

# An Experimental Test of the Predictive Power of Dynamic Ambiguity Models

Konstantinos Georgalos\*

January 14, 2019

## Abstract

In this paper we report results from an economic experiment where we investigate the predictive performance of dynamic ambiguity models in the gains domain. Representing ambiguity with the aid of a transparent and non-manipulable device (a Bingo Blower) and using two-stage allocation questions, we gather data that allow us to estimate particular parametric forms of the various functionals and compare their relative performance in terms of *out-of-sample* fit. Our data show that a dynamic specification of Prospect Theory has the best predictive capacity, closely followed by Choquet Expected Utility, while multiple priors theories can predict choice only for a very restricted subset of our subjects.

*JEL classification:* C91, D81, D83, D90

*Keywords:* Ambiguity, Belief Updating, Dynamic Ambiguity Models, Hierarchical Bayes, Experiment

---

\*Department of Economics, Lancaster University Management School, LA1 4YX, Lancaster, U.K.  
✉ k.georgalos@lancaster.ac.uk, ☎ +44 (0)15245 93170.

I am grateful to Glenn Harrison, John Hey, Ivan Paya, Vitalie Spinu and Mike Tsionas for providing helpful comments. The author would like to thank the Editor of this journal W. Kip Viscusi and a referee for very helpful comments that led to significant improvements in both the analysis and the presentation. This research was funded by a Research and Impact Support Fund awarded by the Department of Economics at the University of York (RIS 39). The financial aid of the Greek Scholarships Foundation (IKY) is gratefully recognised. The usual disclaimer applies.

# 1 Introduction

Many everyday economic life activities involve decision making under some form of uncertainty, either objective (risk) or subjective (ambiguity). Moreover, decisions usually are characterised by a dynamic (sequential nature), in the sense that the decision maker needs to make a series of choices, while gradually she acquires additional, partial information about the state of nature, between choices, that allows her to update her prior beliefs. The standard theory of behaviour under uncertainty in a dynamic framework, Subjective Expected Utility (SEU), is based on three main assumptions regarding decision makers' behaviour: (1) they form well-defined subjective beliefs to describe ambiguous situations; (2) they are ambiguity neutral, in the sense that ambiguity has no impact on their choices and; (3) they update their subjective beliefs applying Bayes rule. Nevertheless, the thought experiments proposed in [Ellsberg \(1961\)](#) challenged the two first assumptions. The direct consequence of this was a growing strand of literature to emerge, aiming to theoretically model non-Expected Utility (see [Etner et al. \(2012\)](#) and [Machina and Siniscalchi \(2014\)](#) for a review of the theoretical models) along with the generation of considerable experimental evidence confirming the existence of non-neutral ambiguity attitudes (see [Hey \(2014\)](#) and [Trautmann and van de Kuilen \(2015\)](#) for a review of the experimental literature).

However, [Gilboa and Schmeidler \(1993\)](#) argue that in order for a model to be theoretically valid, it should be able to successfully cope with the dynamic aspect of the choices. Consequently, a complementary line of research appeared in the literature, trying to reconcile ambiguity non-neutral attitudes and conditional preferences. A common feature of all these non-Expected Utility models is that they relax in one way or another one of the principle axioms that [Savage \(1954\)](#) used in order to characterise preferences, the famous postulate  $P_2$ , also known as the *Sure-Thing Principle*. When one considers decisions in a dynamic framework, the Sure-Thing Principle is equivalent to two rationality axioms, *dynamic consistency* (the *ex-ante* preferences coincide with the *ex-post*) and *consequentialism* (only residual uncertainty and available outcomes matter). Furthermore, [Ghirardato \(2002\)](#) provides a well-established result, that when both axioms are simultaneously satisfied, preferences are represented by the Subjective Expected Utility model, and conditional beliefs are formed according to Bayes rule.

Consequently, violation of the Sure-Thing principle in a dynamic context implies violation of either the dynamic consistency or the consequentialism axiom, and thus deviations from Bayesian updating. As a result, and since there is still no consensus regarding which axiom should be adopted, two kinds of extensions of the static ambiguity models have been proposed in the literature to capture dynamic choice and conditional beliefs. Those that satisfy

consequentialism (Gilboa and Schmeidler (1993), Pires (2002), Wang (2003), Eichberger et al. (2007), Eichberger et al. (2010) and Siniscalchi (2011) among others) and those that satisfy dynamic consistency (Epstein and Schneider (2003), Klibanoff and Hanany (2007), Hanany and Klibanoff (2009) and Klibanoff et al. (2009) among others)<sup>1</sup>.

It is however interesting, the fact that while the experimental literature comparing the various theoretical models of choice under ambiguity in an atemporal framework is quite developed (see section 2), there is a shortage of empirical evidence comparing dynamic ambiguity models. This is a gap that the present study aspires to bridge. To this end, we report the results from an economic experiment where using a sequential decision task, we compare dynamic ambiguity models.

Representing ambiguity with the aid of a transparent and non-manipulable device, a Bingo Blower and using two-stage allocation questions with an interim stage that allows for updating of beliefs, we are able to compare the predictive (out-of-sample fit) power of various dynamic ambiguity models. The predictive performance of the models is based on structural parameters which are estimated for each model and for each subject, using Hierarchical Bayes estimation techniques on a subset of the data. The out-of-sample fit comparison, tests the ability that each specification has to predict behaviour (allocations) in the subset of the data that has not been included during the estimation of the parameters. We can then rank models based on their relevant predictive power.

We therefore compare the predictive power of the dynamic version of the most commonly used models, including the Savage (1954) Subjective Expected Utility (SEU), which we use as a benchmark model, the Gilboa and Schmeidler (1989) Maxmin Expected Utility (MEU), the Ghirardato et al. (2004)  $\alpha$ -Maxmin Expected Utility ( $\alpha$ -MEU), the Schmeidler (1989) Choquet Expected Utility (CEU), the Chateauneuf et al. (2007) CEU model with neo-additive capacities (CEU<sub>NEO</sub>) and the Tversky and Kahneman (1992) Cumulative Prospect Theory (PT). For all the aforementioned models (with the exception of PT), well-defined update rules have been proposed in the literature in order to define conditional preferences. Hence, our analysis allows us to conduct comparisons across different theoretical models<sup>2</sup>, as well as across different update rules within a particular theoretical model. Adopting the results from the recent experimental literature, in our analysis we do not consider models that preserve the assumption of dynamic consistency. Dominiak et al. (2012) reported extensive violations of the dynamic consistency axiom. Furthermore, our data show that for at least 70% of the observed choices, subjects

---

<sup>1</sup> Siniscalchi (2011) and Machina and Siniscalchi (2014) provide extensive reviews of the theoretical literature of dynamic ambiguity models.

<sup>2</sup>We analytically present all the rules we consider in section 4.

behave in a dynamic inconsistent way and therefore such models would poorly describe or predict behaviour.

To the best of our knowledge, this is the first study to experimentally test the predictive power of dynamic models under ambiguity. Our contribution can be summarised as follows: (1) we provide the first experimental test to compare the various ambiguity models in a dynamic framework; (2) we provide the first experimental comparison on the various rules that have been proposed for the updating of Choquet capacities as well as neo-additive Choquet capacities; (3) we perform the first empirical test of Prospect Theory in a dynamic framework under ambiguity and; (4) we use the estimates obtained by the fitted models in order to test the predictive power of the various theories. The following are our key results: (1) the rank dependent family of models (PT and CEU) performs much better compared to the multi-priors model (MEU and  $\alpha$ -MEU) in terms of their predictive capacity; (2) Prospect Theory wins all other models as far as out-of-sample fit is concerned; (3) within the Choquet Expected Utility model, the Generalised Bayesian updating rule best explains behaviour while the Dempster-Shaffer rule performs better within the neo-additive Choquet capacities specification; (4) within the Prospect Theory family, the specification assuming a [Prelec \(1998\)](#) 2-parameter weighting function provides the best fit and; (5) Choquet Expected Utility and Prospect Theory are better in predicting behaviour compared to the competing models.

## 2 Related Literature

As the objective of our study is to compare dynamic models of ambiguity, we split this section in two parts, one where we survey studies that compare atemporal ambiguity models and one where we review the experimental work on dynamic choice under ambiguity. One of the first and most comprehensive studies to compare ambiguity models, in a static framework, is [Hey et al. \(2010\)](#) who investigate the descriptive and predictive power of several models. They find evidence in favour of the multiple prior models and show that a more complicated theoretical structure does not improve the capacity of the models. [Kothiyal et al. \(2014\)](#) re-analyse the data of [Hey et al. \(2010\)](#) including a version of Prospect Theory that has not been used in the analysis of the original study. They find that Prospect Theory provides the highest predictive power against the CEU model, as well as against the multiple prior ones. [Hayashi and Wada \(2010\)](#) also report evidence against the multiple prior models (MEU and  $\alpha$ -MEU). [Hey and Pace \(2014\)](#) using an allocation decision task, compare the explanatory and predictive ability for the most commonly used ambiguity models. They find supporting evidence for the Vector Expected Utility model and for SEU and to lower degree for  $\alpha$ -MEU and CEU.

Ahn et al. (2014), using a portfolio choice problem, compare two main families of models, the kinked (multiple prior models) and the smooth, reporting evidence in favour of the former. On the contrary, Cubitt et al. (2014) investigating preferences for randomised acts, compare two models, namely the  $\alpha$ -MEU and the smooth model, finding supporting evidence for the latter. Recently, Baillon and Bleichrodt (2015) compare the in-sample capacity of various ambiguity models, in both the gains and the losses domain, using matching probabilities to elicit beliefs. They find supporting evidence for Prospect Theory and to a lesser extent for  $\alpha$ -MEU and CEU.

The experimental literature on dynamic choice under ambiguity, on the other hand, is quite limited. Cohen et al. (2000) and Dominiak et al. (2012), both investigate behaviour in the dynamic, 3-colour Ellsberg paradox, using a similar experimental design. The former study assumes multiple prior preferences and investigates which updating rule, between the two that have been proposed for this family of models (Maximum Likelihood Updating and Full Bayesian Updating) best explains behaviour. The authors assume that participants satisfy consequentialism and their analysis provides support to the Full Bayesian updating rule. Dominiak et al. (2012) relax the assumption of consequentialism and they indeed study whether subjects behave according to dynamic consistency or consequentialism. They report extensive violation of dynamic consistency and they also replicate the dominance of the Full Bayesian rule. Corgnat et al. (2013) in an experimental asset market under ambiguity, study how traders react when dividend information is sequentially revealed. They find that the role of ambiguity cannot explain financial anomalies. However, they do not consider any particular preference functional or updating rule in their analysis. Two recent studies Georgalos (2016) and De Filippis et al. (2016) report extensive violations of Bayesian updating in different contexts. The former investigates heterogeneity in the planning strategies of decision makers, while the latter studies behaviour in a social learning experiment. Both assume some kind of multiple prior preferences along with their corresponding updating rules<sup>3</sup>.

It is difficult to directly compare our study with the aforementioned experimental work, since none of these studies performs a comparison between competing theories and updating rules, as they either assume a specific preference functional to represent preferences (e.g. MEU) or they do not assume any preferences at all. As a consequence, the only comparison that has been done to date, is between the two updating rules that have been proposed for the MEU model, namely the Full Bayesian and the Maximum Likelihood updating rule (see subsection 4.2). Our analysis differs from these studies in that using structural econometric techniques,

---

<sup>3</sup>Two recent studies (Baillon et al. (2017) and Peysakhovich and Karmarkar (2016)) investigate how additional information affects choice in a decision task under ambiguity. Nevertheless, none of these studies explicitly considers updating of ambiguous beliefs.

we consider a battery of preference functionals and updating rules, we test the predictive capacity of these models, we represent ambiguity using a Bingo Blower, thus deviating from the standard Ellsberg-type experiments, we ask a large number of questions per subject and finally we consider heterogeneity, both within and between subjects (see section 5).

### 3 Experimental Design

Our experimental design is an extension of the design used in [Hey and Pace \(2014\)](#) to its dynamic dimension. As was previously noted, we represented ambiguity using a transparent device, a Bingo Blower<sup>4</sup>, and we used a set of two-stage allocation questions. In our experimental protocol, there were three payoff relevant, mutually exclusive states of the world. Inside the transparent Bingo Blower, there were 20 balls in total, of three different colours (pink, blue and yellow), representing the three potential states of the world. The balls were in continuous motion and the actual composition of the Bingo Blower consisted of 4 blue (20%), 6 pink (30%) and 10 yellow (50%) balls out of the total 20. The advantage of this device is that although there exist objective probabilities, known only to the experimenters, the subjects lack this information and they somehow form subjective beliefs, regarding the probabilities of the three states of the world, that may not be known to them in a conscious way. That is, while they can observe that there is at least one ball of each colour (lower bound probability), it is almost impossible to be able to identify the exact composition of the Bingo Blower, generating in this way genuine ambiguity. During the experiment, the subjects could observe the physical Bingo Blower placed in the middle of the lab, as often as they wished, and they could also consult live streaming of the blower projected in two large screens in the front of the lab.

We asked subjects a series of two-stage allocation questions. Each allocation question consists of an experimental income  $m$ , expressed in tokens, and an exchange rate  $e_s$  for each of the possible states of the world, with  $s \in \{blue, pink, yellow\}$ . At the first stage, the subjects were asked to allocate their experimental income between the three colours, knowing that at the end of the second stage, one ball would be extracted from the Bingo Blower, and the payoff would be the product between the tokens allocated to this colour and its respective exchange rate. Then, there was an intermediate stage, where the subjects would obtain partial information that a ball was extracted, but they wouldn't learn the exact colour of the ball. For example, the subjects would learn that "the ball is not blue". Then, all the tokens allocated to the blue colour were lost and the subject was given the opportunity to make a new allocation, using

---

<sup>4</sup>A similar Bingo cage has been used by [Andersen et al. \(2012\)](#) and the Bingo Blower has been used by [Hey et al. \(2010\)](#) and [Hey and Pace \(2014\)](#), all in *static* choice problems.

the remaining experimental income, to the two available colours, pink and yellow<sup>5</sup>. This was repeated for a total of sixty, independent allocation questions, where each question included a different experimental income ranging from 9 to 110 tokens, and different exchange rates for the three colours, ranging from 0.1 to 1.8.<sup>6</sup> All subjects faced the same set of allocation questions, presented in a randomised order for each subject, in an effort to eliminate potential order effects. The subjects realised their allocations, using sliders that were interconnected with each other. Consequently, subjects were forced to spend their whole experimental income to the three colours and there was no possibility to invest on a safe asset (keep money on hand), save for the next round or borrow. All the allocations were recorded in integer values in an effort to reduce noise in the data that could have been caused due to difficulties that the subjects might have while operating the software.

Since we are interested in the updating process of subjective beliefs, we needed a way to elicit these beliefs. Rather than asking subjects to state their beliefs, using some kind of scoring rule or matching probabilities, we infer those, based on the *revealed* allocations of the subjects. Asking subjects two-stage allocation questions, allows to observe both the *ex-ante* and the *ex-post* allocations. Generally, it is claimed in the literature, that data from allocation tasks can be richer regarding the provided information, compared to discrete choice data. Allocation type questions were pioneered in Loomes (1991) and have been afterwards exploited in various contexts<sup>7</sup>. Asking subjects to make decisions in a continuous action space, allows us to estimate beliefs (and therefore test updating rules) using a structural econometric model (see section 5).

In total, fifty-eight subjects (52% females) were recruited from a standard student experimental population, using the ORSEE recruitment system (Greiner (2015)). The experiment was conducted at the EXEC lab at the University of York (U.K.). The experiment lasted for

---

<sup>5</sup>As was mentioned before, our design is an extension of Hey and Pace (2014) to its dynamic dimension. Nevertheless, apart for the difference regarding the time of periods of each allocation task, there is another fundamental difference that we should mention. Hey and Pace (2014) use two types of problems. In type 1 they ask subjects to allocate experimental income between two of the colours (implying a zero allocation to the third colour). In type 2, they asked subjects to allocate the tokens between one of the three colours and the other two. These type of questions reduce the problem to a two-way allocation problem. Since we need to introduce an interim stage, and compare allocations *ex-ante* and *ex-post*, we rather ask subjects to explicitly make allocations between the three colours.

<sup>6</sup>The set of questions was chosen after extensive Monte Carlo simulations that were conducted in order to verify both the possibility to discriminate between different theoretical models, and also to confirm that it is possible to recover the value of the underlying behavioural parameters that were used for these simulations (see Wilcox (2007) and Wilcox (2008)). We deliberately varied both the experimental income and the exchange rates in an effort to increase the identification power and robustness of our econometric estimation. Varying the income allows a better identification of the shape of the utility function, while varying the exchange rates, ensures that the ranking between outcomes is not the same for all the problems, which can potentially provide more information regarding the subjects' preferences. All the relevant information was provided via the experimental interface to ensure that confusion is minimised.

<sup>7</sup>See Choi et al. (2007), Charness and Gneezy (2010), Hey and Panaccione (2011), Ahn et al. (2014), Hey and Pace (2014), Loomes and Pogrebna (2014). See Loomes and Pogrebna (2014) for an extensive discussion on the allocation procedure.



less than one hour and the subjects were paid privately and in cash directly after the end of the experiment. Subjects gained on average £14.16 including a *show-up* fee of £2.5. The maximum payment was £25.5. We collected the data using a custom developed for the experiment software, written using the Python programming language.<sup>8</sup>

Subjects were paid according to a random incentive mechanism, where one of the allocation questions was played for real. In order to prevail subjects from learning the actual composition of the Bingo Blower<sup>9</sup>, we informed participants, that no actual draws will take place during the experiment, but instead, hypothetical draws, and subsequently hypothetical partial information would be announced to subjects<sup>10</sup>. A draw was realised individually for each subject at the end of the experiment to determine the payoffs, in a randomly chosen question. If for example, for the allocation question that was played for real, the information “the ball is not blue” was revealed during the experiment, the subject would continue drawing balls, till an either pink or yellow ball was expelled. This ball defined the “winning” colour, and the subject was paid the amount of tokens that she had allocated to this colour multiplied by the corresponding exchange rate. Recent theoretical studies (Oechssler and Roomets (2014), Bade (2015), Baillon et al. (2015)) challenge the incentive compatibility of the random incentive mechanism, arguing that an ambiguity averse agent may not reveal her true preferences in an effort to hedge against ambiguity. Nevertheless, this result is based on a series of assumptions, such that ambiguity is represented by the Ellsberg urn, the choice task is discrete and the subjects know *ex-ante* what the payoffs will be. However, it is not straightforward how this could affect our design, since in our experiment, there was no way for subjects to contemplate what the next allocation question will be, and therefore there was no space for hedging.

## 4 Theoretical Models Under Investigation

In this section we present all the preference functionals that we include in our analysis. All the models that we consider share a common feature, that is, they separate tastes (utility) from beliefs. Consequently, they differ in two main characteristics, the representation of beliefs and

---

<sup>8</sup>Python Software Foundation. Python Language Reference, version 2.7. Available at <http://www.python.org>.

<sup>9</sup>Conducting actual draws from the Bingo Blower, would change the nature of the research question, as in this case, subjects would continuously obtain information on the actual distribution and instead of updating, we would need to focus on learning models under ambiguity (see Marinacci (2002), Epstein and Schneider (2007), Epstein et al. (2010)). We leave this for future research.

<sup>10</sup>The computer was programmed to draw i.i.d. virtual balls from a uniform distribution. Then it would announce that the ball is not one of the two remaining colours with equal chances. For instance, if a ball was blue, it would announce that the ball is not pink with probability 50% and not yellow with the residual probability. In a pilot session we implemented virtual draws from the actual distribution. Nevertheless, this could allow subjects to *learn* the actual composition based on the empirical distribution. To avoid this kind of effects, we resorted to the uniform distribution. The information that the subjects received on average was “not blue” 34.54%, “not yellow” 34.36% and “not pink” 31.09%.



the way that these beliefs are updated in the light of new information. For all the models, we follow [Al-Najjar and Weistein \(2009\)](#) and we assume that our subjects perform *fact based* updating<sup>11</sup>. This assumption is in line with the available experimental evidence ([Dominiak et al. \(2012\)](#) and the present study, where subjects extensively violate dynamic consistency.<sup>12</sup> In our experimental design, there are three mutually exclusive potential states of nature and the decision maker faces ambiguity regarding which state will happen. We represent this ambiguity with a state-space  $S$ , and we define subsets  $E$  of  $S$  to represent events. In our experiment there are potentially three events (the ball is either blue, yellow or pink) and we denote these events as  $E_b, E_y$  and  $E_p$ . We also define the three possible unions between two events. For instance,  $E_b \cup E_y$ , indicates the event where the ball is either blue or yellow. To keep the analysis aligned to the spirit of our experimental design, the latter union event can be also written as  $E_{\neg p}$ , implying that the ball is not pink. Finally, in all the subsequent preference functionals, utility is assumed to be represented by a monotonic, concave and strictly increasing utility function  $u(\cdot)$ , the shape of which we specify in section 5.

In all the preference functionals, subjects are assumed to maximise

$$w(E_b) u(e_b \times x_b) + w(E_y) u(e_y \times x_y) + w(E_p) u(e_p \times x_p) \quad (1)$$

$$\text{s.t. } x_b + x_y + x_p = m \quad (2)$$

at the first stage, where  $w(E_s)$  is the subjective decision weight attached to state  $s$ ,  $e_s$  is the exchange rate of state  $s$  and  $x_s$  is the fraction of the experimental income  $m$  that is allocated to state  $s$ . The solution of this optimisation program provides the first stage optimal demands which are denoted as the allocation vector  $x^* = (x_b^*, x_y^*, x_p^*)$ . Then, given that the subjects obtain some form of partial information (e.g.  $E = E_{\neg y}$ ), they maximise

$$w'(E_b) u(e_b \times x'_b) + w'(E_p) u(e_p \times x'_p) \quad (3)$$

$$\text{s.t. } x'_b + x'_p = m - x_y^* \quad (4)$$

where  $w'(E_s)$  denotes the updated decision weights  $w(E_b|E_{\neg y})$  and  $w(E_p|E_{\neg y})$  and  $x'_i$  the conditional demands. The first order conditions of Equation 1 require the following equality to

---

<sup>11</sup>[Al-Najjar and Weistein \(2009\)](#) note that fact-based updating consists of two parts: (1) when comparing two actions  $f$  and  $g$ , given the event  $E$ , the conditional preference  $\succsim_E$  places no weight on the consequences of these acts at states that are now excluded and; (2) the conditional preference depends only on the information set  $E$  and not on inconsequential aspects of the decision problem (e.g. *ex-ante* optimal plan). As a result, we exclude from our analysis, all these axiomatic models that retain dynamic consistency and reject consequentialism (e.g. the [Klibanoff et al. \(2005\)](#) smooth ambiguity model).

<sup>12</sup>In our allocation framework, violation of dynamic consistency means different *ex-ante* and *ex-post* allocations to a state  $s$ . We have in total data from 3480 2-stage allocations. Out of these allocations, 72% violate dynamic consistency for at least 1 token, 51% for at least 2 tokens and 41% for at least three tokens. These percentages are independent of the kind of information that was revealed to the subjects.

hold

$$w(E_b) \frac{\partial u(e_b \times x_b)}{\partial x_b} = w(E_y) \frac{\partial u(e_y \times x_y)}{\partial x_y} = w(E_p) \frac{\partial u(e_p \times x_p)}{\partial x_p} \quad (5)$$

Similarly, in the second stage, the optimisation conditions require that

$$w'(E_b) \frac{\partial u(e_b \times x'_b)}{\partial x'_b} = w'(E_p) \frac{\partial u(e_p \times x'_p)}{\partial x'_p} \quad (6)$$

Since each of the theoretical models we consider characterises these decision weights in a different way, we devote the rest of the section to describe how these weights are represented, as well as the way that updating takes place for each of the preference functionals.

#### 4.1 Subjective Expected Utility

The benchmark model *Subjective Expected Utility*, assumes that there exists a probability measure  $P$  over  $S$  representing beliefs, such that  $\pi(s) = P(E_s)$  for  $s \in \{b, y, p\}$  and  $\sum \pi(s) = 1$ . In Equation 1,  $w(s)$  is substituted by  $\pi(s)$ . Upon the arrival of information, these subjective beliefs are updated according to the Bayes rule. So for instance, given the information  $E = E_{-y}$ , the conditional beliefs in Equation 3 for outcome 1 are given by

$$w(E_b|E_{-y}) = \frac{P(E_b)}{P(E_{-y})} = \frac{P(E_b)}{P(E_b \cup E_p)} = \frac{\pi(b)}{\pi(b) + \pi(p)}$$

and for outcome 3 by the residual probability  $1 - w(E_b|E_{-y})$ . Substituting these conditional beliefs to Equation 6, it is straightforward to show that the allocation at the first stage should coincide with the conditional allocation, satisfying both dynamic consistency and consequentialism.<sup>13</sup>

#### 4.2 Multiple Prior Theories: $\alpha$ -Maxmin & Maxmin Expected Utility

In *multiple* prior theories, ambiguity is captured by a closed and convex set  $C$  of priors which is a subset of all the possible probability measures  $P$  over the state space  $S$ . In the  $\alpha$ -MEU model, the objective function to maximise can be written as a convex combination between the maximum and the minimum utility that a decision maker can obtain. Denoting by  $z_s$  the

---

<sup>13</sup>The first order conditions, at the first stage, require that  $\pi(b) \frac{\partial u(e_b \times x_b)}{\partial x_b} = \pi(p) \frac{\partial u(e_p \times x_p)}{\partial x_p}$ . Substituting the conditional probabilities to Equation 6, gives  $\frac{\pi(b)}{\pi(b) + \pi(p)} \frac{\partial u(e_b \times x'_b)}{\partial x'_b} = \frac{\pi(p)}{\pi(b) + \pi(p)} \frac{\partial u(e_p \times x'_p)}{\partial x'_p}$  which simplifies to  $\pi(b) \frac{\partial u(e_b \times x'_b)}{\partial x'_b} = \pi(p) \frac{\partial u(e_p \times x'_p)}{\partial x'_p}$ . Therefore, the solution of the maximisation problems will lead to the same optimal allocation and consequently, the unconditional and conditional allocations will coincide ( $x_s^* = x_s^{j*}$ ), satisfying dynamic consistency. Since for the rest of the theoretical models that we consider, beliefs are updated in a non-Bayesian way, the above condition will be occasionally violated, resulting to violations of dynamic consistency, and consequently to different *ex-ante* and *ex-post* allocations.

payoff  $e_s \times x_s$  of the agent, the optimisation problem is given by

$$\alpha \min_{P \in \mathcal{C}} \left[ \sum_s p(s) u(z_s) \right] + (1 - \alpha) \max_{P \in \mathcal{C}} \left[ \sum_s p(s) u(z_s) \right] \quad (7)$$

with  $P = \{p(s) : p(s) \geq \underline{p}(s)\}$ , with  $s \in \{b, y, p\}$  and  $\sum_s \underline{p}(s) < 1$ . This representation requires some form of specification regarding the non-additive lower bounds  $\underline{p}(s)$  in order to be able to fit parametric models. We adopt the representation first proposed in [Hey et al. \(2010\)](#) and then broadly used in the literature (see [Hey and Pace \(2014\)](#), [Kothiyal et al. \(2014\)](#), [Burghart et al. \(2015\)](#)) where beliefs for the three states of the world are represented in a 2-dimension simplex, known from the decision theory literature under risk as the Marschak-Machina Triangle (MMT). The horizontal axis in [Figure 1](#) represents the probability for state  $b$  while the vertical axis for state  $p$ . This kind of models, depend on the relative ranking between outcome payoffs. Assuming the ranking  $z_b > z_y > z_p$  the minimum expected utility is obtained when the probability for the best outcome is minimised (or the probability of the worst outcome is maximised). This happens at point 3 where the decision maker attaches probabilities  $\underline{\pi}(b), \underline{\pi}(y), 1 - \underline{\pi}(b) - \underline{\pi}(y)$  to the three states of the world respectively. In the same spirit, the maximum expected utility is obtained at the point where the probability for the best outcome is maximised (and for the worst is minimised), which happens at point 1. At this point, the subject attaches probabilities  $1 - \underline{\pi}(y) - \underline{\pi}(p), \underline{\pi}(y), \underline{\pi}(p)$  to the three states. Since the decision maker maximises a convex combination of the minimum and the maximum utility, it is easy to show that the weights will be a convex combination between the lower and the upper bound probability for each state. Therefore, the weights for [Equation 1](#) are given by

$$\begin{aligned} w(E_b) &= \alpha \underline{\pi}(b) + (1 - \alpha)(1 - \underline{\pi}(y) - \underline{\pi}(p)) \\ w(E_y) &= \alpha \underline{\pi}(y) + (1 - \alpha) \underline{\pi}(y) \\ w(E_p) &= \alpha(1 - \underline{\pi}(b) - \underline{\pi}(y)) + (1 - \alpha) \underline{\pi}(p) \end{aligned}$$

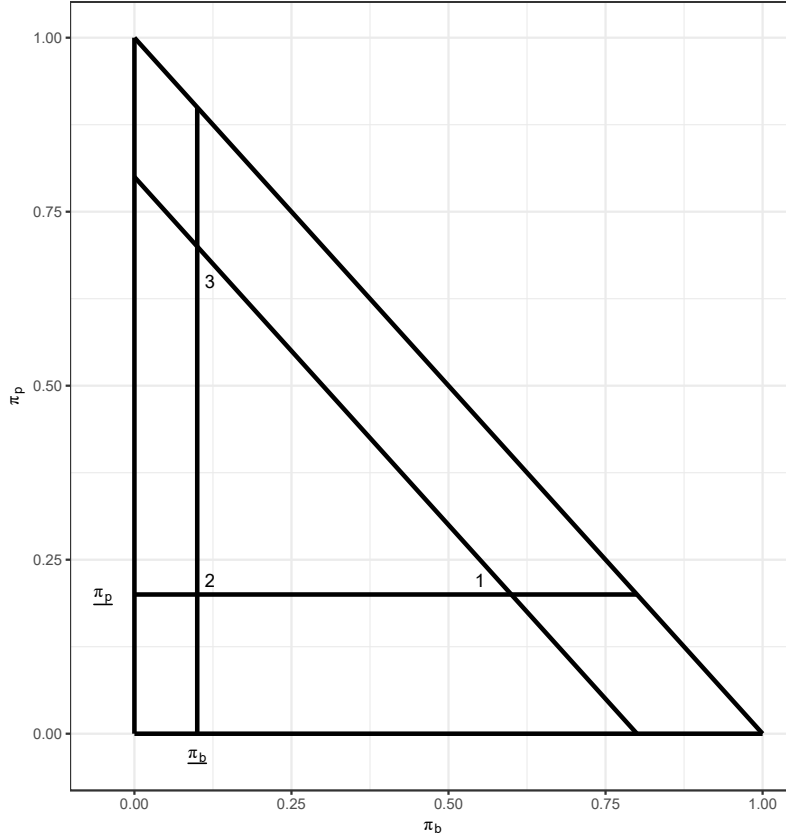


Figure 1: Prior Beliefs

In the second stage, the subject updates her beliefs conditional to the provided partial information. In the literature there have been suggested two types of updating rules for the multiple prior family, those that satisfy *consequentialism* (Gilboa and Schmeidler (1993), Pires (2002)) and those that satisfy *dynamic consistency* (Epstein and Schneider (2003), Klibanoff et al. (2009)). The two most prominent rules of the former category, those that we are using in the present study, are the *Maximum Likelihood Updating* rule (MLU) and the *Full Bayesian Updating* (FBU). The former rule dictates to the decision maker to update only those priors that maximise the probability of the conditional event while according to the latter, all the sets are updated in a prior by prior way. In both cases, all priors are updated according to the Bayesian rule. For both rules, the priors that generate the minimum and the maximum expected utilities are updated to generate the conditional weights. For instance, consider the partial information  $E = E_{\neg y}$  and the conditional ranking of outcomes  $z_b > z_p$ . For the MLU rule, the probability of the event  $E_{\neg y}$  is maximised at points 1 and 3 in Figure 1 and is equal to  $1 - \underline{\pi}(y)$ . Then, given the ranking of the outcomes, the conditional minimum expected utility is given when the priors of point 3 are used while the maximum expected utility is obtained at point 1. Thus,

the conditional weight attached to event 1 is given by

$$w(E_b|E_{-y}) = \alpha \frac{\pi(b)}{1 - \pi(y)} + (1 - \alpha) \frac{1 - \pi(y) - \pi(p)}{1 - \pi(y)}$$

and for the worst outcome, weight equal to  $1 - w(E_b|E_{-y})$  is attached.

Note that in our context, with three ambiguous outcomes, the predictions of both the MLU and the FBU rules are behaviourally *indistinguishable*<sup>14</sup>. According to FBU, the subject updates all the available priors and uses those posteriors that generate the minimum and maximum expected utility. In practice, in our example the subject updates priors at points 1 and 3 in Figure 1 as was the case in MLU, and also updates beliefs at point 2. Nevertheless, [Georgalos \(2016\)](#) shows that given the ranking  $z_b > z_p$ , posteriors at 2 can generate neither the minimum nor the maximum expected utility. Consequently, these values are obtained at points 3 and 1 respectively, which leads to the same predictions as the MLU rule.

We also consider the MEU model which is a special case of the  $\alpha$ -MEU when  $\alpha=1$ . Therefore, the same logic applies in calculating the minimum expected utility, and the weighting formulas presented above are used to optimise Equations 1 and 3, since the coefficient attached to the maximum utility is now equal to 0.

### 4.3 Choquet Expected Utility

In the Choquet Expected Utility model, beliefs are represented by non-additive probability measures (capacities)  $v$  that should satisfy the normalisation conditions  $v(\emptyset) = 0, v(S) = 1$  as well as monotonicity where for  $A \subset B$  it must hold  $v(A) \leq v(B)$ . The decision weights strictly depend on the relative ranking of the various outcomes so first we need to order the payoffs in a descending way. For example, for the ranking  $z_b > z_y > z_p$ , and according to the definition of the Choquet integral, the respective decisions weights are given by

$$\begin{aligned} w(E_b) &= v(E_b) \\ w(E_y) &= v(E_b \cup E_y) - v(E_b) \\ w(E_p) &= 1 - v(E_b \cup E_y) \end{aligned}$$

so the subject maximises Equation 1 given these weights. Regarding the conditional state, in the literature there have been suggested various ways to update non-additive capacities. [Eichberger et al. \(2007\)](#) and [Eichberger et al. \(2010\)](#) summarise the three most commonly used updating rules, namely the *optimistic* updating rule ([Gilboa and Schmeidler, 1993](#)), the *Dempster-Shafer* rule ([Dempster \(1968\)](#), [Shafer \(1976\)](#)) and the Generalised Bayesian Updating

<sup>14</sup>However, this is not the case in the 3-colour Ellsberg urn with one risky and two ambiguous states where MLU and FBU lead to distinct predictions.

rule (Dempster, 1967). In their general form, the three updating rules for a state  $A$  given the event  $E$  are given by the following formulas

$$\begin{aligned} v^{OPT}(A|E) &= \frac{v(A \cap E)}{v(E)} \\ v^{DS}(A|E) &= \frac{v(A \cup E^c) - v(E^c)}{1 - v(E^c)} \\ v^{GB}(A|E) &= \frac{v(A \cap E)}{v(A \cap E) + 1 - v(A \cup E^c)} \end{aligned}$$

where  $E^c$  denotes the complement of  $E$ . Applying this to our framework and using the example that the available information is that  $E = E_{-y}$ , and the ranking  $z_b > z_p$ , the updated capacity for  $E_b$  for each of the updating rules is given by

$$\begin{aligned} v^{OPT}(E_b|E_{-y}) &= \frac{v(E_b \cap E_{-y})}{v(E_{-y})} = \frac{v(E_b)}{v(E_b \cup E_p)} \\ v^{DS}(E_b|E_{-y}) &= \frac{v(E_b \cap E_{-y}^c) - v(E_{-y}^c)}{1 - v(E_{-y}^c)} = \frac{v(E_b \cup E_y) - v(E_y)}{1 - v(E_y)} \\ v^{GB}(E_b|E_{-y}) &= \frac{v(E_b \cup E_{-y})}{v(E_b \cap E_{-y}) + 1 - v(E_b \cup E_{-y}^c)} = \frac{v(E_b \cap (E_b \cup E_p))}{v(E_b \cap (E_b \cup E_p)) + 1 - v(E_b \cup E_y)} \\ &= \frac{v(E_b)}{v(E_b) + 1 - v(E_b \cup E_y)} \end{aligned}$$

Notice that all the above formulas collapse to the Bayes rule  $\frac{v(E_b)}{v(E_b \cup E_p)}$  whenever the capacities are additive. In the second stage and given that the initial ranking of outcomes holds (i.e.  $z_b > z_p$ ) the decision maker is attaching weight  $v^l(E_b|E_{-y})$  to the best outcome and  $1 - v^l(E_b|E_{-y})$  to the worst, in order to maximise Equation 3, with  $l \in \{OPT, DS, GB\}$ . Eichberger et al. (2010) summarise the behavioural characteristics of the three update rules as follows. The optimistic rule puts more weight to the good outcome while the Dempster-Shafer rule adjusts its weight on the bad outcome. On the contrary, the Generalised Bayesian rule balances the weight between the good and the bad outcome. However, these updating rules of the CEU models are based on a representation that requires uniform ambiguity attitudes. In the next subsection we present a model that accommodates both optimistic and pessimistic attitudes towards ambiguity.

#### 4.4 Neo-additive Choquet Expected Utility

Chateauneuf et al. (2007) propose a generalisation of the SEU model that can accommodate both optimistic and pessimistic attitudes towards ambiguity, the so called CEU with *neo-additive*

capacities. A neo-additive capacity is defined as a convex combination between an additive capacity and a special capacity (Hurwicz capacity) that distinguishes between whether an event is impossible, possible or certain. In its general form, for a given set of null events  $\mathcal{N} \subset S$ , the capacity consists of a 3-tuple  $v(E_i) = v(\pi(i), \delta, \alpha)$ , with  $\delta, \alpha \in [0, 1]$  such that

$$v(E_i) = \begin{cases} 0 & \text{if } E_i \in \mathcal{N} \\ \delta\alpha + (1 - \delta)\pi(E_i) & \text{if } E_i \notin \mathcal{N} \text{ and } S \setminus E_i \notin \mathcal{N} \\ 1 & \text{if } S \setminus E_i \in \mathcal{N} \end{cases}$$

Applying the notion of neo-additive capacities to our three-way allocation problem, the weights for the three outcomes in Equation 1, given the ranking  $z_b > z_y > z_p$ , are given by

$$\begin{aligned} w(E_b) &= \alpha\delta + (1 - \delta)\pi(b) \\ w(E_y) &= \alpha\delta + (1 - \delta)\pi(b \cup y) - \alpha\delta - (1 - \delta)\pi(b) \\ &= (1 - \delta)\pi(y) \\ w(E_p) &= 1 - \alpha\delta - (1 - \delta)\pi(b \cup y) \\ &= 1 - \alpha\delta - (1 - \delta)(1 - \pi(p)) \\ &= \delta(1 - \alpha) + (1 - \delta)\pi(p) \end{aligned}$$

Substituting these weights to Equation 1, the objective function to maximise can be rewritten as

$$(1 - \delta) \sum_s \pi(s)u(z_s) + \delta(\alpha u(z_b) + (1 - \alpha)u(z_p)) \quad (8)$$

which is a convex combination between the expected utility of the three outcomes and a weighted sum between the utility received by the best and the worst outcome. The parameter  $(1 - \delta)$  expresses the degree of *confidence* that the decision maker has to her subjective probability distribution over outcomes (likewise the value of  $\delta$  expresses the lack of confidence) while the parameter  $\alpha$  is an indicator of the relative degree of *optimism* which is put to the best outcome while the residual weight  $(1 - \alpha)$  is assigned to the worst outcome. For a particular combination of different values of  $\delta$  and  $\alpha$ , Equation 8 collapses to several standard decision criteria. When  $\delta = 0$ , the representation becomes equivalent to SEU. When  $\delta > 0$  and  $\alpha = 0$  pure pessimism is expressed while when  $\alpha = 1$  pure optimism. Finally, when  $\delta = 1$  and  $\alpha \in (0, 1)$ , Equation 8 becomes the *Hurwicz* criterion where only the utility values of the best and the worst outcome are essential for the choice.

At the second stage, the subject updates her neo-additive capacities in order to assign weights to the available outcomes. More specifically, [Eichberger et al. \(2010, Proposition 1\)](#) prove that for every  $\pi(E_i) > 0$ , every neo-additive capacity that is updated applying one of the three rules from subsection 4.3, the conditional capacity is also neo-additive. For the event  $E = E_{-y}$



and  $z_b > z_p$ , the conditional neo-additive capacities are given by

$$\begin{aligned} v^{OPT}(E_b|E_{-y}) &= \frac{v(E_b \cap E_{-y})}{v(E_{-y})} = \frac{(1-\delta)\pi(E_b \cap E_{-y}) + \alpha\delta}{(1-\delta)\pi(E_{-y}) + \alpha\delta} \\ &= \frac{(1-\delta)\pi(E_b) + \alpha\delta}{(1-\delta)\pi(E_b \cup E_p) + \alpha\delta} \end{aligned}$$

$$\begin{aligned} v^{DS}(E_b|E_{-y}) &= \frac{v(E_b \cup E_{-y}^c) - v(E_{-y}^c)}{1 - v(E_{-y}^c)} \\ &= \frac{(1-\delta)\pi(E_b \cup E_{-y}^c) + \alpha\delta - (1-\delta)\pi(E_{-y}^c) - \alpha\delta}{1 - ((1-\delta)\pi(E_{-y}^c) + \alpha\delta)} \\ &= \frac{(1-\delta)\pi(E_b \cap E_{-y})}{(1-\delta)\pi(E_{-y}) + \delta(1-\alpha)} = \frac{(1-\delta)\pi(E_b)}{(1-\delta)\pi(E_b \cup E_p) + \delta(1-\alpha)} \end{aligned}$$

$$\begin{aligned} v^{GB}(E_b|E_{-y}) &= \frac{v(E_b \cap E_{-y})}{1 - v(E_b \cup E_{-y}^c) + v(1 \cap E_{-y})} \\ &= \frac{(1-\delta)\pi(E_b \cap E_{-y}) + \alpha\delta}{1 - [(1-\delta)\pi(E_b \cup E_{-y}^c) + \alpha\delta] + (1-\delta)\pi(E_b \cap E_{-y}) + \alpha\delta} \\ &= \frac{(1-\delta)\pi(E_b \cap E_{-y}) + \alpha\delta}{1 - (1-\delta)\pi(E_{-y}^c)} = \frac{(1-\delta)\pi(E_b \cap E_{-y}) + \alpha\delta}{(1-\delta)\pi(E_{-y}) + \delta} \\ &= \frac{(1-\delta)\pi(E_b) + \alpha\delta}{(1-\delta)\pi(E_b \cup E_p) + \delta} \end{aligned}$$

As previously, the weight  $1 - v^l(E_b|E_{-y})$  is assigned to the worst outcome, with  $l \in \{OPT, DS, GB\}$  in order to maximise Equation 3.

## 4.5 Prospect Theory

The last model that we include in our comparison is a parsimonious version of the [Tversky and Kahneman \(1992\)](#) Prospect Theory. We follow [Kothiyal et al. \(2014\)](#) and we specify the model based on the source method, as this was recently developed and used in the analysis of experimental data in [Abdellaoui et al. \(2011\)](#). A source of uncertainty concerns a group of events that is generated by a specific mechanism of uncertainty and therefore, different sources of uncertainty can be treated in a different way. In our framework, it is reasonable to assume that the Bingo Blower is a source of ambiguity, as the various events and their respective representation are created in a particular way. A subject  $i$  in a given treatment, holds subjective, additive beliefs for the various states of the world that can be represented by a vector  $\mathcal{P}$  over  $\mathcal{S}$ . In that sense, subjects are assumed to satisfy *probabilistic sophistication* as this is defined in [Machina and Schmeidler \(1992\)](#). Then, there exists a source function  $\bar{w}(\cdot)$ <sup>15</sup>, which is a mapping in the

<sup>15</sup>We denote the source weighting function with  $\bar{w}(\cdot)$  to distinguish it from the weight  $w(\cdot)$  attached to each event in Equation 1.

interval  $[0,1]$ , satisfying the continuity and strict monotonicity properties, such that  $\bar{w}(0) = 0$ ,  $\bar{w}(1) = 1$  and  $w(E) = \bar{w}(P(E))$ , that transforms these subjective probabilities in a pessimistic or optimistic way, expressing in that way ambiguity aversion or seeking respectively. Due to the cumulative nature of this model, the weights assigned to the various outcomes depend on their relative ranking of payoffs. Similarly to the CEU case, given the ranking  $z_b > z_y > z_p$ , the weights in Equation 1 for the unconditional state are substituted by

$$\begin{aligned} w(1) &= \bar{w}(p(E_b)) \\ w(2) &= \bar{w}(p(E_b \cup E_y)) - \bar{w}(p(E_b)) \\ w(3) &= 1 - \bar{w}(p(E_b \cup E_y)) \end{aligned}$$

As source functions we consider some of the most commonly used specifications in the literature, namely the [Prelec \(1998\)](#) family with one ( $PRL_1$ ) and two parameters ( $PRL_2$ ), the [Tversky and Kahneman \(1992\)](#) weighting function (TK) and the [Goldstein and Einhorn \(1987\)](#) function (GE). The functional forms of the source functions are presented below:

$$\begin{aligned} \bar{w}(p) &= \exp(-(-\ln(p))^\gamma), \quad PRL_1 \\ \bar{w}(p) &= \exp(-(-\ln(p))^\gamma)^{\gamma_2}, \quad PRL_2 \\ \bar{w}(p) &= \frac{p^\gamma}{(p^\gamma + (1-p)^\gamma)^{1/\gamma}}, \quad TK \\ \bar{w}(p) &= \frac{\gamma_2 p^\gamma}{\gamma_2 p^\gamma + (1-p)^\gamma}, \quad GE \end{aligned}$$

The functions are linear when  $\gamma = 1$  and  $\gamma_2 = 1$ , implying SEU preferences. For  $\gamma < 1$  the source function tends to an inverse S-shape, while for  $\gamma > 1$  the function becomes S-shaped. The shape of the source function, along with the values of the parameters  $\gamma$  and  $\gamma_2$ , express the degree of *likelihood insensitivity* (the ability of the decision maker to “understand” probabilities), as well as *attitude towards ambiguity* (see [Wakker \(2010\)](#), [Abdellaoui et al. \(2011\)](#)). The parameter  $\gamma$  captures insensitivity for the  $PRL_1$ ,  $PRL_2$  and GE weighting functions, while  $\gamma_2$  acts as an index of ambiguity aversion (for the  $PRL_2$  and the GE functions). In the case of TK source function, the parameter  $\gamma$  acts as both an index of insensitivity and ambiguity aversion.

PT has been extensively used to model atemporal choice under ambiguity (see [Wakker \(2010\)](#)). In addition, while PT has been recently applied to model behaviour in dynamic frameworks in dynamic contexts under risk ([Barberis \(2012\)](#), [Ebert and Strack \(2015\)](#)), this theoretical model has not been yet used on dynamic decision making under ambiguity. As a result, there is a gap in the literature of axiomatised updating rules for a dynamic version of PT. As [Stomper and Vierø \(2015\)](#) highlight, neither of the two aforementioned studies adopts actual updating but they rather assume that probability weighting is applied to probabilities that have already

been updated according to Bayes rule. We follow [Stomper and Vierø \(2015\)](#) and specify an updating process which involves updating of the probability weighting function. Given the conditional information  $E = E_{-y}$  and the relative ranking of outcomes  $z_b > z_p$ , the conditional probability for outcome 1 is given by

$$w(E_b|E_{-y}) = \frac{\bar{w}(p(E_b))}{\bar{w}(p(E_{-y}))} = \frac{\bar{w}(\pi(b))}{\bar{w}(\pi(b) + \pi(p))}$$

and the residual  $1 - w(E_b|E_{-y})$  is assigned to the worst outcome 3 in order to maximise Equation 3. Since the experiment involves only positive outcomes (gains domain), the PT specification presented above is identical to the Source Choquet Expected Utility in [Kothiyal et al. \(2014, page 6\)](#).

## 5 Econometric Analysis

In order to be able to econometrically fit and obtain estimated parameters of the various preference functionals, we need to make several structural assumptions regarding the representation of the utility function, as well as the stochastic part of the decision making process (noise). Following [Wilcox \(2008\)](#), we define a *specification* as the combination between a deterministic choice model, a utility functional and a stochastic choice statistical model. Regarding the utility function, the extensive degree of non-boundary allocations observed throughout the experiment, implies risk-averse attitudes<sup>16</sup>, thus we need to assume a concave utility function. We adopt a power utility function that is characterised by constant relative risk aversion (CRRA) whose shape is given by<sup>17</sup>

$$u(x) = \begin{cases} \frac{x^{1-r}}{1-r} & \text{if } r \neq 1 \\ \ln(x) & \text{if } r = 1 \end{cases}$$

with  $r$  being the power utility coefficient. Under this assumption, we are able to optimise Equations 1 and 3 and obtain closed form formulas of the optimal constrained allocations, for any given allocation tuple  $(e_b, e_y, e_p, m)$ , of the following form

$$x_b^* = \frac{m e_y e_p (w(E_b) e_b)^{1/r}}{e_y e_p (w(E_b) e_b)^{1/r} + e_b e_p (w(E_y) e_y)^{1/r} + e_b e_y (w(E_p) e_p)^{1/r}}$$

$$x_y^* = \frac{m e_b e_p (w(E_y) e_y)^{1/r}}{e_y e_p (w(E_b) e_b)^{1/r} + e_b e_p (w(E_y) e_y)^{1/r} + e_b e_y (w(E_p) e_p)^{1/r}}$$

<sup>16</sup>[Antoniou et al. \(2015\)](#) show how taking into consideration risk aversion alters inferences on Bayesian updating.

<sup>17</sup>The justification of using a power utility is twofold: (1) it is a common finding in the literature that the CRRA utility provides good fit in empirical applications (see [Wakker \(2008\)](#), [Stott \(2006\)](#), [Balcombe and Fraser \(2015\)](#)) and; (2) the power utility function naturally satisfies out non-negativity constraint as the optimal allocations are bound to be always positive under CRRA.

with  $x_p^*$  being calculated as the residual  $m - x_b^* - x_y^*$  and  $w(\cdot)$  being the weights as those calculated in section 4. Similarly, the conditional allocations, based on some partial information (e.g.  $E = E_{-y}$  are given by

$$x_b^{-y*} = \frac{m^{-y} e_p (w(E_b|E_{-y}) e_b)^{1/r}}{e_p (w(E_b|E_{-y}) e_b)^{1/r} + e_b (w(E_p|E_{-y}) e_p)^{1/r}}$$

with  $x_b^{-y*}$  being the conditional allocation,  $m^{-y}$  the conditional remaining income ( $m - x_y^*$ ) and  $w(E_s|E_{-y})$  the corresponding conditional decision weights, updated according to the way that each theoretical model defines (section 4). We call these allocations, the optimal allocations. Note that as the calculation of the optimal allocation depends on the relative ranking between the three outcomes, the decision weights  $w(E_i)$  depend on the outcomes  $z_i$  in a non-linear and discontinuous fashion. Therefore, the first order conditions in Equation 5, are valid only within the comonotonic sets (where events are consistently ranked). Our estimation code takes into consideration all the possible rankings, including strict inequalities (e.g.  $z_b > z_y > z_p$ ) as well as weak inequalities (e.g.  $z_b = z_y > z_p$  or  $z_b > z_y = z_p$ ). All the combinations of weak and strict inequalities provide in total 13 possible rankings for the first stage and 3 rankings for the second stage (conditional), including the case of equal allocation between all events. Following Ahn et al. (2014) and Hey and Pace (2014), we calculate allocations for all the possible rankings, and define the optimal allocation as the one that simultaneously maximises the utility of the subject and satisfies the respective ranking constraints.

Then, in order to specify the likelihood function, we need to make some assumptions regarding the stochastic part of the subjects' decision making process (noise). As the allocations are bounded to the interval  $[0, m]$ , a convenient way to model noise in choices is to assume that the ratios  $x_s/m$  at a specific allocation question are distributed according to a *Dirichlet* distribution. The Dirichlet distribution is a continuous probability distribution over multinomials, which are  $m$ -tuples  $\mathbf{x} = (x_1, \dots, x_m)$  that sum to unity. A simple parametrisation of the Dirichlet is given by setting

$$\sigma = \sum_{k=1}^K \beta_k$$

and

$$\mathbf{K} = \left( \frac{\beta_1}{\sigma}, \dots, \frac{\beta_K}{\sigma} \right)$$

with the vector  $\mathbf{K}$  summing up to unity. In our context with  $K = 3$  the shape parameters of the distribution are defined as

$$\beta_s = \frac{x_s^*}{m} \sigma$$

where  $x_s^*$  is the theoretical optimal allocation to state  $s$ . This specification has the nice property that the mean of the distribution is centered to the optimal allocation, since the expression for the mean of the distribution is given by:

$$E\left(\frac{x_s}{m}\right) = \frac{\beta_s}{\sum_k \beta_k} = \frac{\frac{x_s^*}{m} \sigma}{\sigma} = \frac{x_s^*}{m}$$

$\sigma$  is an indicator of the precision of the distribution of the random variable  $\frac{x_s}{m}$  also known as the precision of the Dirichlet and needs to be estimated. The higher the value of  $\sigma$ , the lower the noise in the data.

To specify the likelihood function for all the problems, for a particular subject, we need to take into consideration the two stages of each allocation question, with 3-way allocations in the first stage and 2-way in the second. For a particular allocation problem the unconditional allocation  $x_b, x_y, x_p$  to blue, yellow and pink and assume the conditional state  $\neg y$  that will lead to the conditional allocation  $x_b^{-y}, x_p^{-y}$  and the conditional income  $\hat{m}^{-y}$ . Using the allocations at the first stage, we assume that  $x = (\frac{x_b}{m}, \frac{x_y}{m}, \frac{x_p}{m})$  is Dirichlet distributed with the appropriate shape parameters that satisfy the properties above. The contribution to the likelihood function by the first stage allocation is given by:

$$g_1(\pi, k, r, e, m, \sigma, X) = \ln(\Psi(\beta, \sigma))$$

where  $\pi$  represents beliefs for every preference functional that we consider,  $k$  is a vector of the behavioural parameters of each model (e.g.  $\alpha$  for the  $\alpha$ -MEU model,  $\alpha$  and  $\delta$  for the CEU<sub>NEO</sub> and so on),  $e$  are the exchange rates for that problem,  $m$  is the available income,  $X$  is the vector of the actual unconditional allocations,  $\Psi$  is the density function of the Dirichlet distribution,  $\beta$  are the shape parameters of the distribution and  $\sigma$  is the precision parameter.

In the second stage, there are two available allocations to be made and the dimension of the distribution is equal to 2 (it becomes a standard Beta distribution). In our example, we assume that  $x^{-y} = (\frac{x_b^{-y}}{\hat{m}^{-y}}, \frac{x_p^{-y}}{\hat{m}^{-y}})$  is Dirichlet (Beta) distributed, subject to the suitable shape parameters. The contribution to the likelihood function by the first stage allocation is given by:

$$g_2(\pi, k, r, e, m^{-y}, \sigma, X^{-y}) = \ln(\Psi(\beta^{-y}, \sigma))$$

with  $m^{-y}$  and  $X^{-y}$  denoting the conditional income and allocation respectively. The total contribution to the likelihood function of a particular problem is given by  $g_1 + g_2$ . We consider the remaining two conditional states in a symmetric way. The log-likelihood function, conditional to a particular model specification over the 60 allocation problems and over all 58 participants, is defined as:

$$\ln(\mathcal{L}(\pi, k, r, \sigma)) = \sum_{n=1}^{58} \sum_{i=1}^{60} (g_1 + g_2) \quad (9)$$

We now turn to the estimation method. The most common approach to estimate structural decision making models, is the use of either population or subject-level maximum likelihood estimation techniques (MLE). Nevertheless, MLE suffers from few drawbacks. As MLE considers each participant as a separate unit of analysis, if there is a lack of a large number of observations, it may generate noisy and unreliable estimates and therefore can produce extreme estimates for some of the subjects. Second, MLE is prone to overfitting and adjust mostly noise rather than the actual preferences of the subject, leading to very poor predictive performance of the models. To mitigate these drawbacks, we adopt Hierarchical Bayesian estimation techniques (see [Balcombe and Fraser \(2015\)](#) and [Ferecatu and Öncüler \(2016\)](#) for some recent applications of Hierarchical Bayesian models for choice models under risk). The key aspect of hierarchical modelling is that even though it recognises individual variation, it also assumes that there is a distribution governing this variation (individual parameter estimates originate from a group-level distribution). An hierarchical Bayes model, simultaneously estimates the individual level parameters, along with the hyper-parameters of the group level distributions. In typical hierarchical models, the estimates of the low level parameters are pulled closer together than they would in the absence of a higher-level distribution, leading to the so called *shrinkage* of the estimates.

Given a dataset  $y$  and a set of parameters  $\theta$ , Bayesian inference focuses on the posterior distributions of the parameters of each preference functional by combining the likelihood in Equation 9 with the priors of the individual and the group-level parameters. The posterior distribution is given by

$$P(\theta|y) \propto P(y|\theta) \times P(\theta)$$

with  $P(y|\theta)$  being the likelihood and  $P(\theta)$  the priors for all parameters in the set  $\theta$ . Prior distributions need to be specified for all the parameters of the model. For all the parameters (with the exception of the decision weights), we assume that individual parameters  $\theta_n$  are normally distributed ( $\theta_n \sim N(\mu_\theta, \sigma_\theta)$ ), while for the hyper-parameters we assume normal priors for the mean  $\mu_\theta$  and uninformative priors (uniform) for  $\sigma_\theta$ . We also follow the standard procedure and transform all the parameters to their exponential form to ensure that they lie within the appropriate bounds. As far as the decision weights are concerned, given that they need to be jointly distributed, we set Dirichlet priors. This kind of priors guarantees that the sum of the decision weights is equal to unity for the SEU, PT and CEU<sub>NEO</sub> specifications, while a re-parametrisation is required for the MEU,  $\alpha$ -MEU and CEU models so that the sum of the decision weights lies in the range  $[0,1]$ .

Monte Carlo Markov Chains (MCMC) were used to estimate all the specifications. The

estimation was implemented in JAGS (Plummer, 2017)<sup>18</sup>. The posterior distribution of the parameters is based on draws from two independent chains, with 200,000 MCMC draws each, for all the specifications. Due to the high level of non-linearity of the models, there was a burn-in period of 100,000 draws, while to reduce autocorrelation on the parameters, the samples were thinned by 10 (every tenth draw was recorded). Convergence of the chains was confirmed by computing the  $\hat{R}$  statistic (Gelman et al., 1992). Our inference is based on the point estimates for all model parameters  $\pi, k, r, \sigma$  which are obtained from the posterior distribution (posterior mode).

We complete this section by presenting the total number of parameters that we need to estimate, at the individual subject level, for the various preference functionals, as well as the constraints that we need to impose during the estimation. Table 1 summarises the number of parameters for each specification.

Table 1: Number of parameters for each model

Model	Parameters	Total
$\alpha$ -Maxmin Expected Utility ( $\alpha$ -MEU)	$\underline{\pi}(b), \underline{\pi}(y), \underline{\pi}(p), \alpha, r, \sigma$	6
Choquet Expected Utility (CEU)	$v(b), v(y), v(p), v(b \cup y), v(y \cup p), v(b \cup p), r, \sigma$	8
Neo-additive Choquet Expected Utility ( $CEU_{NEO}$ )	$\pi(b), \pi(y), \alpha, \delta, r, \sigma$	6
MaxMin Expected Utility (MEU)	$\underline{\pi}(b), \underline{\pi}(y), \underline{\pi}(p), r, \sigma$	5
Prospect Theory ( $PT_1$ )	$\pi(b), \pi(y), \gamma, r, \sigma$	5
Prospect Theory ( $PT_2$ )	$\pi(b), \pi(y), \gamma, \gamma_2, r, \sigma$	6
Subjective Expected Utility (SEU)	$\pi(b), \pi(y), r, \sigma$	4

Notes:  $CEU_{NEO}$  stands for the CEU specification with *neo-additive* capacities,  $PT_k$  stands for the PT specification with  $k$  indicating the number of parameters of the weighting function.  $\pi(s)$  stands for the additive probabilities in SEU,  $CEU_{NEO}$  and PT,  $\underline{\pi}(s)$  stands for the lower bounds of the probabilities for MEU and  $\alpha$ -MEU,  $v(s)$  stands for the individual and joint capacities for the CEU.

For estimation reasons we need to apply several box constraints (lower and upper bounds) to the various parameters on top of inequality or equality constraints. The bounds of these parameters are dictated by either feasibility constraints, the experimental design or by the relevant literature, while the inequality constraints by the particular model. In all the theoretical models we consider, all probabilities (either additive or lower bounds) and all the capacities are constrained to the interval  $[0,1]$ . The additivity constraint  $\sum_{s=1}^3 \pi(E_s) = 1$ , is imposed to all models that assume probabilistic sophistication (i.e. SEU,  $CEU_{NEO}$  and PT) while the weak inequality constraint is applied to those models that do not (i.e. MEU,  $\alpha$ -

<sup>18</sup>The estimation codes are available upon request.



MEU and CEU). Moreover, the capacities in the CEU model are constrained to be convex (i.e.  $v(E_i \cup E_j) \geq v(E_i) + v(E_j)$ , for all states  $i, j$ ). The power utility parameter  $r$  is constrained to the interval  $[0,5]$ .  $r = 0$  denotes a risk neutral subject, while  $r > 0$  a risk averse one.<sup>19</sup> The upper bound is set to 5, value that implies high levels of risk aversion<sup>20</sup>. The  $\alpha$  parameter in the  $\alpha$ -MEU model is constrained to the unit interval, with  $\alpha = 1$  implying extreme aversion to ambiguity (the model collapses to MEU) and  $\alpha = 0$ , denoting ambiguity loving attitude. The parameters  $\alpha$  and  $\delta$  in the CEU<sub>NEO</sub> model are also constrained to the interval  $[0,1]$ . Regarding the parameters of the weighting functions for the PT specifications,  $\gamma$  and  $\gamma_2$ , both are constrained to the interval  $[0,2]$ , with the only exception the  $\gamma$  parameter of the TK source function which is constrained to the interval  $[0.27,2]$ . This range of values guarantees that the functions are both monotonically increasing on the probabilities  $\pi(\cdot)$  and that they exhibit the S and inverse S shapes. Finally, the precision parameter  $\sigma$  is constrained to the interval  $[1,100]$ . The maximum value indicates very high precision and therefore, low noise.

## 6 Results

Our methodology is that of prediction. Each subject made in total 60 2-stage allocations. Using a subset of the data, the *training* set, namely 45 randomly drawn questions out of the total 60, we first estimated the parameters for each of the theoretical models presented above. Therefore, for each subject and for each theoretical model, we obtain estimates for the subject's beliefs, their power utility coefficient, their precision parameter and additional behavioural coefficients that some models require. Then, using the estimated parameters, we simulate choices in order to predict behaviour (allocations) on the *test* set. Each set of parameters, generates a deterministic allocation, conditional on the model under consideration and the allocation problem. The precision parameter  $\sigma$  allows us to generate simulated allocations which are centered on the optimal allocations implied by the estimated parameters. For each subject, for each preference functional, and for each of the 15 problems of the test set, we generate 100 simulated datasets and we calculate the mean squared deviation between the simulated and the actual data from the test set. The model with the highest predictive capacity should minimise the deviation score. The selection of the estimation and prediction questions was based

<sup>19</sup>This lower bound is imposed by the requirement of non-negative portfolios, since a risk neutral subject would allocate everything to the colour with the highest product between the rate of return and the probability for this colour, while a risk loving subject ( $r < 0$ ) would be willing to allocate negative amounts. This constraint does not impose problems to our analysis, since only two subjects selected boundary portfolios and not for the whole range of problems. The lack of risk loving subjects is also confirmed in [Gneezy et al. \(2015\)](#).

<sup>20</sup>At very large values of  $r$  it becomes impossible to observationally distinguish the different allocations, as they all the allocations tend to equalise the payoffs at each state.

on a random sampling without replacement from the whole set of the 60 allocation questions. While each subject faced the questions in a randomised order during the experiment, we deliberately used the same set of questions as the training and test sets, so that both the estimation and the prediction are based on the same questions for all subjects.

As is common in these studies (see [Stahl \(2014\)](#), [Hey and Pace \(2014\)](#)), we expect large heterogeneity across and within subjects. We allow for heterogeneity across subjects in the econometric comparison of the various specifications by analysing the data in both the group and the subject level, while regarding within heterogeneity we adopt a suitable stochastic error story.

## 6.1 Predictive Power of the Models

We now turn to the predictive accuracy of the various theoretical models. We rank the models based on the value of the mean squared deviation which serves as a measure of prediction. The lower the deviation (the closest to zero), the better the predictive capacity of the theory. The value of the deviation is directly comparable between models and there is no need to use information criteria in order to correct for the different degrees of freedom.

In order to choose which models will participate to the overall comparison, we first perform horse race comparisons within each family of models, that is we include the best specification of the PT family, the best of the CEU and so on. We include in total 6 specifications. [Table 2](#) reports the results of the within CEU family comparison. More specifically, we make the assumption that our subjects are either SEU or one of the three CEU types. Then for each subject we rank the models from the best to the worst according to their predictive capacity. The first column of the Table reports the percentage of subjects for which each row model is classified as best. The second column reports the cumulative percentage of subjects for which the column model is classified as best or second best and so on. It appears that the  $CEU_{GB}$  specification being best for 36% of the subjects outperforms both the other two CEU specifications and SEU which is classified first for 29% of the subjects, while  $CEU_{DS}$  can predict the choices of a very restricted subset.

[Table 3](#) reports the same kind of ranking for the PT family against again SEU. In this case, SEU is the winner by being the best model for 22% of the subjects, closely followed by  $PT_{PRL2}$  with 21% which marginally wins all the remaining 3 PT specifications which explain behaviour for roughly 1/5 of the population each. The within  $CEU_{NEO}$  family comparison is reported in [Table 4](#). Here there is a clear winner, the  $CEU_{NEO/DS}$  which predicts best for 29% of the subjects followed by SEU with 26%. In all cases, the SEU predictive capacity is restricted to 22-29% of

the total sample.

We finally perform an overall ranking of the models, including all the winning specifications from the within-families comparison. We include in total 6 models, as these are presented in the rows of Table 5. The overall winner is a dynamic specification of PT,  $PT_{PRL2}$  closely followed by  $CEU_{GB}$  which together may predict 57% of the choices of our sample. SEU and  $CEU_{NEO/DS}$  follow with 14% each, while the multiple priors families have the worst performance overall.

Table 6 reports the mean squared deviations, as well as the trimmed means at 5 and 10% levels to mitigate the impact of outliers, the median value and the standard deviation for each of the theoretical specifications. Overall,  $CEU_{GB}$  outperforms, on average, all the competing model, according to all but one measures that we report, with  $PT_{PRL2}$  in the second place. The multi-priors family models (MEU and  $\alpha$ -MEU) perform relatively bad, while SEU is ranked last. The size of the standard deviations indicates that there is considerable variation across subjects and that while some models perform well in the subject level analysis (e.g.  $PT_{PRL2}$ ), they fail to perform well on average due to this variation, which provides evidence in favour of the individual level inference.

## 6.2 Estimated Parameters

To conclude, we briefly report the estimated values of the parameters for the best specifications of each family in Table 7. The Table reports the mean, median and the standard deviations of all the parameters. The large values of the standard deviations capture the extensive level of heterogeneity between subjects, within each preference functional. From this Table we summarise two important findings: (1) the overweight of low probabilities and underweight of high probabilities, a result commonly reported in studies of choice under ambiguity; and (2) the non-additivity of beliefs for the models that allow for violation of probabilistic sophistication. The latter is in line with the findings in Baillon et al. (2017), where extensive violation of probabilistic sophistication was reported. The over/under-weighting of probabilities finding, is also confirmed by Figure 2 where the boxplots of all the parameters representing beliefs, for all the specifications are illustrated. The solid vertical lines correspond to the actual probabilities of the three colours which were known to the experimenters but unknown to the subjects (0.2 for  $E_b$ , 0.5 for  $E_y$  and 0.3 for  $E_p$ ). Nevertheless, it is clear that subjects do not exhibit extreme likelihood insensitivity (failure to distinguish the different probabilities of the three events and therefore resort to a uniform probability distribution) since in all the specifications, the ranking

$\pi(E_y) > \pi(E_p) > \pi(E_b)$  is on average satisfied<sup>21</sup>.

Figure 5 presents the boxplots for the parameters of the utility function ( $r$ ) and precision ( $\sigma$ ), which are common in all models, as well as the boxplot for the parameters  $\gamma$  and  $\gamma_2$  of the PRL2 weighting function. In the first panel of the Figure, it is obvious that subjects are considerably risk averse with the median value of the coefficient being between 0.833 and 1.013, finding which is in line with experiments that use a similar experimental design (Hey and Pace (2014), Ahn et al. (2014)). The noise in the data (second panel in Figure 5) seems to follow a uniform pattern across the different specifications, where the mean and the median values are almost identical for all the models. The last panel in Figure 5 illustrates the boxplot of the estimates for the PRL2  $\gamma$  and  $\gamma_2$  parameter. These parameters are key ones, since they simultaneously capture ambiguity attitudes and likelihood insensitivity. The large standard deviation confirms that the subject-level estimates are significantly different than 1. This large variation is also confirmed by Figure 4, where the Kernel density plots of the estimated values of  $\gamma$  and  $\gamma_2$  are presented. Finally, it worths commenting on the estimated parameters of the CEU<sub>NEO/DS</sub> model since it has been widely used in the literature. The value of the parameter  $\delta$ , which expresses lack of confidence to the additive subjective beliefs, is on average quite low ( $\delta = 0.139$ ) indicating that roughly, only 14% of the weight is put to the best and the worst outcome, with the higher volume of the weight being assigned to the subjective beliefs. Moreover, the parameter  $\alpha$  that captures optimism, is estimated close to 0.5 ( $\alpha = 0.419$ ). Kernel distribution plots, verify that the  $\delta$  parameter follows a skewed to the right distribution across subjects, while the parameter  $\alpha$  is mostly concentrated on the 0.5 value for a majority of the subjects indicating that the subjects are neither extremely pessimistic ( $\alpha = 0$ ) or extremely optimistic ( $\alpha = 1$ ), with pessimism prevailing marginally.

## 7 Discussion and Concluding Remarks

This paper experimentally tested the out-of-sample predictive power of dynamic models under ambiguity. Employing a within subjects experimental design, using a 2-stage allocation decision task, and representing ambiguity with a transparent, non-manipulable device (a Bingo blower), we gather data that allow us to run a horse race between the dynamic versions of the most commonly used theoretical models of choice under ambiguity. We include in total 13 different theoretical specifications. We conduct comparisons between theories as well as within theories, when various different updating rules have been proposed in the literature. Our key results are summarised as follows: (1) overall, Prospect Theory outperforms all the

---

<sup>21</sup>This also holds for the capacities  $v(\cdot)$  as well as for the lower bounds of the probabilities  $\underline{\pi}(\cdot)$ .

other models regarding out-of-sample fit capacity closely followed by the Choquet Expected Utility model; (2) within the Prospect Theory family, the [Prelec \(1998\)](#) 2-parameter weighting function (PRL2) best describes behaviour; (3) within the Choquet Expected Utility model, the Generalised Bayesian rule provides the best fit, while and; (4) within the neo-additive CEU family, the Dempster-Shaffer rule.

As a general conclusion, we find that a dynamic specification of Prospect Theory has the highest predictive capacity closely followed by the CEU model. Both PT and CEU differ in the way they represent beliefs and therefore they differ in their assumptions on the assumptions behind these beliefs, the way they form decision weights and the way these decision weights are updated. Given the strong performance of the PT model in predicting behaviour calls for further theoretical advances and axiomatisation of the model in its dynamic version, while the strong performance of CEU, offers a strong candidate model for empirical applications.

[Wakker \(2010, Appendix C\)](#) referring to non-Expected Utility models and dynamic decisions, concludes: “I do not consider nonexpected utility to be normative and think that no normatively satisfactory method for updating or implementation in dynamic decisions will be found. [...] Which method of updating is most appropriate for descriptive applications is an interesting topic for empirical research”. Given the increasing popularity of dynamic ambiguity models in modelling financial decision making ([Easley and O’ Hara \(2009\)](#), [Mele and Sangiorgi \(2015\)](#)) or in empirical research using field data ([Thimme and Völkert \(2015\)](#), [Jeong et al. \(2015\)](#), [Li et al. \(2016\)](#)), it is crucial to understand which theoretical model is the most appropriate to apply in empirical analysis.

The present study is a first step towards understanding choices and updating of ambiguous beliefs in a dynamic decision making problem. To keep the experiment as simple as possible, we used an urn and balls type experimental design, with a 2-period decision task and three possible states of nature, constrained to the gains domain. Given the advances in the experimental literature of static choice under ambiguity, it would be interesting to extend the current analysis and explore dynamic choice in richer contexts such as addressing multiple-period decision problems, study behaviour in the losses domain, allowing learning under ambiguity or consider natural events as the source of ambiguity. We leave these for future research.

Table 2: Ranking of CEU specifications based on the mean squared predicted deviation

Model	1	1-2	1-3	1-4
$CEU_{DS}$	0.09	0.38	0.69	1.00
$CEU_{GB}$	0.36	0.57	0.81	1.00
$CEU_{OPT}$	0.26	0.62	0.78	1.00
$SEU$	0.29	0.43	0.72	1.00

Notes: All values represent the cumulative percentages. CEU stands for the Choquet Expected Utility model, with the subscript indicating the respective updating rule, DS the Dempster-Shafer, GB the Generalised Bayesian and OPT the Optimistic rule. SEU stands for the Subjective Expected Utility model.

Table 3: Ranking of PT specifications based on the mean squared predicted deviation

Model	1	1-2	1-3	1-4	1-5
$PT_{GE}$	0.19	0.43	0.57	0.72	1.00
$PT_{PRL_1}$	0.19	0.33	0.57	0.83	1.00
$PT_{PRL_2}$	0.21	0.40	0.57	0.79	1.00
$PT_{TK}$	0.19	0.38	0.64	0.86	1.00
$SEU$	0.22	0.47	0.66	0.79	1.00

Notes: All values represent the cumulative percentages. PT stands for the Prospect Theory model, with the subscript indicating the respective probability weighting function, GE for the [Goldstein and Einhorn \(1987\)](#),  $PRL_1$  and  $PRL_2$  for the 1 and 2-parameter [Prelec \(1998\)](#) and TK for the [Tversky and Kahneman \(1992\)](#) weighting function. SEU stands for the Subjective Expected Utility model.

Table 4: Ranking of  $CEU_{NEO}$  specifications based on the mean squared predicted deviation

Model	1	1-2	1-3	1-4
$CEU_{NEO/DS}$	0.29	0.50	0.83	1.00
$CEU_{NEO/GB}$	0.24	0.53	0.76	1.00
$CEU_{NEO/OPT}$	0.21	0.52	0.66	1.00
$SEU$	0.26	0.45	0.76	1.00

Notes: All values represent the cumulative percentages.  $CEU_{NEO}$  stands for the Choquet Expected Utility model with neo-additive capacities, with the subscript indicating the respective updating rule, DS the Dempster-Shafer, GB the Generalised Bayesian and OPT the Optimistic rule. SEU stands for the Subjective Expected Utility model.

Table 5: Ranking of all specifications based on the mean squared predicted deviation

Model	1	1-2	1-3	1-4	1-5	1-6
$\alpha$ -MEU	0.05	0.29	0.47	0.60	0.84	1.00
CEU <sub>GB</sub>	0.28	0.43	0.57	0.71	0.79	1.00
CEU <sub>NEO/DS</sub>	0.14	0.29	0.48	0.76	0.95	1.00
MEU	0.10	0.28	0.53	0.69	0.86	1.00
PT <sub>PRL2</sub>	0.29	0.41	0.53	0.60	0.69	1.00
SEU	0.14	0.29	0.41	0.64	0.86	1.00

Notes: All values represent the cumulative percentages.  $\alpha$ -MEU stands for the  $\alpha$ -Maxmin Expected Utility model, CEU<sub>GB</sub> stands for the Choquet Expected Utility with a Generalised Bayesian updating rule, CEU<sub>NEO/DS</sub> stands for the Choquet Expected Utility with neo-additive capacities and a Dempster-Shafer updating rule, MEU stand for the Maxmin, PT<sub>PRL2</sub> stands for the Prospect Theory with a 2-parameter [Prelec \(1998\)](#) weighting function and SEU stands for the Subjective Expected Utility model.

Table 6: Mean predicted squared deviations

	Mean	Mean.1	Mean.05	Median	s.d.
$\alpha$ -MEU	1.59	1.37	1.48	0.68	2.31
CEU <sub>GB</sub>	1.53	1.30	1.42	0.66	2.29
CEU <sub>NEO/DS</sub>	1.57	1.35	1.46	0.65	2.33
MEU	1.59	1.37	1.48	0.68	2.32
PT <sub>PRL2</sub>	1.69	1.49	1.59	0.63	2.46
SEU	1.72	1.51	1.61	0.69	2.49

Notes: The mean, the 5 and 10% trimmed means, the median and the standard deviation of the mean predicted squared deviations are reported for each specification. A lower value of each measure implies a better predictive capacity of the model on average.  $\alpha$ -MEU stands for the  $\alpha$ -Maxmin Expected Utility model, CEU<sub>GB</sub> stands for the Choquet Expected Utility with a Generalised Bayesian updating rule, CEU<sub>NEO/DS</sub> stands for the Choquet Expected Utility with neo-additive capacities and a Dempster-Shafer updating rule, MEU stand for the Maxmin, PT<sub>PRL2</sub> stands for the Prospect Theory with a 2-parameter [Prelec \(1998\)](#) weighting function and SEU stands for the Subjective Expected Utility model.

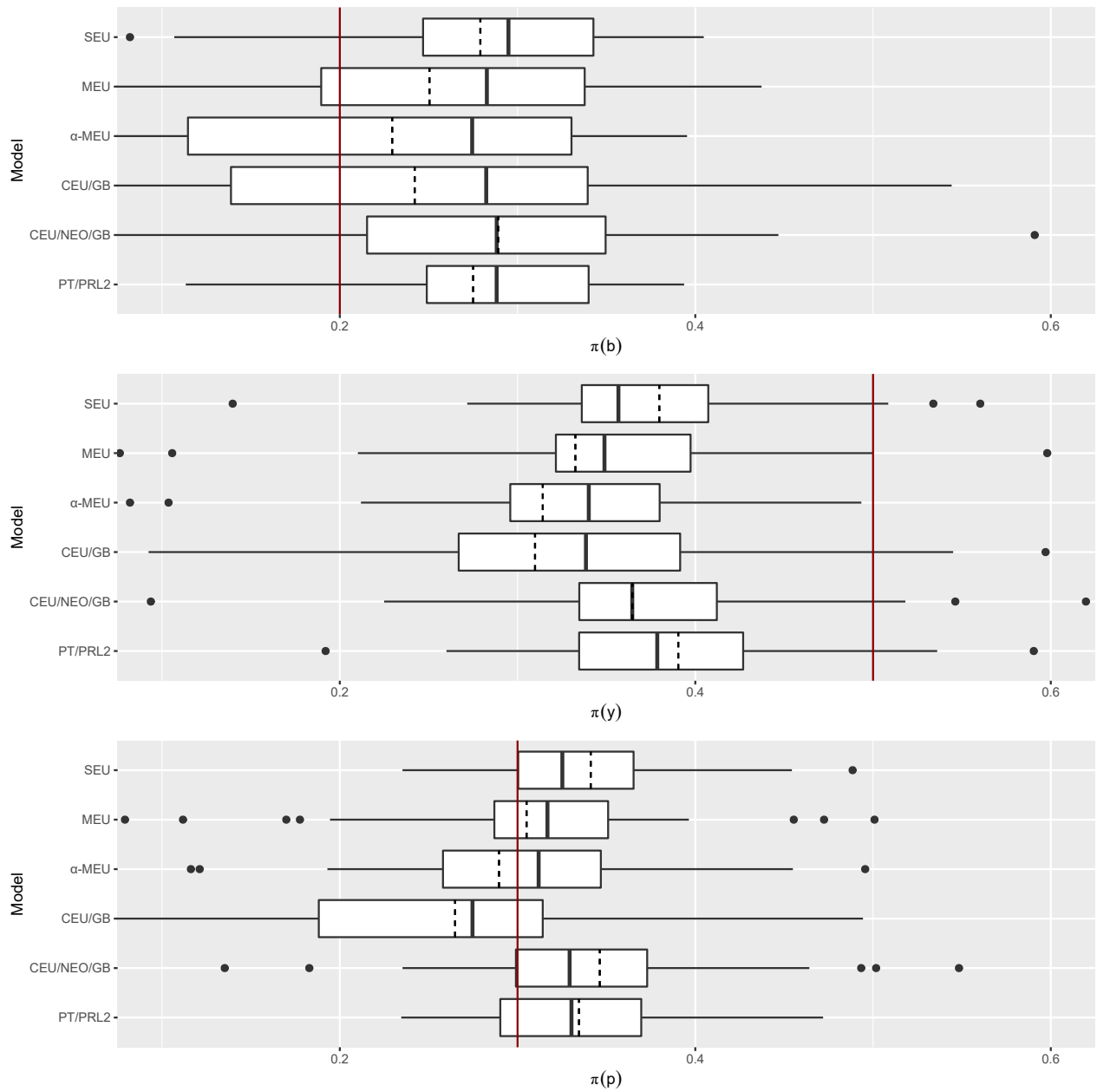


Table 7: Descriptive Summary of Estimates

Theory	Parameter	Mean	Median	s.d
$\alpha$ -MEU	$\underline{\pi}(b)$	0.229	0.274	(0.115)
	$\underline{\pi}(y)$	0.314	0.340	(0.123)
	$\underline{\pi}(p)$	0.290	0.312	(0.097)
	$\alpha$	0.527	0.563	(0.302)
	$r$	1.047	1.011	(0.491)
	$s$	41.576	17.362	(54.054)
CEU <sub>GB</sub>	$v(b)$	0.242	0.282	(0.128)
	$v(y)$	0.310	0.338	(0.142)
	$v(p)$	0.265	0.275	(0.130)
	$v(b \cup y)$	0.595	0.660	(0.165)
	$v(y \cup p)$	0.675	0.695	(0.146)
	$v(b \cup p)$	0.530	0.521	(0.126)
	$r$	0.983	0.951	(0.400)
	$s$	41.882	18.57	(53.028)
CEU <sub>NEO/DS</sub>	$\pi(b)$	0.291	0.292	(0.152)
	$\pi(y)$	0.364	0.362	(0.105)
	$\pi(p)$	0.345	0.328	(0.130)
	$\alpha$	0.419	0.441	(0.206)
	$\delta$	0.139	0.022	(0.248)
	$r$	0.997	0.931	(0.478)
	$s$	41.493	17.381	(54.065)
MEU	$\underline{\pi}(b)$	0.250	0.283	(0.112)
	$\underline{\pi}(y)$	0.333	0.349	(0.130)
	$\underline{\pi}(p)$	0.305	0.317	(0.107)
	$r$	0.935	0.883	(0.430)
	$s$	40.946	17.043	(53.309)
PT <sub>PRL2</sub>	$\pi(b)$	0.275	0.288	(0.085)
	$\pi(y)$	0.390	0.379	(0.079)
	$\pi(p)$	0.335	0.330	(0.069)
	$\gamma$	0.971	0.982	(0.168)
	$\gamma_2$	1.125	1.104	(0.247)
	$r$	1.075	0.954	(0.623)
	$s$	43.727	19.766	(55.114)
SEU	$\pi(b)$	0.279	0.295	(0.086)
	$\pi(y)$	0.380	0.357	(0.076)
	$\pi(p)$	0.341	0.325	(0.073)
	$r$	1.360	1.013	(0.985)
	$s$	43.102	16.833	(56.875)

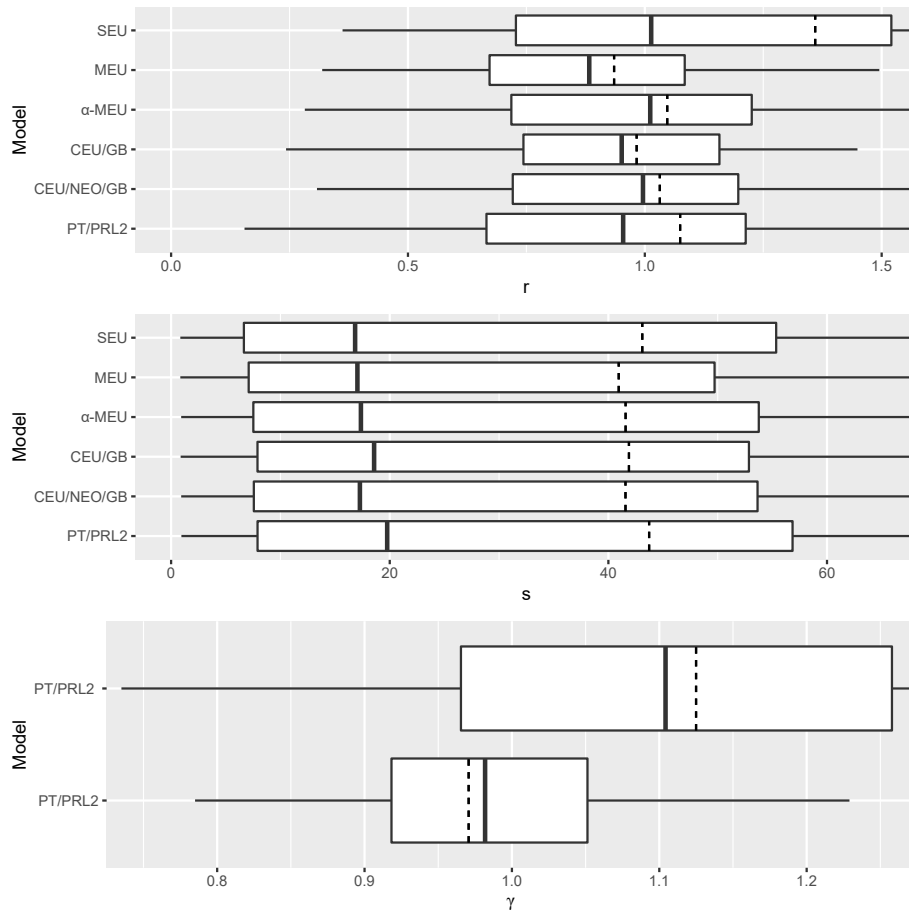
Notes:  $\alpha$ -MEU stands for the  $\alpha$ -Maxmin Expected Utility model, CEU<sub>GB</sub> stands for the Choquet Expected Utility with a Generalised Bayesian updating rule, CEU<sub>NEO/DS</sub> stands for the Choquet Expected Utility with neo-additive capacities and a Dempster-Shafer updating rule, MEU stand for the Maxmin, PT<sub>PRL2</sub> stands for the Prospect Theory with a 2-parameter [Prelec \(1998\)](#) weighting function and SEU stands for the Subjective Expected Utility model.

Figure 2: Boxplots of all the parameters representing beliefs



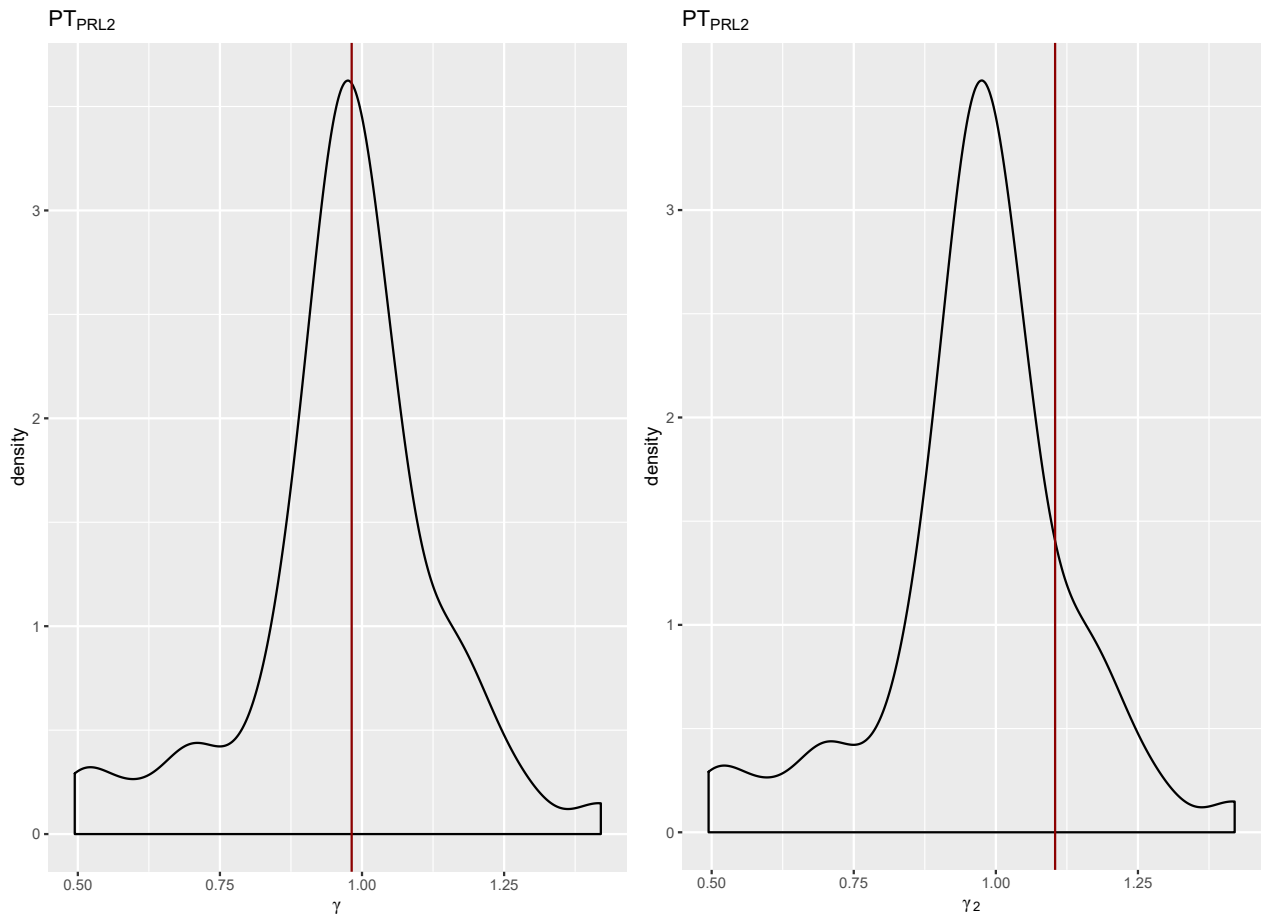
The vertical lines correspond to the actual probabilities of the three colours, 0.2 for colour 1, 0.5 for colour 2 and 0.3 for colour 3. The dashed line inside each box illustrates the mean value, while the continuous line corresponds to the median.

Figure 3: Boxplots of risk, precision and weighting coefficients



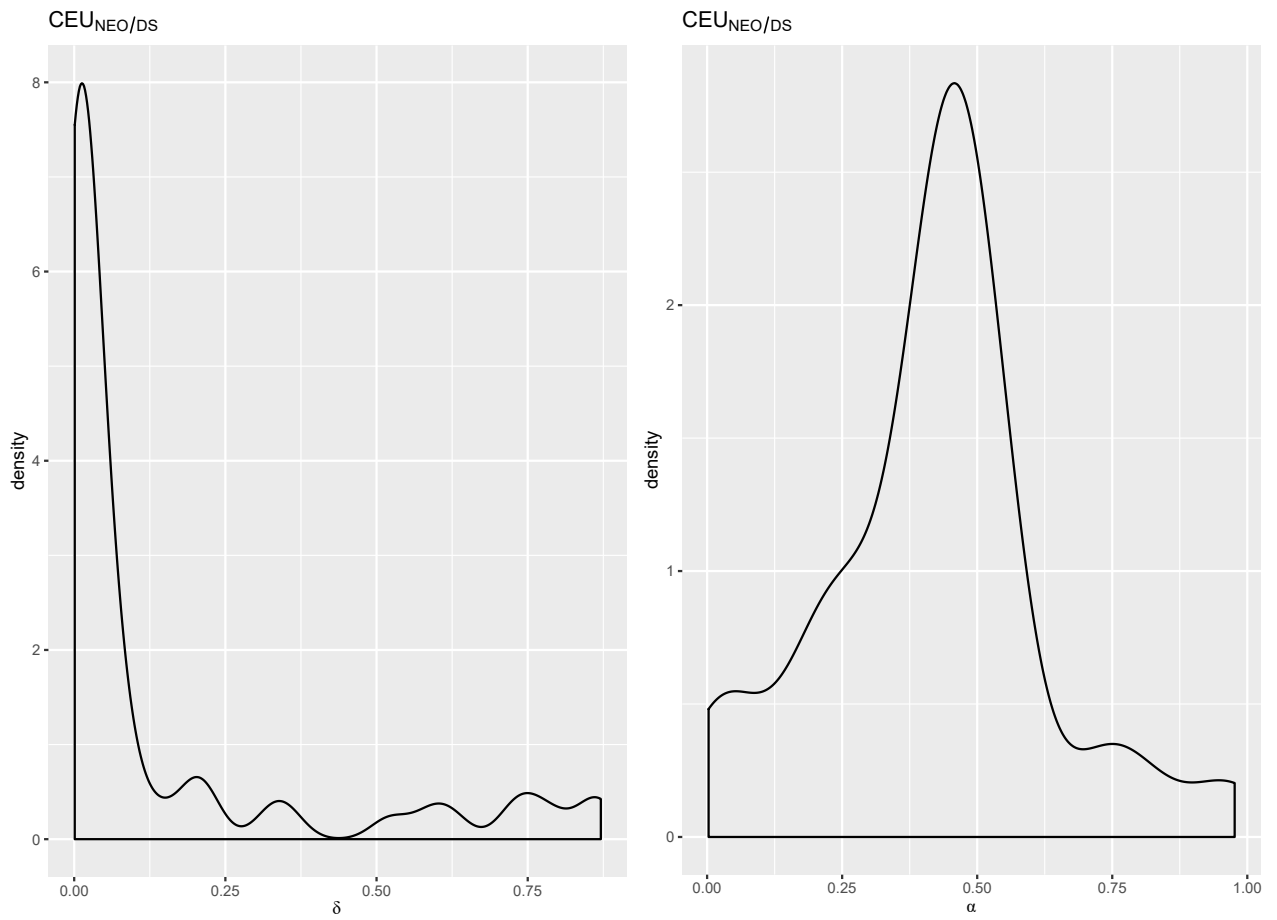
Notes: The dashed line inside each box illustrates the mean value, while the continuous line corresponds to the median.

Figure 4: Kernel density plot of the  $\gamma$  and  $\gamma_2$  parameters



Notes: The Figure presents the density plot of the estimates of the parameters  $\gamma$  and  $\gamma_2$  for the [Prelec \(1998\)](#) weighting function with two parameters. The vertical line illustrates the median value.

Figure 5: Kernel densities plots for the  $CEU_{NEO,DS}$  model



Notes: The parameters  $\delta$  (left panel) and  $\alpha$  (right panel) are illustrated for the  $CEU_{NEO,DS}$  model.

## References

- Abdellaoui, M., A. Baillon, L. Placido, and P. Wakker (2011), "The Rich Domain of Uncertainty: Source Functions and Their Experimental Implementation." *American Economic Review*, 101, 695–723.
- Ahn, D., S. Choi, D. Gale, and S. Kariv (2014), "Estimating Ambiguity Aversion in a Portfolio Choice Experiment." *Quantitative Economics*, 5, 195–223.
- Al-Najjar, N. and J. Weinstein (2009), "The Ambiguity Aversion Literature: A Critical Assessment." *Economics and Philosophy*, 25, 249–284.
- Andersen, S., J. Fountain, W. G. Harrison, A. Hole, and E. Rutstrom (2012), "Inferring Beliefs as Subjectively Imprecise Probabilities." *Theory and Decision*, 73, 161–184.
- Antoniou, C., . Harrison, G., M. Lau, and D. Read (2015), "Subjective Bayesian Beliefs." *Journal of Risk and Uncertainty*, 50, 35–54.
- Bade, S. (2015), "Randomization Devices and the Elicitation of Ambiguity-averse Preferences ." *Journal of Economic Theory*, 159,, 221 – 235.
- Baillon, A. and H. Bleichrodt (2015), "Testing Ambiguity Models through the Measurement of Probabilities for Gains and Losses." *American Economic Journal: Microeconomics*, 7, 77–100.
- Baillon, A., H Bleichrodt, U. Keskin, O. L'Haridon, and C. Li (2017), "Learning Under Ambiguity: An Experiment Using Initial Public Offerings on a Stock Market." *Management Science*, Articles in advance.
- Baillon, A., Y. Halevy, and C. Li (2015), "Experimental Elicitation of Ambiguity Attitude using the Random Incentive System." Microeconomics working papers, Vancouver School of Economics.
- Balcombe, K. and I. Fraser (2015), "Parametric Preference Functionals under Risk in the Gain Domain: A Bayesian Analysis." *Journal of Risk and Uncertainty*, 50, 161–187.
- Barberis, N. (2012), "A Model of Casino Gambling." *Management Science*, 58, 35–51.
- Burghart, D., T. Epper, and E. Fehr (2015), "The Ambiguity Triangle: Uncovering Fundamental Patterns of Behavior Under Uncertainty." Discussion Paper 9150, IZA.
- Charness, G. and U. Gneezy (2010), "Portfolio Choice and Risk Attitudes: an Experiment." *Economic Inquiry*, 48, 133–146.
- Chateauneuf, A., J. Eichberger, and S. Grant (2007), "Choice under Uncertainty with the Best and Worst in Mind: Neo-additive Capacities." *Journal of Economic Theory*, 137, 538 – 567.
- Choi, S., R. Fisman, D. Gale, and S. Kariv (2007), "Consistency and heterogeneity of individual behavior under uncertainty." *American Economic Review*, 97, 1921–1938.
- Cohen, M., I. Gilboa, and D. Schmeidler (2000), "An Experimental Study of Updating Ambi-

- guous Beliefs." *Risk, Decision and Policy*, 5 (2), 123–133.
- Corgnet, B., P. Kujal, and D. Porter (2013), "Reaction to Public Information in Markets: How much does Ambiguity Matter?" *Economic Journal*, 123, 699–737.
- Cubitt, R., G. van de Kuilen, and S. Mukerji (2014), "Discriminating between Models of Ambiguity Attitude: A Qualitative Test." Technical Report 692, University of Oxford, Discussion Paper Series.
- De Filippis, R., A. Guarino, P. Jehiel, and T. Kitagawa (2016), "Updating Ambiguous Beliefs in a Social Learning Experiment." Technical Report CWP 18/16, cemmap working paper.
- Dempster, A. P. (1967), "Upper and Lower Probabilities Induced by a Multivalued Mapping." *The Annals of Mathematical Statistics*, 38 (2), 325–339.
- Dempster, A. P. (1968), "A Generalization of Bayesian Inference." *Journal of the Royal Statistical Society. Series B (Methodological)*, 30, pp. 205–247.
- Dominiak, A., P. Dürsch, and J. Lefort (2012), "A Dynamic Ellsberg Urn Experiment." *Games and Economic Behavior*, 75, 625–638.
- Easley, D. and M. O' Hara (2009), "Ambiguity and Nonparticipation: The Role of Regulation." *Review of Financial Studies*, 22, 1817–1843.
- Ebert, S. and P. Strack (2015), "Until the Bitter End: On Prospect Theory in a Dynamic Context." *American Economic Review*, 105, 1618–33.
- Eichberger, J., S. Grant, and D. Kelsey (2007), "Updating Choquet Beliefs." *Journal of Mathematical Economics*, 43, 888–899.
- Eichberger, J., S. Grant, and D. Kelsey (2010), "Comparing Three Ways to Update Choquet Beliefs." *Economics Letters*, 107, 91 – 94.
- Ellsberg, D. (1961), "Risk, Ambiguity and the Savage Axioms." *Quarterly Journal of Economics*, 75, 643–669.
- Epstein, L., J. Noor, and A. Sandroni (2010), "Non-Bayesian Learning." *The B.E. Journal of Theoretical Economics*, 10, 1–20.
- Epstein, L. and M. Schneider (2003), "Recursive Multiple-Priors." *Journal of Economic Theory*, 113, 1–31.
- Epstein, L. and M. Schneider (2007), "Learning Under Ambiguity." *Review of Economic Studies*, 74, 1275–1303.
- Etner, J., M. Jeleva, and J.M. Tallon (2012), "Decision Theory Under Ambiguity." *Journal of Economic Surveys*, 26(2), 234–270.
- Ferecatu, A. and A. Öncüler (2016), "Heterogeneous Risk and Time Preferences." *Journal of Risk and Uncertainty*, 53, 1–28.



- Gelman, A., , and D. Rubin (1992), "Inference from Iterative Simulation Using Multiple Sequences." *Statistical Science*, 7, 457–472.
- Georgalos, K. (2016), "Dynamic Decision Making under Ambiguity: a Portfolio Choice Experiment." Working Papers 2016/004, Lancaster University Management School, Economics Department.
- Ghirardato, P. (2002), "Revisiting Savage in a Conditional World." *Journal of Economic Theory*, 20, pp. 83–92.
- Ghirardato, P., F. Maccheroni, and M. Marinacci (2004), "Differentiating Ambiguity and Ambiguity Attitude ." *Journal of Economic Theory*, 118, 133 – 173.
- Gilboa, I. and D. Schmeidler (1989), "Maxmin Expected Utility with non-Unique Prior." *Journal of Mathematical Economics*, 18, 141–153.
- Gilboa, I. and D. Schmeidler (1993), "Updating Ambiguous Beliefs." *Journal of Economic Theory*, 59, 33–49.
- Gneezy, U., A. Imas, and J. List (2015), "Estimating Individual Ambiguity Aversion: A Simple Approach." Technical Report 20982, NBER Working Paper.
- Goldstein, W. and H. Einhorn (1987), "Expression Theory and the Preference Reversal Phenomena." *Psychological Review*, 236–254.
- Greiner, B. (2015), "Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE." *Journal of the Economic Science Association*, 114–125.
- Hanany, E. and P. Klibanoff (2009), "Updating Ambiguity Averse Preferences." *The B.E. Journal of Theoretical Economics*, 9 (1), 1–53.
- Hayashi, T. and R. Wada (2010), "Choice with Imprecise Information: an Experimental Approach." *Theory and Decision*, 69, 355–373.
- Hey, J. (2014), "Chapter 14 - Choice Under Uncertainty: Empirical Methods and Experimental Results." In *Handbook of the Economics of Risk and Uncertainty* (M. Machina and K. Viscusi, eds.), volume 1 of *Handbook of the Economics of Risk and Uncertainty*, 809 – 850, North-Holland.
- Hey, J., G. Lotito, and A. Maffioletti (2010), "The Descriptive and Predictive Adequacy of Theories of Decision Making under Uncertainty / Ambiguity." *Journal of Risk and Uncertainty*, 41, 81–111.
- Hey, J. and N. Pace (2014), "The Explanatory and Predictive Power of Non Two-Stage Probability Models of Decision Making Under Ambiguity." *Journal of Risk and Uncertainty*, 49, 1–29.
- Hey, J. and L. Panaccione (2011), "Dynamic Decision Making: What Do People Do?" *Journal of*

- Risk and Uncertainty*, 42, 85–123.
- Jeong, D., H. Kim, and J. Park (2015), “Does Ambiguity Matter? Estimating Asset Pricing Models with a Multiple-priors Recursive Utility.” *Journal of Financial Economics*, 115, 361 – 382.
- Klibanoff, P. and E. Hanany (2007), “Updating Preferences with Multiple Priors.” *Theoretical Economics*, 2 (3), 261–298.
- Klibanoff, P., M. Marinacci, and S. Mukerji (2005), “A Smooth Model of Decision Making under Ambiguity.” *Econometrica*, 73, 1849–1892.
- Klibanoff, P., M. Marinacci, and S. Mukerji (2009), “Recursive Smooth Ambiguity Preferences.” *Journal of Economic Theory*, 144, 930–976.
- Kothiyal, A., V. Spinu, and P. Wakker (2014), “An Experimental Test of Prospect Theory for Predicting Choice under Ambiguity.” *Journal of Risk and Uncertainty*, 48, 1–17.
- Li, W., A. Tiwari, and L. Tong (2016), “Investment Decisions Under Ambiguity: Evidence from Mutual Fund Investor Behavior.” *Management Science*, forthcoming.
- Loomes, G. (1991), “Evidence of a New Violation of the Independence Axiom.” *Journal of Risk and Uncertainty*, 4, 91–108.
- Loomes, G. and G. Pogrebna (2014), “Measuring Individual Risk Attitudes when Preferences are Imprecise.” *The Economic Journal*, 124, 569–593.
- Machina, M. and D. Schmeidler (1992), “A More Robust Definition of Subjective Probability.” *Econometrica*, 60, 745–80.
- Machina, M. and M. Siniscalchi (2014), “Chapter 13 - Ambiguity and Ambiguity Aversion .” In *Handbook of the Economics of Risk and Uncertainty* (M. Machina and K. Viscusi, eds.), volume 1 of *Handbook of the Economics of Risk and Uncertainty*, 729 – 807, North-Holland.
- Marinacci, M. (2002), “Learning from Ambiguous Urns.” *Statistical Papers*, 43, 145–151.
- Mele, Antonio and Francesco Sangiorgi (2015), “Uncertainty, Information Acquisition, and Price Swings in Asset Markets.” *The Review of Economic Studies*, 82, 1533–1567.
- Oechssler, J. and A. Roomets (2014), “Unintended Hedging in Ambiguity Experiments.” *Economics Letters*, 122, 243 – 246.
- Peysakhovich, A. and U. Karmarkar (2016), “Asymmetric Effects of Favorable and Unfavorable Information on Decision Making Under Ambiguity.” *Management Science*, 62, 2163–2178.
- Pires, C. (2002), “A Rule For Updating Ambiguous Beliefs.” *Theory and Decision*, 53, 137–152.
- Plummer, M. (2017), “JAGS Version 4.3.0 User Manual.” Technical report.
- Prelec, D. (1998), “The Probability Weighting Function.” *Econometrica*, 66, pp. 497–527.
- Savage, L. (1954), *The Foundations of Statistics*. Wiley, New York.

- Schmeidler, D. (1989), "Subjective Probability and Expected Utility Without Additivity." *Econometrica*, 57, 571–587.
- Shafer, G. (1976), *A Mathematical Theory of Evidence*. Princeton University Press, Princeton.
- Siniscalchi, M. (2011), "Dynamic Choice under Ambiguity." *Theoretical Economics*, 6, 379–421.
- Stahl, D. (2014), "Heterogeneity of Ambiguity Preferences." *The Review of Economics and Statistics*, 96, 609–617.
- Stomper, A. and M. Vierø (2015), "Iterated Expectations under Rank-Dependent Expected Utility and Model Consistency." Working Papers 1228, Queen's University, Department of Economics.
- Stott, H. (2006), "Cumulative Prospect Theory's Functional Menagerie." *Journal of Risk and Uncertainty*, 32, 101–130.
- Thimme, J. and C. Völkert (2015), "Ambiguity in the Cross-Section of Expected Returns: An Empirical Assessment." *Journal of Business & Economic Statistics*, 33, 418–429.
- Trautmann, S. and G. van de Kuilen (2015), *Wiley Blackwell Handbook of Judgment and Decision Making*, chapter Ambiguity Attitudes, 89–116. Blackwell.
- Tversky, A. and D. Kahneman (1992), "Advances in Prospect Theory: Cumulative Representation of Uncertainty." *Journal of Risk and Uncertainty*, 5, 297–323.
- Wakker, P. (2008), "Explaining the Characteristics of the Power (CRRA) Utility Family." *Health Economics*, 17, 1329–1344.
- Wakker, P. (2010), *Prospect Theory*. Cambridge University Press.
- Wang, T. (2003), "Conditional Preferences and Updating." *Journal of Economic Theory*, 286–321.
- Wilcox, N. (2007), "Predicting Risky Choices out-of-context: A Monte Carlo Study." Technical report, University of Houston.
- Wilcox, N. (2008), "Stochastic Models for Binary Discrete Choice Under Risk: A Critical Primer and Econometric Comparison." In *Research in Experimental Economics* (J. Cox and G. Harrison, eds.), volume 12, 41–196, Emerald Group Publishing Limited.