Technical University of Denmark

Master Thesis

Normative accounts of Probability Weighting in Decision Science

Benjamin Skjold Frederiksen

supervised by

Oliver J. Hulme; DRCMR
Tobias Andersen; DTU Compute

January 3^{rd}, 2021

Abstract

Probability Weighting is a concept that originated in Prospect Theory (PT) that describes a distortion of probabilities empirically shown in decisions under risk. Since its introduction, probability weighting has been considered an irrational cognitive bias, i.e. an error of judgment by the decision maker. Recently, though, theories seeking to provide normative accounts for this distortion of probabilities have been seen to emerge. It is found that probability weighting can be explained simply as a decision maker having greater uncertainty about the world than an observer. And such greater uncertainty arises naturally in decisions from experience, if a decision maker, who must infer the probabilities from his/her experience, recognizes the fact that whenever the sampling is finite, the decision maker is expected to experience rare outcomes less often than the probabilities prescribe. Here, I present a modeling framework where I investigate qualitatively different models of probability weighting, by studying the predictive mass of the models, as well as the effects of increasing knowledge (indicated by time) in risky choices with respect to probability weighting. Through a simulation study, I show that the framework can recover the 'ground truth' parameters and effectively discriminate between Cumulative Prospect Theory (CPT) and a time dependent normative account of probability weighting.

Keywords: Decision Theory, Probability weighting, Prospect Theory, Bayesian Modelling

Preface

This paper is the product of the final project written to fulfill the requirements to obtain a master's degree in Mathematical Modelling and Computation from the Technical University of Denmark (DTU) by stud. polyt Benjamin Skjold Frederiksen.

The work was assigned a workload of 30 ECTS points ($\sim$ 840 working hours) and has been conducted from Aug. 3^{rd} 2020 to Jan. 3^{rd} 2021 at the Department of Applied Mathematics and Computer Science (DTU COMPUTE) in collaboration with the Danish Research Centre for Magnetic Resonance at Hvidovre Hospital, Denmark.

Benjamin Skjold Frederiksen

Acknowledgements

Thank you to my supervisors Tobias Andersen (DTU compute) and Oliver J. Hulme (Danish Research Centre for Magnetic Resonance), for providing guidance and feedback throughout this project. Also, thanks to other personnel at the Danish Research Center for Magnetic Resonance for ad hoc assistance. Lastly, a thank you to Mark Kirstein (London Mathematical Laboratory) for great discussions on the concept probability weighting, and the evolvement of decision theory, in general.

Contents

CHAPTER 1

Introduction

The fundamental principles of probability date back to the mid-1600s, when mathematicians Blaise Pascal and Pierre Fermat dealt with various gambling problems in an exchange of letters.[1] This discussion eventually gave rise to the concept of mathematical expectation, now known as *Expected Value* (EV) - in modern notation defined as:

$$EV = \Sigma p_i x_i, \tag{eq. 1.1}$$

where p and x are the probability and the amount associated with each possible outcome, respectively.

It was believed that any rational decision would optimize the EV, thus in a choice between gambles one must choose the option with the higher EV. To this day EV still serves as the normative null model in the prevailing (behavioral) economic models.[2] However, the notion "being rational" is not easily defined. This was shown with multiple examples - most famously by an example from Nicholas Bernoulli, now known as the St. Petersburg paradox [5]. In this experiment, the researcher tosses a (fair) coin. The initial stake is USD2 and for every time 'heads' appear the experiment goes to the next stage, where the experimenter repeats the toss, and the stake is doubled. The experiment ends the first time 'tails' appear and the player wins the stake that is in the pot at that stage. Formalizing this into the notion of eq. 1.1 yields $p_i = \frac{1}{2^i}$ and $x_i = 2^i$ and thus $\mathrm{EV} = \Sigma \frac{1}{2^i} \cdot 2^i = \Sigma 1 = \infty$.

[1] The meaning of probability is subject to a large philosophical debate. I do not seek to summarize this, nor by any means claim any interpretation more correct than others, but do (unless otherwise specified) simply use the *classical interpretation* and refer to Stanford Encyclopedia of Philosophy for a detailed overview [1].

[2] I note that recent developments in decision science have seen the emergence of models predicated on the concept of ergodicity, thus proposing a different null model. This discussion is, though, not deemed necessary in this context, but I refer to the work done by Peters and colleagues [2, 3, 4].

In other words, the EV of this experiment is infinitely large, so when a player is asked how much he/she is willing to pay to play, EV predicts a player to be willing to pay any amount. However, multiple studies, as well as common intuition, indicates that a player is willing to pay only a (relatively small) finite amount. I, therefore, introduce the term *seemingly rational* and define it by a person that behaves as expected according to the general consensus.

The solution that was widely accepted in the field was to keep the core of EV, but, in line with the theory of differential sensitivity [6], replace the objective outcomes with a subjective utility; a non-linear increasing function with respect to the outcome. This is known as *Expected Utility* (EU) - in modern notation defined as:

$$EU = \Sigma p_i u(x_i), \qquad \text{(eq. 1.2)}$$

where $u(x)$ defines the subjective utility function.

In the mid-1900s EU was axiomatized by Von Neumann and Morgenstern, which quickly made it the dominant normative model in decision science.[3] However, as with EV, researchers discovered inconsistencies between the predictions of EU and empirical evidence of behavior emerging from seemingly rational people; for example, why do people purchase both insurance policies (risk aversion) and lottery tickets (risk seeking) [9, 10]. Before long researchers revealed systematic violations of EU's axioms. Most importantly, the *independence axiom*, according to which outcomes common to all prospects (and with known probabilities) should have no influence on the decision [7, 8]. Most famously shown by the experimental work by Allias (1953) [11], which is now known as the *Allias paradox*. In these experiments, the decision maker was faced with two consecutive choices, with fixed outcomes, but changing probabilities. For example, consider:

Choice 1: Choose between

A:	\$300 with probability 1.00;	B:	\$400 with probability	.80,
			\$0 with probability	.20.

Choice 2: Choose between

A:	\$300 with probability .25,	B:	\$400 with probability	.20,
	\$0 with probability .75;		\$0 with probability	.80.

Here EU does not predict whether to choose A or B in each choice (this depends on the subjective value associated with USD300 and USD400, respectively). However, as the (potential) outcomes for A and B are fixed between choice 1 and 2, and the fact that the change in probabilities is a linear transformation (all probabilities in choice 2 can be written as the respective probabilities in choice A divided by 4), EU predicts that if a decision maker prefers A in choice 1, then he/she must also prefer A in choice 2 (and opposite if

[3] It is up for discussion whether EU can be classified as a normative model as the utility function by definition is subjective - this discussion is though not deemed relevant here. For a detailed description of the axiomatization, I refer to Fishburn (1992) [7] and Mas-Colell et al. (1995) [8].

B is preferred). Allias (1953) [11], though, show that the majority of the decision makers systematically violate this, preferring A in choice 1, but B in choice 2, and thus conclude "[...] the difference between probabilities of [say] 0.99 and 1.00 has more impact on preferences than the difference between [say] 0.10 and 0.11" [11]. A large number of models have since been developed to account for the regularities in decisions inconsistent with the predictions from EU. The vast majority of these adopt the same strategy as Bernoulli, to keep the general structure of EV, and again increase the model structure [12, 13].

1.1 Introducing probability weighting

One such attempt was to introduce what is known as probability weights (also referred to as *decision weights*). The concept of probability weighting originated in Prospect theory (PT; Kahnemann & Tversky (1979) [14]).[4] This was later updated to Cumulative Prospect Theory (CPT; Tversky & Kahneman (1992) [16]), which is arguably the most prominent model in the field today. The general idea of CPT is that not only the outcomes but also the probabilities are translated (potentially) non-linearly into subjective representations. Thus, CPT describes human decision making as maximizing a utility function, which for a gamble with outcomes x and corresponding probabilities p is given by:

$$U = \Sigma v(x_i) w(p_i), \qquad \text{(eq. 1.3)}$$

where $v(x)$ represents the subjective value (relative to a reference point) given by a *value function*, as known from EU. In CPT the value function is parameterized with two parameters; α and λ, which characterize outcome sensitivity and loss aversion, respectively:

$$v(x; \alpha, \gamma) = \begin{cases} (x)^\alpha & \text{if } x \geq 0 \\ -\lambda |(x)|^\alpha & \text{if } x < 0 \end{cases}. \qquad \text{(eq. 1.4)}$$

However, for this paper I intend to focus solely on the *probability weighting*, $w(x)$, and therefore do not go into further details of the value function, but refer to the original paper (Tversky and Kahneman (1992) [16], or for a detailed description, see Wakker (2010) [13]). Neither, do I consider any other aspects that are suggested in the literature, such as recency parameters (see Fox et al. (2006) [17] and Rakkow et al. (2008) [18]), or different gambling dynamics (see peters et al. (2019) [3] and Meder et al. (2019) [4]). However, I encourage any reader to draw on the framework I present and increase the complexity to investigate possible cross-effects.

Probability weighting defines a distortion of probabilities in the subjective representation that describes the discrepancies found within EU. As opposed to subjective utility, the consensus of expected behavior was, however, not changed, and thus seemingly rational people should not deploy probability weighting; "no theory of choice can be both normatively adequate and descriptively accurate" [19].

Several functional forms of the probability weighting function have been proposed. Here,

[4] It is noted that there is literature dating back to the late 1940s, where transformations of probability within decision science is considered - see e.g. Preston & Baratta (1948) [15]. The modern theories are, though, predicated on different ground, and it is therefore not commented further upon.

I only consider the two-parameter function originally proposed by Goldstein and Einhorn (1987) [20] (For others, see e.g. Tversky and Kahneman (1992) [16] or Prelec (1998) [21]):

$$w(p;\delta,\gamma) = \frac{\delta p^\gamma}{\delta p^\gamma + (1-p)^\gamma}, \qquad \text{(eq. 1.5)}$$

where γ is a *probability-sensitivity* parameter that primarily controls the curvature and δ is an *elevation* parameter, which can be interpreted as an indicator of the decision makers optimism/pessimism regarding gambles [22] - see Fig. 1.1 for a visual representation with different parameter settings.

1.2 Experiments - evaluating w(p)

After the publication of PT (and CPT), many experiments were conducted to estimate the parameters of the probability weighting function (and the value function). The general setup in most experiments has been to present the decision maker with a gamble of simple prospects with monetary outcomes and stated probabilities (known as *decisions from description*) and then ask if the decision maker would accept the gamble rather than a sure gain of various amounts. Thus inferring how highly the decision maker valued that specific gamble (see Tversky & Kahneman (1992) [16], Camerer & Ho (1994) [23], Tversky & Fox (1995) [22], Gonzales & Wu (1996) [24] Gonzales & Wu (1999) [25], see also Birnbaum (1996) [26] for a different approach). While the experiments present slightly different estimates of the parameters, the general trend of the results are very similar. Namely, that unlikely events are overweighted and likely events underweighted; resulting in the famous inverse-S shape (see Fig. 1.1, red/dashed line). However, within the last decades researchers have begun investigating a new experimental paradigm, where probabilities are derived from experience rather than described explicitly.[5] In many cases this is considered to better resemble choices made outside the laboratory, where "people often must make choices without a description of possible choice outcomes, let alone their probabilities." [28]. While this follows Knight's (1921) [29] terminology of distinguishing between *a priori probabilities*, where the probabilities of each event are explicit (decisions from description), and *statistical probabilities*, where the probabilities cannot be calculated exactly, but are accessed empirically (in modern litterateur known as *decisions from experience*),[6] it came as a surprise in the field when several studies revealed observed choices, which indicated not overweighting of rare events, but rather the opposite - as visualized by the blue/dash-dotted line in Fig 1.1 (see Barkan, Zohar, & Erev (1998) [30]; Barron & Erev (2003) [31]; Weber, Shafir & Blais (2004) [32]).

Hertwig et al. (2004) [28] formulates this with the same four-fold pattern introduced by Kahnemann and Tversky (1979, 1992) [14, 16], proposing that for decisions from experience rare events are underweighted rather than overweighted. This discrepancy between choices

[5] Deriving probabilities from experience was not per se a new approach (in fact this was much used in the 1950s and 1960s; for review see Luce & Suppes (1965) [27]), but is referred to as new as more modern researchers had turned away from this transient of learning for decades and the theories, therefore, were predicated solely on decisions from description.

[6] Knight (1921)[29] consider a third distinction: *estimates* which refer to unique situations, where there is no meaningful way of applying or empirically deriving the probability. This distinction is though not deemed important in this context.

emerging from decisions from description and decisions from experience is known as the *description-experience gab* [33]. A meta-analysis conducted by Wulff et al. (2018) [34], including nearly 70.000 decisions made from experience, report that the description-experience gab is a robust finding. However, in line with the criticism of the description-experience gab (see e.g. Fox and Hadar (2006) [17] and Rakow et al. (2008) [18]), Wulff et al. show that limitations of the experiments, such as reliance on small samples (and thus a high risk of sampling error) is an important contributor and further discuss how acquiring information in decisions from experience is subject to individual decisions from the decision maker, which might lead to bias and thus contribute to the gab. Lastly, Wulff et al. (2018) [34] show that the overall pattern of the probability weighting emerging from decisions from experience is very mixed: Hau et al. (2008) [35], Camilleri and Newell (2009) [36], Ungemach et al. (2009) [37] and Hau et al. (2010) [38], for example, report regular-S shaped functions, while Abdellaoui et al. (2011) [39], Glöckner et al. (2016) [40] and Kellen et al. (2015) [41] report Inverse-S. The discrepancies between the studies are heavily discussed, but can be summarized to be attributed to methodological differences; for example, there exist various paradigms within decisions from experience (sampling, partial feedback, and full feedback; see Hertwig et al. (2009) [33] for a description of the paradigms), and even within each paradigm there can be (and are) discrepancies. This includes, but is not limited to, the probabilities of the rare event, the problem's placement in the sequence, and the problem structure [34].

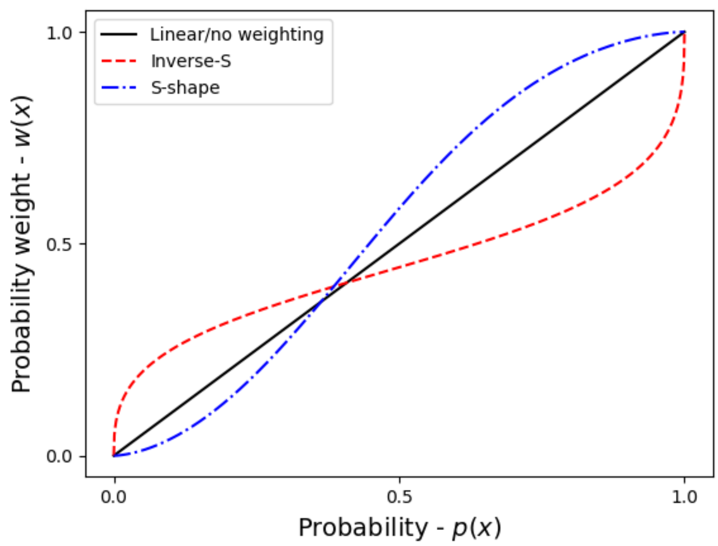


Fig. 1.1. Different shapes of the probability weighting function: red/dashed line corresponding to inverse-S (eq. 1.5; $\delta = 0.8, \gamma = 0.4$), black/solid line corresponding to linear/no weighting and blue/dash-dotted line corresponding to regular-S shape (eq. 1.5; $\delta = 1.4, \gamma = 1.6$)

1.3 Normative accounts for w(p)

In recent years, researchers have gone beyond this discussion and begun questioning the consensus that "no theory of choice can be both normatively adequate and descriptively accurate" [19]. This has led to the emergence of new models in which the non-linear transformation of probability is explained normatively. I here present one of these: Peters et al. (2020) [42], but highlight that other approaches have also been proposed (see e.g. Stewart, Chater & Brown (2006) [43] or Steiner & Stewart (2016) [44]).

Peters et al. (2020) propose merely that "[...] probability weighting means that a decision maker has greater uncertainty about the world than the observer". Thus, explaining probability weighting by the limited information that naturally occurs in a world where people neither have access to all information, have perfect memory nor infinite computational precision. They consider a *disinterested observer* (referred to as the observer), such as an researcher or theorist, and a *decision maker*, such as a subject in a study, who both for a given event have a probability distribution, which is denoted *probabilities* (expressed by $p(x)$), and *decision weights* (expressed by $w(x)$), respectively.

I divide the theory into two parts: 1) conceptualizing probability weighting as a disagreement of models between the observer and decision maker, and 2) how this disagreement naturally occurs in decisions from experience and how it can be modeled.

The conceptual claim is that probability weighting can be described simply by a *scale*, *location* and *shape*, which I, as they, illustrate this using a Gaussian distribution (the shape), where the location is the mean and the squared scale is the variance:[7] Assume an observer models a potential outcome x as a Gaussian with mean μ and variance σ^2 and a decision maker models the same potential outcome x, also as a Gaussian with same mean μ, but a different variance, $(\alpha\sigma)^2$ (here simplified to be proportional to the observer's scale by a factor α). With this example the distributions can be written explicitly:

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[\frac{-(x-\mu)^2}{2\sigma^2}\right], \quad w(x) = \frac{1}{\sqrt{2\pi(\alpha\sigma)^2}} \exp\left[\frac{-(x-\mu)^2}{2(\alpha\sigma)^2}\right], \qquad \text{(eq. 1.6)}$$

which yields the following expression for the decision weight as a function of probability:

$$w(p) = p^{\frac{1}{\alpha^2}} \frac{(2\pi\sigma^2)^{\frac{1-\alpha^2}{2\alpha^2}}}{\alpha}. \qquad \text{(eq. 1.7)}$$

Figure 1.2 illustrates the relationship between the observer and the decision maker's models, where the decision maker assumes a greater scale, $\alpha = 2$. In the left panel, the probability density function (PDF) of each model (eq. 1.6) is plotted, where the direct relationship of the PDFs of the models (eq. 1.7) is plotted in the right panel. Apparent from both panels, the decision maker's model over-estimates low probability outcomes and underestimates high probability outcomes with respect to the observer's model, illustrated by the vertical

[7] It is noted that this generalizes to arbitrary distributions, and is neither constrained to the observer and decision maker using the same distribution - illustrated by the authors using Student's t-distribution [42, pp.7-8].

arrows.[8]

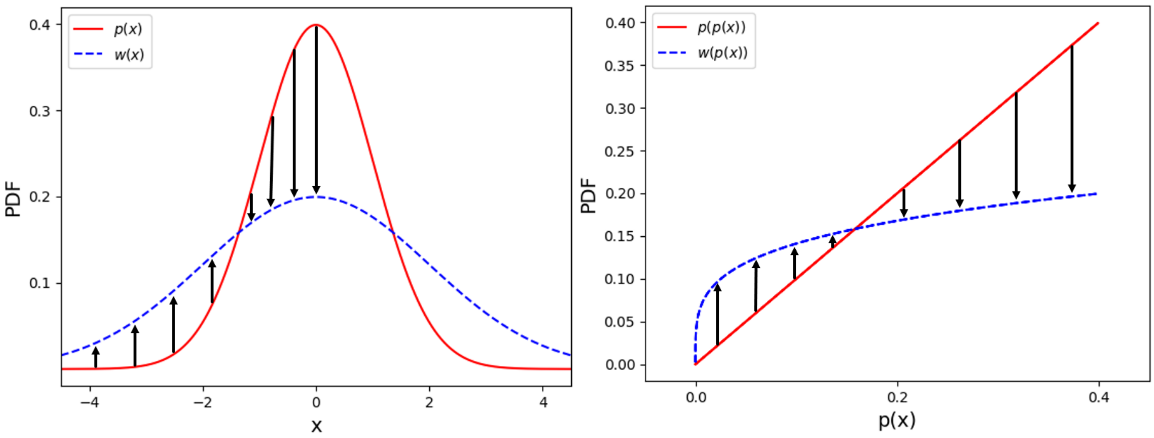


Fig. 1.2. PDF mapping of the relationship between the observer's model and the decision maker's model of probability: Left: Probability PDF ($p(x)$; red/solid), estimated by the observer, and decision weight PDF ($w(x)$; blue/dashed), estimated by the decision maker. Each representing what the observer and decision maker, respectively, assumes to be the true frequency distribution. Comparing the curves, the decision maker with a greater scale ($\alpha > 1$; here $\alpha = 2$) seems to over-estimate low probabilities and under-estimate large probabilities, with respect to the observer - indicated by the vertical arrows. Right: The same PDFs as in the left panel, but expressed by directly plotting (for any value of x), both against the probability PDF ($p(x)$).

An alternative way of expressing the same models is by their cumulative probability distributions (CDF). This is illustrated in Fig. 1.3, where (as in Fig. 1.2) the left and right panel corresponds to eq. 1.6 and eq. 1.7, respectively. This again illustrates an over-estimation of low probability outcomes and underestimation of high probability outcomes for the decision maker's model with respect to the reference model in both panels (again, illustrated by vertical arrows). It is also seen that an inverse-S shape emerges when investigating the CDF of the decision maker with respect to the CDF of the observer (right panel), which show that the findings that decision makers tend to overweigh small probability events and overweigh high probability events can be explained simply by the point of view. I note that for the remainder of this paper, I do not, per se, distinguish between the observer and the decision maker, but simply view all results from the observer's point of view.

Secondly, Peters et al. relate their concept of probability weighting into a testable model. They first define probability as the relative frequency of an event in an infinitely long time series of observations, though by no means stating that this, in general, is the correct definition of probability (see footnote 1). However, since real time series (by nature) have finite length, probabilities defined this way must be latent parameters in a model, which cannot actually be observed. But, from real time series, the best values as inputs to the model can be estimated, by counting the frequency of a given event. Thus, this model requires a time series of events as produced in decisions from experience.

As the probability of an event decreases, so does the number of times we see it (the frequency), but in series of finite length this relationship is not proportional; e.g. in a Bernoulli

[8] For $\alpha < 1$ the results are easily deduced. However, this represents the decision maker having less uncertainty than the observer, which is only applicable for very few and specific situations and is therefore not commented upon further.

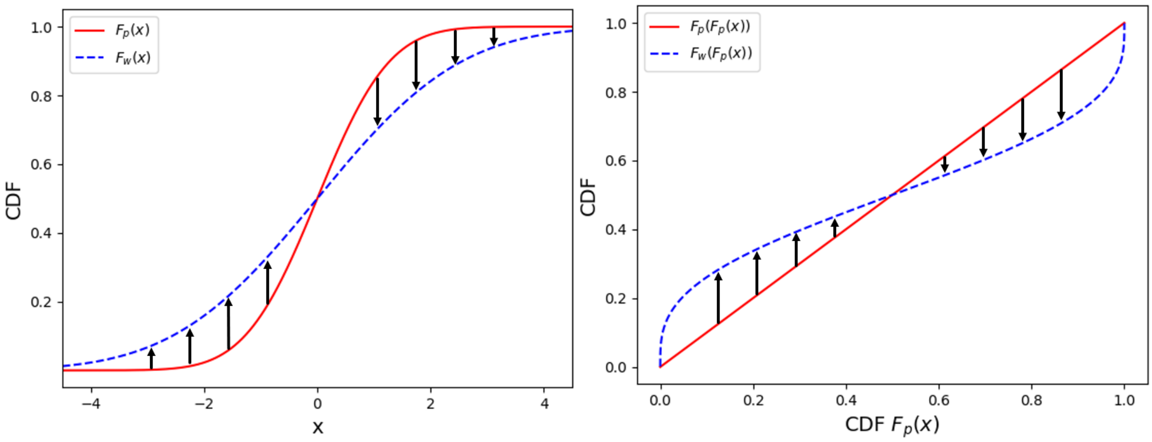


Fig. 1.3. CDF mapping of the relationship between the observer's model and the decision maker's model of probability: Left: Probability CDF ($F_p(x)$; red/solid), estimated by the observer, and decision weight CDF ($F_w(x)$; blue/dashed), estimated by the decision maker. Each representing the same models assumed in Fig. 1.2. Right: The same CDFs as on the left, but here plotted, not against x, but against the decision weight CDF $F_p(x)$, (representing the observer's worldview). Trivially, F_p plotted against it-self ($F_p(F_p(x))$) is the diagonal; the CDF F_w now displays the inverse-S shape known from, e.g. CPT. Note the curves coincide at 0.5 (both left and right panel), because no difference in location is assumed (I refer to Peters er al. (2020) [42, pp.6-7] for a discussion different scales and locations).

trial (which each individual gamble considered here is distributed as) the number of times we see an event is right-skewed and this skewness increases with smaller probability (and/or smaller sample size). Wulff et al. (2018) [34] studies this in decisions from experience and report a very clear tendency that the experienced frequencies are lower/higher than the true probabilities for low/high probability events. A decision maker, who must estimate probabilities from observations, should therefore acknowledge that rare events may be rather more common than data suggests. Such caution may dictate that the decision maker for rare events assigns higher probabilities than his/her estimate suggests.

Formalizing this yields that for an asymptotic probability density $p(x)$, the number of events $n(x)$ that is seen in an interval of T observations, is proportional to $n(x)$ and T. Using that such counts, for example, in a simple Poissonian case, are random variables whose uncertainty scale like $\sqrt{n(x)}$, an estimate for $p(x)$, can be defined as:

$$\hat{p}(x) \equiv \frac{n(x)}{T}, \qquad \text{(eq. 1.8)}$$

and the uncertainty;

$$\varepsilon[\hat{p}(x)] \equiv \sqrt{\frac{\hat{p}(x)}{T}} = \sqrt{\frac{n(x)}{T^2}}. \qquad \text{(eq. 1.9)}$$

Assuming the decision maker defines probability by the relative frequency and further is aware of this uncertainty and therefore adds the error to the estimate of the probability, the decision weight can be written as:

$$w(x) \quad = \frac{\hat{p}(x)+\varepsilon[\hat{p}(x)]}{\int_{-\infty}^{\infty} \hat{p}(s)+\varepsilon[\hat{p}(s)]ds} \qquad \text{(eq. 1.10a)}$$

$$= \frac{\frac{n(x)}{T}+\sqrt{\frac{n(x)}{T^2}}}{\int_{-\infty}^{\infty} \frac{n(s)}{T}+\sqrt{\frac{n(s)}{T^2}}ds}, \qquad \text{(eq. 1.10b)}$$

where the denominator is just a normalizing factor ensuring $w \in [0, 1]$.

An important observation in eq. 1.10 is that the error term is parameterized by T and thus scales like the number of observations. As T grows, $w(x)$ therefore approaches the asymptotic probability density, $p(x)$. In contrast to CPT, this model is thus dynamic with respect to time, suggesting that perfect information leads to linear/no weighting - illustrated in Fig. 1.4.

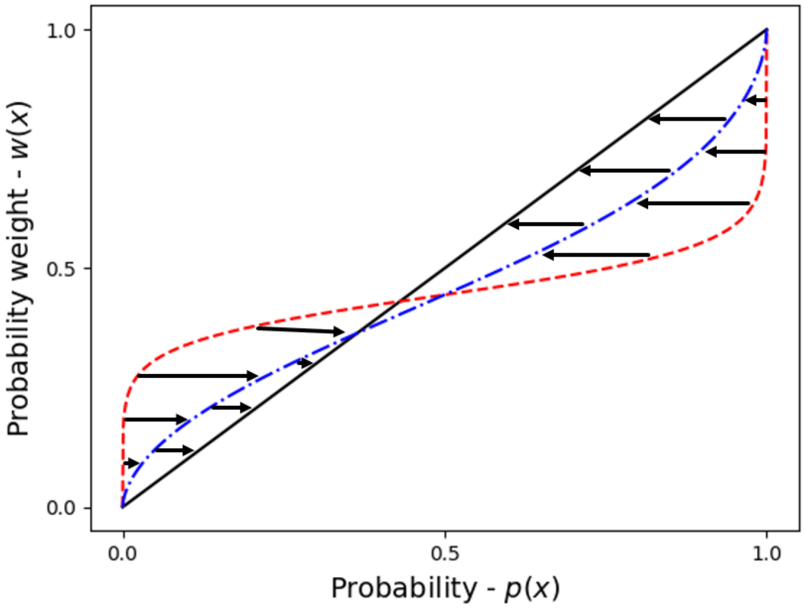


Fig. 1.4. Probability weighting function as predicted by Peters et al. (2020) [42] for increasing number of observations: The PWF plotted for different values of T; the red/dashed line representing small T, the blue/dashed-dotted line representing a larger T and the black/solid line representing (the theoretically) perfect information ($T = \infty$).

1.4 Thesis objective

The goal of this paper is to provide insights into the state of the art literature within probability weighting, by providing a modeling framework that effectively discriminates between qualitatively different models of probability weighting. To that end, I draw on the work done by Peters et al. (2020) (referred to as the LML model; [42]) as the normative model and the most prominent descriptive model in the field *Cumulative Prospect Theory* (CPT; [16]), which both characterize choice behavior using transformations of probability. I use the experimental paradigm of *decisions from experience* [28],[9] which leads me to address two sub-questions: First, do the subjective representations of probabilities in experience-based choices behave non-linearly - and if so how? Second, does this subjective representation change over time with respect to the actually experienced frequencies - and if so how?

As described, several studies have attempted to answer the first question. However, the studies have produced mixed results, because of methodological heterogeneity and/or limitations of the experiments, such as comparing subjective representations with the objective probabilities versus those experienced, aggregation of data, and reliance on small samples. In this paper, I provide an experimental framework that improves on these limitations, with no prior assumptions on the shape of the weighting function. To my knowledge, there is no systematic research on the second question. The prominent models all assume that the parameters are stable over time. However, this is by no means clear and in fact, one prediction of the LML-model is that the weighting function is dynamic over time. By exploiting the nature of how decisions in decisions from experience are made and drawing on the properties of Bayesian modeling, I seek to provide a modeling framework that captures the evolvement of the weighting function over time.

The paper is organized as follows: Firstly, I present the experimental setup. I then introduce the different components of the CPT- and LML-model. Finally, I present a simulation study that reveals systematic differences in the choices based on the different models, which ultimately ought to lead to an experiment with human subjects.

[9] For a normative account of probability weighting for *decisions from description*, which aligns with the conceptual claim presented in Peters et al. (2020) [42], I refer to Juechems et al. (2020) [45].

CHAPTER 2

Methods

2.1 Experimental design

I use the decisions from experience experimental paradigm of full feedback, which has the advantage of producing many decisions for each problem, as opposed to decision from description and the sampling paradigm, and the information available to the DM in each trial is independent of previous choices, as opposed to the partial feedback paradigm.

The payoff distributions behind each option in a choice are initially unknown. In each trial, the agents are presented with a choice between two options and after choosing, the agent is informed of the outcome for both options. Thus, each trial can be considered as an individual experiment for the decision maker, but each trial also provides information about the potential outcomes and respective relative frequencies the decision maker can expect later - this is illustrated in Fig. 2.1. For this experiment, I use an additive gambling dynamic, such that the decision makers wealth at time $i+1$ equals the wealth at time i plus the outcome at time i, i.e. $W_{i+1} = W_i + x_i$, with no endowment, i.e. $W_0 = 0$.

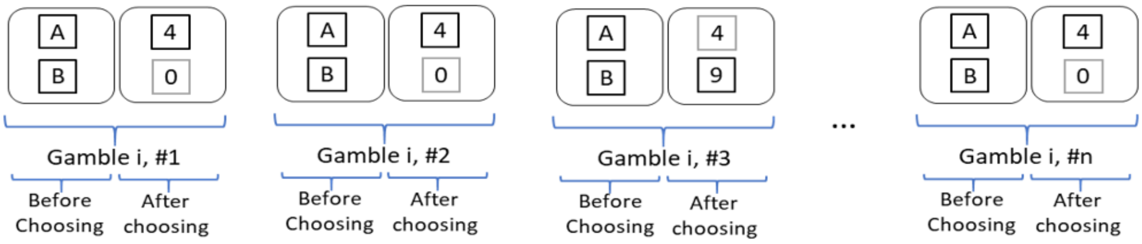


Fig. 2.1. Experimental design: An example sequence of n choices for an underlying gamble (here $(9, 0.2; 4)$), where in each trial the agent must choose between A and B, and after choosing is presented with the outcomes for both choices.

2.1.1 Gambles

I consider 24 qualitatively different gamble set, meaning 24 gamble set with different underlying probability distributions, defined as set of two options: Option A being a gamble and option B the EV of that gamble (rounded to the nearest integer). Here, I define a gamble to be a simple prospect, consisting of two (monetary) outcomes x_1, x_2 with corresponding underlying probabilities $p_1, 1 - p_1$. Written in the form $(x_1, p_1; x_2)$. Tab. 2.1 illustrates the 24 gamble set that are considered. These are chosen, as the gambles are the exact gambles used when developing the CPT-model - the gambles are all the gains only gambles[1] (see Tversky and Kahnemann (1992) [16, Tab. 3]).

For each of the 24 qualitatively different gamble set, the gambles are transformed into a series of gambles of length $n_T = 100$; in each iteration (n_T times) drawing a (pseudo-)random number from a uniform distribution to generate an outcome for option A, O_A. Note that as option B is defined as the EV of O_A and thus is just a series of $EV[O_A]$ with length n_T. To ensure the underlying distributions are well represented by the series, I check that the difference between underlying probability and the frequency distribution is within a threshold set to 0.01. This is repeated for each agent, such that all agents face different sequences with the same underlying probability distribution. I, though, note that I consider multiple *species*, which can be interpreted as different agent-types behaving according to different models (see. section 2.3.1 for further explanation). Across these species the sequence for agent i is identical.

	Probability								
Outcomes	~~.01~~	.05	.10	.25	.50	.75	.90	.95	~~.99~~
(0,50)			5		25		45		
(0,100)		5		25	50	75		95	
(0,200)	~~2~~		20		100		180		~~198~~
~~(0,400)~~	~~4~~								~~396~~
(50,100)			55		75		95		
(50,150)		55		75	100	125		145	
(100,200)		105		125	150	175		195	

Table 2.1: **Gamble set:** Sure option (option B) for all respective gambles (option A). The two outcomes of option A are given in the left-hand side of each row; the probability of the second (i.e. higher) outcome is given by the corresponding column. For example, the value of 5 in the upper left corner is option B corresponding to (50, 0.1;0) as option A. Note these are retrieved from Tversky & Kahnemann (1992) [16] - crossed out gambles are the four left out.

[1] except the gambles where $p_1 = 0.01$ or $p_1 = 0.99$, as these resulted in relative error between the frequency distribution and underlying probability distribution outside an acceptable range.

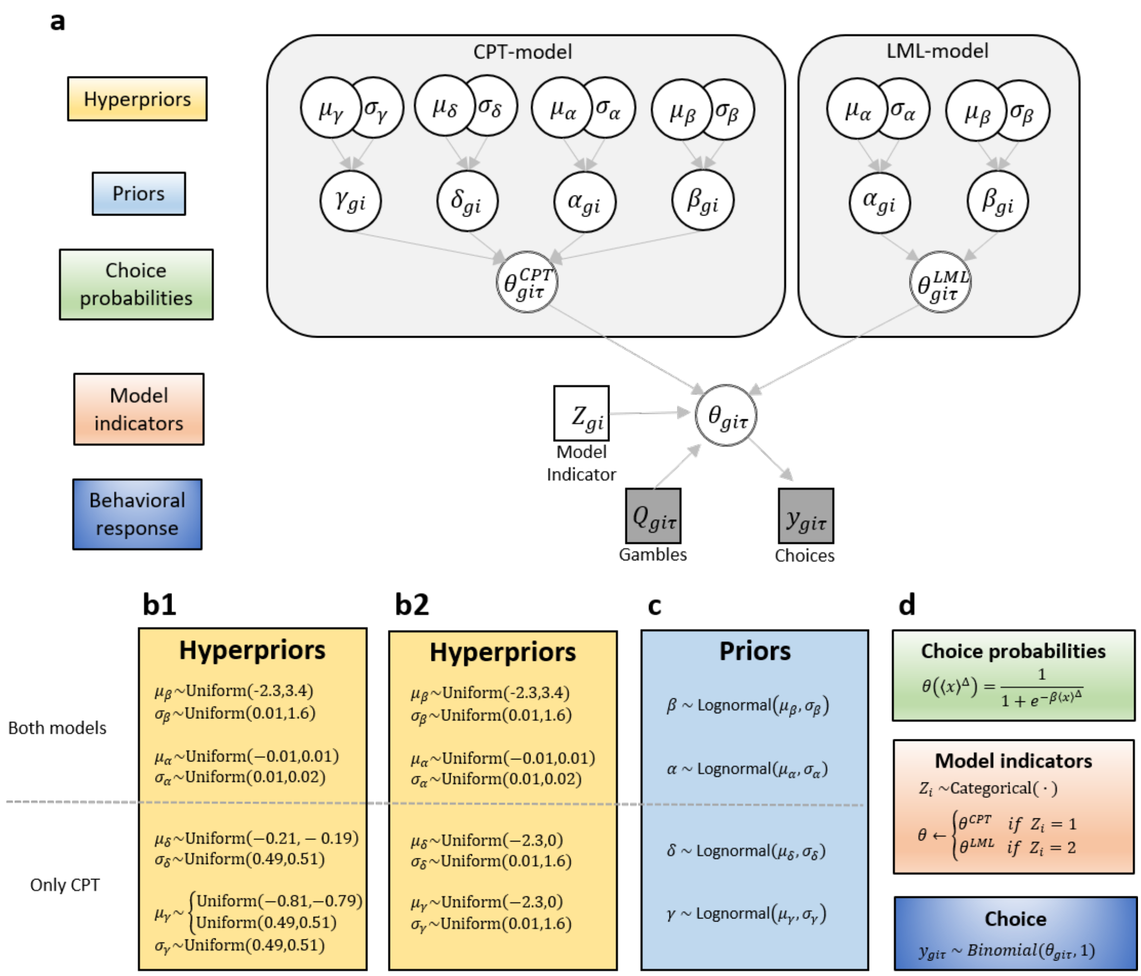


Fig. 2.2. Hierarchical Bayesian latent mixture model. a, graphical representation of hierarchical Bayesian model for estimating latent mixtures for two different models. Circular nodes denote continuous variables, square nodes discrete variables; shaded nodes denote observed variables, unshaded nodes unobserved variables; single bordered nodes denote stochastic variables, double bordered nodes deterministic variables. Labels on the left describe the role of each parameter in the model. **b1,** hyperprior distributions used when simulating choices; interval in uniform distributions set very narrow, thus acting as Dirac delta functions. **b2,** hyperprior distributions used for model- and parameter recovery. **c,** prior distributions for both simulating choices, model recovery and parameter recovery. **d,** choice functions and choice generating distributions.

2.2 Models

2.2.1 Model space

The following models are described by specifying the probability weighting function, as well as a stochastic choice function [42, 46].[2] Through the following sections, by probability, I mean the observed frequency and thus define $\hat{p}(x) = \frac{n(x)}{T}$.

Cumulative Prospect Theory (CPT) weighting function as presented in eq. 1.5 with two-parameters:

$$w(x;\delta,\gamma) = \frac{\delta\hat{p}(x)^\gamma}{\delta\hat{p}(x)^\gamma + (1-\hat{p}(x))^\gamma}, \quad \text{(eq. 2.1)}$$

where γ is a probability-sensitivity parameter and δ an elevation parameter.

LML-model (LML) as presented in eq. 1.10:

$$w(x) = \frac{\hat{p}(x) + \varepsilon[\hat{p}(x)]}{\hat{p}(x) + \varepsilon[\hat{p}(x)] + (1 - (\hat{p}(x) + \varepsilon[\hat{p}(x)]))}. \quad \text{(eq. 2.2)}$$

Note that this weighting function does not have any latent parameters.

Expected value. The expected utility as presented in eq. 1.3 (with linear utility) is calculated for each option (done identically for each of the two models), by:

$$\langle x^A \rangle = w(\hat{p}_1^A) \cdot x_1^A + w(1-\hat{p}_1^A) \cdot x_2^A, \quad \text{(eq. 2.3)}$$

and equivalent for option B.

I denote the difference between the option A and option B by Δ and thus the difference for each choice can be written as:

$$\langle x \rangle^\Delta = \langle x^A \rangle - \langle x^B \rangle. \quad \text{(eq. 2.4)}$$

Stochastic choice function. The stochastic choice function is likewise identical for the two models and is comprised of a logistic function with sensitivity parameter β:

$$\theta\left(\langle x \rangle^\Delta\right) = \frac{1}{1+e^{-\beta\langle x \rangle^\Delta}}, \quad \text{(eq. 2.5)}$$

where θ evaluates to the probability of choosing option A.

2.2.2 Sampling procedure.

The latent (i.e. unobserved) parameters are estimated by hierarchical Bayesian modeling, which affords computation of full probability distributions rather than point estimates that ignore the uncertainty in the estimates [47]. Further, via a hierarchical structure,

[2] I deploy linear utility to focus solely on probability weighting and note that within the literature it is argued that the probability weighting is qualitatively unaffected by the utility function [16].

individual parameter estimates are provided from group level distributions, such that estimations of individuals are informed from the group and extreme values that might be estimated from uncertainties are constrained [48]. The sampling is done with Monte-Carlo Markov Chain (MCMC) sampling via JAGS (v4.03), called from MATLAB™ (v9.5.0.944444 (R2018b), Mathworks®, `mathworks.com`) via the interface MATJAGS (v1.3, `psiexp.ss.uci.edu/research/programs_data/jags`). For all simulations (i.e. simulating choices, model recovery, and parameter recovery), I estimate the two models using a hierarchical latent mixture (HLM) model. Fig. 2.2a illustrates a graphical representation of the HLM model. In Fig 2.2b-d the distributional and structural equations are listed. These are, though, subject to change depending on whether it is for simulating choices, model recovery, or parameter recovery and are thus described in the respective sections. I note that the CPT- and LML-model technically are only sub-models of the HLM model, but refer to them simply as models for simplicity and consistency. For simulating choices, I use just one chain and one sample resulting in one choice per agent per simulated trial. Where I, for model recovery and parameter recovery, use four independent chains, 10^3 samples per chain and include a burn-in > 500. Convergence was established via monitoring R-hat values between 0.99 to 1.01 [49, 50].

2.3 Simulation study

2.3.1 Simulating choices

To evaluate whether the model estimation methods are capable of recovering the models, as well as the parameters, I simulate choices from synthetic agents. These agents come from group level distributions, for which 'ground truth' parameter values are set a priori using Dirac delta functions.[3] I simulate three different species, each containing 10 agents; two species that has the CPT-model as the 'ground truth' model, where one typically shows the inverse-S weighting function and the other regular-S, and one species that has the LML-model as the 'ground truth' model. This is obtained by running the HLM model three times with no choice input and prior distribution on the indicator variable, Z, set to $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$, respectively (see Fig. 2.2d).[4] For the sensitivity parameter used in the stochastic choice function, I follow Nilsson et al. [48] and assume the group mean for β to lie in the interval $(0, 30)$ and therefore set $\mu_\beta \sim \text{Uniform}(-2.3, 3.4)$ and the standard deviation to $\sigma \sim \text{Uniform}(0.01, 1.6)$. For the weighting function parameters (only in the CPT model), I fix the prior distributions such that for species one the mean values follow approximations from the literature providing a inverse-S shaped weighting function [24, 25] and the species two I fix the mean, such that a regular-S shaped weighting function is expected [35]. Hence, I assume $\delta = 0.8$, $\gamma_1 = 0.5$ and $\gamma_2 = 1.5$, and therefore set $\mu_\delta = -0.2$, $\mu_{\gamma_1} = -0.8$, $\mu_{\gamma_2} = 0.5$ and further set $\sigma_\delta = \sigma_{\gamma_1} = \sigma_{\gamma_2} = 0.5$ - Fig. 2.2b1.

[3] In practice the Dirac delta functions are achieved using very narrow intervals on uniform distributions.

[4] Note that the only difference between the two species with the CPT-model as the 'ground truth' model is the prior on the mean of the sensitivity parameter, γ. Thus, they are treated one species for recovery.

2.3.2 Model and parameter Recovery

I perform both model and parameter recovery by estimating the HLM model on the basis of the simulated choices. For both, I follow Nilsson et al. [48] and set weakly informed hyperpriors. As when simulating choices, I assume β to have common group level distribution across models, while δ and γ are only applied to the CPT-model. I note that the parameter-space for all priors lie in $\mathbb{R}^+$ and thus I assume each to come from uninformative log-normal distributions with uninformative priors on group mean, μ, and variance, σ^2, which are both drawn from uniform distributions. I assume group mean for β to lie in the interval $(0, 30)$ and therefore set $\mu_\beta \sim \text{Uniform}(-2.3, 3.4)$ and the standard deviation to $\sigma_{beta} \sim \text{Uniform}(0.01, 1.6)$. For both the elevation parameter, δ, and sensitivity parameter, γ, I assume the group means to lie in the interval $(0, 2)$; $\mu_\delta \sim \mu_\gamma \sim \text{Uniform}(-2.3, 0)$, $\sigma_\delta \sim \sigma_\gamma \sim \text{Uniform}(0.01, 1.6)$ - Fig. 2.2b2.

Parameter recovery is performed separately for each species. In all instances, though, I estimate the CPT parameters, as the CPT weighting function is predicatively uninformative and ranges the full parameter-space, i.e. all shapes from very regular-S shape to very inverse-S shape (see Fig. 1.1). To evaluate the potential evolvement of the weighting function over time, I estimate the parameters for three experiments of various length (I refer to these as chunks 1-3), such that chunk 3 is the full experiment of $n = 100$ trials, chunk 2 is a subset consisting of the first $m^* = 50$ trials, and chunk 1 is a further subset consisting of only the first $m = 10$ trials. By estimating the parameters independently for each chunk, parameters are estimated for a various number of trials and thus I am able to evaluate the evolvement of the weighting function over time.

Model recovery is performed by monitoring the latent mixtures. These are modeled via indicator variables, which allow comparison between the qualitatively different models with the HLM model, by integrating over all parameters in each model (with respect to the respective priors) and penalizing for increasing model structure [47]. For model recovery, the indicator variable is set with uninformed uniform priors but is free to vary across agents. The likelihood for an agent to be best represented by model i is calculated with Bayes Theorem: $p(Z = i|D) \propto p(D|Z = i)p(Z = i)$ [47]. As the indicator variable is set uninformed, i.e. $p(Z = i) = p(Z = j)$, this gives the likelihood for an agent to be best represented by model i to be proportional to the posterior distribution of the indicator variable. The posterior model probabilities and Bayes factors are estimated via the Variational Bayesian Analysis toolbox (`mbb-team.github.io/VBA-toolbox/`) [51].

2.3.3 Data and code availability.

All datasets and code needed to replicate the findings presented here are available in my 'Master-thesis' repository: `github.com/benj1003/Master-thesis`.

CHAPTER 3

Results

3.1 Parameter recovery

Maximum a posteriori probability estimates are strongly correlated with 'ground truth'/predicted parameters. Initially, I investigated if the HLM model was able to recover the 'ground truth'/predicted parameters for both models. For the CPT-agents, the 'ground truth' parameters for all agents are known, as the choices for the agents are simulated based on the same parameters; δ_g, γ_g. I, therefore, examined the correlation between the 'ground truth' parameter and the corresponding maximum a posteriori probability (MAP) estimates; δ_e, γ_e. Fig. 3.1 illustrates this in a scatter plot of the two weighting function parameters (left panel: elevation parameter, δ; right panel: sensitivity parameter, γ) for each of the 3 chunks, represented by different colors. Here, it is evident that the MAP estimates and the 'ground truth' parameters are strongly correlated (Pearson correlation coefficient of: $\mathrm{pc}_\delta^{(1)} = 0.997, \mathrm{pc}_\delta^{(2)} = 0.997$, $\mathrm{pc}_\delta^{(3)} = 0.999, \mathrm{pc}_\gamma^{(1)} = 0.967, \mathrm{pc}_\gamma^{(2)} = 0.978, \mathrm{pc}_\gamma^{(3)} = 0.994$), which is strong evidence that the HLM model can recover the parameters for the CPT-agents.

Alternatively, the estimation error can be expressed by investigating the difference between the 'ground truth'/predicted weighting function (which I denote $\bar{w}(x)$) and the estimated weighting function, $w(x)$. An advantage of this approach is that it is possible to evaluate how well the parameters are recovered for the LML-agents, by evaluating against the predicted weighting function (eq. 1.10). Thus, I transform all estimated (and 'ground truth') parameters into a weighting function by eq. 1.5 and evaluate $\bar{w}(x) - w(x)$. Fig. 3.2 illustrates the average estimation error (including 95 % confidence interval) for all agents within each 'ground truth' model. It is seen that there is a systematic error for both models, for

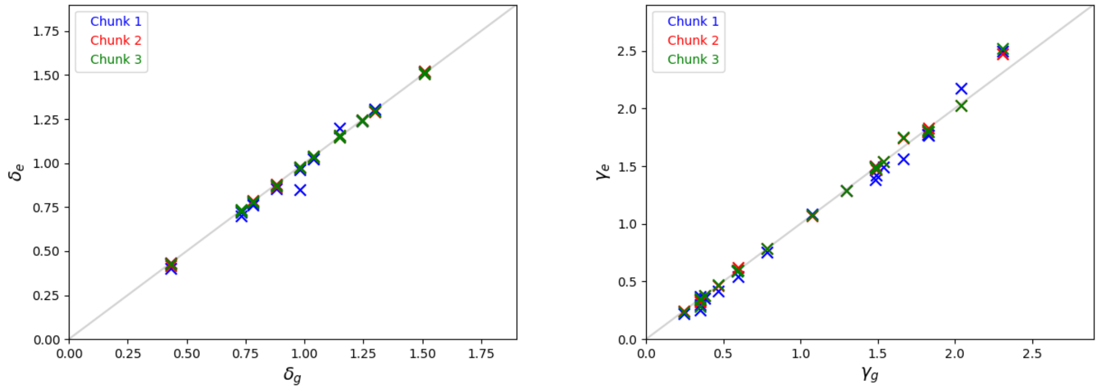


Fig. 3.1. Estimated parameters against 'ground truth' parameters for CPT-agents: left: Estimated elevation parameter (δ) plotted against the 'ground truth' parameter. Right: Estimated sensitivity parameter (γ) plotted against the 'ground truth' parameter. Both: light-grey line representing perfect correlation; different chunks represented by different colors; agents 1-4 and 11-14 highlighted by different markers.

which the origin, though, is unknown. However, the error in both models is of such low magnitude that it is deemed neglectable for the purpose of this paper. Therefore this is considered evidence that the HLM model is able to recover the 'ground truth'/predicted parameters for both the CPT- and LML-model. In Appendix A.1 and A.2 the weighting functions for all agents are plotted individually, which confirm that the model recovers the parameters with low estimation error for all agents.

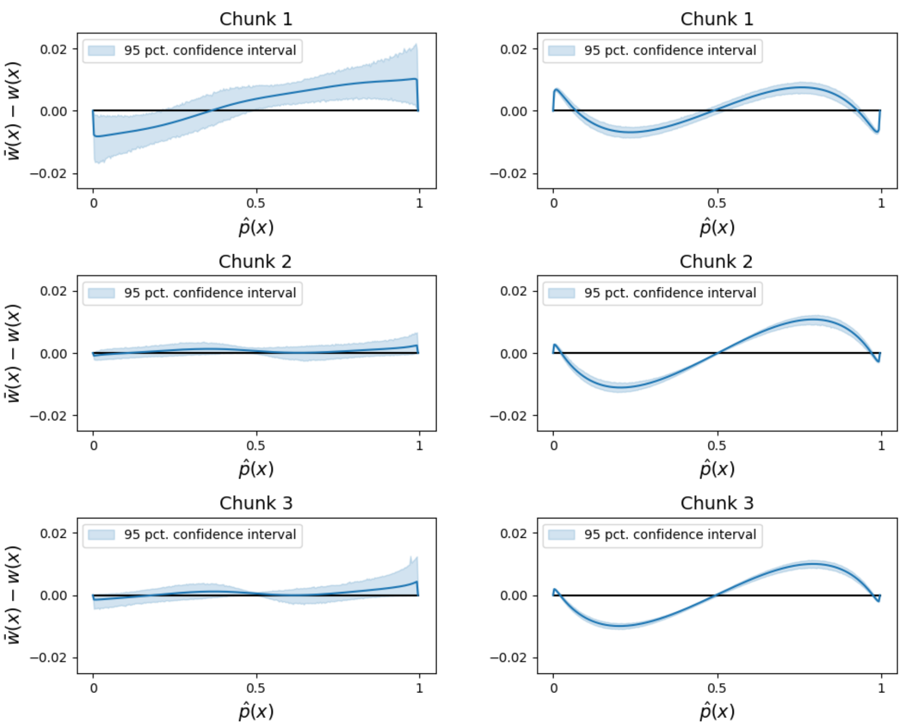


Fig. 3.2. Estimation error as a function of $\hat{p}(x)$: Left: Difference between estimated weighting function and 'ground truth' weighting function averaged over the 20 CPT-agents, including 95 % confidence interval. Right: Difference between estimated weighting function and predicted weighting function averaged over the 10 LML-agents, including 95 % confidence interval.

Parameter estimation reveal static weighting functions for CPT-agents and dynamic weighting functions for LML-agents. I next examined the estimated parameters in weighting function parameter space, i.e. γ_e against δ_e (eq. 1.5) - this is illustrated in Fig. 3.3 (Left panel: CPT-agents; Right panel: LML-agents). Again, each chunk is colored by different colors, and further, the first 4 agents within each species are highlighted by different shapes, such that one can examine these agents individually. Here, two important observations are made: 1) the CPT-model (left panel) predicts a course of variety of agents, while the LML-model (right panel) has a very narrow predictive mass, and 2) when investigating the parameters for each agent, it is clear that for the CPT-agents (left panel zoom) there is no systematic bias in the parameters across the different chunks, while for the LML-agents (right panel zoom) there is a clear tendency for the sensitivity parameter, γ, to increase towards 1, which is equivalent to 'moving' towards linear/no weighting, as illustrated in Fig. 1.4. This shows that the framework provided is able to capture the differences in model structure between the two models, as well as the evolvement of parameters and thus weighting function over time.

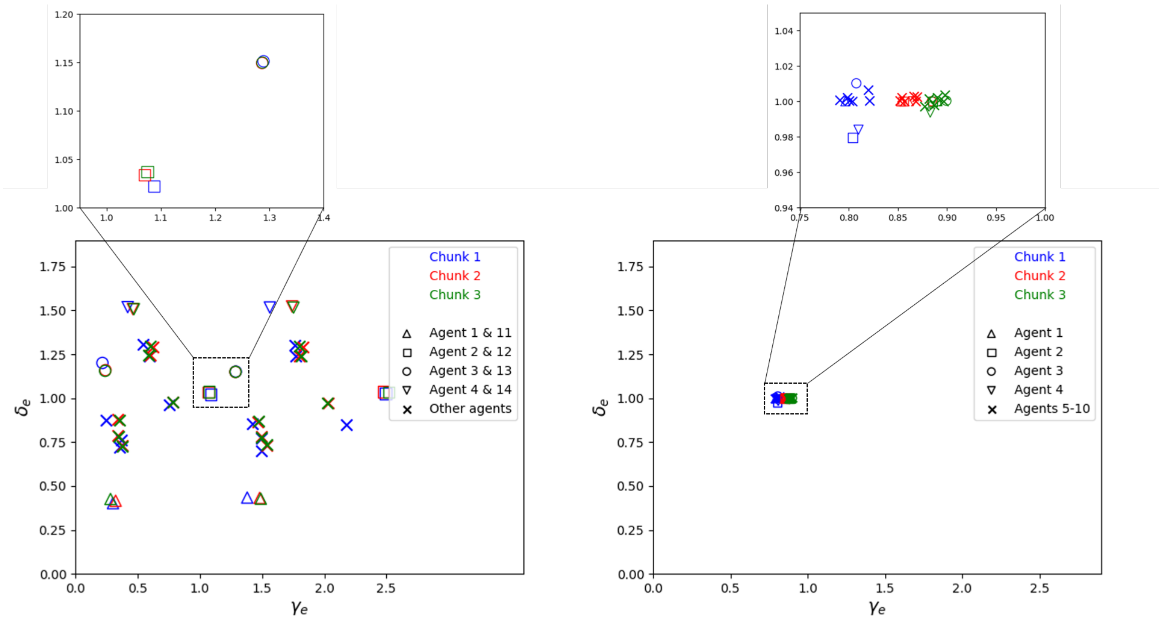


Fig. 3.3. Weighting function (eq. 1.5) parameter space for both models: Left: Estimated parameters for all 20 CPT-agents, showing a course of variety of agents being predicted; zoom showing no systematic bias across chunks for two arbitrary agents. Right: Estimated parameters for all 10 LML-agents, showing very narrow predictive mass; zoom showing a clear systematic bias of $\gamma \to 1$ for $T \to \infty$ as expected, illustrated in Fig. 1.4.

3.2 Model recovery

Bayesian model selection correctly identifies the 'ground truth' model. Lastly, I compare the predictive accuracy of the two models on the basis of the simulated choices. MCMC sampling of the HLM model (see Fig. 2.2a) results in posterior frequencies for the model indicator variable, which is interpreted as posterior probabilities for each model, estimated for each agent. Evident from Fig. 3.4, all agents have all of their probability mass located over their 'ground truth' model ($BF_{\text{CPT-CPT}}$=41, $BF_{\text{LML-LML}}$=21). Thus, there is strong evidence that the framework provided correctly identifies the 'ground truth' model [52].

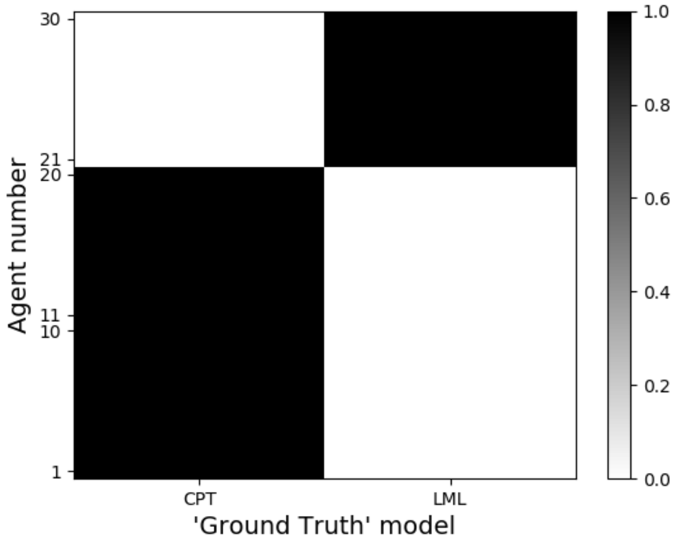


Fig. 3.4. Posterior model probabilities for each model based on the indicator variable; 'Ground truth' model for Agent 1-10 being CPT (inverse-S); Agent 11-20 CPT (regular-S); Agent 21-30 LML

CHAPTER 4

Discussion

By manipulating the fundamental properties of the probability weighting function, I show that the modeling framework provided can effectively discriminate between different models of probability weighting and further is able to capture the potential systematic bias in the weighting function over time. Switching between the CPT- and the LML-model in a simulation study reveals the expected behavior of high model structure and stable parameters over time for the CPT-model, and low model structure and dynamic parameters over time for the LML-model.

Much criticism in decisions from experience focuses on sampling error and/or lack of independence between trials. I eliminate these limitations by 1) investigating the weighting function as a function of the observed frequencies, rather than the objective probabilities, and 2) using the full feedback paradigm, where the agents are presented with all possible outcomes after each trial.

The length of the gamble sequences ($n_T = 100$) was constrained, such that a wide range of underlying gambles ($n_g = 24$) could be tested and still provide results in this simulation setting that are compatible with results from an experiment with human subjects. This has the drawback that extreme probability events had to be discarded, as the relative error between the frequency distribution and underlying probability would not be within an acceptable range (for intuitive example of this issue I refer to Peters et al. (2020) [42, p.10]). This might be the reason for the systematic error that is seen in Fig. 3.2 and might be worth investigating further. However, for the purpose of this paper, it was deemed that the potential increase in accuracy these gambles would add would not outweigh the cost associated with increasing the number of trials. Likewise, it was deemed sufficient, to only consider the MAP estimates, as this suffice for a proof of concept framework. However, Bayesian modeling affords full a posterior probability distributions for all parameters. It would therefore be an extension to the work presented here to investigate the full distributions, which potentially could provide further insights. Lastly, Peters et al. (2020) [42] distinguish between an observer and a decision maker, who both models the potential outcomes with uncertainty. For simplicity and to get results directly compatible with the prominent litterateur, it was decided to not make this distinction, by modeling the observer without any uncertainty. However, in a finite and dynamic world, perfect information is reserved for very few cases and thus it is deemed a very interesting way to model probabilities (and probability weighting), which could be experimentally tested for example by considering two agents in an experiment, where one agent receives more information (the observer) than the other (the decision maker).

The choice of not including various aspects of utility theory that are heavily investigated in the literature (e.g. non-linear utility and/or loss aversion) is not seen as a limitation of the modeling framework, but rather a way to investigate probability weighting using the simplest possible setting. One limitation in the provided framework is, though, that it is assumed that agents have perfect memory over all n_T trials, which must be considered unrealistic in an experiment with human subjects. It is therefore deemed necessary to include a memory loss/recency parameter to provide an accurate modeling framework for experiments with human subjects. This could for example be done as it is often seen in marketing modelling [53].

Models of probability weighting are predominantly developed under the assumption that non-linear weighting functions are irrational and typically discussed based on their ability to describe, rather than explain, observed behavior. This framework allows researchers to go beyond discussing the exact parameters of some arbitrary function, and instead the actual reasons behind the concept of probability weighting. This should motivate the need for further investigations into normative accounts of probability weighting, as well as stringent experimental studies of these theories.

References

[1] Alan Hájek. “Interpretations of Probability”. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Fall 2019. Metaphysics Research Lab, Stanford University, 2019.

[2] Ole Peters and Alexander Adamou. *Ergodicity Economics*. Lecture notes - `https://ergodicityeconomics.com/lecture-notes/`. 2017.

[3] Ole Peters. “The ergodicity problem in economics”. In: *Nature Physics* 15.12 (Dec. 2019), pp. 1216–1221.

[4] David Meder et al. *Ergodicity-breaking reveals time optimal economic behavior in humans*. preprint at `https://arxiv.org/abs/1906.04652`. 2019. arXiv: `1906.04652 [econ.GN]`.

[5] Martin Peterson. “The St. Petersburg Paradox”. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Fall 2020. Metaphysics Research Lab, Stanford University, 2020.

[6] George A Gescheider. *Psychophysics : the fundamentals / George A. Gescheider*. eng. 3rd ed. Mahwah, N.J.: L. Erlbaum Associates, 1997.

[7] Peter C. Fishburn. *The foundations of expected utility*. Vol. 31. Theory and Decision Library. Kluwer Boston, Inc., 1982.

[8] Andreu Mas-Colell, Michael D. Whinston, and Jerry R. Green. *Microeconomic Theory*. OUP Catalogue 9780195102680. Oxford University Press, 1995.

[9] Milton Friedman and L. J. Savage. “The Utility Analysis of Choices Involving Risk”. In: *Journal of Political Economy* 56.4 (1948), pp. 279–304. eprint: `https://doi.org/10.1086/256692`.

[10] Harry Markowitz. “The Utility of Wealth”. In: *Journal of Political Economy* 60.2 (1952), pp. 151–158. eprint: `https://doi.org/10.1086/257177`.

[11] M. Allais. “Le Comportement de l’Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l’Ecole Americaine”. In: *Econometrica* 21.4 (1953), pp. 503–546.

[12] R. Luce. *Utility of Gains and Losses*. New York: Psychology Press, 2000.

[13] Peter P. Wakker. *Prospect Theory: For Risk and Ambiguity*. Cambridge University Press, 2010.

[14] Daniel Kahneman and Amos Tversky. "Prospect Theory: An Analysis of Decision under Risk". In: *Econometrica* 47.2 (1979), pp. 263–291.

[15] Malcolm G. Preston and Philip Baratta. "An experimental study of the auction-value of an uncertain outcome." In: *The American Journal of Psychology* 61 (1948), pp. 183–193.

[16] A. Tversky and D. Kahneman. "Advances in prospect theory: Cumulative representation of uncertainty." In: *Journal of Risk and Uncertainty* 5 (1992), pp. 297–323.

[17] Craig R. Fox and Liat Hadar. ""Decisions from experience" = sampling error + prospect theory: Reconsidering Hertwig, Barron, Weber & Erev (2004)." In: *Judgment and Decision Making* 1.2 (2006), pp. 159–161.

[18] Tim Rakow, Kali A. Demes, and Ben R. Newell. "Biased samples not mode of presentation: Re-examining the apparent underweighting of rare events in experience-based choice". In: *Organizational Behavior and Human Decision Processes* 106.2 (2008), pp. 168–179.

[19] Amos Tversky and Daniel Kahneman. "Rational Choice and the Framing of Decisions". In: *The Journal of Business* 59.4 (1986), S251–S278.

[20] William M. Goldstein and Hillel J. Einhorn. "Expression theory and the preference reversal phenomena." In: *Psychological Review* 94.2 (1987), pp. 236–254.

[21] Drazen Prelec. "The Probability Weighting Function". In: *Econometrica* 66.3 (1998), pp. 497–527.

[22] Amos Tversky and Craig R. Fox. "Weighing Risk and Uncertainty". In: *Psychological Review* 102.2 (1995), pp. 269–283.

[23] Colin F. Camerer and Teck-Hua Ho. "Violations of the betweenness axiom and nonlinearity in probability". In: *Journal of Risk and Uncertainty* 8.2 (Mar. 1994), pp. 167–196.

[24] Richard Gonzalez and George Wu. "Curvature of the Probability Weighting Function". In: *Management Science* 42.12 (1996), pp. 1676–1690.

[25] Richard Gonzalez and George Wu. "On the Shape of the Probability Weighting Function". In: *Cognitive Psychology* 38.1 (1999), pp. 129–166.

[26] Michael H. Birnbaum and William Ross McIntosh. "Violations of Branch Independence in Choices between Gambles". In: *Organizational Behavior and Human Decision Processes* 67.1 (1996), pp. 91–110.

[27] R. D. Luce and P. Suppes. "Preference, utility, and subjective probability". In: *R. D. Luce, R. R. Bush, & E. Galanter (Eds.)* Vol. 3. New York, NY: Wiley: Handbook of mathematical psychology, 1965. Chap. 19, pp. 249–410.

[28] Ralph Hertwig et al. "Decisions from Experience and the Effect of Rare Events in Risky Choice". In: *Psychological Science* 15.8 (2004). PMID: 15270998, pp. 534–539. eprint: `https://doi.org/10.1111/j.0956-7976.2004.00715.x`.

[29] Frank H. Knight. *Risk, uncertainty and profit*. Boston, New York, Houghton Mifflin Company, 1921.

[30] Rachel Barkan, Dov Zohar, and Ido Erev. "Accidents and Decision Making under Uncertainty: A Comparison of Four Models". In: *Organizational Behavior and Human Decision Processes* 74.2 (1998), pp. 118–144.

[31] Greg Barron and Ido Erev. "Small feedback-based decisions and their limited correspondence to description-based decisions". In: *Journal of Behavioral Decision Making* 16.3 (2003), pp. 215–233. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/bdm.443.

[32] E. Weber, S. Shafir, and Ann-Renée Blais. "Predicting Risk-Sensitivity in Humans and Lower Animals: Risk as Variance or Coefficient of Variation". In: *Psychological review.* 2004.

[33] Ralph Hertwig and Ido Erev. "The description–experience gap in risky choice". In: *Trends in Cognitive Sciences* 13.12 (2009), pp. 517–523.

[34] Dirk U. Wulff, Max Mergenthaler-Canseco, and Ralph Hertwig. "A meta-analytic review of two modes of learning and the description-experience gap." In: *Psychological Bulletin* 144.2 (2018), pp. 140–176.

[35] Robin Hau et al. "The description–experience gap in risky choice: the role of sample size and experienced probabilities". In: *Journal of Behavioral Decision Making* 21.5 (2008), pp. 493–518. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/bdm.598.

[36] A. Camilleri and B. Newell. "The role of representation in experience-based choice." In: *Judgment and Decision Making* 4.9 (2009), pp. 518–529.

[37] Christoph Ungemach, Nick Chater, and Neil Stewart. "Are Probabilities Overweighted or Underweighted When Rare Outcomes Are Experienced (Rarely)?" In: *Psychological Science* 20.4 (2009), pp. 473–479.

[38] Robin Hau, Timothy J. Pleskac, and Ralph Hertwig. "Decisions from experience and statistical probabilities: Why they trigger different choices than a priori probabilities". In: *Journal of Behavioral Decision Making* 23.1 (2010), pp. 48–68. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/bdm.665.

[39] Mohammed Abdellaoui, Olivier L'Haridon, and Corina Paraschiv. "Experienced vs. Described Uncertainty: Do We Need Two Prospect Theory Specifications?" In: *Management Science* 57.10 (2011), pp. 1879–1895. eprint: ttps://doi.org/10.1287/mnsc.1110.1368.

[40] Andreas Glöckner et al. "The reversed description-experience gap: Disentangling sources of presentation format effects in risky choice". In: *Journal of Experimental Psychology: General* 145 (Apr. 2016), pp. 486–508.

[41] David Kellen, Thorsten Pachur, and Ralph Hertwig. "How (in)variant are subjective representations of described and experienced risk and rewards?" In: *Cognition* 157 (Dec. 2016), pp. 126–138.

[42] Ole Peters et al. "What are we weighting for? A mechanistic model for probability weighting". In: *arXiv e-prints*, arXiv:2005.00056 (Apr. 2020), arXiv:2005.00056. arXiv: 2005.00056 [econ.TH].

[43] Neil Stewart, Nick Chater, and Gordon D. A. Brown. *Decision by sampling.* 2006.

[44] Jakub Steiner and Colin Stewart. "Perceiving Prospects Properly". In: *The American Economic Review* 106.7 (2016), pp. 1601–1631.

[45] Keno Juechems et al. *Optimal utility and probability functions for agents with finite computational precision.* Jan. 2020.

[46] Henry P. Stott. “Cumulative prospect theory’s functional menagerie”. In: *Journal of Risk and Uncertainty* 32.2 (Mar. 2006), pp. 101–130.

[47] Michael D. Lee and Eric-Jan Wagenmakers. *Bayesian cognitive modeling: A practical course.* Bayesian cognitive modeling: A practical course. New York, NY, US: Cambridge University Press, 2013, pp. xiii, 264–xiii, 264.

[48] Håkan Nilsson, Jörg Rieskamp, and Eric-Jan Wagenmakers. “Hierarchical Bayesian parameter estimation for cumulative prospect theory”. In: *Journal of Mathematical Psychology* 55.1 (2011). Special Issue on Hierarchical Bayesian Models, pp. 84–93.

[49] Steve Brooks et al. “Inference from Simulations and Monitoring Convergence”. In: *Handbook of Markov Chain Monte Carlo.* CRC press, 2011. Chap. 6.

[50] Andrew Gelman, Jessica Hwang, and Aki Vehtari. “Understanding predictive information criteria for Bayesian models”. In: *Statistics and Computing* 24.6 (Nov. 2014), pp. 997–1016.

[51] Jean Daunizeau, Vincent Adam, and Lionel Rigoux. “VBA: a probabilistic treatment of nonlinear models for neurobiological and behavioural data”. eng. In: *PLoS computational biology* 10.1 (Jan. 2014). 24465198[pmid], e1003441–e1003441.

[52] Payam Piray et al. “Hierarchical Bayesian inference for concurrent model fitting and comparison for group studies”. In: *PLOS Computational Biology* 15.6 (June 2019), pp. 1–34.

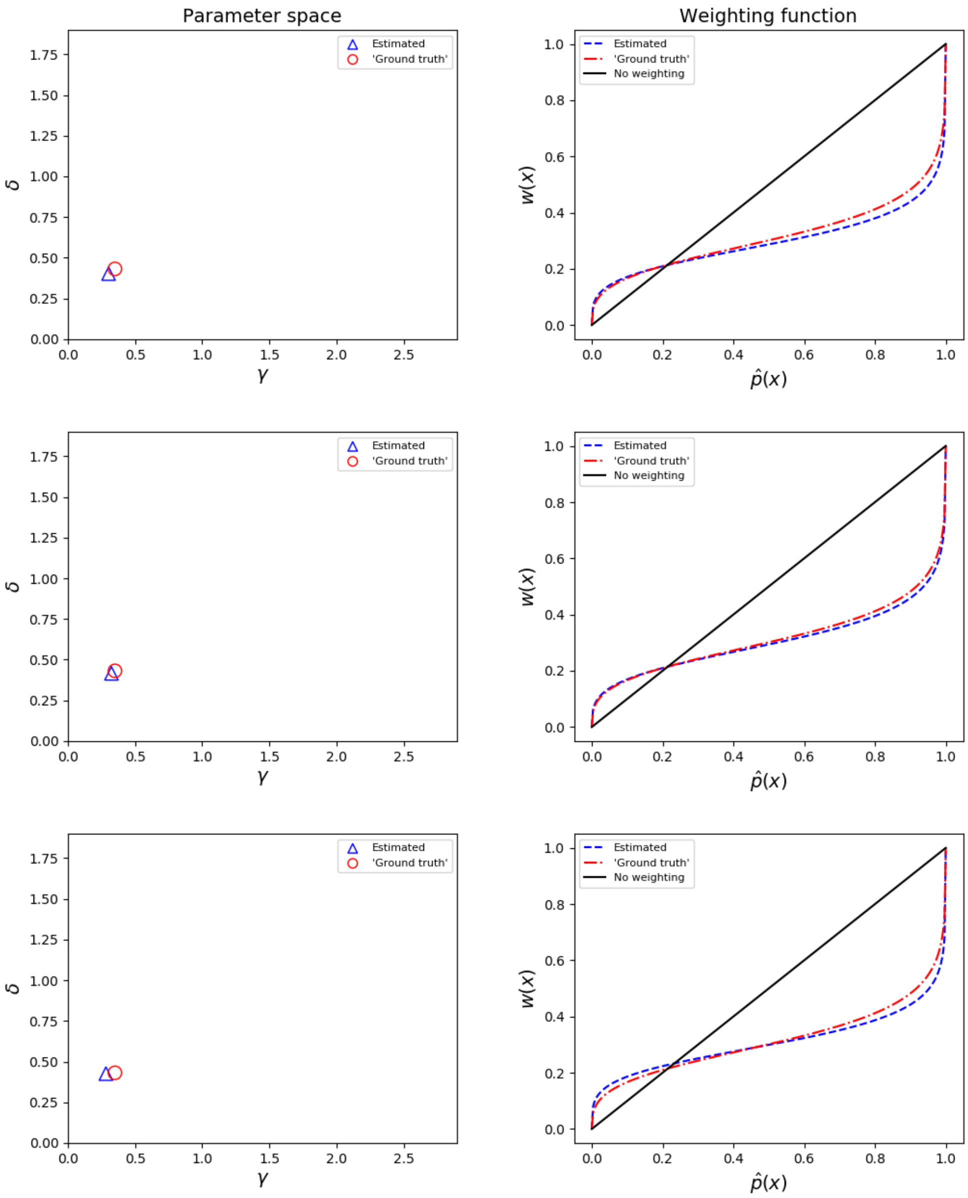
[53] Richard Colombo and Weina Jiang. “A stochastic RFM model”. In: *Journal of Interactive Marketing* 13.3 (Jan. 1999), pp. 2–12.

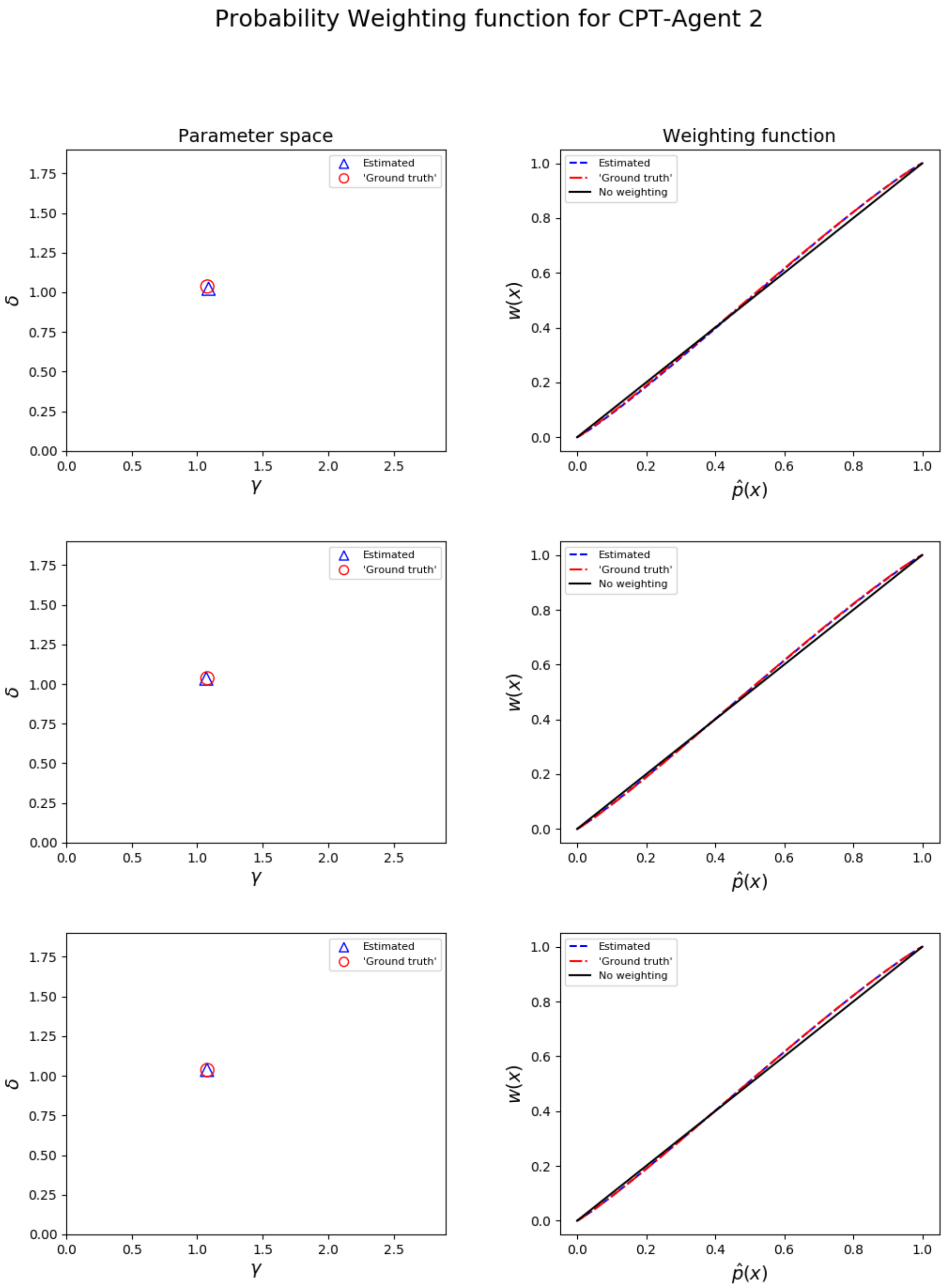
APPENDIX A

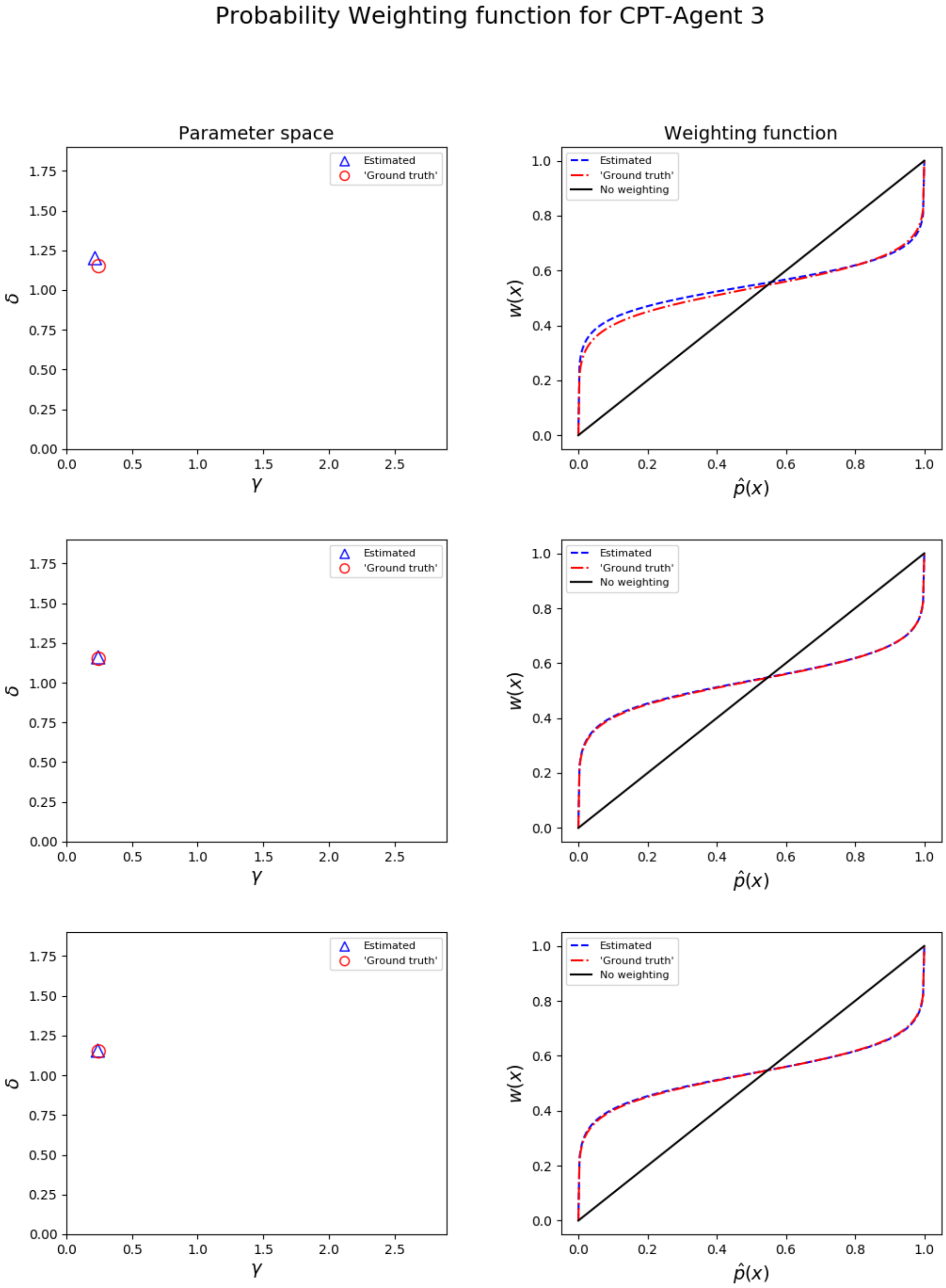
Individual PW functions for all agents

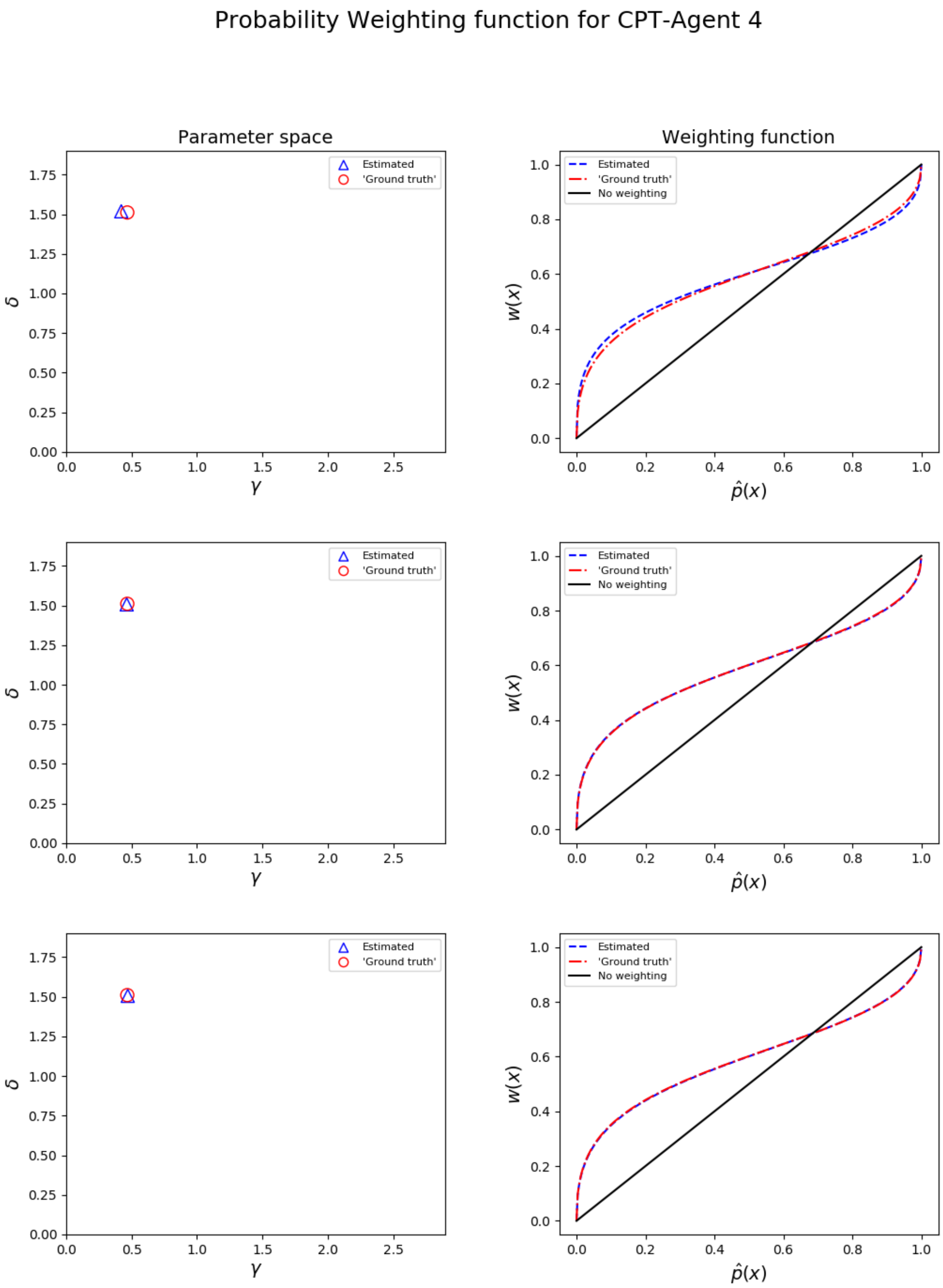
A.1 CPT-agents

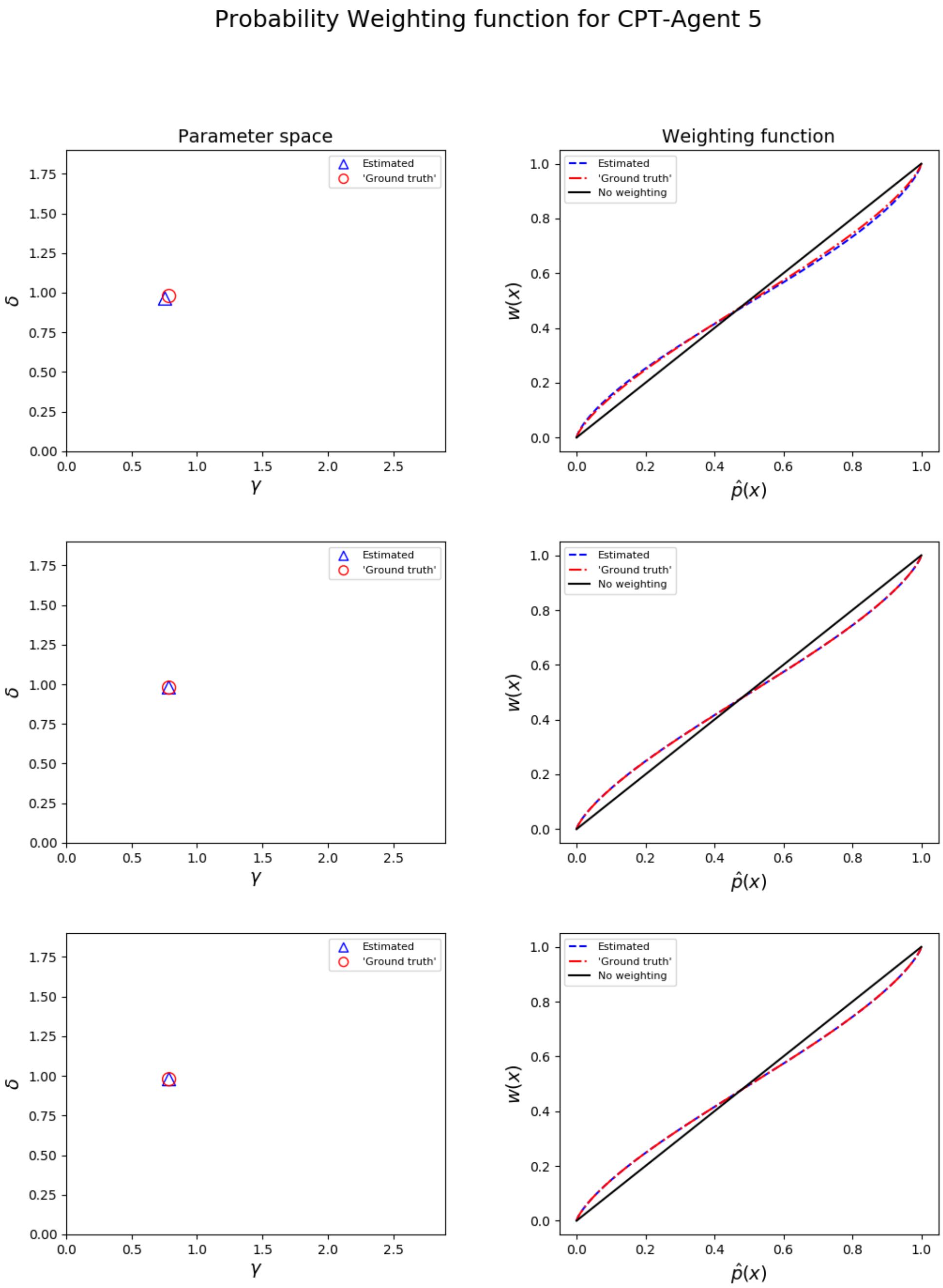
Probability Weighting function for CPT-Agent 1

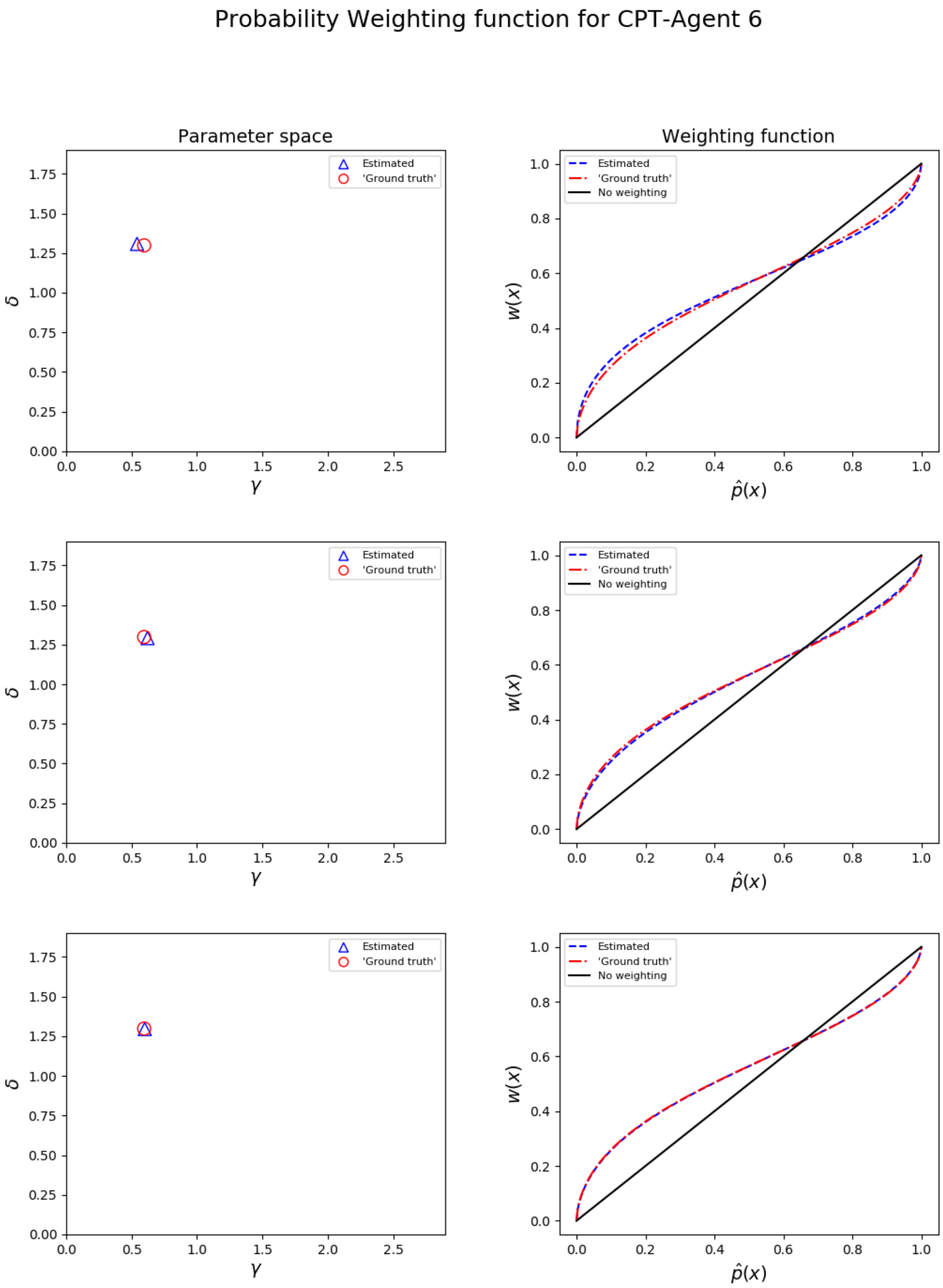



Probability Weighting function for CPT-Agent 2
Parameter space
Weighting function
Estimated
'Ground truth'
No weighting
δ
γ
w(x)
p̂(x)
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
0.0
0.5
1.0
1.5
2.0
2.5
0.2
0.4
0.6
0.8

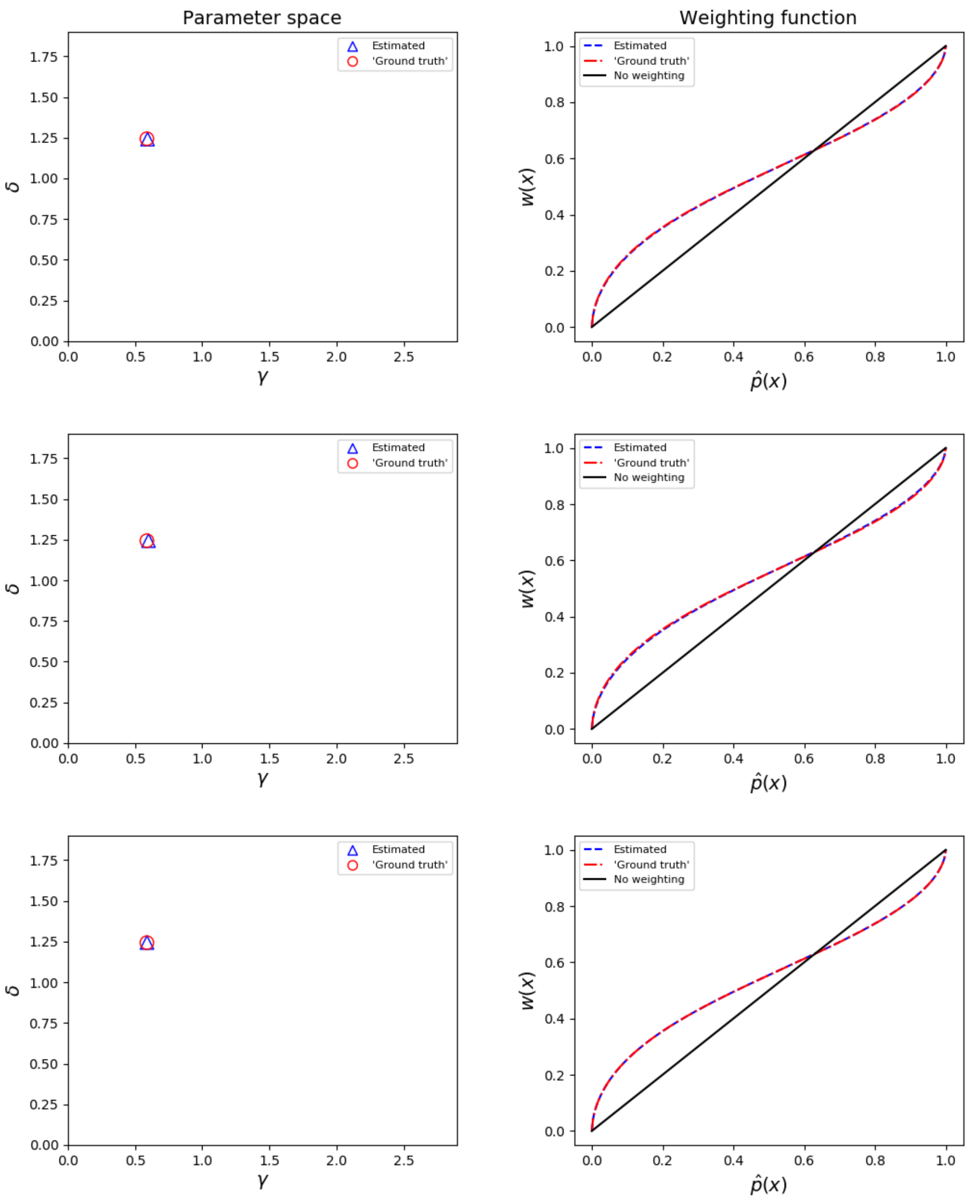

Probability Weighting function for CPT-Agent 3
Parameter space
Weighting function
Estimated
'Ground truth'
No weighting
δ
γ
$w(x)$
$\hat{p}(x)$

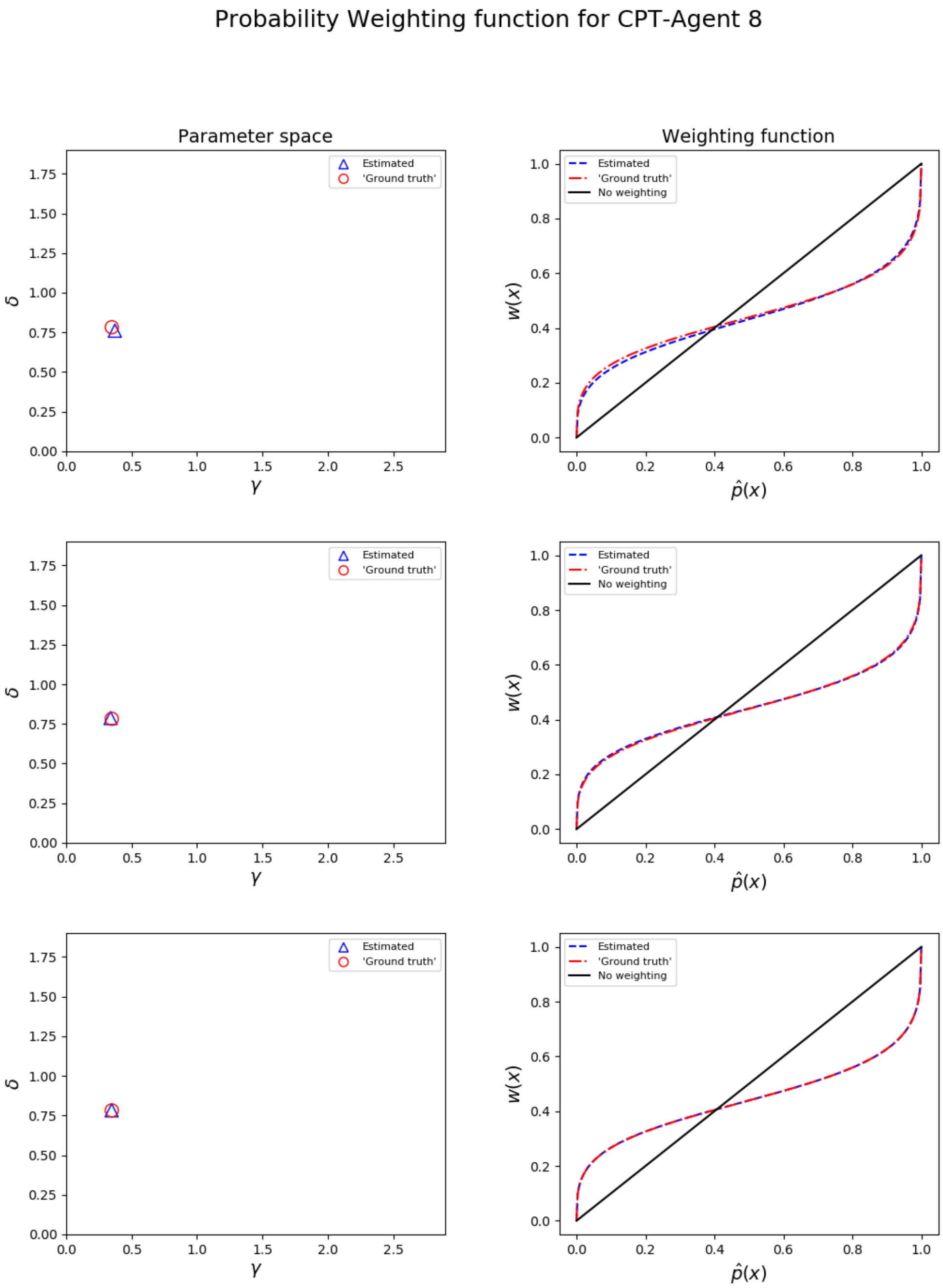

Probability Weighting function for CPT-Agent 4
Parameter space
Weighting function
Estimated
'Ground truth'
No weighting
δ
γ
w(x)
p̂(x)

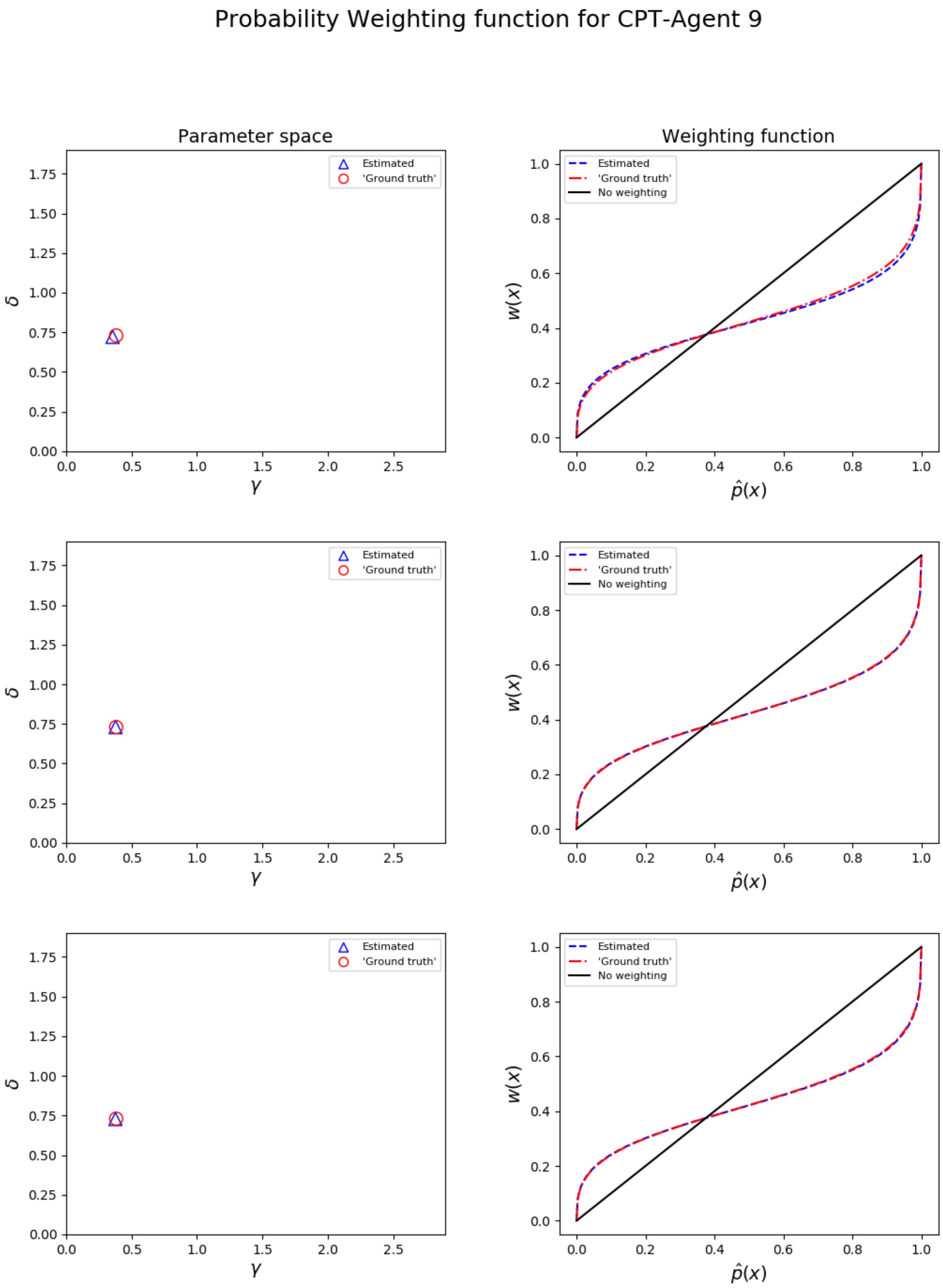

Probability Weighting function for CPT-Agent 5
Parameter space
Weighting function
Estimated
'Ground truth'
No weighting
δ
γ
w(x)
p̂(x)
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
0.0
0.5
1.0
1.5
2.0
2.5
0.0
0.2
0.4
0.6
0.8
1.0

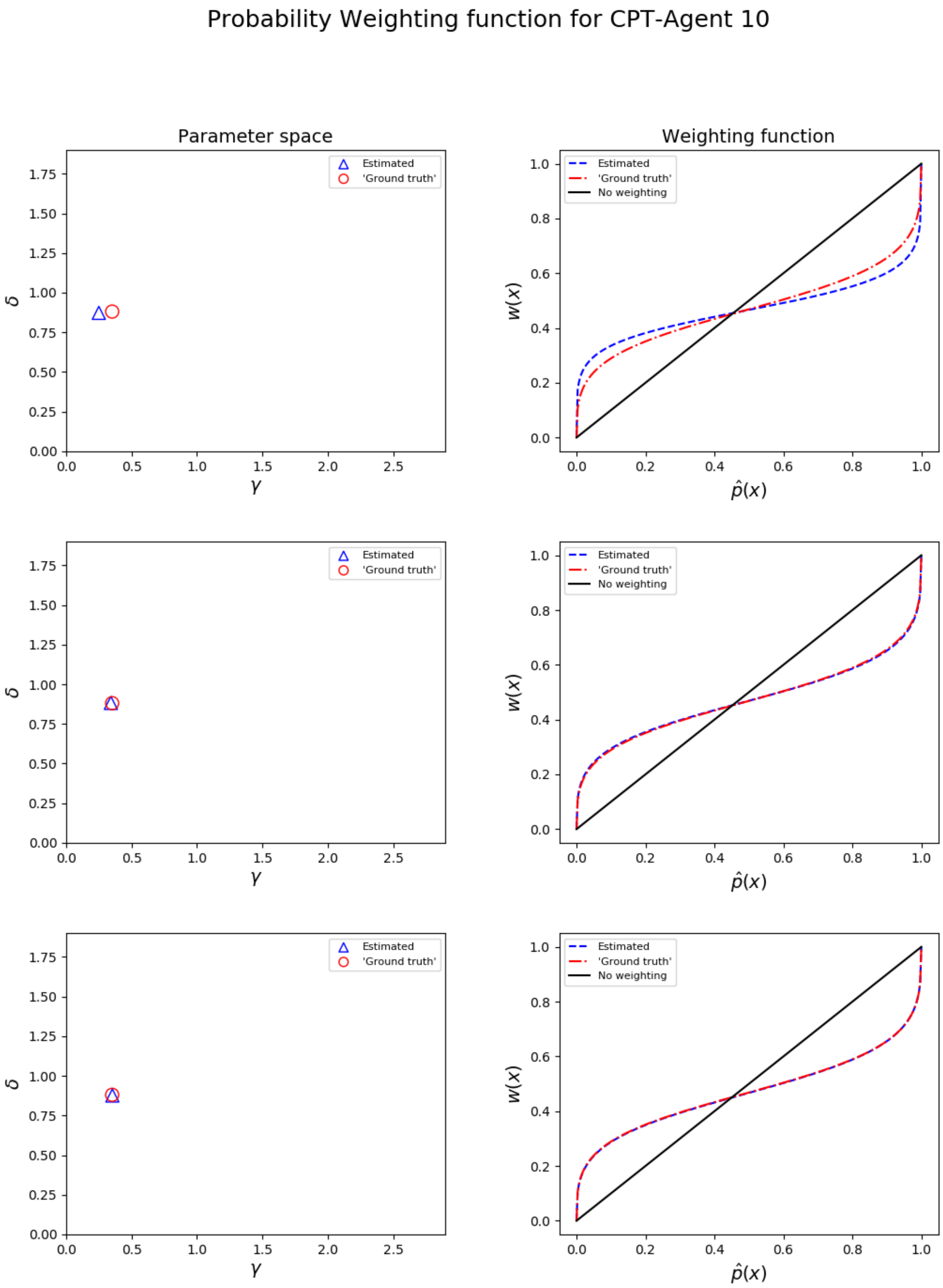

Probability Weighting function for CPT-Agent 6
Parameter space
Weighting function
Estimated
'Ground truth'
No weighting
γ
δ
w(x)
p̂(x)

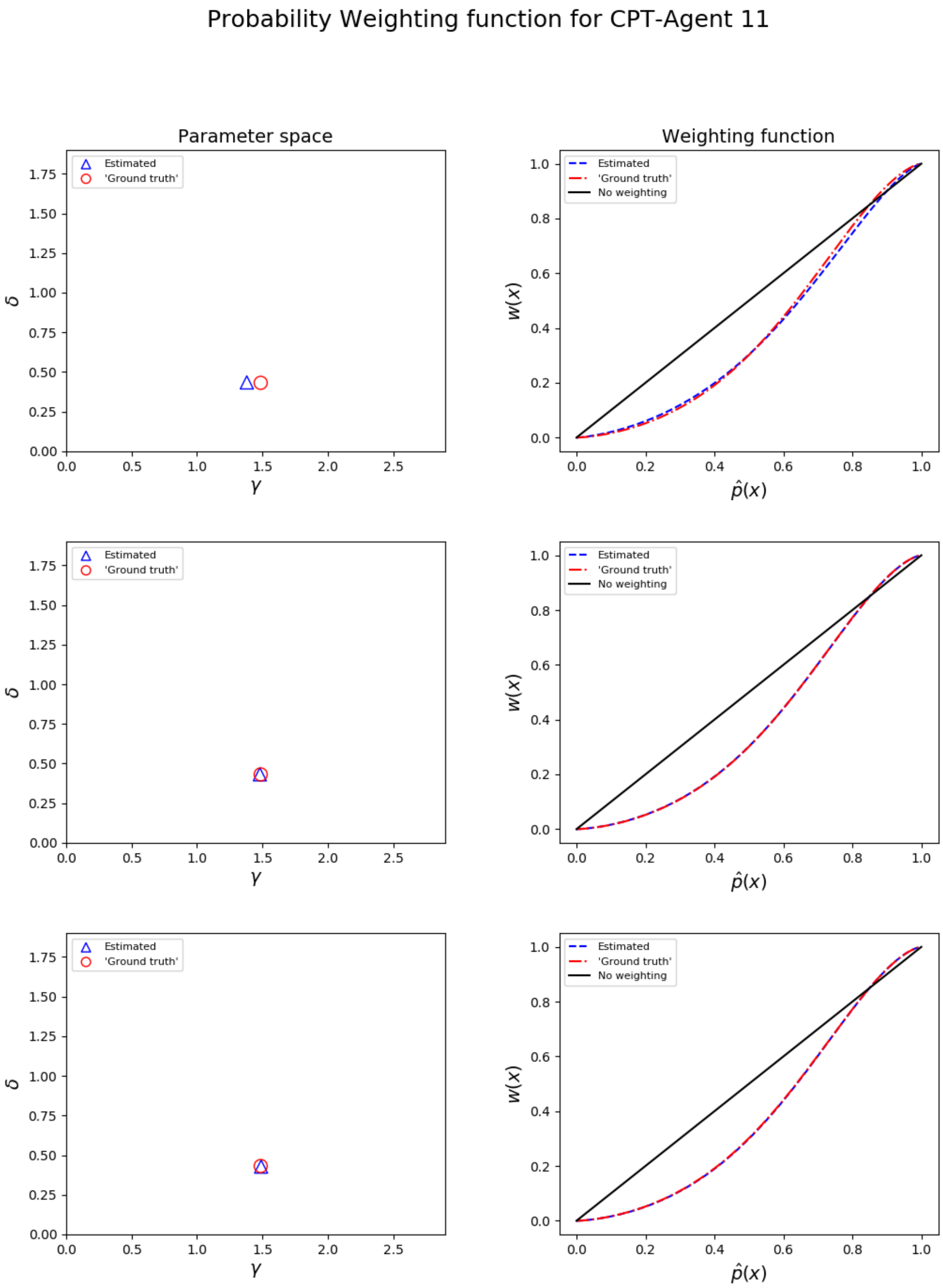
Probability Weighting function for CPT-Agent 7

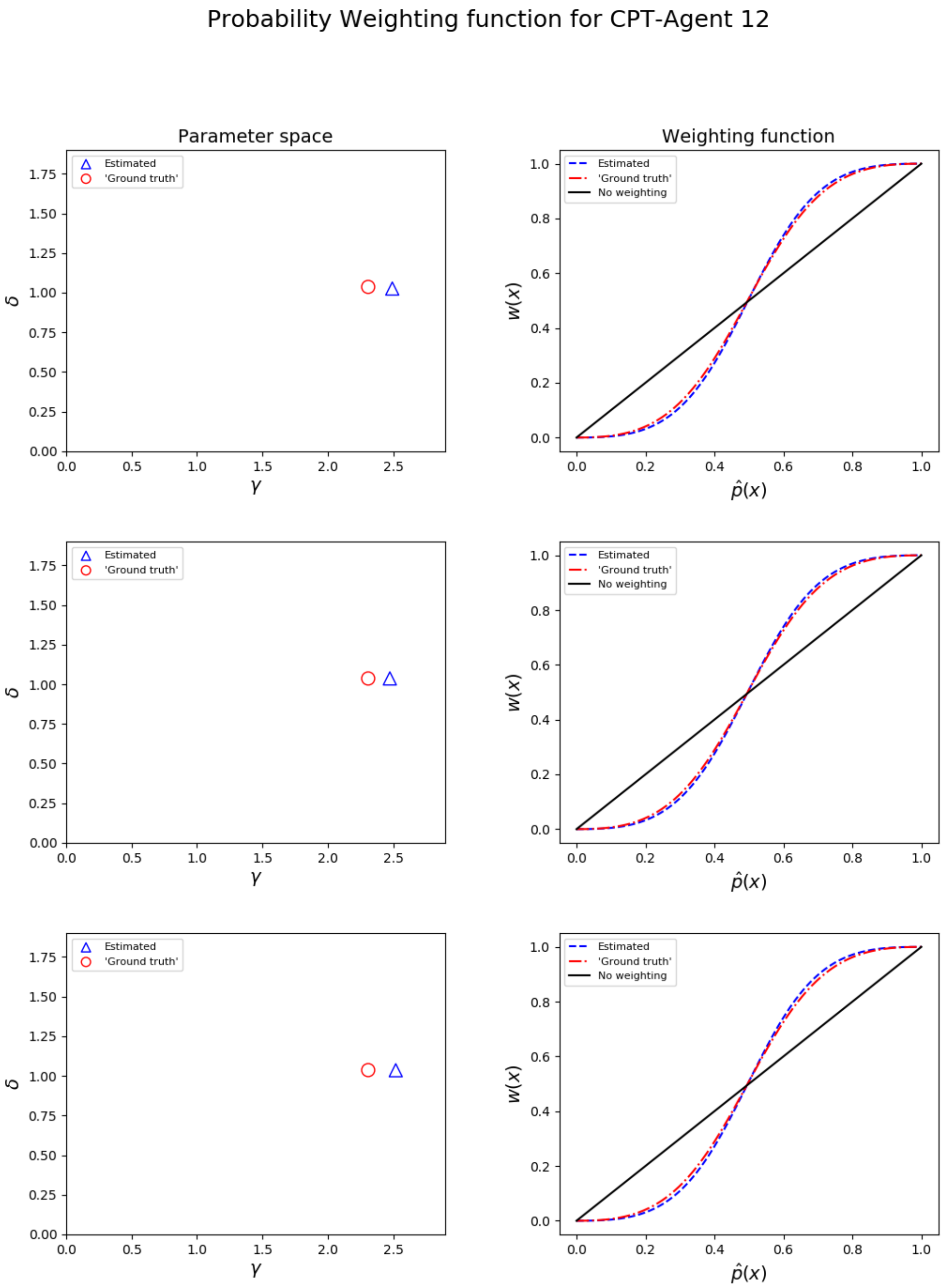


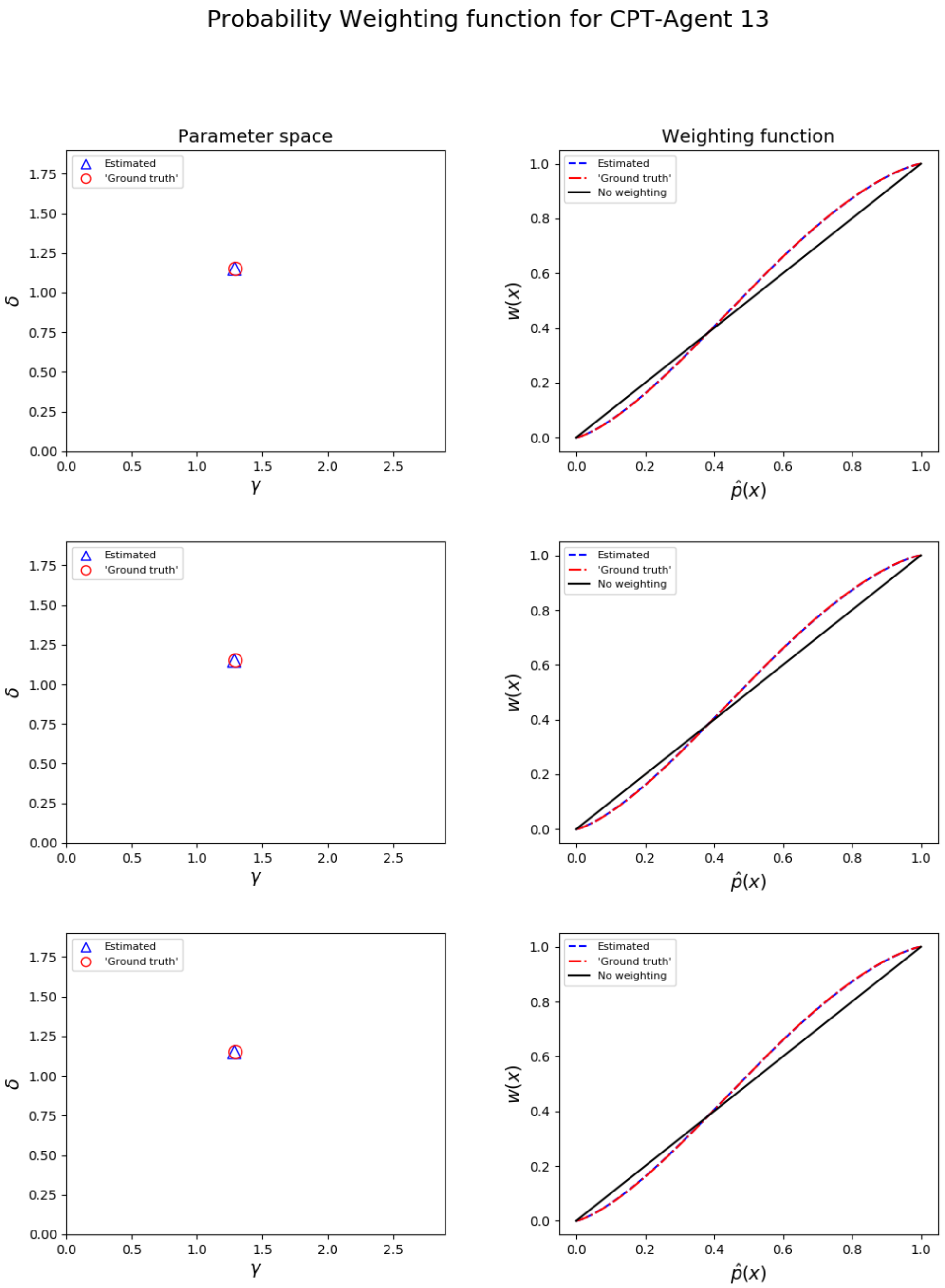

Probability Weighting function for CPT-Agent 8
Parameter space
Weighting function
Estimated
'Ground truth'
No weighting
δ
γ
w(x)
p̂(x)
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
0.0
0.5
1.0
1.5
2.0
2.5
0.2
0.4
0.6
0.8

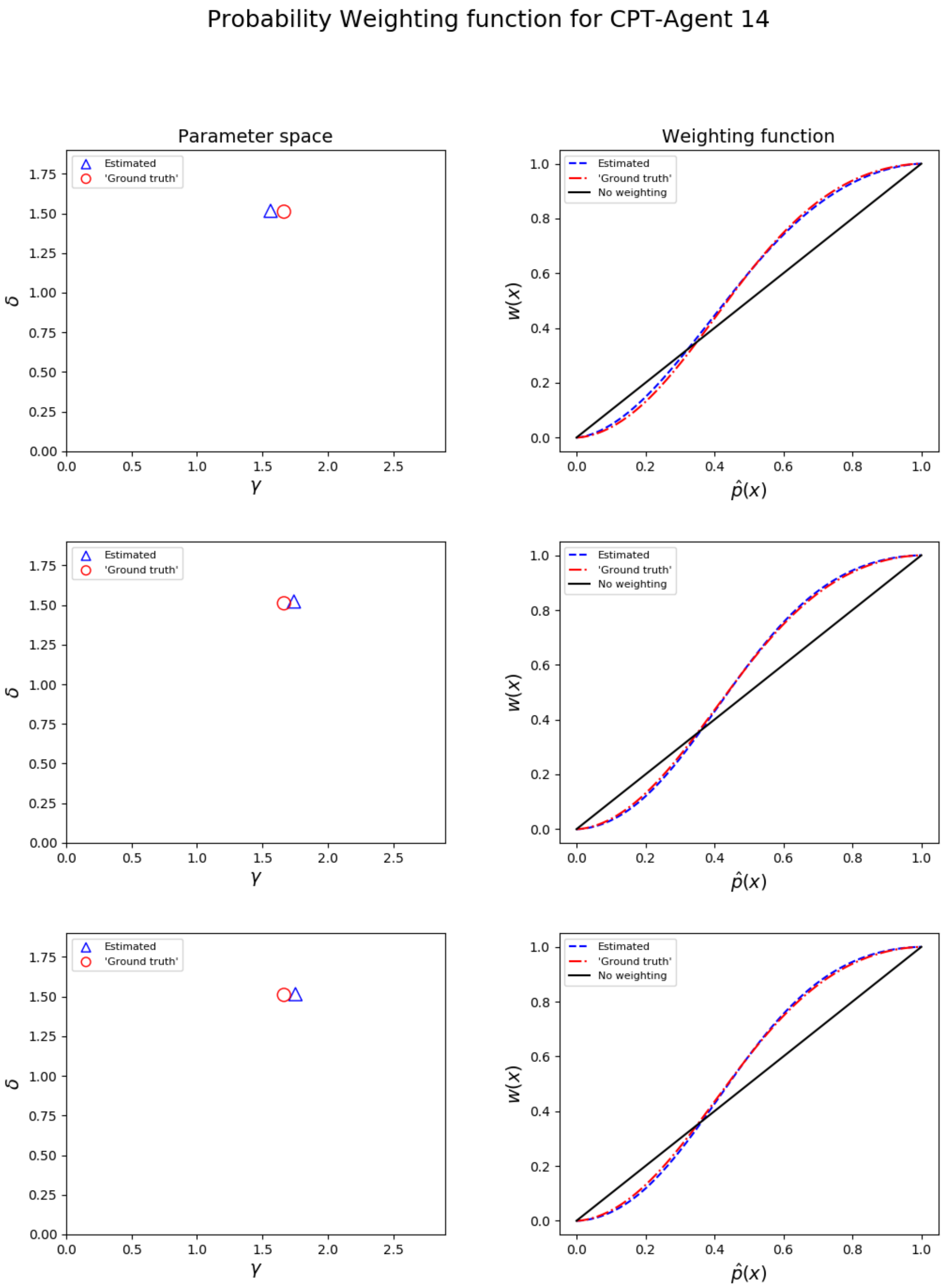

Probability Weighting function for CPT-Agent 9
Parameter space
Weighting function
Estimated
'Ground truth'
No weighting
δ
γ
$w(x)$
$\hat{p}(x)$

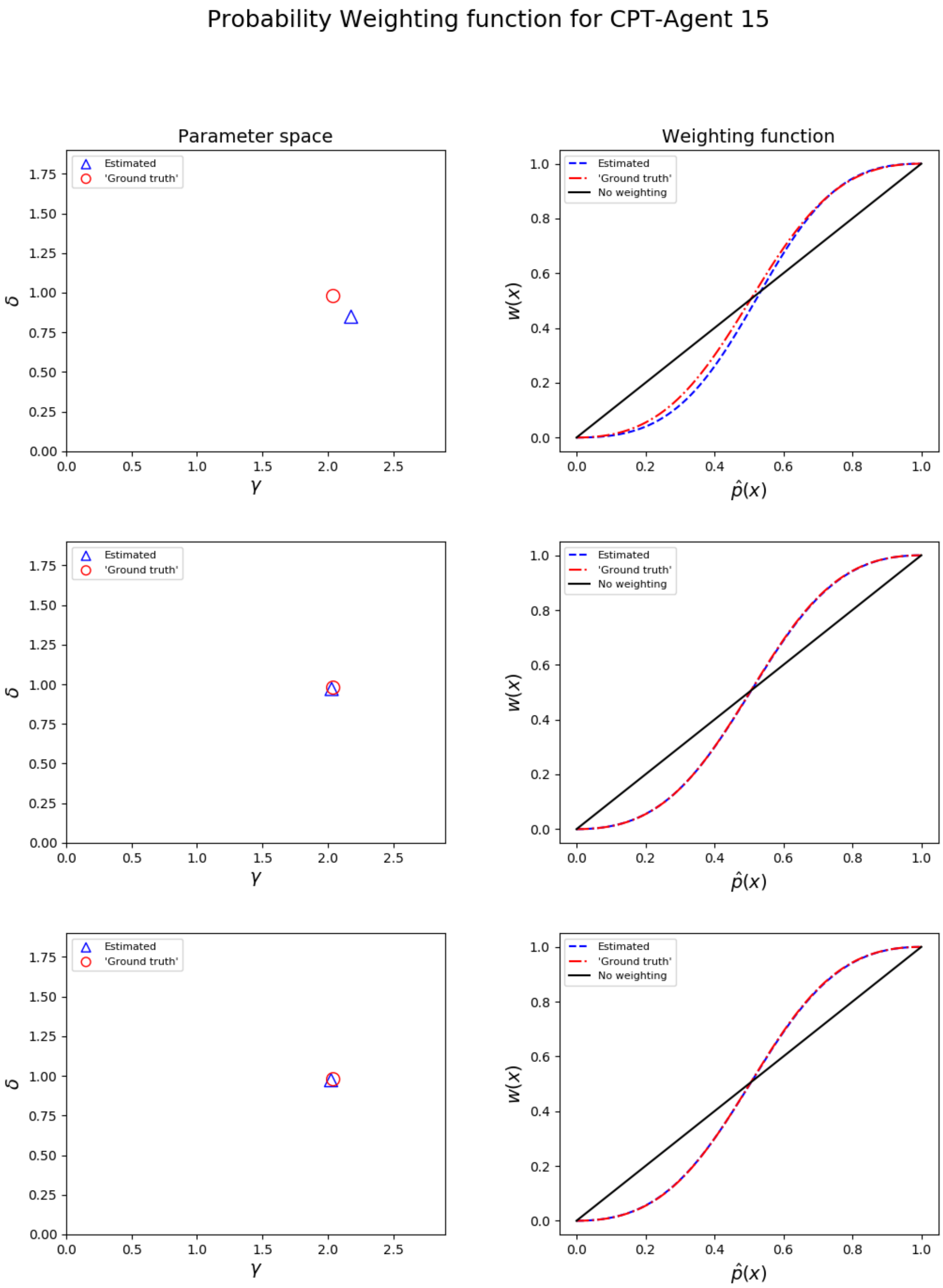

Probability Weighting function for CPT-Agent 10
Parameter space
Weighting function
Estimated
'Ground truth'
No weighting
δ
γ
w(x)
p̂(x)
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
0.0
0.5
1.0
1.5
2.0
2.5
0.2
0.4
0.6
0.8

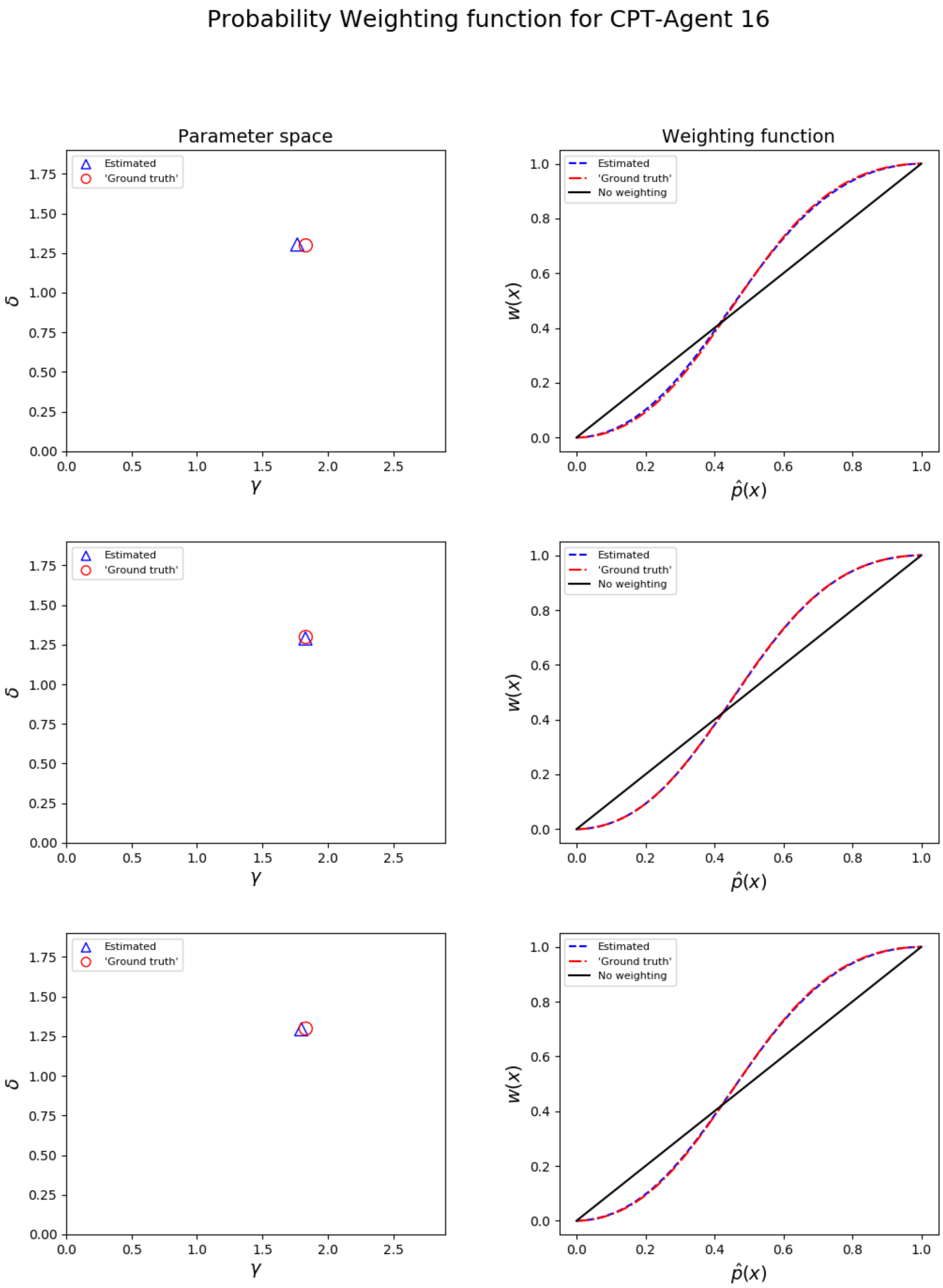

Probability Weighting function for CPT-Agent 11
Parameter space
Weighting function
Estimated
'Ground truth'
No weighting
δ
γ
w(x)
p̂(x)
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
0.0
0.5
1.0
1.5
2.0
2.5
0.0
0.2
0.4
0.6
0.8
1.0

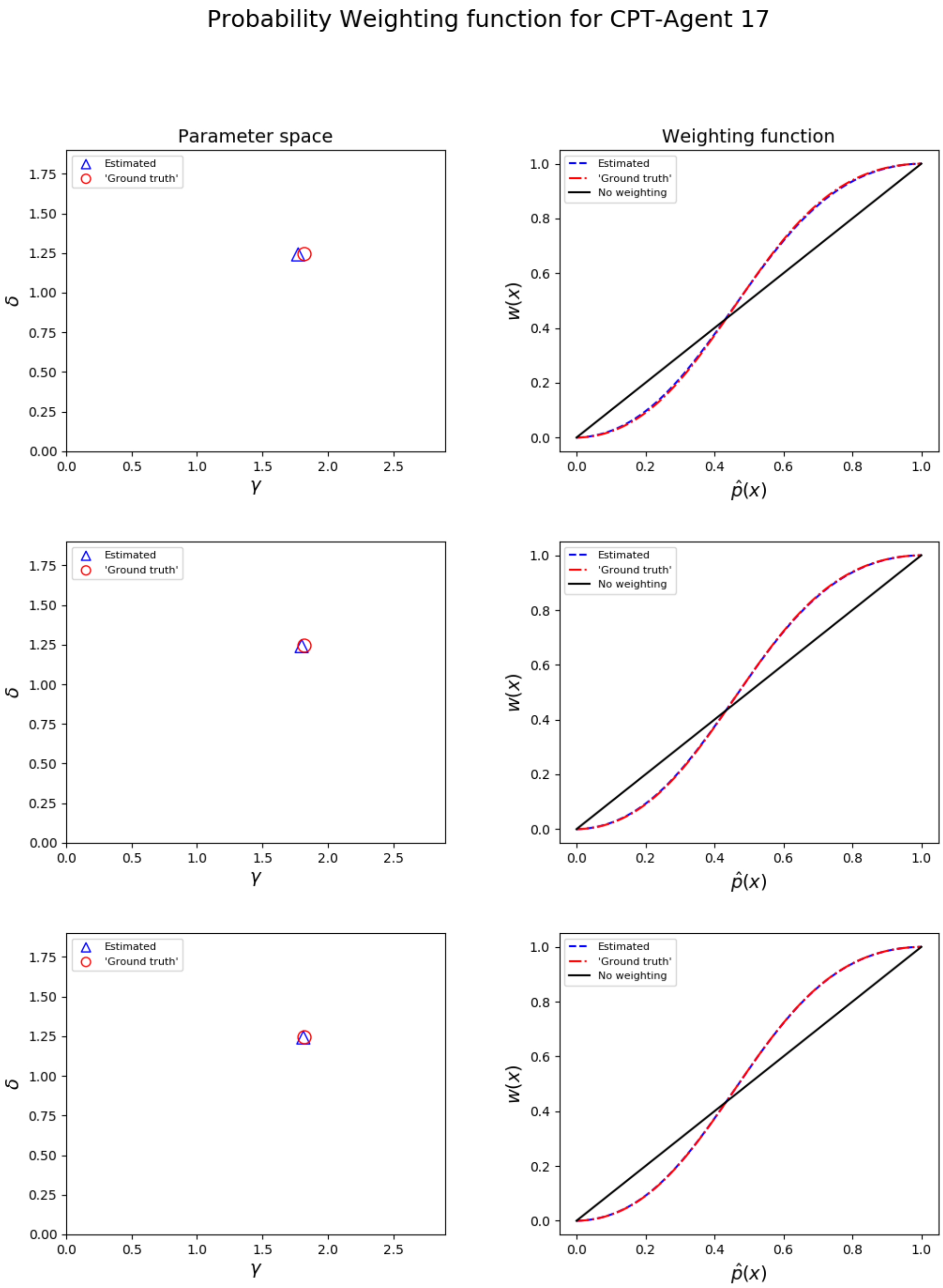

Probability Weighting function for CPT-Agent 12
Parameter space
Weighting function
Estimated
'Ground truth'
No weighting
δ
γ
w(x)
p̂(x)

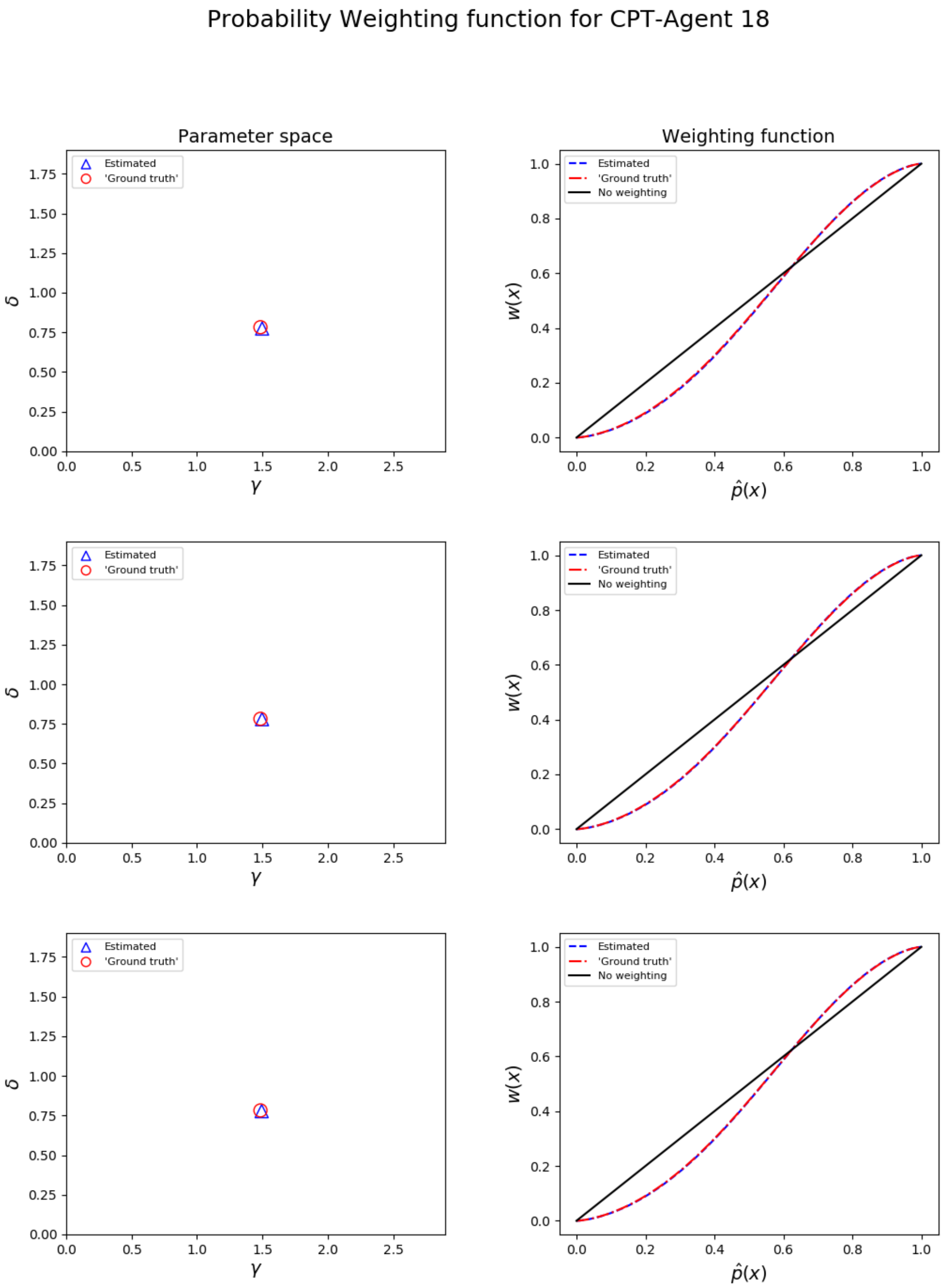

Probability Weighting function for CPT-Agent 13
Parameter space
Weighting function
Estimated
'Ground truth'
No weighting
δ
γ
w(x)
p̂(x)

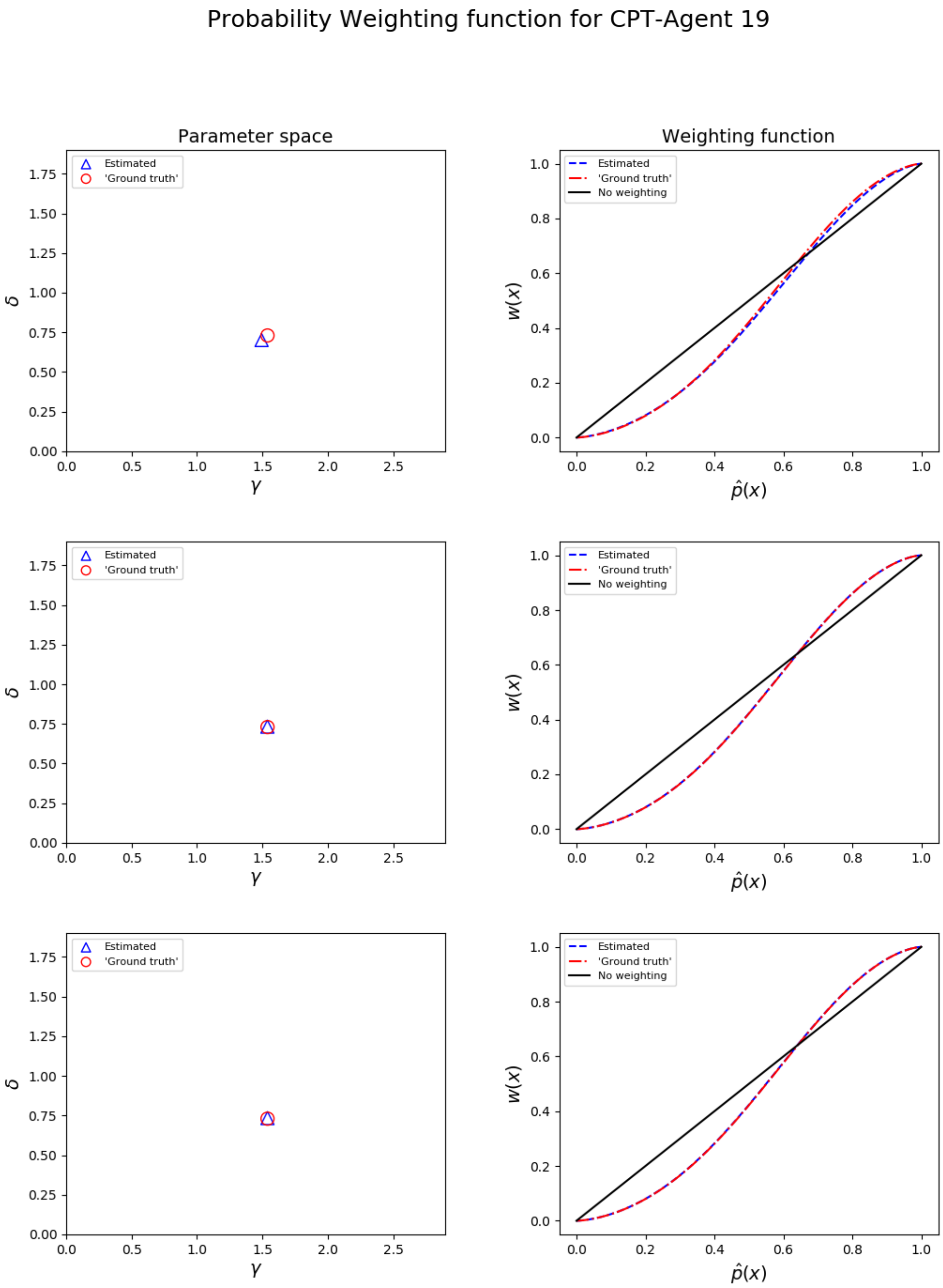

Probability Weighting function for CPT-Agent 14
Parameter space
Weighting function
Estimated
'Ground truth'
No weighting
δ
γ
w(x)
p̂(x)

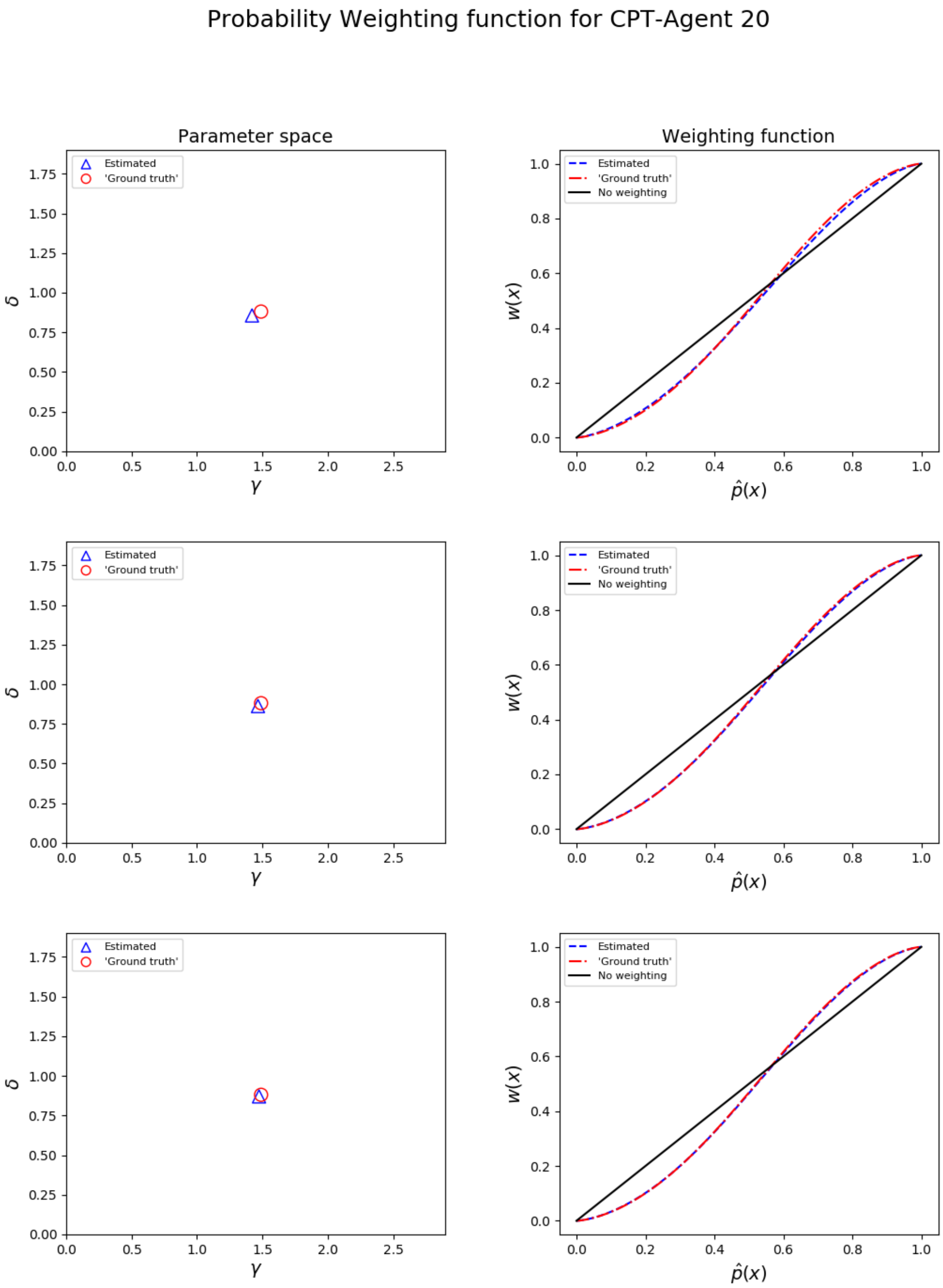

Probability Weighting function for CPT-Agent 15
Parameter space
Weighting function
Estimated
'Ground truth'
No weighting
δ
γ
w(x)
p̂(x)
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
0.0
0.5
1.0
1.5
2.0
2.5
0.2
0.4
0.6
0.8


Probability Weighting function for CPT-Agent 16
Parameter space
Weighting function
Estimated
'Ground truth'
No weighting
δ
γ
w(x)
p̂(x)

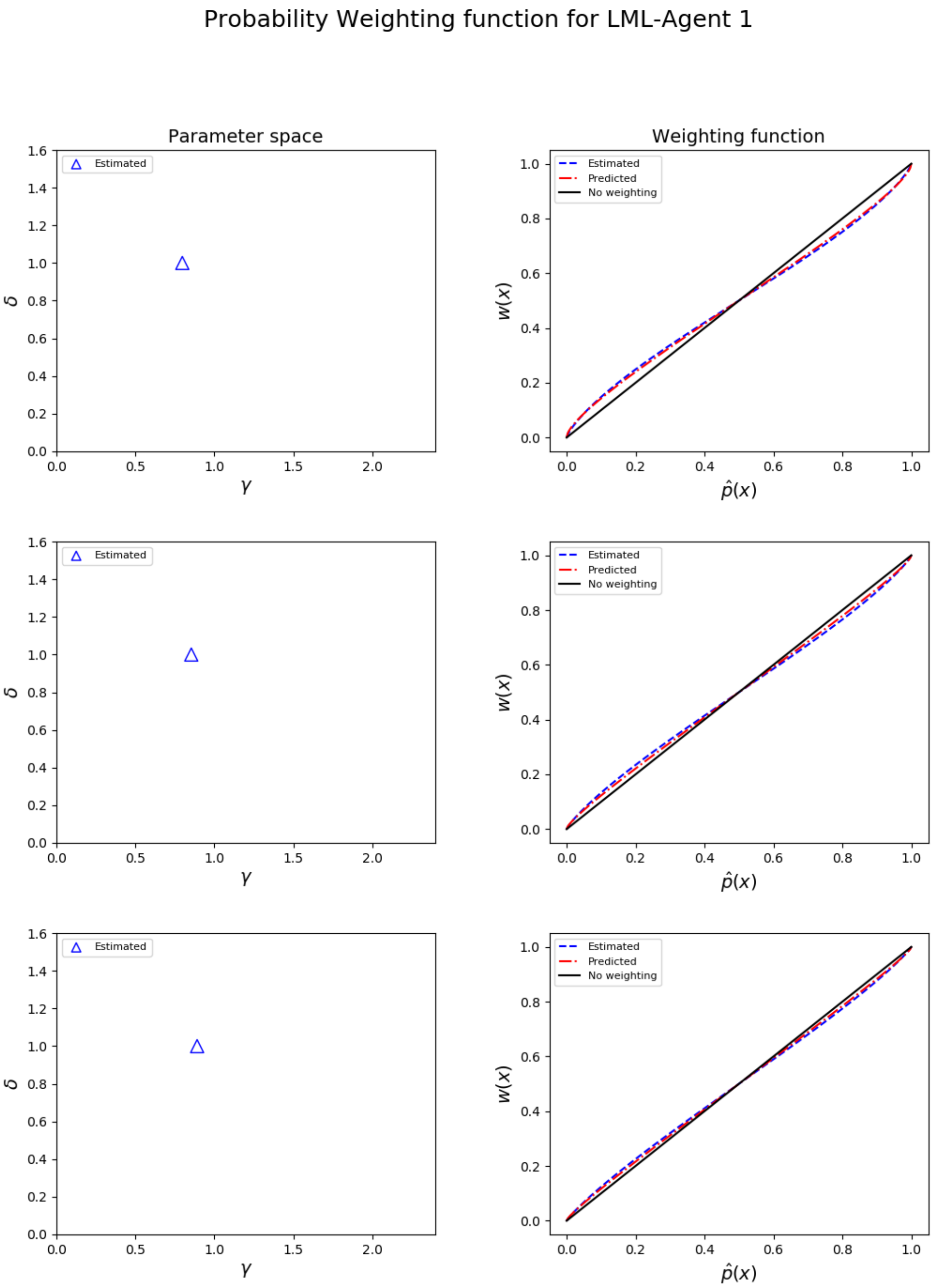

Probability Weighting function for CPT-Agent 17
Parameter space
Weighting function
Estimated
'Ground truth'
No weighting
δ
γ
$w(x)$
$\hat{p}(x)$

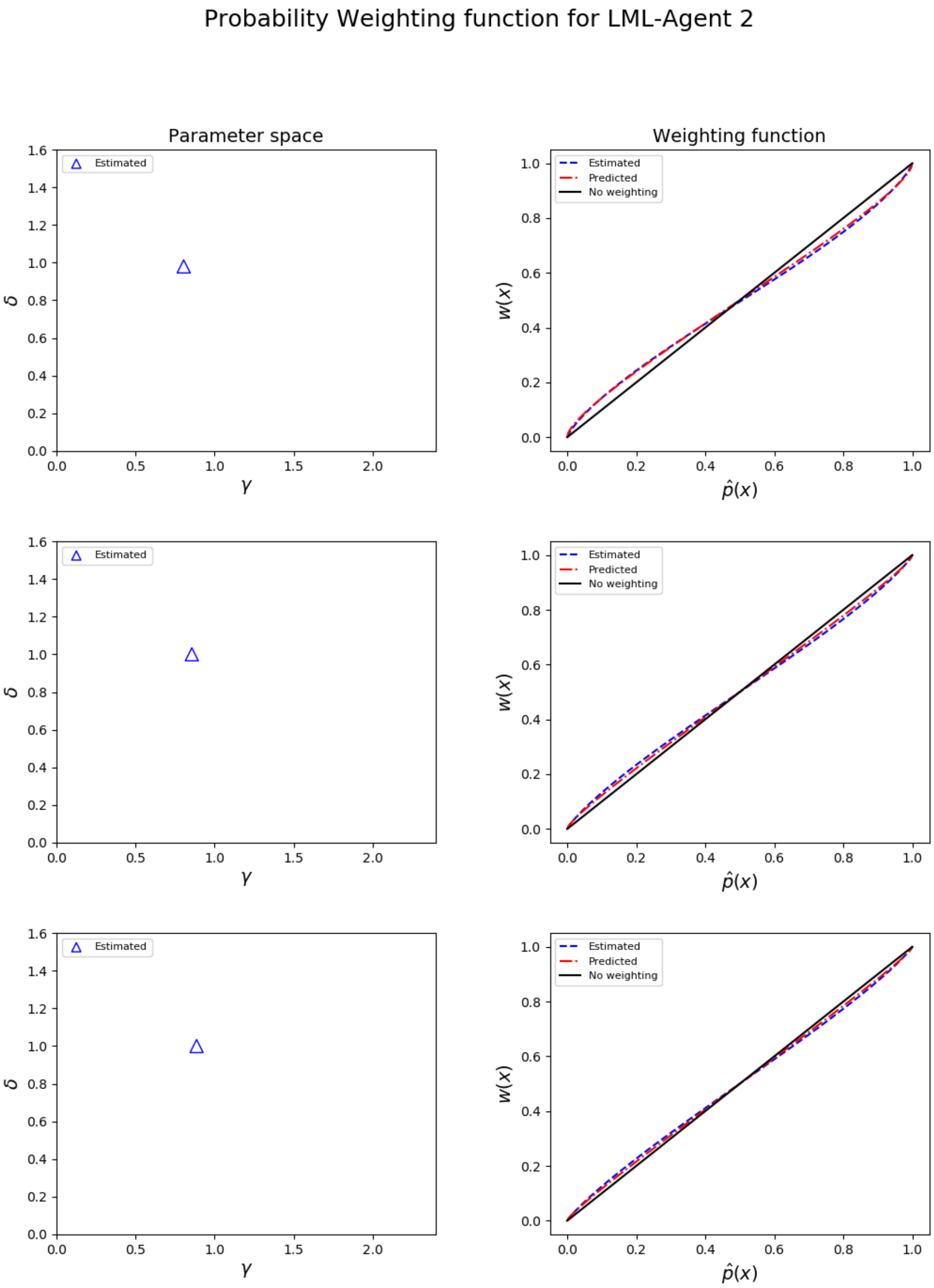

Probability Weighting function for CPT-Agent 18
Parameter space
Weighting function
Estimated
'Ground truth'
No weighting
δ
γ
w(x)
p̂(x)

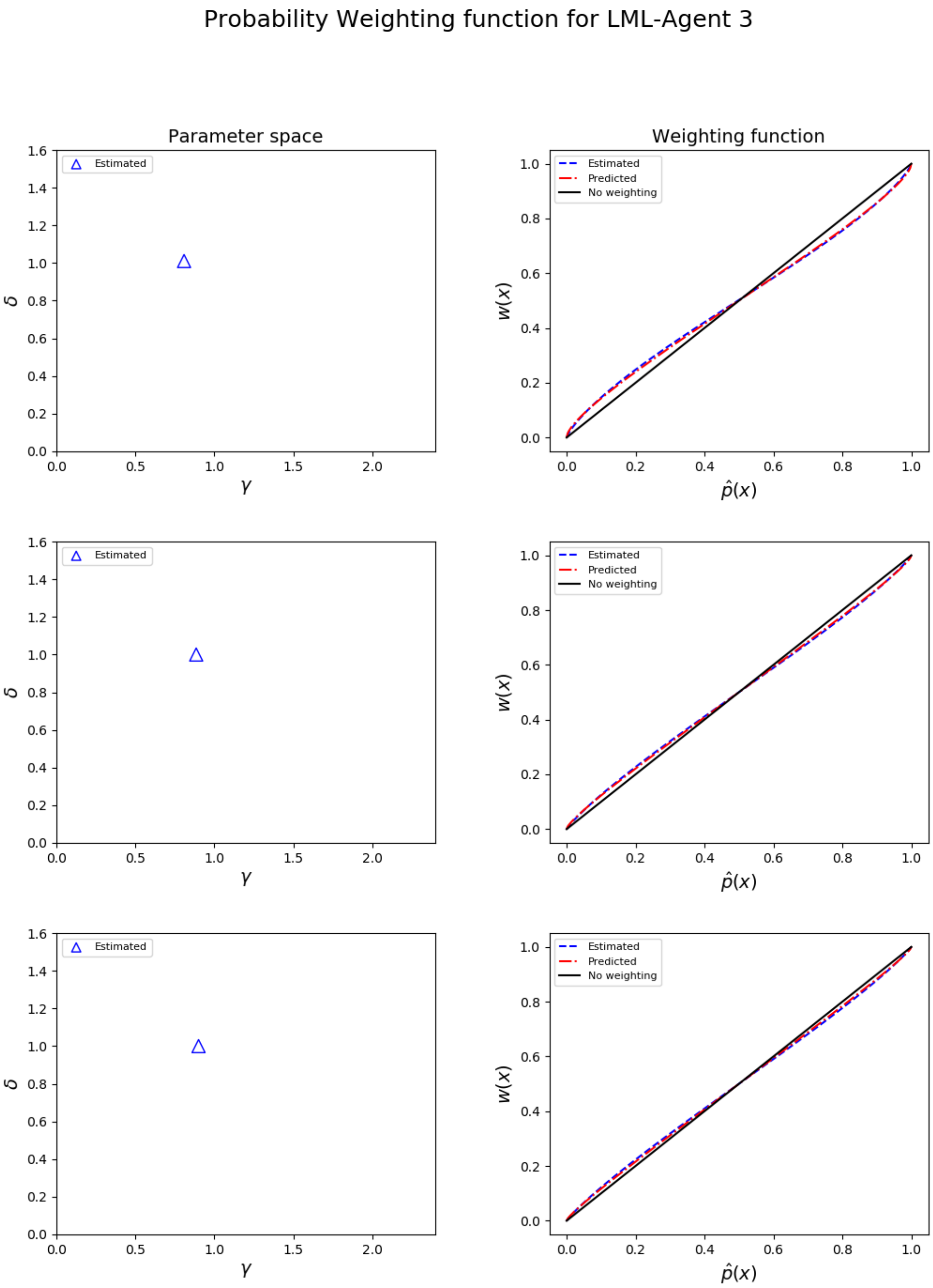

Probability Weighting function for CPT-Agent 19
Parameter space
Weighting function
Estimated
'Ground truth'
No weighting
δ
γ
w(x)
p̂(x)

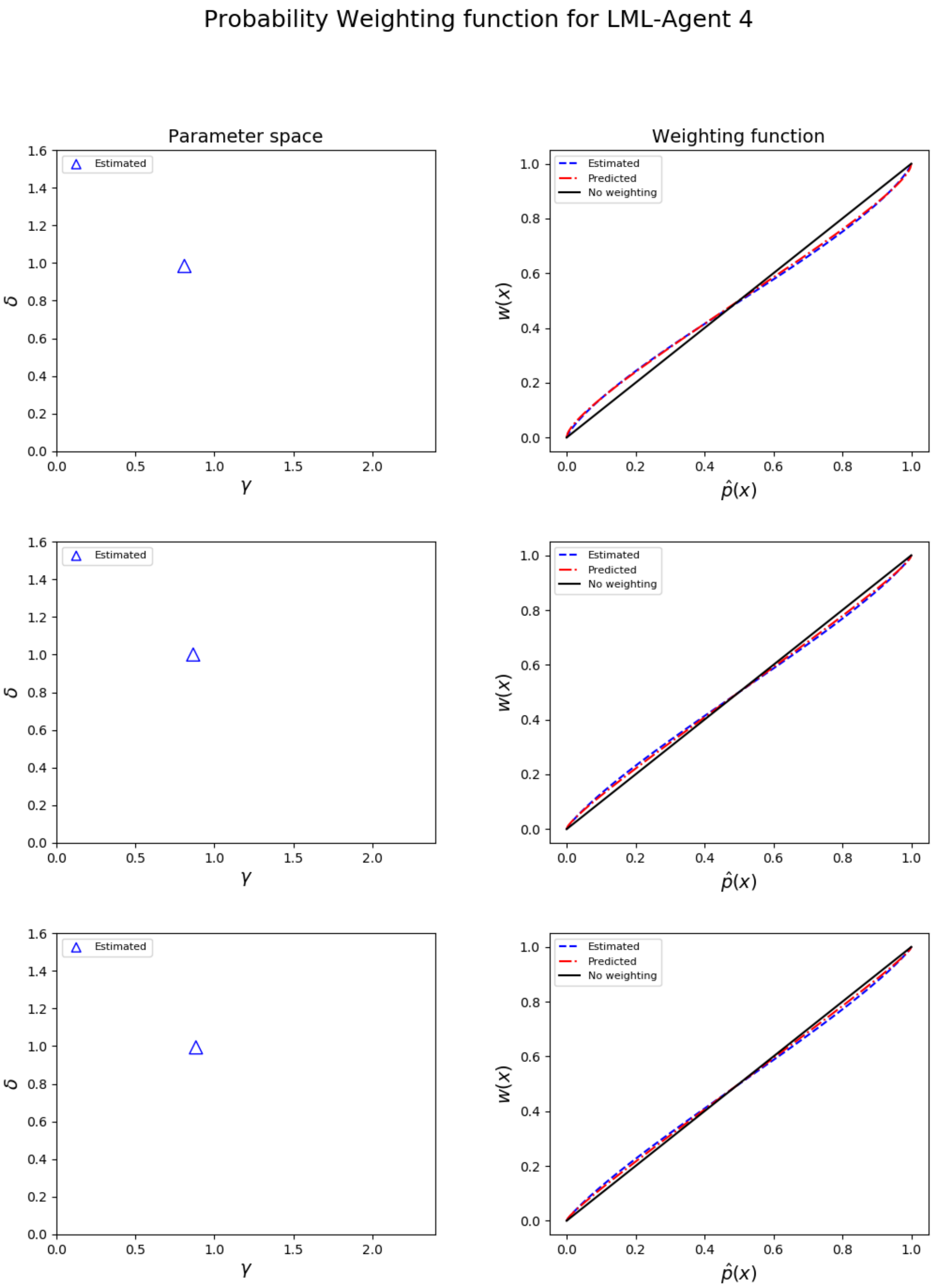

Probability Weighting function for CPT-Agent 20
Parameter space
Weighting function
Estimated
'Ground truth'
No weighting
δ
γ
w(x)
p̂(x)
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
0.0
0.5
1.0
1.5
2.0
2.5
0.0
0.2
0.4
0.6
0.8
1.0

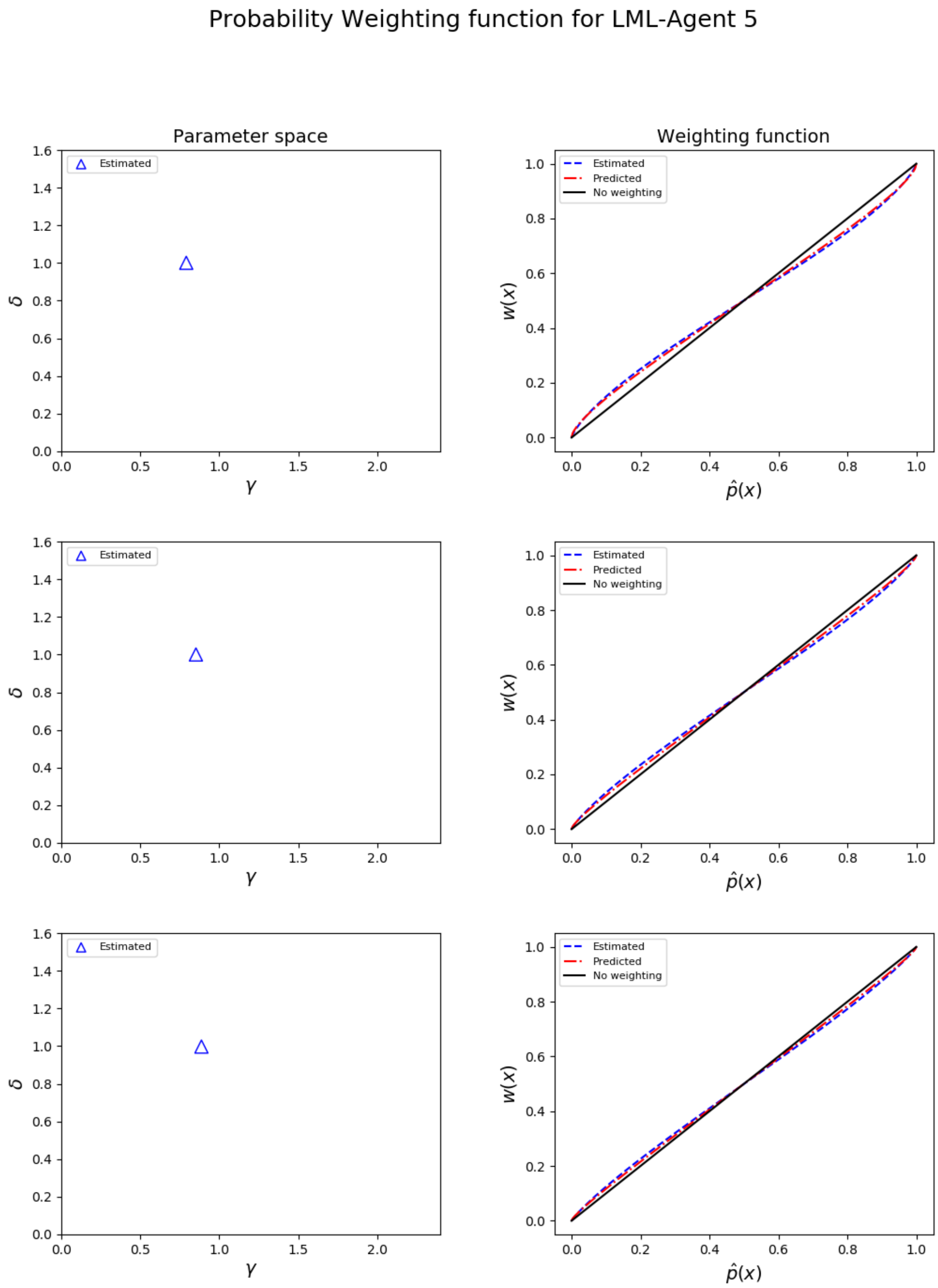
A.2 LML-agents



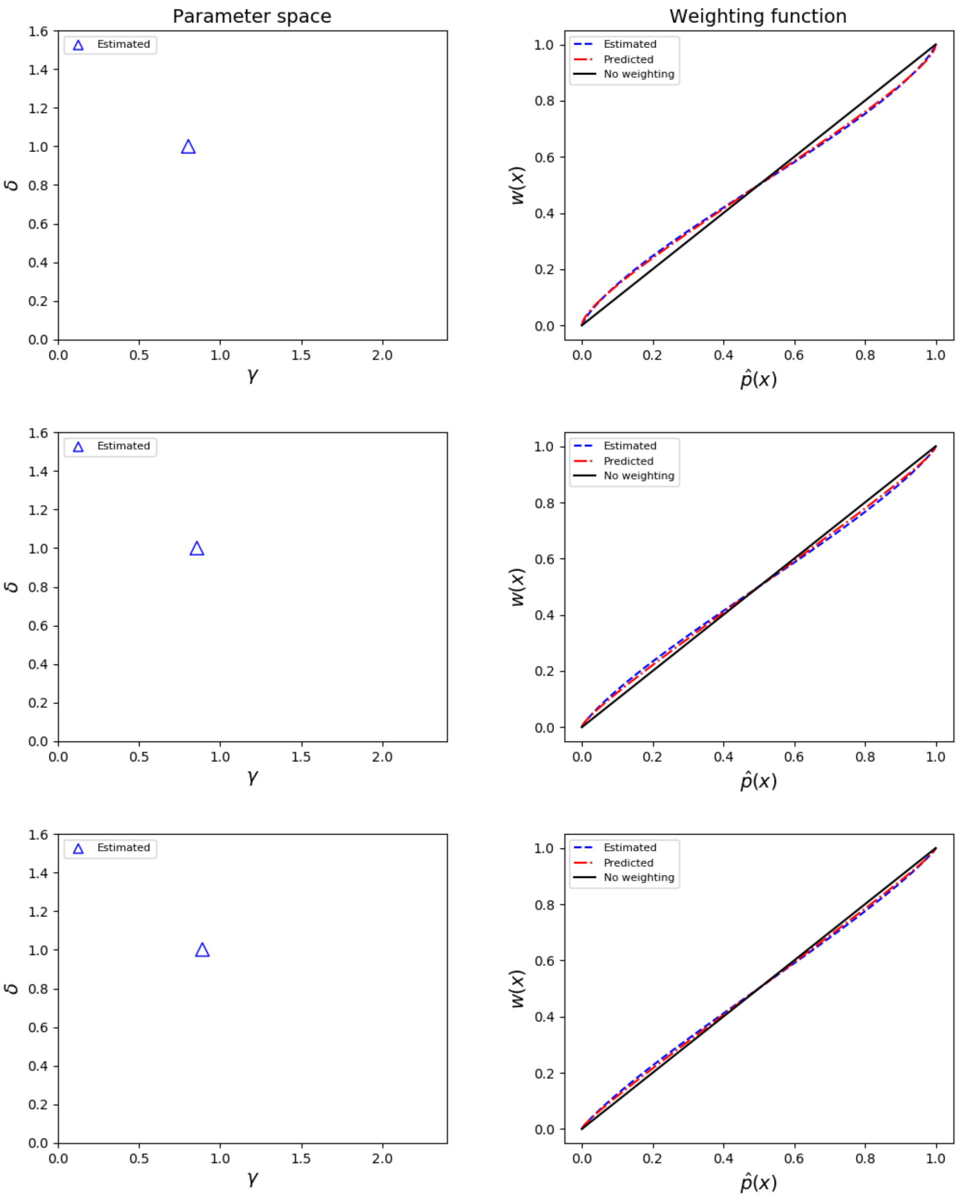

Probability Weighting function for LML-Agent 2
Parameter space
Weighting function
Estimated
Predicted
No weighting
δ
γ
w(x)
p̂(x)

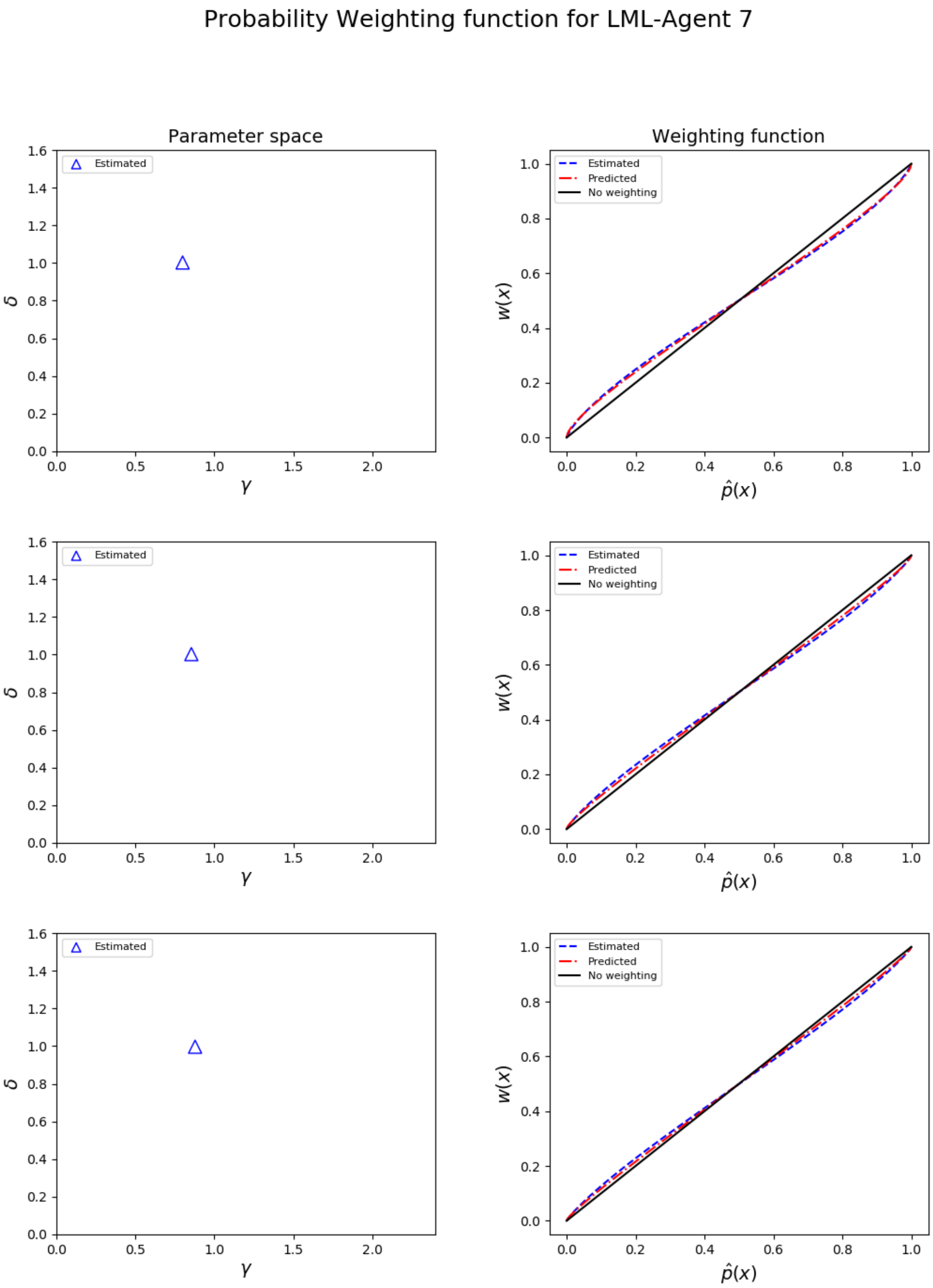

Probability Weighting function for LML-Agent 3
Parameter space
Weighting function
Estimated
Predicted
No weighting
δ
γ
w(x)
p̂(x)

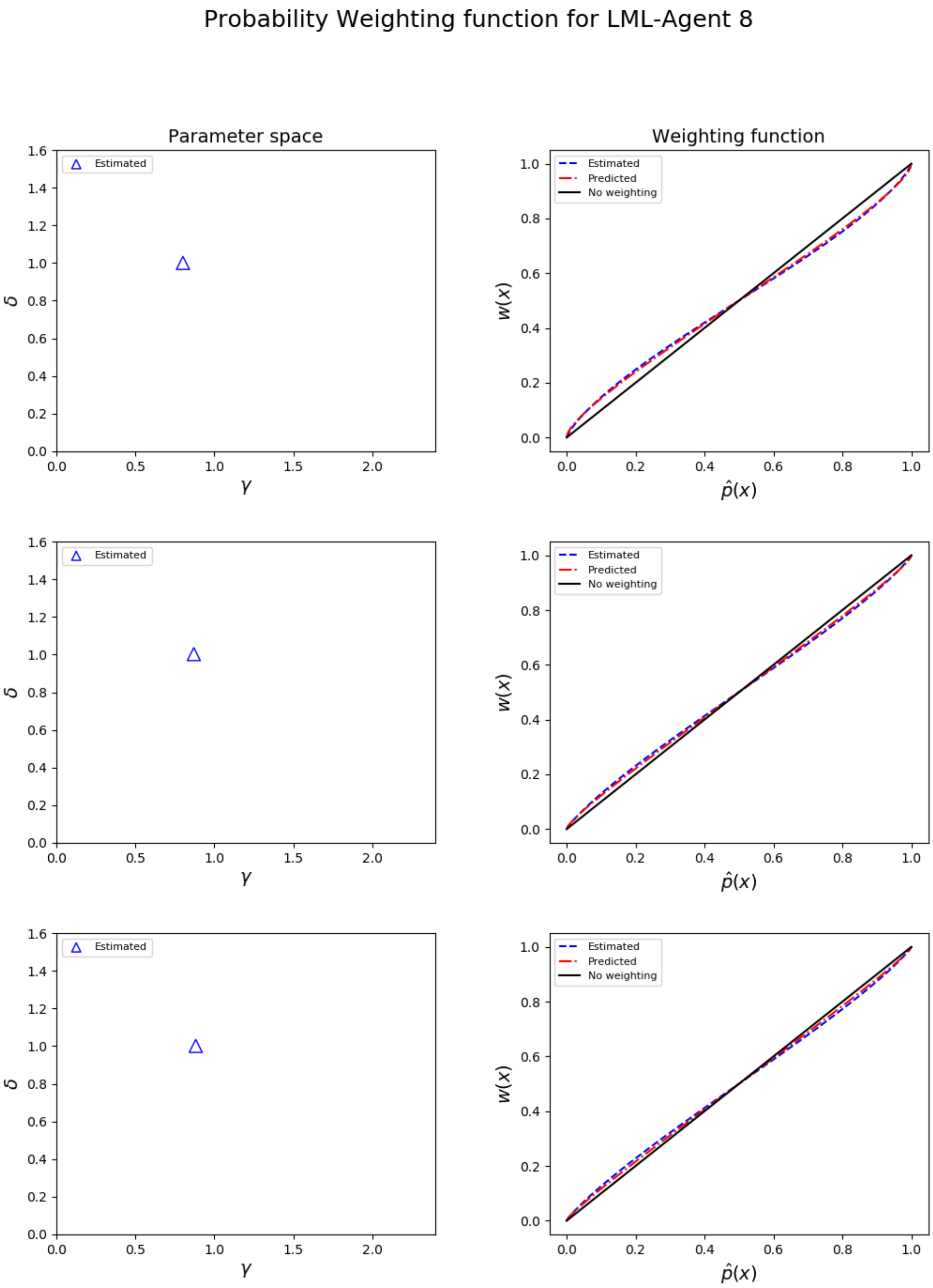

Probability Weighting function for LML-Agent 4
Parameter space
Weighting function
Estimated
Predicted
No weighting
δ
γ
w(x)
p̂(x)
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
0.5
1.5
2.0

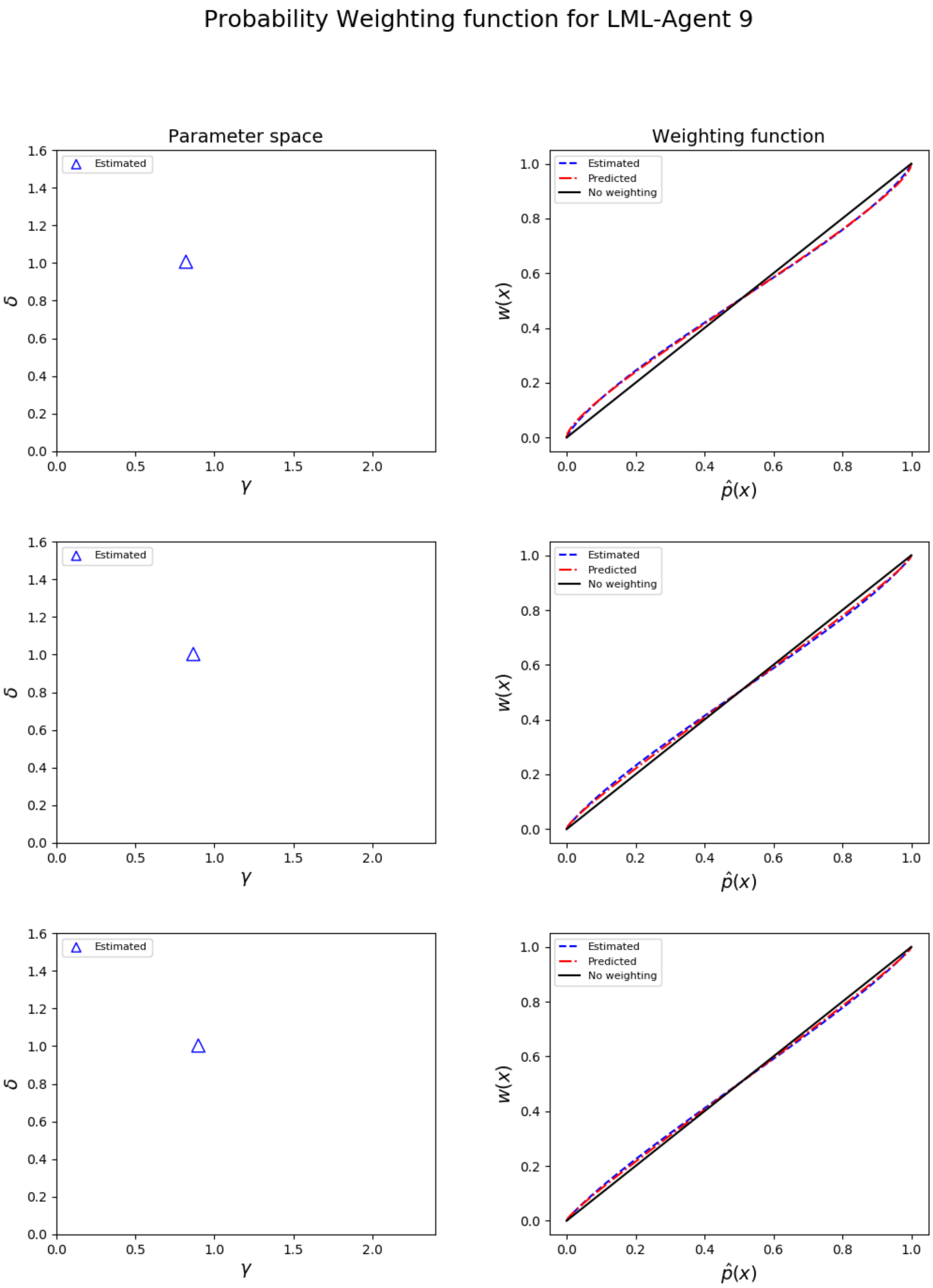

Probability Weighting function for LML-Agent 5
Parameter space
Weighting function
Estimated
Predicted
No weighting
δ
γ
$w(x)$
$\hat{p}(x)$

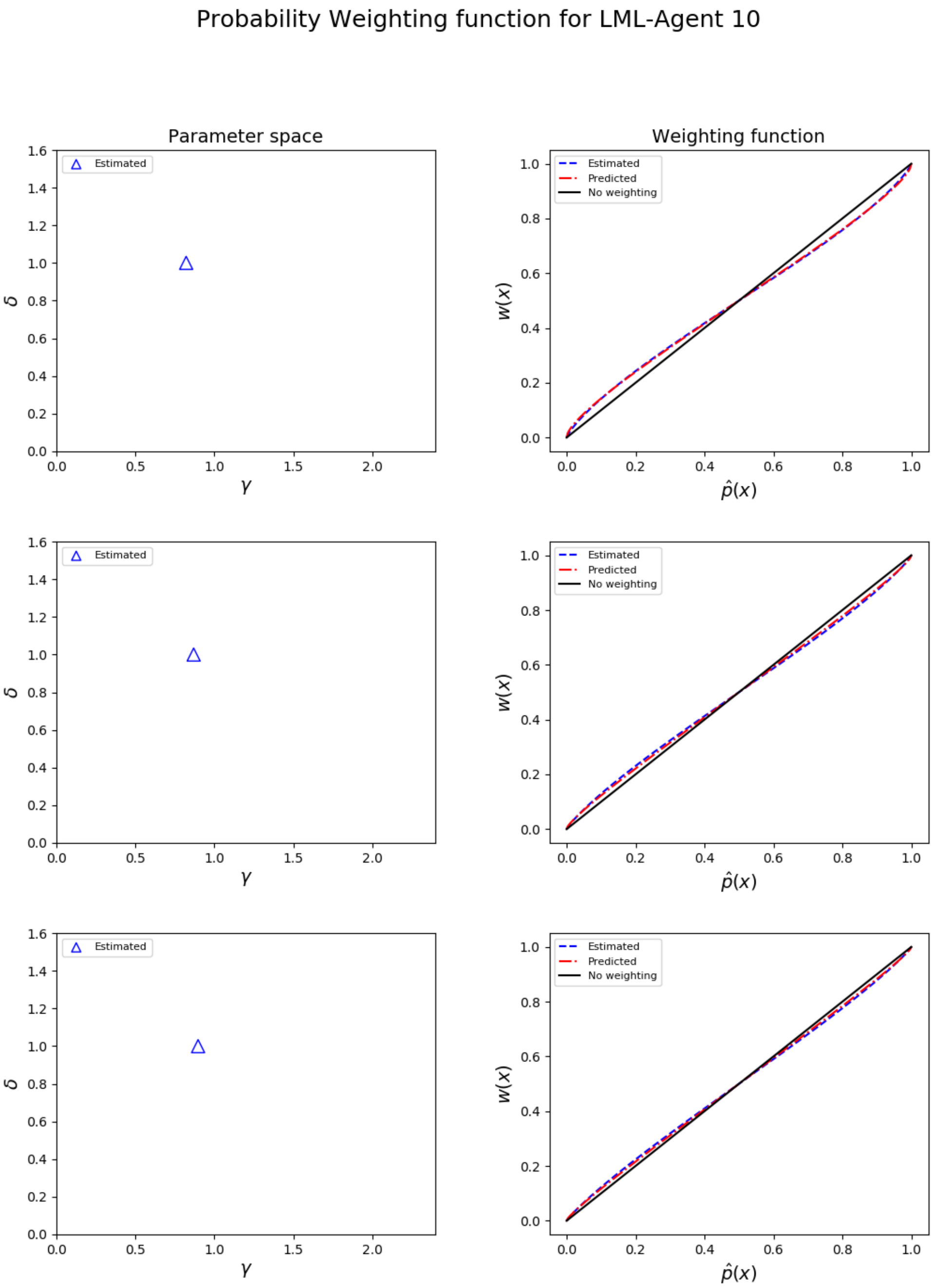
Probability Weighting function for LML-Agent 6




Probability Weighting function for LML-Agent 7
Parameter space
Weighting function
Estimated
Predicted
No weighting
δ
γ
$w(x)$
$\hat{p}(x)$


Probability Weighting function for LML-Agent 8
Parameter space
Weighting function
Estimated
Predicted
No weighting
δ
γ
w(x)
p̂(x)


Probability Weighting function for LML-Agent 9
Parameter space
Weighting function
Estimated
Predicted
No weighting
δ
γ
w(x)
p̂(x)


Probability Weighting function for LML-Agent 10
Parameter space
Weighting function
Estimated
Predicted
No weighting
δ
γ
$w(x)$
$\hat{p}(x)$
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0.5
1.5
2.0