.org/10.1017/eec.2025.10012 Published online by Cambridge University Press

Table 2. Parameter distributions for sample individuals for Mdl3

	Sd	Min	5%	25%	50%	75%	95%	Max
α	0.091	0.471	0.587	0.683	0.751	0.798	0.888	0.971
β	0.105	0.678	0.804	0.910	0.985	1.064	1.124	1.217
λ	0.434	0.398	0.443	0.522	0.611	0.696	0.864	1.114
ρ_G	0.256	0.403	0.479	0.628	0.760	0.968	1.282	1.556
ρ_L	0.215	0.196	0.250	0.364	0.473	0.640	0.856	1.614
ρ_M	0.237	0.192	0.324	0.466	0.584	0.804	1.038	1.418
δ_1	0.415	0.551	0.779	1.344	1.673	1.890	2.184	2.605
δ_2	0.385	0.495	0.652	0.942	1.222	1.498	1.911	2.237
γ_1	0.280	0.184	0.409	0.696	0.900	1.080	1.366	1.547
γ_2	0.260	0.618	0.679	0.967	1.117	1.323	1.546	1.716
π_G	0.058	0.765	0.841	0.895	0.938	0.976	1.026	1.080
π_L	0.076	0.645	0.730	0.800	0.853	0.904	0.976	1.069
η	0.054	-0.188	-0.135	-0.072	-0.037	-0.009	0.052	0.106

Notes: Sd - Standard Deviation; Min - Minimum value of parameter distribution; Max - Maximum value of parameter distribution.

the probability weightings weakly conform with the inverse-S shape in the gain domain. However, this is not the case in the loss domain which tends to overweight the mid values and underweight the end values. For example, whereas previous literature suggests that for a two-payoff example a "small-probability-large-loss" would be given excessive weight by people, the result here is strongly in the opposite direction. It is more consistent with people largely ignoring low probability bad outcomes.

5.1. Heterogeneity and model results

Table 2 gives the percentile values for the key model parameters along with the associated minimum, maximum, and standard deviation for model *Mdl*3.

Comparing the median values in Table 2 with the mean estimates in Table 1, we can see that these are broadly in accordance. Each estimate for the individual (i.e., the values that correspond to α_j, β_j etc.) is the mean of the posterior for that individual. Figure 6 presents histograms that show the distribution across the sample for the individual-specific parameters. Figure 5 gives the cumulative histograms using the distributions depicted in Figure 6 so that the portions lying above and below key values can be assessed. In HBMs individual distributions can be bimodal or multimodal or highly skewed even though the underlying prior is normal. While there is some skewness in some of the distributions, the individual estimates are broadly in accordance with the assumptions behind the model.

In most respects, the individual estimates underline some of the findings already discussed about the representative agent values. In relation to the value function, many individuals have concave value functions in the loss domain. From the cumulative distributions in Figure 6, we see that around 60% of respondents have sub-unity estimates for β. Thus, a substantial minority of individuals are not behaving in a way that is generally characterized by PT models. If the probability weighting functions were linear, this would suggest that many individuals are risk averse in the loss domain, and over and above this have a bias toward choosing the sure-thing. In contrast, when we consider the gain domain, 100% of respondents had sub-unity values for α that is consistent with PT models.

From Figures 6 and 7, it can also be observed that nearly all individuals had estimates for π_L that were sub-unity, and although it cannot be seen from the results, all respondents with concave

https://doi.org/10.1017/eec.2025.10012 Published online by Cambridge University Press

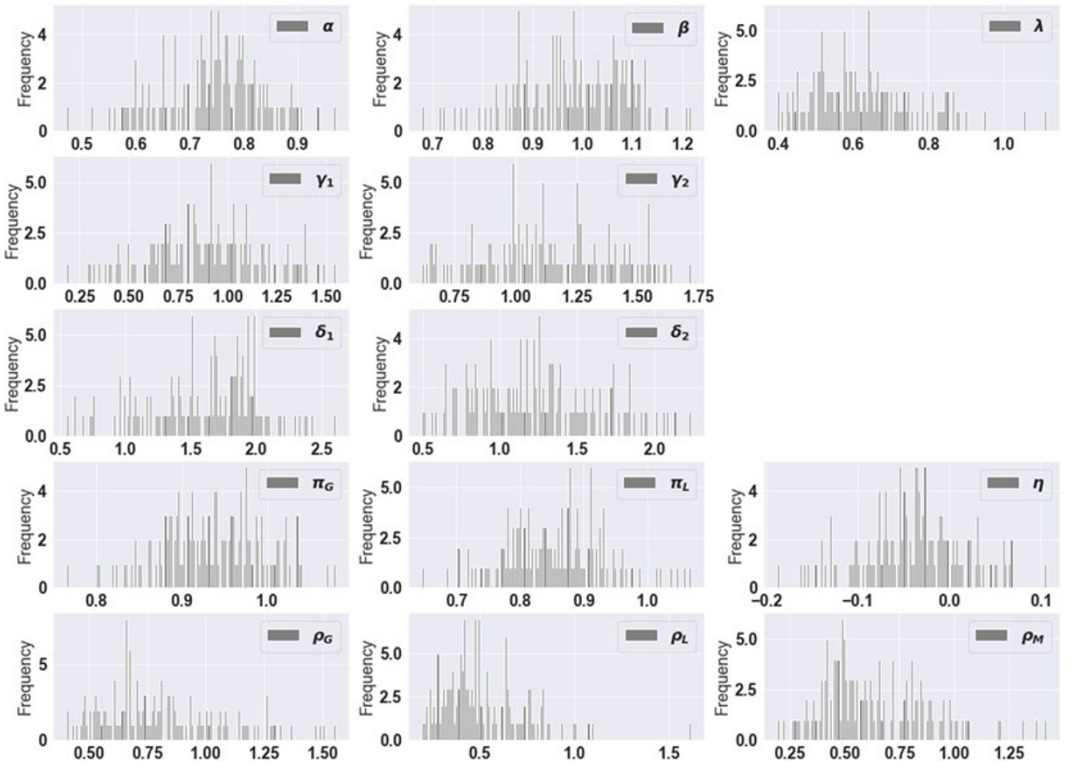

Figure 6. Histogram for individual specific parameters

value functions in the loss domain also had a certainty effect towards choosing the sure-thing (except one).

The conclusion that a substantial proportion of people are risk-averse in the loss domain must be tempered by considering the probability weighting functions. At the representative agent values, both two-payoff and three-payoff prospects in the loss domain, the largest loss is underweighted to a greater degree than the smallest loss. Thus, the probability weightings are suggesting a tendency towards choosing options that underweight the largest loss, which is arguably a form of risk-seeking behaviour.

Next, Figure 8 breaks down the probability weighting functions into quadrants based on whether the Prelec II parameters are above or below one (recalling that where both are unity the probability weighting function will be linear).

In Figure 8 the cumulative distributions for each individual are plotted and the numbers of people occurring within the quadrant are within the title of each panel. As with the previous plots these can be understood as the how the probability of the larger payoff (in absolute terms) is transformed for a two payoff prospect. As can be seen from these plots, the distribution of weighting types is more diffuse in the gain domain (left-hand figures for γ), where only a minority of respondents have the inverse S type weighting (as in the bottom-right panel where $n = 36$) though the majority overweighting low probability large gains.

By contrast, the weightings for the loss domain (right-hand panel for δ) are more similar across respondents but with the dominant type being the S type weighting (as in top-left panel where $n = 94$). Naturally, these broadly mirror the tendency reflected for the representative agent results discussed earlier, and do not generally conform to what we would regard as the consensus. That is,

https://doi.org/10.1017/eec.2025.10012 Published online by Cambridge University Press

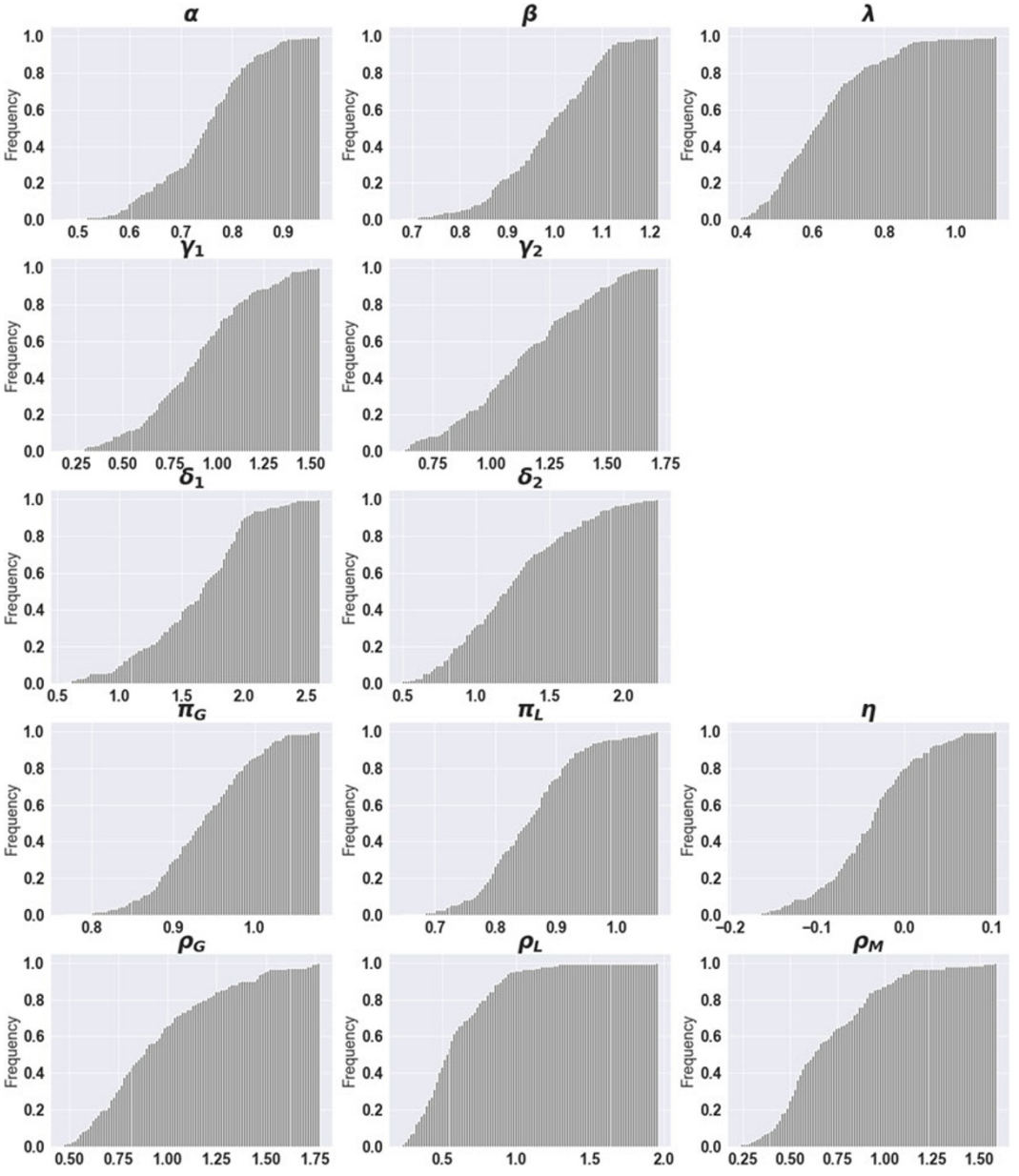

Figure 7. Cumulative histograms for individual-specific parameters

here the substantive majority underweight small probability larger losses but conversely overweight high probability larger losses.

6. Conclusions

We have introduced and presented HBMs to parametrically estimate a flexible PT model that goes beyond those typically found in the literature. Our method adds to a growing literature employing HBMs to estimate and recovery individual respondent risk parameters (e.g., Balcombe & Fraser,

https://doi.org/10.1017/eec.2025.10012 Published online by Cambridge University Press

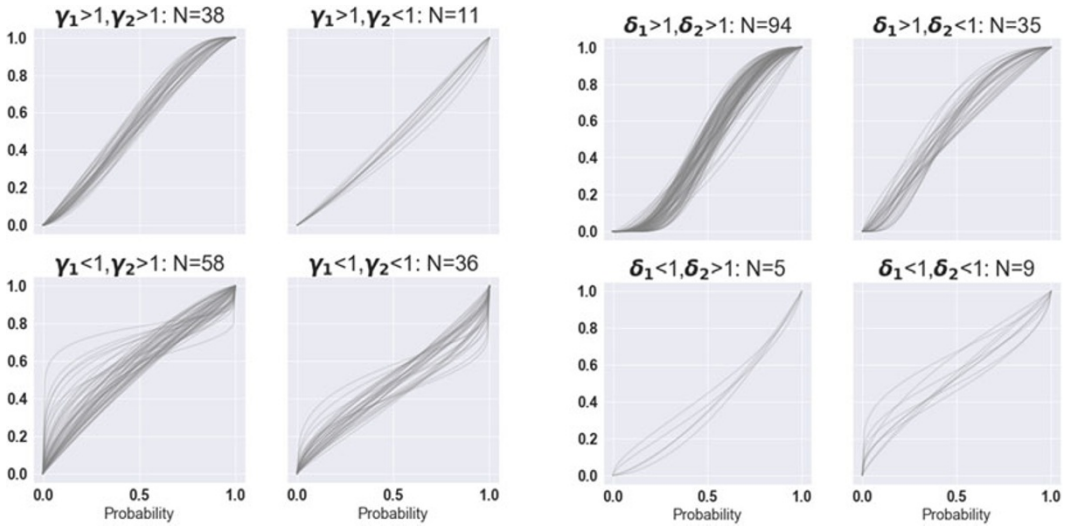

Figure 8. Individual probability weightings

2015, Gao et al., 2023). In particular, the HBMs not only allow for individual respondent estimates to be generated but also the incorporation of significant additional flexibility in the value function to capture differences in behaviour between domains (gain, loss and mixed). The model is also able to capture aspects of lottery design that are frequently employed in an effort to reduce task complexity. However, it is conceivable that employing an even more flexible functional form or the adoption of non-parametric methods could further improve model performance. Both of options are feasible given the use of HBMs and the flexibility that they offer for applied econometrics.

In terms of the impact of experimental design and how explicitly this is taken into account, we recommend that researchers explicitly account for these features in their model specifications. The use of a sure-thing option in risk experiments is typically justified as a means to reduce task complexity. Given earlier research (e.g., Kahneman & Tversky, 1986), we expected a priori that there would be a certainty effect. However, our findings did not align completely with this expectation. Instead, our results indicate that respondents did exhibit a certainty effect towards sure-things, but not necessarily consistently towards choosing the sure-thing option (or simpler options in general), as we initially hypothesized. Positive and negative sure-things had their utilities reduced in absolute terms meaning that there was a certainty effect towards negative sure-things but not for positive sure-things. Taking account of the certainty effect for negative sure-things lead to a slightly convex value function in the loss domain for a majority of respondents, whereas without this adaptation the value functions were predominately concave in the loss domain. However, allowing for the certainty effect did not substantially change the values of the probability warping parameters of the weighting function.

Turning to the payoff format employed in this lottery instrument, our findings provide limited support for the notion that individuals tend to prefer three-payoff prospects over two-payoff prospects. While the evidence is not conclusive, it suggests that this type of task complexity may have some influence on decision making, highlighting the importance of understanding individuals' perceptions of different types of task format. Indeed, it may be that choosing between a three-payoff prospect and a sure-thing is not significantly more cognitively challenging than deciding between a two-payoff prospect and a sure-thing. While it is not surprising that our research identified domain-specific behavioral differences, this specific finding underscores the need for further investigation. In particular, it is crucial to carefully consider what individuals perceive as a complicated task when examining potential types of behaviour in decision making.

https://doi.org/10.1017/eec.2025.10012 Published online by Cambridge University Press

Another interesting feature of the results revealed "as a result of employing HBM" relates to the evidence we find of probability warping. Although probability warping was present (i.e., the weighting function parameter estimates were different from unity) our parameter estimates did not universally reflect the values that are typically reported in the literature. Notably, we observed a tendency among participants to assign disproportionate weight to the middle values of prospects within the loss domain. This suggests that individuals may exhibit a specific behaviour towards valuing prospects in this particular context.

Finally, our respondents demonstrated higher predictability when making choices in the gain domain compared to choices in the loss or mixed domains. It is possible that individuals find it easier to make decisions in the gain domain, which may explain their tendency to display a certainty effect when facing losses.

In terms of future research "given the HBMs" developed an obvious extension would be to consider other data sets and to see if any other aspects of experimental design, once taken into account, yield improvements in model performance. Another related extension would be to consider individuals' subjective assessments of task complexity (could be derived for various aspects of the experimental design) which may provide insights into their decision-making processes. Related research would be to link attentional weighting (i.e., which option attracts most attention during the decision-making process) and probability weighting using eye tracking. Zilker and Pachur (2021) provide results examining the relationship between these constructs. They indicate that the probability weighting function might also be capturing aspects of how respondents acquire and process information.

Another avenue of future research would be to consider making other functional forms that are employed in the literature more flexible. This option has not been pursued here simply to avoid a "horse race" between model specifications. However, the use of Bayesian model averaging (see Balcombe & Fraser, 2015) provides a means by which model selection could be undertaken and may prove a useful exercise. Similarly, there are additional means by which functional form flexibility as well as individual level parameters could be introduced with the HBM context. For example, we could in principle, estimate a semi-parametric model giving a distinct utile for every outcome, and a distinct decision weight for each probability. This exercise could further highlight the benefits of employing HBMs in comparison to existing research such as Gonzalez and Wu (1999). In their work they needed 100+ certainty equivalents to achieve this type of task. Could we even do away with all functional form assumptions?

Of course, there may also be good reason not to introduce ever more flexible functional forms. For example, the introduction of four additional parameters to capture wobbles in utility may not be necessary if the wobbles are not the same for gain and loss domain prospects compared to the mixed domain prospects. This may be the case if the data has few mixed prospects which is typical of many data sets used in the literature. However, as we have noted when excluding the mixed prospects from our data, we still find that evidence of significant utility wobble especially in the loss domain. This finding is important and warrants further examination in future research.

Turning to the generation of experimental data Gao et al. (2023) demonstrate that use of HBMs allows for key parameter recovery to occur using a sub-set of the overall experimental design without the loss of significant parameter accuracy. In the example they provide, the set of lotteries are randomly assigned into sub-sets. The experimental approach used by Balcombe and Fraser (2024) could in principle be used to enable a more effective design of sub-sets that in turn would help with preference parameter estimation.

It is also the case the link between experimental data generation and the impact this has on parameter estimates requires more research. The generation of the experimental data provided by Balcombe and Fraser (2024) has several features that attempt to reduce task complexity and reduce the noise inherent in the data. Given the emerging literature on complexity (Oprea, 2022, Oprea, 2024) there is a need for more explicit consideration of how complexity in task design impacts respondent engagement. This need also links back to the issue of how noise in experimental data is modelled. The

https://doi.org/10.1017/eec.2025.10012 Published online by Cambridge University Press

traditional approach of imposing a noise structure as opposed to viewing the tasks as inherently noisy, as examined by Vieider (2024) requires more consideration. At the same time, the way in which the use of RUM implicitly incorporates noise needs to be reconsidered so that the apparent gap between deterministic and stochastic choice can be reconciled.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/eec.2025.10012.

Data availability statement. The replication material for the study is available at: DOI 10.17605/OSF.IO/A9EWS.

Acknowledgements. This research was part funded by British Academy/Leverhulme Small Research Grants (SRG 170874) "Improving the estimates of attitudes towards risk". We thank Andreas Markoulakis for providing research assistance with data collection. We also thank an associate editor and two reviewers for insightful and positive comments on an earlier versions of the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

References

Abdellaoui, M. (2000). Parameter-Free elicitation of utility and probability weighting functions. *Management Science*, *46*(11), 1497–1512.

Alam, J., Georgalos, K., & Rolls, H. (2022). Risk preferences, gender effects and bayesian econometrics. *Journal of Economic Behavior & Organization*, *202* October, 168–183.

Allenby, G. M., & Rossi, P. E. (2006). Hierarchical Bayes Models. In R. Grover & M. Vriens (Ed.) *The Handbook of Marketing Research: Uses, Misuses, and Future Advances*. pp. 418–440 Sage Publishing Ltd.

Amador-Hidalgo, L., Brañas-Garza, P., Espín Martín, A. M., García Muñoz, T. M., & Hernández, A. (2021). Cognitive abilities and risk-taking: Errors, not preferences. *European Economic Review*, *134*, 103694.

Andersen, S., Harrison, G. W., Lau, M., & Rutström, E. (2010). Behavioral econometrics for psychologists. *Journal of Economic Psychology*, *31* 4, 553–576.

Andersson, O., Holm, H. J., Tyran, J. R., & Wengström, E. (2016). Risk aversion relates to cognitive ability: Preferences or noise?. *Journal of the European Economic Association*, *14*(5), 1129–1154.

Baillon, A., Bleichrodt, H., & Spinu, V. (2020). Searching for the reference point. *Management Science*, *66*(1), 93–112.

Balcombe, K. G., Bardsley, N., Dadzie, S., & Fraser, I. M. (2019). Estimating parametric loss aversion with prospect theory: Recognising and dealing with size dependence. *Journal of Economic Behaviour and Organisation*, *162* June, 106–119.

Balcombe, K. G., & Fraser, I. M. (2015). Parametric preference functionals under risk in the gain domain: A Bayesian analysis. *Journal of Risk and Uncertainty*, *50*, 161–187.

Balcombe, K. G., Fraser, I. M. (2024) "A note on an alternative approach to experimental design of lottery prospects," Memo. Available at MPRA.

Bernheim, B. D., Royer, R., & Sprenger, C. (2022). Robustness of rank independence in risky choice. *American Economic Review*, *112*, 415–420 Bernheim, B.D., Royer, R. and Sprenger, C., 2022, May. Robustness of rank independence in risky choice. In AEA Papers and Proceedings (Vol. 112, pp. 415-420). 2014 Broadway, Suite 305, Nashville, TN 37203: American Economic Association..

Bernheim, B. D., & Sprenger, C. (2020). On the empirical validity of cumulative prospect theory: Experimental evidence of rank-independent probability weighting. *Econometrica*, *88* 4, 1363–1409.

Bouchouicha, R., & Vieider, F. M. (2017). Accommodating stake effects under prospect theory. *Journal of Risk and Uncertainty*, *55*(1), 1–28.

Brown, A. L., Imai, T., Vieider, F. M., & Camerer, C. F. (2024). Meta-analysis of empirical estimates of loss aversion. *Journal of Economic Literature*, *62*(2), 485–516.

Bruhin, A., Fehr-Duda, H., & Epper, T. (2010). Risk and rationality: Uncovering heterogeneity in probability distortion. *Econometrica*, *78* 4, 1375–1412.

Buschena, D. E., & Atwood, J. A. (2011). Evaluation of similarity models for expected utility violations. *Journal of Econometrics*, *162* 1, 105–113.

Buschena, D. E., & Zilberman, D. (1999). Testing the effects of similarity on risky choice: Implications for violations of expected utility. *Theory and Decision*, *46* June, 251–276.

Chapman, J., Snowberg, E., Wang, S. W., & Camerer, C. (2024). Looming large or seeming small? Attitudes towards losses in a representative sample. *The Review of Economic Studies*, p.rdae093.

Enke, B., & Shubatt, C. (2023). *Quantifying lottery choice complexity*. National Bureau of Economic Research, No. w31677.

Etchart-Vincent, N., & L'Haridon, O. (2011). Monetary incentives in the loss domain and behaviour toward risk: An experimental comparison of three reward schemes including real losses. *Journal of Risk and Uncertainty*, *42*(1), 61–83.

Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., & Sunde, U. (2018). Global evidence on economic preferences. *The Quarterly Journal of Economics*, *133*(4), 1645–1692.

Fehr-Duda, H., Bruhin, A., Epper, T., & Schubert, R. (2010). Rationality on the rise: Why relative risk aversion increases with stake size. *Journal of Risk and Uncertainty*, *40*(2), 147–180.

Frydman, C., & Jin, L. J. (2022). Efficient coding and risky choice. *The Quarterly Journal of Economics*, *137*(1), 161–213.

Gao, X. S., Harrison, G. W., & Tchernis, R. (2023). Behavioral welfare economics and risk preferences: A Bayesian approach. *Experimental Economics*, *26*(2), 273–303.

Goldstein, W. M., & Einhorn, H. J. (1987). Expression theory & the preference reversal phenomena. *Psychological Review*, *94*(2), 236.

Gonzalez, R., & Wu, G. (1999). On the shape of the probability weighting function. *Cognitive Psychology*, *38* 1, 129–166.

Harrison, G. W., & Swarthout, T. Cumulative prospect theory in the laboratory: A reconsideration. Experimental Economics Center, Andrew Young School of Policy Studies, Georgia State University, Experimental Economics Center Working Paper Series 2016-04.

Hoffman, M. D., & Gelman, A. (2014). The No-U-Turn Sampler: Adaptively setting path lengths in hamiltonian monte carlo. *Journal of Machine Learning Research*, *15* 1, 1593–1623.

Holzmeister, F., & Stefan, M. (2021). The risk elicitation puzzle revisited: Across-methods (in) consistency?. *Experimental Economics*, *24* June, 593–616.

Johnson, R. J., Boyle, K. J., Adamowicz, W., Bennett, J., Brouwer, R., Cameron, T. A., Hanemann, W. M., Hanley, N., Ryan, M., Scarpa, R., Tourangeau, R., & Vossler, C. A. (2017). Contemporary guidance for stated preference studies. *Journal of the Association of Environmental and Resource Economists*, *4*(2), 319–405.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–291.

Kahneman, D., & Tversky, A. (1986). Rational choice and the framing of decisions. *Journal of Business*, *59*(4), 251–278.

Kobberling, V., & Wakker, P. P. (2005). An index of loss aversion. *Journal of Economic Theory*, *122*(1), 119–131.

Kpegli, Y. T., Corgnet, B., & Zylbersztejn, A. (2023). All at once! A comprehensive and tractable semi-parametric method to elicit prospect theory components. *Journal of Mathematical Economics*, *104*, 102790.

l'Haridon, O., & Vieider, F. M. (2019). All over the map: A worldwide comparison of risk preferences. *Quantitative Economics*, *10*(1), 185–215.

Loomes, G., & Pogrebna, G. (2017). Do preference reversals disappear when we allow for probabilistic choice?. *Management Science*, *63*(1), 166–184.

Murphy, R. O., & ten Brincke, R. H. (2018). Hierarchical maximum likelihood parameter estimation for cumulative prospect theory: Improving the reliability of individual risk parameter estimates. *Management Science*, *64*(1), 308–326.

Neilson, W., & Stowe, J. (2002). A further examination of cumulative prospect theory parameterizations. *Journal of Risk and Uncertainty*, *24* January, 31–46.

Nilsson, H., Rieskamp, J., & Wagenmakers, E. J. (2011). Hierarchical bayesian parameter estimation for cumulative prospect theory. *Journal of Mathematical Psychology*, *55*(1), 84–93.

Oprea, R. (2022). Simplicity equivalents. *Working Paper*, Available at: https://www.econ.queensu.ca/sites/econ.queensu.ca/files/Ryan%20Oprea%20Paper.pdf.

Oprea, R. (2024). Decisions under risk are decisions under complexity. *American Economic Review*, *114*(12), 3789–3811.

Pfeiffer, J., Meissner, M., Brandstatter, E., Riedl, R., Decker, R., & Rothlauf, F. (2014). On the influence of context-based complexity on information search patterns: An individual perspective. *Journal of Neuroscience, Psychology, and Economics*, *7*(2), 103–124.

Regier, D. A., Watson, V., Burnett, H., & Ungar, W. J. (2014). Task complexity and response certainty in discrete choice experiments: An application to drug treatments for juvenile idiopathic arthritis. *Journal of Behavioral and Experimental Economics*, *50* June, 40–49.

Rieger, M., Wang, M., & Hens, T. (2017). Estimating cumulative prospect theory parameters from an international survey. *Theory and Decision*, *82*(4), 567–596.

Rossi, P., Allenby, G., & McCulloch, R. (2005). *Bayesian statistics and marketing*. Wiley.

Saha, A. (1993). Expo-power utility: A 'flexible' form for absolute and relative risk aversion. *American Journal of Agricultural Economics*, *75*(4), 905–913.

Stott, H. P. (2006). Cumulative prospect theory's functional menagerie. *Journal of Risk and Uncertainty*, *32* March, 101–113.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*, 453–458.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*(4), 297–323.

Vehtari, A., Gelman, A., Simpson, D., Carpenter, B., & Bürkner, P. C. (2019) Rank-normalization, folding, and localization: An improved R-hat for assessing convergence of MCMC. arXiv preprint arXiv:1903.08008, arXiv preprint arXiv:1903.08008.

Vieider, F. M. (2024). Decisions under uncertainty as Bayesian inference on choice options. *Management Science 70*(12), 9014–9030.

https://doi.org/10.1017/eec.2025.10012 Published online by Cambridge University Press

Vieider, F. M., Lefebvre, M., Bouchouicha, R., Chmura, T., Hakimov, R., Krawczyk, M., & Martinsson, P. (2015). Common components of risk and uncertainty across contexts and domains: Evidence from 30 countries. *Journal of the European Economic Association*, *13*(3), 421–452.

Wakker, P. P. (2010). *Prospect Theory: For risk and ambiguity*. Cambridge University Press.

Wakker, P. P., & Zank, H. (2002). A simple preference foundation of cumulative prospect theory with power utility. *European Economic Review*, *46*(7), 1253–1271.

Walasek, L., Mullett, T. L., & Stewart, N. (2024). A meta-analysis of loss aversion in risky contexts. *Journal of Economic Psychology*, *103*, 102740.

Watanabe, S. (2013). WAIC and WBIC are information criteria for singular statistical model evaluation, *Proceedings of the Workshop on Information Theoretic Methods in Science and Engineering*, 90–94.

Zeisberger, S., Vrecko, D., & Langer, T. (2012). Measuring the time stability of Prospect Theory preferences. *Theory and Decision*, *72*(3), 359–386.

Zilker, V., & Pachur, T. (2022). Nonlinear probability weighting can reflect attentional biases in sequential sampling. *Psychological Review*, *129*(5), 949.

Zrill, L. (2024). Non-Parametric recoverability of preferences and choice prediction. *The Review of Economics and Statistics*, *106*(1), 217–229.

Cite this article: Balcombe, K., & Fraser, I. (2025). Flexible estimation of parametric prospect models using hierarchical bayesian methods. *Experimental Economics*, 1–25. https://doi.org/10.1017/eec.2025.10012