Violations of coalescing in parametric utility measurement

Andreas Glöckner[1,2] · Baiba Renerte[3,4] · Ulrich Schmidt[5,6,7]

Abstract

The majority consensus in the empirical literature is that probability weighting functions are typically inverse-S shaped, that is, people tend to overweight small and underweight large probabilities. A separate stream of literature has reported event-splitting effects (also called violations of coalescing) and shown that they can explain violations of expected utility. This leads to the questions whether (1) the observed shape of weighting functions is a mere consequence of the coalesced presentation and, more generally, whether (2) preference elicitation should rely on presenting lotteries in a canonical split form instead of the commonly used coalesced form. We analyze data from a binary choice experiment where all lottery pairs are presented in both split and coalesced forms. Our results show that the presentation in a split form leads to a better fit of expected utility theory and to probability weighting functions that are closer to linear. We thus provide some evidence that the extent of probability weighting is not an ingrained feature, but rather a result of processing difficulties.

Keywords Decision making under uncertainty · Cumulative prospect theory · Expected utility theory · Violations of coalescing · Event-splitting effects

1 Introduction

Experiments on decision making under risk mostly employ a *coalesced* presentation of lotteries, i.e., branches which lead to the same consequences are combined and the respective probabilities are added up. However, presenting gamble pairs in a canonical *split* form makes them easier to compare and process for the decision maker since, in the case of binary choice, both gambles involve the same set of probabilities. For illustration, consider the classic paradox of Allais (1953), also termed common consequence effect, where M$ denotes millions of dollars. Figure 1 presents the Allais paradox in the commonly used coalesced form. Here, subjects

---

Extended author information available on the last page of the article

tend to choose option A in Choice 1 and option B' in Choice 2, which constitutes a violation of expected utility.

Birnbaum (2004) showed that violations of expected utility (EU) in the common ratio effect can be substantially reduced if the gamble pairs are presented in their canonical split form as depicted in Fig. 2. In the canonical split form, which is a commonly known way of splitting, both lotteries are split such that there are equal probabilities on corresponding ranked branches and the numbers of branches are equal in both gambles and minimal. The presentation in Fig. 2 makes it more transparent that both gambles in each choice have an 89% chance of a common outcome, which should be ignored when determining the preferred option under EU.

Also other typical violations of EU, like the common ratio effect or violations of transitivity, are less frequently observed for split than for coalesced presentation of gambles (Humphrey 2001; Schmidt and Seidl 2014; Birnbaum et al. 2017). The fact that coalesced and split presentation of gambles can lead to systematically different choice behavior has already been discussed by Starmer and Sugden (1993) and Humphrey (1995) under the term *event-splitting effects*. Nevertheless, the implications of these effects have remained largely unexplored in the economics and management literature. The present paper focuses on two questions in this context.

Question 1 is devoted to the shape of the probability weighting function. Many descriptive alternatives to the EU, like prospect theory (Kahneman and Tversky 1979; Tversky and Kahneman 1992) or rank-dependent utility (Quiggin 1982), integrate a non-linear distortion of probabilities formalized by a probability

Choice 1			
A:	100% to win 1 M$	B:	10% to win 5 M$
			89% to win 1 M$
			1% to win 0 M$
Choice 2			
A':	11% to win 1 M$	B':	10% to win 5 M$
	89% to win 0 M$		90% to win 0 M$

Fig. 1 Common consequence effect in a coalesced form

Choice 1			
A:	10% to win 1 M$	B:	10% to win 5 M$
	89% to win 1 M$		89% to win 1 M$
	1% to win 1 M$		1% to win 0 M$
Choice 2			
A':	10% to win 1 M$	B':	10% to win 5 M$
	1% to win 1 M$		1% to win 0 M$
	89% to win 0 M$		89% to win 0 M

Fig. 2 Reduced common consequence effect in a split form

weighting function in their utility representation. Nowadays, the majority consensus in the literature is that this function is typically inverse-S shaped, i.e., small probabilities are overweighted, whereas large ones are underweighted (Wu and Gonzalez 1996; Gonzalez and Wu 1999; Abdellaoui 2000; Bleichrodt et al. 2001; although not unanimously, see, e.g., Hertwig 2012 for critique). The evidence has mainly been derived using coalesced presentation of the lotteries. We analyze how the shape of the probability weighting function differs if the gamble pairs are presented in a split form instead. Since probabilities are easier to compare in split form, an absence or diminished extent of the typical non-linear shape in this form could indicate that the previous evidence mostly reflects difficulties in processing probabilities instead of an ingrained non-linear weighting of probabilities.

Question 2 is related but more general. As violations of EU decrease under the split form, it could be more suitable for elicitation of von Neumann–Morgenstern utility functions in EU. The split form could thus improve prescriptive decision analysis, as the assessment of von Neumann–Morgenstern functions is central in this context (von Winterfeldt and Edwards 1986; McCord and de Neufville 1986; Fischoff 1991; Bleichrodt et al. 2001).

Our analyses of both questions are based on parametric analysis with fitting the parameters of EU and rank-dependent utility (RDU)—with the latter corresponding to the gain-domain parameters of cumulative prospect theory (CPT). Firstly, to answer Question 2, we ask whether the fit of EU improves if we use choice data from split lotteries. We then extend this analysis from the EU to the RDU. To keep the analysis manageable, we restrict our attention to pure gain gambles.

Up to date, only limited work has been done to examine what intelligible impact, if any, failing to account for splitting effects shows in the RDU framework. Indeed, quite little is known about the impact of the splitting effects on the estimated values of the central parameters of RDU or the model fit. Real-life gambles do not always occur in a split form, but explicitly presenting them in a split form could improve the fit of RDU. Moreover, it could alter the features of the probability weighting function, namely, the magnitude of the usually observed non-linear shape, which brings us back to Question 1. And this is indeed a result that we find: The split form significantly improves prescriptive decision analysis.

The paper is organized as follows. We discuss the theoretical background and related literature in Sect. 2. We lay out the experimental design and estimation approaches in Sect. 3 and present the results in Sect. 4. Finally, we discuss the limitations and implications for future research in Sect. 5 and 6.

2 Background

2.1 Expected utility, rank-dependent utility and cumulative prospect theory

We consider a set of real-valued outcomes X. $\mathcal{P}$ denotes the set of all gambles or lotteries over X. A gamble $P \in \mathcal{P}$ satisfies the axioms of Kolmogoroff (1933), i.e., $0 \leq p(x_i) \leq 1 \, \forall \, x_i \in X$ and $p(X) = 1$. Preferences of the decision are formalized by a binary relation $\succcurlyeq \, \subseteq \mathcal{P} \times \mathcal{P}$. A function $V : X \to \nabla$ represents $\succcurlyeq$ on $\mathcal{P}$ if and only if

$P \succcurlyeq Q \Leftrightarrow V(P) \geq V(Q)$. For a gamble with n possible outcomes, the preferences in EU can be represented by

$$V(P) = \sum_{i=1}^{n} u(x_i)p(x_i), \tag{1}$$

where u is the von Neumann–Morgenstern utility function. In parametric analysis, u is commonly assumed to be a power function $u(x_i) = x_i^{\alpha}$, where $1 - \alpha$ is the coefficient of relative risk aversion.

In the CPT framework, outcomes of a gamble are ordered in an increasing order $x_1 \leq \cdots \leq x_k \leq 0 \leq x_{k+1} \leq \cdots \leq x_n$ and preferences can be represented by a sum of two RDU functionals

$$V(P) = \sum_{i=1}^{k} \pi_i^- v(x_i) + \sum_{j=k+1}^{n} \pi_j^+ v(x_j), \tag{2}$$

where sign dependence, reference dependence and rank dependence are all satisfied (Tversky and Kahneman 1992). In parametric analysis, the value function is mostly taken as a two-part power function:[1]

$$v(x) = \begin{cases} x^{\alpha} & x \geq 0 \\ -\lambda(-x)^{\beta} & x < 0 \end{cases}. \tag{3}$$

It assigns a number $v(x)$ to each outcome x to describe the subjective value of the outcome relative to a reference point. The reference-dependent S-shaped value function with $v(0) = 0$ firstly exhibits diminishing sensitivity to gains and losses, such that the function is concave (with $0 < \alpha < 1$ exhibiting risk aversion for gains) or convex (with $0 < \beta < 1$ exhibiting risk seeking behavior for losses), respectively. Secondly, the value function implies loss aversion (when $\lambda > 1$) in that a loss of a given amount has more impact on the attractiveness of a prospect than a gain of an equivalent amount: $-v(-x) > v(x)$ for all $x > 0$ (Kahneman and Tversky 1979, Tversky and Kahneman 1991). Note, however, that we focus exclusively on the gain domain in our analyses.

The decision weights π^+ for the cumulative probabilities of positive outcomes in (1) are defined by

$$\pi_n^+ = w^+(p_n), \tag{4}$$

with

$$\pi_j^+ = w^+(p_j + \cdots + p_n) - w^+(p_{j+1} + \cdots + p_n), \; k < j < n. \tag{5}$$

The probability weighting function w^+ is strictly increasing and continuous. It is defined for the whole probability domain $[0, 1]$ and satisfies $w^+(0) = 0$ and $w^+(1) = 1$.

---

[1] Please note that several other specifications of the functional forms for the value function have also been proposed (for example, exponential or logarithmic forms, see Stott (2006) for a comprehensive review). Yet, according to the author, the power function has proven to be the most favorable empirically.

In their work, Tversky and Kahneman (1992) propose fitting the data to the following single-parameter functional form of the probability weighting function:

$$w^+(p) = \frac{p^{\gamma^+}}{\left(p^{\gamma^+} + (1-p)^{\gamma^+}\right)^{1/\gamma^+}}. \tag{6}$$

The typically observed inverse-S shape of this weighting function (henceforth TKW) exhibits overweighting of small probabilities (up to the crossover point where $w(p) = p$) and underweighting of large probabilities.

Furthermore, to contribute to the empirical tractability of the model, we consider additional parametric specifications of the probability weighting function. For example, the two-parameter linear-in-log-odds specification introduced by Goldstein and Einhorn (1987) has been claimed to be the most commonly used specification of the probability weighting function (Booij et al. 2010). This function (henceforth GEW) is given by

$$w(p) = \frac{\delta p^\gamma}{\delta p^\gamma + (1-p)^\gamma}, \tag{7}$$

where *two* parameters independently, instead of one, explain the shape of the weighting function. Namely, the γ parameter (which is usually assumed $0 < \gamma < 1$ to maintain the inverse-S shape) allows controlling for the curvature of the function and thus serves as a discriminability indicator, while the δ parameter[2] ($\delta > 0$) explains the elevation of the function and thus serves as an attractiveness indicator (Tversky and Kahneman 1992; Gonzalez and Wu 1999). Given that the two properties often do not covary, this specification can offer a substantial advancement in relation to the Tversky and Kahneman (1992) specification outlined above (Gonzalez and Wu 1999; Booij et al. 2010). We should note, however, that the exact values of the named weighting parameter estimates are subject to possible interaction effects,[3] e.g., between δ and α, which both take into account some part of risk aversion (Nillson et al. 2011; Glöckner and Pachur 2012).

Another two-parameter specification enabling similar interpretation of the γ and δ parameters as GEW was introduced by Prelec (1998) (henceforth P2W):

$$w(p) = \mathrm{e}^{-\delta(-\ln(p))^\gamma}. \tag{8}$$

A special case of the P2W with $\delta = 1$ is expressed in the single-parameter form (henceforth P1W):

---

[2] Although generally there tend to be no further restrictions imposed on δ, Glöckner and Pachur (2012) propose restricting the δ parameter values (in GEW) between 0 and 4.

[3] If a relatively large number of free parameters are being estimated for choice data on proportionally fewer games, some interactions might occur due to negative correlations of δ and α. Generally, the functional relationships between the CPT model's parameters are still subject to debate (Nillson et al. 2011).

$$w(p) = e^{-(-\ln(p))^{\gamma}}. \tag{9}$$

It can reportedly outperform the other weighting functions presented above if used in combination with the power value function in (3) (Stott 2006).

Previous studies have delivered mixed results regarding which of the four functional forms of the probability weighting function provides the best fit (see, e.g., Wu and Gonzalez 1996; Gonzalez and Wu 1999; Sneddon and Luce 2001). In this paper, we thus consider the power value function in combination with all four functional forms of the weighting function presented above: the one-parameter TKW and P1W and the two-parameter GEW and P2W. In addition, we introduce a fifth weighting function attributable to the EU: $w(p) = p$ (restricting the γ parameter to the value of unity such that no weighting taking place, henceforth denoted EUW) for benchmark purposes as a special linear case of the one-parameter weighting function.

2.2 Event-splitting effects

By definition, an event-splitting effect (also called violation of coalescing) occurs when a reversal of preference arises in response to a coalesced- versus split-form change of the same choice. By using the term *split form* of a gamble throughout the paper, we refer to the canonical split form, for which both gambles of a choice are split in a way that allows the corresponding ranked branches to have equal probabilities while keeping the number of branches minimal (also implying the same number of branches in both gambles).[4] For example, a choice between $A = (€40, 0.1; €40, 0.1; €2, 0.8)$ and $A' = (€98, 0.1; €2, 0.1; €2, 0.8)$ is called the canonical split form of the choice between the coalesced $B = (€40, 0.2; €2, 0.8)$ and $B' = (€98, 0.1; €2, 0.9)$ (Birnbaum and Navarrete 1998).

Kahneman (2003) suggests that "most decision makers will spontaneously transform the former prospect into the latter and treat them as equivalent in subsequent operations of evaluation and choice" (p. 727). This observation largely coincides with the definition of coalescing, which refers to an assumption that any two or more branches leading to the same outcome can be combined by adding their probabilities without affecting the utility of the gamble, such that $A = (x, p; x, q; y, 1 - p - q) \sim B = (x, p + q; y, 1 - p - q)$ and additionally $C = (x, p; y, q; y, 1 - p - q) \sim D = (x, p; y, 1 - p)$. Coalescing implies that $A \succ C$ if and only if $B \succ D$. From above, $A \sim B$ and $C \sim D$. Therefore, $A \succ C$ if and only if $B \sim A \succ C \sim D$. Thus, by transitivity, $B \succ D$ (Birnbaum et al. 2017). Because coalescing and transitivity should be satisfied within the CPT framework with any

---

[4] Before examining coalescing in more detail, it is important to define the elements called *branches* that comprise any gamble in question. A branch is a probability-consequence pair that is distinct in the presentation to the decision maker. To illustrate this, consider two gambles: a three-branch gamble $A = (x, p; x, q; y, 1 - p - q)$, for which the three branches lead to outcomes x, x and y with probabilities p, q and $(1 - p - q)$, respectively; and a two-branch gamble $B = (x, p + q; y, 1 - p - q)$, for which the two branches are $(x, p + q)$ and $(y, 1 - p - q)$ for outcomes x and y respectively. A keen observer will notice that the two gambles appear objectively equivalent, as the combination of outcomes for the first two branches of gamble A, (x, p) and (x, q) lead to the same result as the first branch of gamble B, $(x, p + q)$. Thus gamble B is the two-branch coalesced form of the three-branch gamble A; and conversely, gamble A is the split form of gamble B.

$w(p)$ function, there should also be no splitting effects (Birnbaum and Navarrete 1998; Luce 1998; Birnbaum 2008; see Appendix 1 for a proof illustrating how the CPT's assumption of rank dependency implies the satisfaction of coalescing).

In the meantime, there exists abundant evidence showing that people do not treat coalesced-form and split-form gambles as equal (e.g., Conlisk 1989; Starmer and Sugden 1993; Humphrey 1995). In fact, it has been shown that splitting the branch with the highest available outcome can increase the attractiveness of a gamble in comparison to a coalesced form of the same gamble. Conversely, splitting the branch with the lowest available outcome decreases the attractiveness of the gamble (Starmer and Sugden 1993; Humphrey 1995, 2001). Splitting both the highest and lowest branches in a binary gamble with two equiprobable positive branches tends to make the gamble worse, in compliance with loss aversion (Birnbaum 2008).

Interestingly, it appears that violations of coalescing cannot be attributed to lack of knowledge, as they are persistent even in people with doctoral degrees who are familiar with the literature on decision making (Birnbaum 1999). Neither can the splitting effects be explained by errors, as they are still persistent when errors are factored out (Birnbaum et al. 2017), nor can the effects be attributed to the particular format[5] used for presenting, or framing, the gambles (Birnbaum 2004, 2006; Birnbaum et al. 2008). Decision heuristics, like anchoring and adjustment, cannot account for the observed splitting effects either (Humphrey 1996). Meanwhile, the results regarding the effects of specific learning and experience are still mixed (see, e.g., Humphrey 2006; Birnbaum and Schmidt 2015).

Indeed, it seems that people simply "do not obey coalescing" (Birnbaum 2007, p. 171). The assumption that coalesced and split forms of the same gamble would be treated equivalently is thus "empirically false" (Birnbaum 2008, p. 464). And yet, CPT is still argued by many to be "the "best", if imperfect, description of decision making under risk and uncertainty" (Birnbaum 2008, p. 463). In light of this controversy, a question arises regarding what intelligible impact, if any, the splitting effects have on the conclusions drawn by the "imperfect" CPT framework.

We examine which of the two gamble presentation forms (coalesced versus split) leads to more normatively accurate, or rational, results. In this context, the term *rationality* is used in the sense proposed by von Neumann and Morgenstern (1947). It implies being in line with the normative preference axioms of the EU and, most notably, the substitution axiom.

We hypothesize that the normative EU explains data comprised of split-form gambles better than data comprised of coalesced-form gambles, while it is unclear whether this is to be the case for the descriptive RDU model.[6] A couple of previous

---

[5] We use the term *format* to denote the way how the gambles are displayed (in arrangement, position, color etc.) to the experiment subjects, as opposed to the term *form* that merely refers to the coalesced- or split-form branches of a gamble. While the form should have no effect on preferences in the CPT, the format (or framing) can have a significant effect (Tversky and Kahneman 1986).

[6] It is also likely that violations of coalescing would cause systematic deviations in the CPT parameter estimates themselves. Given the evidence that splitting effects come about following from the way the probability information in a gamble is compared and processed (e.g., Humphrey 1995; Schmidt and Seidl 2014), it is likely that the splitting of gambles influences the estimated CPT weighting function parameter values. Related previous studies indicate that the weighting function for split-form gambles might, in fact, exhibit behavior that contrasts the classic CPT assumption of a non-linear inverse-S shaped weighting

studies have already focused on the CPT's fit to data that indirectly test coalescing (e.g., Birnbaum and Chavez 1997; Birnbaum and Navarrete 1998), but little is known about how well selected data comprised of coalesced- versus split-form gamble pairs fit an RDU model with varying functional specifications.

3 Methodology

3.1 Experimental procedure

We examine these questions using data from a pairwise choice experiment (see Birnbaum et al. 2017), conducted with 54 student subjects at the University of Kiel in Germany (all undergraduate students, 61% in the economics and business administration programs; 22.0 years old on average; of them, 21 female). The experiment is based on a random-lottery incentive mechanism, which is a commonly used one-step choice-based elicitation approach that lets subjects face multiple pairs of gambles in a sequence and choose a preferred gamble for each of the pairs (see, e.g., Hey and Orme 1994; Wu and Gonzalez 1996).

At the end of the experiment, one pair is chosen randomly and played out for real. Each pair consists of a risky gamble $R = (x_1, p_1; x_2, p_2; x_3, p_3; x_4, p_4)$ and a safe gamble $S = (y_1, q_1; y_2, q_2; y_3, q_3; y_4, q_4)$ with two to four outcomes x_i, y_i and respective probabilities p_i, q_i that are systematically varied. The choices between the gambles are presented in a pseudo-random order and the outcomes x_i, y_i are ordered from the lowest to the highest within each gamble (see Appendix 2 for an example of the presentation format and the experiment instructions).

The dataset comprises 28 gamble pairs (14 of them presented in a split form, see Appendix 3 for an overview of all gambles and the respective descriptive statistics of gamble choices), implying $4 \times 28 = 112$ choice situations faced by each subject over four repetitions, that is, $54 \times 4 \times 28 = 6048$ choice situations observed in total, 3024 of them in a split form. Note, however, that the gambles in our study are relatively specific in that no certain outcomes are included and the values and probabilities of high outcomes are relatively similar between the gambles within a decision. The subjects received, on average, a €19.1 cash reward (including a €5.0 show-up reward) for an approximately 90-minute session, leading to an average reward of €12.8 per hour.

3.2 Structural modeling

We apply structural methods to jointly estimate several core parameters of the EU and RDU frameworks. In particular, we use maximum likelihood estimation (MLE) to determine the most likely parameter values to have generated the given dataset

---

Footnote 6 continued

function (e.g., Birnbaum and McIntosh 1996; Birnbaum and Chavez 1997). Hence, we additionally hypothesize finding significant differences in γ (and δ) parameters between coalesced- and split-form gambles, respectively. We expect that this result would remain robust also when controlling for additional variables such as the numbers of branches or outcomes in a gamble as well as demographic information.

within the specification bounds. In addition, the log-likelihood of the MLE allows us to measure the goodness of fit of the respective frameworks.

RDU includes subjective values of outcomes and subjective weights of probabilities. For the subjective values, we use a power value function in all models. We thus add the α parameter as the first one in the list of estimable parameters. For the probability weighting, we examine four weighting function specifications given that, firstly, the probability weighting function appears to be central for considering the coalesced- versus split-form data and, secondly, the explanatory power of the RDU model depends on the function specifications and the corresponding interaction effects (e.g., Stott 2006). We thus add the γ (and δ, where relevant) parameter to the list of estimable parameters (see Appendix 7 for a full list of parameters in each model).

In addition, we extend the RDU to accommodate stochastic behavior by applying an exponential specification of the choice rule[7] of Luce (1959) (see also Rieskamp 2008). The exponential specification of Luce's choice rule is defined as

$$p(R,S) = \frac{e^{\phi v(R)}}{e^{\phi v(R)} + e^{\phi v(S)}}, \tag{10}$$

where $p(R, S)$ stands for the probability of choosing the risky gamble R over the safe gamble S and the sensitivity parameter $\phi > 0$ specifies how sensitively the model reacts to differences between the subjective values $V(R)$ and $V(S)$ of the gambles R and S, respectively (Rieskamp 2008; Nillson et al. 2011). Consequently, given that it is feasible to structurally estimate all specified parameters jointly with MLE, we add the ϕ argument to the list of estimable parameters.

Using the choice rule, we attempt to quantify the goodness of fit of the RDU model predictions given the actual choices between the gambles. A useful tool for the purpose is the deviation measure G^2, expressed as

$$G^2 = -2 \sum_{i=1}^{N} \ln[f_i(y|\theta)], \tag{11}$$

with i denoting the choice among gambles and N denoting the total number of gambles. $f_i(y|\theta)$ expresses the probability that the RDU model with its parameter values θ predicts a choice y, such that $f_i(y|\theta) = p_i(R,S)$ if the gamble R is chosen and $f_i(y|\theta) = 1 - p_i(R,S)$ if the gamble S is chosen. Low values of G^2 are indicators of good choice predictions and, hence, a good fit of the RDU model (Rieskamp 2008). A directly related measure of fit is the Akaike information criterion[8] (AIC) that additionally adjusts for the complexity of the model (namely, the number of parameters in a given specification) and thus allows comparing the explanatory power of differing models. The AIC is defined as

---

[7] Stott (2006) argues that this general specification by Luce "performs well" (p. 102) and is "the most psychologically motivated rule tested" (p. 109).

[8] According to Burnham and Anderson (2002), the AIC is theoretically superior to the Bayesian information criterion (BIC) measure (sometimes referred to as the Schwarz's Bayesian criterion), although they are generally equivalent. Accordingly, only the AIC measure will be relevant for the scope of this paper.

$$AIC = G^2 + 2n, \tag{12}$$

where n stands for the number of free parameters in a model (Akaike 1973). As a rule, an AIC difference of $\Delta AIC > 10$ for two given models strongly favours the model with the lowest AIC measure over the other (Burnham and Anderson 2002).

4 Results

4.1 Fitting coalesced- and split-form data to RDU and EU

For the main hypothesis, we analyze the AIC measures for the EU specification using the coalesced-form and split-form data separately. This approach allows us to consider whether, in some settings, one could find grounds for preferring the one form of presenting the gambles over the other. The linear case of the RDU one-parameter weighting function (with restricting the γ parameter to unity, thus attaining EUW) shows clear evidence that the split-form gamble pairs provide a better fit than coalesced-form gambles in the EU specification (coalesced form $AIC^{EUW} = 3870$, split form $AIC^{EUW} = 3816$, see ΔAIC in Table 1). However, the results also indicate that for the RDU specifications, the coalesced form data provide a better fit than the split-form data. Namely, three out of the four core RDU specifications show a ΔAIC that notably exceeds 10, thus providing a result in favor of the coalesced-form data, while the fourth (TKW) shows a ΔAIC smaller than 10, providing an insufficiently conclusive result.

Interestingly, the opposite holds when examining directly comparable one-type gambles with split highest and lowest branches exclusively (see ΔAIC in Table 2 and the list of gambles in Appendix 3): The model that uses split-form data *outperforms* the same model that uses equivalent coalesced-form data in all four RDU specifications. The question of whether experiments ought to include split-form gamble pairs rather than their coalesced equivalents to ensure a more accurate preference elicitation thus finds some confirmation here but is still open for further examination.

The four RDU specifications outperform the EU specification in terms of the model fit (see ΔAIC^{EUW} in Table 2) for the coalesced-form data. For the split-form data of the directly comparable one-type gambles, the differences between the AIC^{EUW} and the respective measures of fit for the four RDU specifications no longer exhibit significant differences.

Finally, the one-parameter P1W outperforms TKW and the two-parameter P2W outperforms GEW in terms of the fit measures in practically all of the examined models. Therefore, in the following discussion of results, we focus on these best-performing models in particular; see Appendices 4–6 for the full results of all examined RDU specifications and model versions with various control variables. In total, we use four *versions* of the model (referred to as models M1 to M4, see Appendix 7 for summary).

Table 1 Comparison of the model fit with 28 gamble pairs considered

	TKW	GEW	P1W	P2W	EUW	
Coalesced-form data AIC_{coa}	3782	3753	3689	3690	3870	
$\Delta AIC^{EUW} = AIC_{coa}^{EUW} - AIC_{coa}^{w_i}$	+88	+117	+181	+180		Better fit for CPT than EU
Split-form data AIC_{split}	3779	3781	3804	3786	3816	
$\Delta AIC^{EUW} = AIC_{split}^{EUW} - AIC_{split}^{w_i}$	+37	+35	+12	+30		
$\Delta AIC = AIC_{coa} - AIC_{split}$	$< \lvert -10 \rvert$	$> \lvert -10 \rvert$	$> \lvert -10 \rvert$	$> \lvert -10 \rvert$	> 10	

The model that uses coalesced-form data outperforms the model that uses split-form data

The model that uses split-form data outperforms the model that uses coalesced-form data

The four RDU specifications are applied on split-form data as compared to coalesced-form data with the whole dataset considered

Table 2 Comparison of the model fit with 16 one-type gamble pairs considered

	TKW	GEW	P1W	P2W	EUW	
Coalesced-form data AIC_{coa}	2176	2141	2119	2117	2204	Better fit for CPT than EU
$\Delta AIC^{EUW} = AIC_{coa}^{EUW} - AIC_{coa}^{w_i}$	+28	+63	+85	+87		
Split-form data AIC_{split}	2105	2107	2105	2107	2104	
$\Delta AIC^{EUW} = AIC_{split}^{EUW} - AIC_{split}^{w_i}$	-1	-3	-1	-3		No difference between CPT and EU fit
$\Delta AIC = AIC_{coa} - AIC_{split}$	> 10	> 10	> 10	≥ 10	> 10	

The model that uses split-form data outperforms the model that uses coalesced-form data

The four RDU specifications are applied on split form data as compared to coalesced-form data with 16 out of the 28 gamble pairs considered

Table 3 Results of the models M1 and M2 in two best-performing RDU specifications

	Model M1			Model M2	
	P1W	P2W		P1W	P2W
α_{coa}	0.654	0.709			
	(0.022)	(0.054)		0.653	0.729
α_{split}	0.651	0.759**		(0.025)	(0.051)
	(0.033)	(0.052)			
γ_{coa}	0.840	0.808		0.839	0.814
	(0.020)	(0.015)		(0.022)	(0.016)
γ_{split}	0.998**	1.034***		0.999***	0.998***
	(0.029)	(0.064)		(0.026)	(0.055)
δ_{coa}		6.242			6.070
		(0.671)			(0.654)
δ_{split}		3.232***			3.546***
		(0.410)			(0.380)
ϕ	5.063	1.132		5.056	1.147
	(0.483)	(0.068)		(0.469)	(0.067)
AIC	7538	7483		7536	7488
N_{obs}	6048	6048		6048	6048

The symbol ** for the corresponding split parameters indicates a difference between the coalesced- and split-form parameter values at a 95% significance level, *** indicates a difference at a 99% significance level. All differences in this model are due to a binary *splitcoal* variable. All constant terms are significant at a 99% level. As for the constant term representing the ϕ parameter, its values are likely dependent on the α as well as the δ parameters, which explains the differences over the four specifications. Standard errors (in parentheses) are clustered at subject level

4.2 Comparing RDU parameter estimates for the coalesced- and split-form data

The results for the weighting function parameters γ and δ of model M1 in Table 3 are unequivocal. There exist significant differences at a 99% significance level between these parameters when considered in the contexts of coalesced- and split-form gamble pairs. Indeed, also the direction of the differences is consistent over the RDU specifications, with γ_{split} remaining significantly higher than γ_{coa} and δ_{split} remaining significantly lower than δ_{coa} (all at a 99% significance level, according to Wald tests).

These results are, however, worthy of attention not only for their significant differences between the respective coalesced- and split-form γ and δ, but also due to the absolute values of these weighting function parameters. While the coalesced form allows maintaining the γ_{coa} value significantly different (lower) from unity, thus confirming the established RDU predictions (Wald tests, p values < 0.001), the split-form γ_{split} reveals an unusual picture. The estimated parameter value is indifferent from unity in the one-parameter P1W specification and significantly different (higher) from unity in the two-parameter P2W specification (Wald tests,

p values < 0.001). These results thus conflict with the usual results of the RDU framework, namely, the aforementioned $0 < \gamma < 1$.

Model M2 might provide additional insights into this uncommon result. This model, compared to M1, considers the weighting function exclusively, keeping the utility function parameters unchanged. According to Chow tests,[9] the respective weighting parameter values estimated in M1 and M2 do not significantly differ between these model versions for the one-parameter P1W specifications (neither for the corresponding coalesced-, nor split-form parameters), while the two-parameter P2W specification estimates in M2 are different from the estimates in M1 at a 95% significance level for γ_{split}, δ_{coal} and δ_{split}.

The results of M2 confirm the results of M1 and reveal the familiar trend of significantly different weighting function parameters γ and δ in the coalesced-form as compared to the split-form gamble pairs (Wald tests, p values < 0.001), with γ_{coa} significantly different (lower) from unity (Wald tests, p values < 0.001). In this case, the γ_{split} parameters report practically no curvature, which implies that no weighting of probabilities for split-form gamble pairs could be identified in the P1W case (essentially making it equivalent to the EUW case), while for the two-parameter function P2W, the elevation parameter δ_{split} alone assures that the weighting function is curved.[10]

As depicted in Fig. 3 where we plot the results of M1 graphically, the results of the one-parameter weighting functions are admittedly easier to interpret than the two-parameter weighting functions. The implications of splitting—namely, that split-form gambles result in less probability weighting than coalesced-form gambles—hold for both P1W and P2W, but the particular shapes of the split and coalesced P2W should be interpreted with caution, as the δ values are quite high and indicate more pessimism in the split form.

Although there are some limitations to our results, we can indeed conclude that violations of coalescing for split-form gamble pairs in the RDU framework not only explain differences in the model fit, but also affect the subjective weighting function. Namely, the use of the split-form gamble pairs appears to change gamble choices and considerably diminish the weighting as compared to the coalesced-form gambles.

Note that we also check the robustness of these results by considering further weighting function specifications and adjusting the RDU parameter estimates to further variables. Our analysis (see M1, M2 in Appendix 4 and M3, M4 in Appendices 5 and 6) shows that the familiar trend of significantly different weighting function parameters γ and δ for the coalesced- versus split-form gambles remains strong and consistent across all RDU specifications and considered models. Taken together, our results indicate that one source of the more pronounced

---

[9] In the given case when linear-form restrictions are met, a Chow test (for comparisons between, rather than within, models) provides results that are essentially equivalent to the results of a Wald test.

[10] For the elevation parameter, the estimated value and value change between the models exhibit a curious development, which appears to be attributable to the interaction effects between the value function and weighting function specifications. As moderately significant differences in α values are only found in cases where the power value function is combined with a *two-parameter* weighting function, one can reasonably conclude that it is a result of an increased "latitude to interact" (see Stott 2006, p. 102).
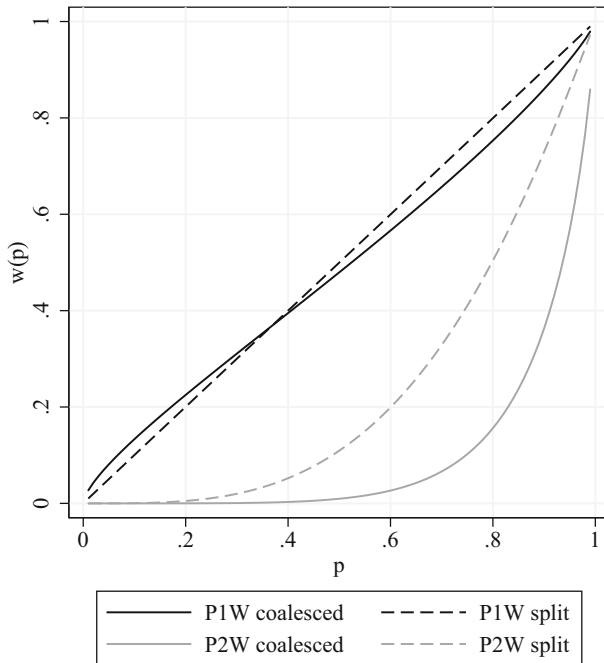
Fig. 3 Illustration of the weighting-function results of M1 in Table 3 above. The solid lines refer to coalesced-form data and the dashed lines refer to split-form data in the one-parameter (black, P1W) and two-parameter (gray, P2W) Prelec weighting functions, respectively

deviation from EU could be comparison difficulties caused by the coalesced form of the presented gambles.

5 Discussion and limitations

The results of this paper are insightful in a number of ways. Firstly, we have shown that the fit of EU is indeed better if the gamble pairs are presented in a split form. This result indicates that the split form improves prescriptive decision analysis. For example, we could conclude that one should rather use split-form than coalesced-form gamble pairs when advising decision making.

Secondly, we have found evidence for significant differences in magnitude between the γ parameters in the RDU weighting functions for coalesced- and split-form gamble pairs in the given dataset. Meanwhile, the somewhat mixed evidence regarding the logically independent elevation parameter δ calls for further examination of this property. Note that the interaction effects between δ and α cannot be ruled out in this setting and are a possible source of the mixed evidence. Still, although the values of δ are quite high (particularly for coalesced gambles), there is no indication for strong effects of cross-parameter compensations: γ and δ

do not vary considerably in concert, and δ values are relatively stable in all model specifications. Whether this result reflects some more fundamental theoretical issues or is merely a method artifact is due to further research.

Thirdly, it appears that presenting gamble pairs in a split form changes gamble choices as to bring the RDU closer to the EU. This is the case not only for the measures of model fit, but also for the properties of the weighting function. With the curvature parameter γ largely closer to unity (and often not significantly different from unity) for split-form than for coalesced-form gambles, the results indicate that the subjects tend to pay less account to subjective probability weighting when evaluating split-form gamble pairs of a certain type.[11] They act comparatively more "normatively" than expected and thus put one of the cornerstones of the RDU (and by implication, CPT) into question.

Note, however, that the gambles in our study were quite specific in that values and probabilities of the high outcomes were relatively similar between the gambles within a decision. Note also that no certain outcomes were included that might drive the typically stronger deviations from linearity of the weighting function. We thus acknowledge the fact that the particular shape of the weighting function might be different with additional gambles and tests for generality are due to further research.

6 Conclusions

The results of this paper invite its readers to carefully rethink RDU and its perspective on the subjective probability weighting or, more particularly, on the stability of the probability weighting function against the editing of lotteries. We have provided some reasons to conclude that the non-linearity in the weighting function might be more pronounced in result of coalescing. That is, probability weighting does not necessarily appear to be an ingrained feature, but rather a result facilitated by processing difficulties.

What do our results imply for utility theory and its applications? Firstly, one could argue in favor of using gamble pairs in a split form and employing EU as a decision criterion, given that it performs rather well for the split-form pairs. However, real-life gambles do not always occur in a split form. Therefore, a second recommendation could be to increasingly employ other theories that imply splitting effects. Other models, like the transfer of attention exchange (TAX) model or the rank-affected multiplicative weights (RAM) model, could potentially be rivals to the EU, RDU and CPT (Birnbaum and Chavez 1997; Birnbaum 1999, 2008).

---

[11] We identify certain contradictions in the estimated CPT parameter values, with some one-parameter models leading to a common inverse-S shape probability weighting function, while others resulting in a linear function for the one-parameter weighting function. Note that the linear shape appears to become less pronounced with addition of further relevant explanatory variables in a model. We thus confirm the proposition by Birnbaum and Navarrete (1998), among others, who argue that the CPT improves its pertinence if the parameters can additionally depend on the number of branches or outcomes in a gamble. Controlling for this has indeed allowed the CPT parameters to be in more compliance with the theory than otherwise. In the meantime, we also find that the importance of the splitting itself is not diminished when these additional variables are added. This result extends the discussion of bounded rationality as a reason behind the observed splitting effects and indicates that the underlying psychological reasoning reaches beyond minimized cognitive expenditures related to the absolute frequency of considered attributes.

Thirdly, one could continue employing RDU (and by implication, CPT), albeit cautiously, knowing that the splitting has a fairly intelligible effect on the weighting function.

Much remains to be done still. Firstly, because this paper concerns coalesced- and split-form gamble pairs in the gain domain exclusively, we advise extending the scope of the forthcoming experiments to also include mixed and loss-only gambles (to examine the splitting effects in the respective parameters for the loss domain, including the loss aversion parameter). Secondly, because the construction of splitting appears to have an influence on the resulting model fit, we advise extending the binary concept of splitting and examining a variety of directly comparable coalesced versus differently split datasets. Thirdly, because the interpretations of the psychological reasons behind the splitting effects are still manifold, we advise gathering further insights from parallel lines of research, such as neuroeconomics and others.

Appendix 1: Proof of coalescing within the framework of CPT

In its most general representation as a part of RDU family, CPT guarantees satisfaction of coalescing (i.e., the equivalence of coalesced and split forms of a given gamble) even if no editing rules are assumed (Birnbaum 2008). In particular, as demonstrated by Birnbaum and Navarrete (1998, p. 57–58), RDU implies coalescing for any $w(p)$ function in

$$V = \sum_{i=0}^{n} v(x_i)[w(p_i) - w(p_{i-1})],\tag{13}$$

where the (decumulative) probability of receiving x_i or more is denoted by p_i, while the probability of receiving more than x_i is denoted by p_{i-1}.

Proof If outcomes $x > y > z > 0$ and probabilities $p + q + r = 1$, then $V(x, p; y, q; z, r) = v(x)w(p) + v(y)[w(p + q) - w(p)] + v(z)[w(1) - w(p + q)]$, by Eq. (13) above.

If outcomes $x = y$, then $V(x, p; x, q; z, r) =$

$$v(x)w(p) + v(x)[w(p + q) - w(p)] + v(z)[1 - w(p + q)]$$
$$= v(x)w(p + q) + v(z)[1 - w(p + q)]$$
$$= V(x, p + q; z, r), \text{ since } r = 1 - p - q.$$

If outcomes $y = z$, then $V(x, p; y, q; y, r) =$

$$v(x)w(p) + v(y)[w(p + q) - w(p)] + v(y)[1 - w(p + q)]$$
$$= v(x)w(p) + v(y)[1 - w(1 - p)]$$
$$= V(x, p; y, 1 - p) = V(x, p; y, q + r), \text{ since } 1 - p = q + r.$$

Appendix 2: Format of the presented gambles

See Fig. 4.

At the beginning of the experiment, the subjects received written instructions that were also read out aloud. Following the instructions, four transparent dominance questions were used to test the subjects' understanding. Only after completing these were the subjects allowed to proceed with the experiment. Four booklets with pairwise choices between presented gambles were then consecutively provided to each of the subjects, with every choice being replicated four times using counter-balanced left–right positioning.

In addition, two transparent stochastic dominance problems were included in each of the booklets to check whether the subjects were being sufficiently attentive. Importantly, both coalesced and split forms of the same gambles were included in the booklets, but the booklets differed in *how* the gambles were included. Namely, while half of the subjects received four booklets with intermixed coalesced- and split-form gambles, the other half received booklets with only the one or the other form gambles.

For further details, please refer to the instructions (translated from German) on the following pages.

Date: _____   **Number:** _____

A:	50% to win 20 Euro	B:	33% to win 10 Euro
	30% to win 30 Euro		34% to win 15 Euro
	20% to win 40 Euro		33% to win 60 Euro

Fig. 4 The presentation design of the gambles. Note: The subjects were asked to mark their preferred gamble by circling it

Instructions

Thank you for your willingness to participate in this study. Your data and personal details will be treated confidentially, will not be passed on to anyone and will not be used for any other purposes than this study.

The study

We want to know how people choose between lotteries, that is, how they choose between different options of winning money.

For this purpose, two lotteries are offered for you to play in each task. Your task is to decide which of the two lotteries you prefer to participate in.

Example 1

Imagine a lottery wheel with 100 lottery tickets. 50 tickets are marked with "0€", whereas the other 50 tickets are marked with "20€". If you draw a ticket from the lottery wheel, you have a 50–50 chance of winning either 0€ or 20€. Both amounts can be drawn with 50% probability.

Imagine now a second lottery wheel, which contains 75 "25€" tickets and 25 "50€" tickets. Would you rather draw a ticket from the first or second lottery wheel? This task is described in the questionnaires as follows:

Example 1: Which lottery would you like to play?

□ A   with a probability of 50% you win 0 €

          with a probability of 50% you win 20 €

OR

□ B   with a probability of 75% you win 25 €

          with a probability of 25% you win 50 €

In this case, it is clear that you should prefer the second lottery B. In this lottery you would win at least 25 €, whereas for lottery A even the maximum prize is only 20 €.

Example 2

In this study, however, it will often depend on your own opinion which lottery you prefer. An example of this is the next task:

Example 2: Which lottery would you like to play?

□ A   with a probability of 50% you win 0 €

          with a probability of 25% you win 5 €

          with a probability of 25% you win 50 €

OR

□ B   with a probability of 50% you win 12 €

          with a probability of 50% you win 14 €

There are three different possible prizes in Lottery A. Lottery B has two different, equally probable prizes. Suppose you choose Lottery A: With 25 lottery tickets you could receive the high prize of €50. However, if you draw one of the other 75 tickets from Lottery A, i.e., €0 or €5, you would have been better off if you had chosen Lottery B, as in Lottery B you will win at least €12. Which lottery would you rather play?

Some would rather play the first lottery A, others would rather play the second lottery B. It is just a matter of opinion. There is no "right" or "wrong" answer! Once you have made your choice, tick the appropriate lottery in the box.

Rules

For all lotteries in this study, you can imagine a lottery wheel with 100 tickets with cash winnings. All tickets can be randomly drawn from the wheel with the same probability. Therefore, the probability of winning a prize corresponds to the number of tickets with exactly this prize in the wheel. In each task, the probability and amount of all possible winnings for both lotteries are given. The probability adds up to 100% in each lottery.

Please answer the questions alone, do not talk to the other participants during the study. If you violate this rule, you will be excluded from the study and will not receive any payment.

The study will take about 1 h and 15 min. We ask all participants to stay until the end and not to leave earlier. There are enough tasks for the whole period. Payment is not made until the end of the study.

Payment

Each participant receives a participation fee of 5 €. In addition, one of the tasks is randomly selected at the end of the study. For this purpose, you are asked to draw a number from a bag. It is checked which lottery you have chosen for this task and this lottery is then played, again by drawing a number.

The lowest possible prize is 0 €. If you draw this prize, you go home with 5 €. The highest possible prize is 60 €. In this case you would receive 65 € in total. The other profit opportunities are in between, the 5 € participation fee is always added. Your earnings are paid out immediately in cash.

Thus, your payment depends on the decisions you make during the study!

Further procedure

On the next page, please answer the general questions first.

On page four you will find four exercises. When you have answered them, please raise your hand. Someone will check the exercises to see if you have understood them.

You will then receive your first task sheet. Each time you have finished answering a sheet, please raise your hand, then you will receive the next task sheet.

Questions?

If you have any questions, raise your hand and someone will come to you. Please don't ask your neighbors.

Appendix 3: Structure of the presented gambles in coalesced and split forms

See Table 4.

Appendix 4: Exploring M1 and M2

See Tables 5 and 6.

One possible interpretation of the one-parameter weighting function estimates in M1 and M2 is as follows. As already noted, a weighting parameter $\gamma < 1$ implies an inverse-S shape of the one-parameter weighting function, with overweighting of small probabilities and underweighting of large probabilities (while weighting the middle outcomes less extremely). If the value of γ in the one-parameter weighting function is larger than unity, however, it rather implies a less[12] conventional S-shaped weighting function that is convex for smaller probabilities and concave for larger probabilities, such that extreme outcomes under RDU are underweighted.[13]

Note that the more outcomes in a single gamble are given, the smaller the probabilities for each outcome logically become, such that effectively only the probability range closer to 0 becomes relevant (Wakker 2010). The explanation could then also be consistent with the reasoning behind bounded rationality, considering that for gambles with three or more outcomes or branches, the smallest probabilities in absolute terms might be disregarded (e.g., Payne 2005). Indeed, some authors argue that these opposite relations of weight ratios observed in split-form gambles here could, in fact, be simply explained by the number of outcomes or branches in a gamble, rather than the splitting itself (Birnbaum and McIntosh 1996; Birnbaum and Chavez 1997; Birnbaum and Navarrete 1998). Please refer to M3 and M4 in Appendices 5 and 6 for further results concerning this proposition.

Appendix 5: Adding the number of outcomes or branches in M3

See Tables 7 and 8.

Humphrey (1999, 2001) suggests that the perception of a split-form gamble could be dependent not only on learning the combined probabilities of the respective split events but also on learning the absolute number of the split events. Indeed, already Estes (1976) argued in favor of a psychological phenomenon termed the *categorical memory for absolute frequency of outcome*, which describes how the encoding of event categories is done in terms of cardinal numbers and not probabilities (see also Einhorn and Hogarth (1978) for related empirical evidence).[14] While Humphrey (1999) find mixed results in favor of this explanation in his study, Birnbaum and Navarrete (1998) find some evidence supporting the importance of the absolute number of branches in a gamble.

---

[12] This shape of probability weighting is less conventional, but not unprecedented in the literature; see, e.g., Jullien and Salanie (2000) and Hertwig (2012).

[13] Note that one-parameter weighting functions have additional limitations if an S-shaped (in contrast to inverse-S shaped) curve is predicted. For example, TKW has a crossover point at $p = 1/e = 0.37$ or $\gamma < 1$ but for $\gamma > 1$ it is at $p = 1 - 0.37 = 0.63$

[14] Furthermore, this reasoning is also in line with the established phenomenon of minimal cognitive effort preference that describes how decision makers tend to remain "economical in their expenditures of cognitive effort" (Humphrey 1999, p. 55). As frequency information is more easily learned and remembered than probability information, this explanation appears intuitively appealing as a plausible cognitive antecedent of splitting effects.

Table 4 The coalesced-form and respective split-form gambles used in the parameter estimation, and descriptive statistics

No.	Safe gambles				Risky gambles				Risky (%)
G1	0.89	0.11			0.90	0.10			96.3
	0	16			0	32			
G3 split	0.89	**0.01**	**0.10**		**0.89**	**0.01**	0.10		87.5
	0	**16**	**16**		**0**	**0**	32		
G5	0.80	0.20			0.90	0.10			55.5
	0	19			0	44			
G7 split	0.80	**0.10**	**0.10**		**0.80**	**0.10**	0.10		29.6
	0	**19**	**19**		**0**	**0**	44		
G9	0.70	0.30			0.80	0.10	0.10		37.0
	0	21			0	21	42		
G11 split	0.70	**0.10**	**0.10**	**0.10**	**0.70**	**0.10**	0.10	0.10	23.6
	0	**21**	**21**	**21**	**0**	**0**	21	42	
G15	0.98	0.02			0.99	0.01			70.8
	0	23			0	46			
G17 split	0.98	**0.01**	**0.01**		**0.98**	**0.01**	0.01		56.5
	0	**23**	**23**		**0**	**0**	46		
G19	0.40	0.60			0.58	0.42			29.2
	0	28			0	44			
G21 split	0.40	**0.18**	**0.42**		**0.40**	**0.18**	0.42		21.3
	0	**28**	**28**		**0**	**0**	45		
G20	0.80	0.20			0.86	0.14			50.0
	0	28			0	44			
G22 split	0.80	**0.06**	**0.14**		**0.80**	**0.06**	0.14		27.3
	0	**28**	**28**		**0**	**0**	45		
G34	0.75	0.23	0.02		0.99	0.01			76.9
	33	34	36		33	60			
G36 split	0.75	0.23	**0.01**	**0.01**	**0.75**	**0.23**	**0.01**	0.01	87.5
	33	34	**36**	**36**	**33**	**33**	**33**	60	
G41	0.60	0.40			0.80	0.20			58.8
	3	10			3	21			
G42 split	0.60	**0.20**	**0.20**		**0.60**	**0.20**	0.20		34.7
	3	**10**	**10**		**3**	**3**	21		
12 gamble pairs with varying splitting structure out of the total of 28:									
G10	0.70	0.20	0.10		0.80	0.20			10.6
	0	21	42		0	42			
G12 split	0.70	**0.10**	**0.10**	0.10	**0.70**	**0.10**	**0.10**	**0.10**	20.8
	0	**21**	**21**	42	**0**	**0**	**42**	**42**	
G13	0.40	0.20	0.40		0.50	0.50			29.6
	0	19	44		0	44			

16 directly comparable one-type gamble pairs out of the total of 28: (above G1)

Table 4 continued

No.	Safe gambles				Risky gambles				Risky (%)
G14 split	*0.40*	***0.10***	***0.10***	*0.40*	***0.40***	***0.10***	***0.10***	***0.40***	50.0
	0	**19**	**19**	44	**0**	**0**	**44**	**44**	
G30	*0.73*	*0.02*	*0.25*		*0.74*	*0.26*			43.5
	0	15	33		0	33			
G32 split	*0.73*	***0.01***	***0.01***	*0.25*	***0.73***	***0.01***	***0.01***	***0.25***	72.2
	0	**15**	**15**	33	**0**	**0**	**33**	**33**	
G29	*0.73*	*0.02*	*0.25*		*0.74*	*0.01*	*0.25*		82.4
	0	15	60		0	33	60		
G31 split	*0.73*	***0.01***	***0.01***	*0.25*	***0.73***	***0.01***	*0.01*	*0.25*	72.2
	0	**15**	**15**	60	**0**	**0**	33	60	
G33	*0.75*	*0.23*	*0.02*		*0.75*	*0.24*	*0.01*		91.2
	1	34	36		1	33	60		
G35 split	*0.75*	*0.23*	***0.01***	***0.01***	*0.75*	***0.23***	***0.01***	*0.01*	90.7
	1	34	**36**	**36**	1	**33**	**33**	60	
G38	*0.40*	*0.60*			*0.20*	*0.80*			62.5
	9	21			3	21			
G39 split	***0.20***	***0.20***	*0.60*		*0.20*	***0.20***	***0.60***		80.6
	9	**9**	21		3	**21**	**21**		

The top numbers indicate the probabilities, the bottom numbers indicate the outcome values for each of the branches. The probabilities of all gambles are presented in italic font. The split branches of the coalesced outcomes are denoted in bold font. The right-side column describes how many choices (in %) the subjects made in favor of the risky gamble for each gamble pair

Birnbaum and Navarrete (1998) demonstrate that, for a more general CPT model, median parameters do indeed differ between two-outcome and three-outcome gambles. The authors argue that the estimated weighting function parameters for two-outcome gambles are consistent with previous numerously replicated results (in that $0 < \gamma < 1$, see, e.g., Tversky and Kahneman 1992; Wu and Gonzalez 1996), whereas for three-outcome gambles the results indicate $\gamma > 1$, consistent also with Birnbaum and McIntosh (1996) and Birnbaum and Chavez (1997). In line with this, the number of branches or outcomes in a gamble should have a significant impact on the CPT weighting function parameter estimates, even when the splitting is controlled for.

We consider these additions to the literature in M3 and find that both weighting function parameter estimates γ and δ are significantly explained by the number of branches as well as the number of outcomes in a gamble in all four weighting function specifications. Furthermore, Chow tests for comparisons between M2 and M3 indicate that adding the *outcomes* variable has also resulted in altering of γ_{coa}, γ_{split}, δ_{coa} and δ_{split} values significantly (at a 99% significance level). Although similar to that in M1 and M2, the familiar significant differences between the weighting function parameters for coalesced- and split-form gambles are still

Table 5 Results of the M1 in four specifications

Model:	M1			
Value f.:	Power			
Weighting f.:	TKW	GEW	P1W	P2W
α_{coa}	0.674	0.717	0.654	0.709
	(0.025)	(0.045)	(0.022)	(0.054)
α_{split}	0.695	0.785**	0.651	0.759**
	(0.042)	(0.052)	(0.033)	(0.052)
γ_{coa}	0.889	0.863	0.840	0.808
	(0.020)	(0.014)	(0.020)	(0.015)
γ_{split}	1.065***	1.096***	0.998**	1.034***
	(0.032)	(0.037)	(0.029)	(0.064)
δ_{coa}		5.909		6.242
		(0.563)		(0.671)
δ_{split}		3.032***		3.232***
		(0.456)		(0.410)
ϕ	4.275	0.776	5.063	1.132
	(0.528)	(0.069)	(0.483)	(0.068)
AIC	7603.14	7538.40	7537.46	7482.88
N_{obs}	6048	6048	6048	6048

The symbol ** for the corresponding split parameters indicates a difference between the coalesced-form and split-form parameter values at a 95% significance level, ***indicates a difference at a 99% significance level. All differences in this model are due to a binary *splitcoal* variable. All constant terms are significant at a 99% level. Standard errors (in parentheses) are clustered at subject level

pinpointed (at a significance level of 99%), the γ_{split} estimates are significantly lower than unity in all four specifications in M3.

Appendix 6: Adding demographic variables in M4

See Tables .

Appendix 7: Overview of the estimated model versions

See Table .

Table 6 Results of the M2 in four specifications

Model:	M2			
Value f.:	Power			
Weighting f.:	TKW	GEW	P1W	P2W
α	0.680	0.745	0.653	0.729
	(0.029)	(0.044)	(0.025)	(0.051)
γ_{coa}	0.894	0.874	0.839	0.814
	(0.021)	(0.014)	(0.022)	(0.016)
γ_{split}	1.055***	1.053***	0.999***	0.998***
	(0.026)	(0.029)	(0.026)	(0.055)
δ_{coa}		5.693		6.070
		(0.584)		(0.654)
δ_{split}		3.450***		3.546***
		(0.397)		(0.380)
ϕ	4.335	0.742	5.056	1.147
	(0.498)	(0.063)	(0.469)	(0.067)
AIC	7602.24	7546.04	7535.49	7487.58
N_{obs}	6048	6048	6048	6048

The symbol ** for the corresponding split parameters indicates a difference between the coalesced-form and split-form parameter values at a 95% significance level, *** indicates a difference at a 99% significance level. All differences in this model are due to a binary *splitcoal* variable. All constant terms are significant at a 99% level. Standard errors (in parentheses) are clustered at subject level

Table 7 Results of the M3 with *outcomes* in four specifications

Model:	M3			
Value f.:	Power			
Weighting f.:	TKW	GEW	P1W	P2W
α	0.692	0.875	0.682	0.860
	(0.030)	(0.018)	(0.025)	(0.019)
γ_{coa}	0.801	0.864	0.795	0.753
	(0.024)	(0.014)	(0.024)	(0.017)
Splitcoal	**0.0978**	**0.035**	**0.109**	**0.044**
Outcomes	**− 0.202**	**− 0.120**	**− 0.177**	**− 0.098**
Constant	**1.772**	**1.440**	**1.645**	**1.226**
γ_{split}	0.851***	0.871***	0.862***	0.773***
	(0.024)	(0.015)	(0.023)	(0.018)
δ_{coa}		10.513		7.944
		(1.429)		(0.878)
Splitcoal		− 0.087		*0.065*
Outcomes		**9.065**		**5.163**
Constant		**− 33.112**		**− 16.902**
δ_{split}		12.579***		9.236***
		(1.674)		(0.999)
ϕ	5.277	0.515	5.398	1.389
	(0.549)	(0.047)	(0.476)	(0.042)
AIC	7371.06	7218.91	7267.36	7138.98
N_{obs}	6048	6048	6048	6048

The symbol ** for the corresponding split parameters indicates a difference between the coalesced-form and split-form parameter values at a 95% significance level, *** indicates a difference at a 99% significance level. Meanwhile, all significant variables (at a 99% level) are presented in bold font. Alongside with the non-linear predictions of the weighting function parameters, the respective regression coefficients are presented in italic font. The binary variable *splitcoal* takes the value of 1 if the corresponding gamble is presented in a split form, 0 otherwise; the *outcomes* variable indicates the total number of outcomes in a gamble pair (taking a value of 4–6, as in the coalesced form); *constant* refers to the constant term. Standard errors (in parentheses) are clustered at subject level

Table 8 Results of the M3 with *branches* in four specifications

Model:	M3			
Value f.:	Power			
Weighting f.:	TKW	GEW	P1W	P2W
α	0.684	0.896	0.676	0.886
	(0.029)	(0.008)	(0.025)	(0.009)
γ_{coa}	0.822	0.864	0.806	0.743
	(0.022)	(0.010)	(0.022)	(0.014)
Splitcoal	*0.399*	*0.286*	*0.399*	*0.243*
Branches	*− 0.121*	*− 0.094*	*− 0.120*	*− 0.080*
Constant	*1.405*	*1.318*	*1.385*	*1.126*
γ_{split}	0.993***	0.973***	0.978***	0.836***
	(0.021)	(0.012)	(0.021)	(0.015)
δ_{coa}		11.476		8.821
		(1.707)		(0.959)
Splitcoal		*− 20.575*		*− 12.743*
Branches		*9.924*		*6.240*
Constant		*− 36.283*		*− 21.210*
δ_{split}		9.632***		7.856***
		(1.530)		(0.888)
ϕ	5.544	0.469	5.718	1.446
	(0.557)	(0.042)	(0.509)	(0.043)
AIC	7445.27	7265.35	7319.97	7121.15
N_{obs}	6048	6048	6048	6048

The symbol ** for the corresponding split parameters indicates a difference between the coalesced-form and split-form parameter values at a 95% significance level, *** indicates a difference at a 99% significance level. Meanwhile, all significant variables (at a 99% level) are presented in bold font. Alongside with the non-linear predictions of the weighting function parameters, the respective regression coefficients are presented in italic font. The binary variable *splitcoal* takes the value of 1 if the corresponding gamble is presented in a split form, 0 otherwise; the *branches* variable indicates the total number of branches in a gamble pair (taking a value of 4–8); *constant* refers to the constant term. Standard errors (in parentheses) are clustered at subject level

Table 9 Results of the M4 with *outcomes* in four specifications

Model:	M4			
Value f.:	Power			
Weighting f.:	TKW	GEW	P1W	P2W
α	0.691	0.874	0.682	0.858
	(0.030)	(0.018)	(0.025)	(0.021)
γ_{coa}	0.802	0.864	0.795	0.752
	(0.038)	(0.020)	(0.038)	(0.022)
Splitcoal	*0.098*	*0.035*	*0.109*	*0.044*
Outcomes	*− 0.199*	*− 0.120*	*− 0.175*	*− 0.101*
Repetitions	*0.005*	*0.000*	*0.004*	*0.000*
Gender	*− 0.044*	*− 0.011*	*− 0.050*	*− 0.012*
Major	*0.001*	*0.010*	*0.004*	*0.012*
Constant	*1.773*	*1.444*	*1.655*	*1.237*
γ_{split}	0.853***	0.871***	0.863***	0.773***
	(0.037)	(0.021)	(0.038)	(0.023)
δ_{coa}		10.468		7.861
		(1.530)		(1.038)
Splitcoal		*− 0.087*		*0.076*
Outcomes		*8.995*		*5.026*
Repetitions		*0.152*		*0.181*
Gender		*− 0.028*		*− 0.395*
Major		*− 0.011*		*− 0.163*
Constant		*− 33.175*		*− 16.437*
δ_{split}		12.518***		9.131***
		(1.751)		(1.129)
ϕ	5.327	0.516	5.434	1.386
	(0.549)	(0.050)	(0.478)	(0.054)
AIC	7361.75	7224.45	7255.40	7141.42
N_{obs}	6048	6048	6048	6048

The symbol ** for the corresponding split parameters indicates a difference between the coalesced-form and split-form parameter values at a 95% significance level, *** indicates a difference at a 99% significance level. Meanwhile, all significant variables (at a 99% level) are presented in bold font. Alongside with the non-linear predictions of the weighting function parameters, the respective regression coefficients are presented in italic font. The binary variable *splitcoal* takes the value of 1 if the corresponding gamble is presented in a split form, 0 otherwise; the *outcomes* variable indicates the total number of outcomes in a gamble pair (taking a value of 4–6, as in the coalesced form); *constant* refers to the constant term. Furthermore, *repetition* takes on a value of 1–4 depending on the filling order; *gender* takes on a value of 1 for male subjects, 0 for female subjects; and *major* takes on a value of 1 for subjects studying economics or business administration, 0 otherwise. Standard errors (in parentheses) are clustered at subject level

Table 10 Results of the M4 with *branches* in four specifications

Model:	M4			
Value f.:	Power			
Weighting f.:	TKW	GEW	P1W	P2W
α	0.683	0.896	0.675	0.886
	(0.029)	(0.008)	(0.025)	(0.009)
γ_{coa}	0.823	0.864	0.807	0.742
	(0.036)	(0.014)	(0.036)	(0.017)
Splitcoal	***0.395***	***0.287***	***0.396***	***0.246***
Branches	***− 0.119***	***− 0.095***	***− 0.119***	***− 0.081***
Repetitions	*0.006*	*− 0.000*	*0.004*	*− 0.000*
Gender	*− 0.045*	*− 0.006*	*− 0.049*	*− 0.005*
Major	*− 0.002*	*0.006*	*0.002*	*0.007*
Constant	***1.412***	***1.320***	***1.396***	***1.130***
γ_{split}	0.993***	0.973***	0.979***	0.836***
	(0.035)	(0.015)	(0.036)	(0.018)
δ_{coa}		11.467		8.813
		(1.810)		(1.104)
Splitcoal		***− 20.534***		***− 12.703***
Branches		***9.896***		***6.212***
Repetitions		*0.088*		*0.140*
Gender		*− 0.059*		*− 0.373*
Major		*− 0.121*		*− 0.209*
Constant		***− 36.269***		***− 21.085***
δ_{split}		9.612***		7.834***
		(1.684)		(1.073)
ϕ	5.571	0.468	5.746	1.449
	(0.556)	(0.044)	(0.509)	(0.046)
AIC	7434.49	7273.31	7307.40	7126.75
N_{obs}	6048	6048	6048	6048

The symbol ** for the corresponding split parameters indicates a difference between the coalesced-form and split-form parameter values at a 95% significance level, *** indicates a difference at a 99% significance level. Meanwhile, all significant variables (at a 99% level) are presented in bold font. Alongside with the non-linear predictions of the weighting function parameters, the respective regression coefficients are presented in italic font. The binary variable *splitcoal* takes the value of 1 if the corresponding gamble is presented in a split form, 0 otherwise; the branches variable indicates the total number of branches in a gamble pair (taking a value of 4 to 8); *constant* refers to the constant term. Furthermore, *repetition* takes on a value of 1–4 depending on the filling order; *gender* takes on a value of 1 for male subjects, 0 for female subjects; and *major* takes on a value of 1 for subjects studying economics or business administration, 0 otherwise. Standard errors (in parentheses) are clustered at subject level

Table 11 Overview of the four models examined in this paper

Value f.:	Power				
Weight. f.:	TKW	GEW	P1W	P2W	EUW
M1					α: *splitcoal*
	α: *splitcoal*	α: *splitcoal*	α: *splitcoal*	α: *splitcoal*	γ: [fixed]
	γ: *splitcoal*	γ, δ: *splitcoal*	γ: *splitcoal*	γ, δ: *splitcoal*	
M2					
	γ: *splitcoal*	γ, δ: *splitcoal*	γ: *splitcoal*	γ, δ: *splitcoal*	
M3					
	γ: *splitcoal, outcomes* or *branches*	γ, δ: *splitcoal, outcomes* or *branches*	γ: *splitcoal, outcomes* or *branches*	γ, δ: *splitcoal, outcomes* or *branches*	
M4					
	γ: *splitcoal, outcomes* or *branches, repetitions, gender, major*	γ, δ: *splitcoal, outcomes* or *branches, repetitions, gender, major*	γ: *splitcoal, outcomes* or *branches, repetitions, gender, major*	γ, δ: *splitcoal, outcomes* or *branches, repetitions, gender, major*	

The ϕ parameter is estimated in each of the models; the detailed interpretation of it is, however, beyond the scope of this paper. The values of the constant are included in each of the cases by default

References

Abdellaoui, M. (2000). Parameter free elicitation of utilities and probability weighting functions. *Management Science, 46,* 1497–1512.

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. Petrox & F. Caski (Eds.), *Second international symposium on information theory* (pp. 267–281). Budapest: Akademiai Kiado.

Allais, M. (1953). L'extension des théories de l'équilibre économique général et du rendement social au cas du risque. *Econometrica, Journal of the Econometric Society*, 269–290.

Birnbaum, M. (1999). Testing critical properties of decision making on the Internet. *Psychological Science, 10,* 399–407.

Birnbaum, M. (2004). Tests of rank-dependent utility and cumulative prospect theory in gambles represented by natural frequencies: Effects of format, event framing, and branch splitting. *Organizational Behavior and Human Decision Processes, 95,* 40–65.

Birnbaum, M. (2006). Evidence against prospect theories in gambles with positive, negative, and mixed consequences. *Journal of Economic Psychology, 27,* 737–761.

Birnbaum, M. (2007). Tests of branch splitting and branch-splitting independence in Allais paradoxes with positive and mixed consequences. *Organizational Behavior and Human Decision Processes, 102,* 154–173.

Birnbaum, M. (2008). New paradoxes of risky decision making. *Psychological Review, 115,* 463–501.

Birnbaum, M., & Chavez, A. (1997). Tests of theories of decision making: Violations of branch independence and distribution independence. *Organizational Behavior and Human Decision Processes, 71,* 61–194.

Birnbaum, M., Johnson, K., & Longbottom, J. (2008). Tests of cumulative prospect theory with graphical displays of probability. *Judgement and Decision Making, 3,* 528–546.

Birnbaum, M., & McIntosh, W. (1996). Violations of branch independence in choices between gambles. *Organizational Behavior and Human Decision Processes, 67,* 91–110.

Birnbaum, M., & Navarrete, J. (1998). Testing descriptive utility theories: Violations of stochastic dominance and cumulative independence. *Journal of Risk and Uncertainty, 17,* 17–49.

Birnbaum, M., & Schmidt, U. (2015). The impact of learning by thought on violations of independence and coalescing. *Decision Analysis, 12,* 144–152.

Birnbaum, M., Schmidt, U., & Schneider, M. (2017). Testing independence conditions in the presence of errors and splitting effects. *Journal of Risk and Uncertainty, 54,* 61–85.

Bleichrodt, H., Pinto, J., & Wakker, P. (2001). Using descriptive findings of prospect theory to improve the prescriptive use of expected utility. *Management Science, 47,* 1498–1514.

Booij, A., van Praag, B., & van de Kuilen, G. (2010). A parametric analysis of prospect theory's functionals for the general population. *Theory and Decision, 68,* 115–148.

Burnham, K., & Anderson, D. (2002). *Model selection and multimodel inference: A practical information-theoretic approach.* New York: Springer.

Conlisk, J. (1989). Three variants on the Allais example. *American Economic Review, 79,* 392–407.

Einhorn, H., & Hogarth, R. (1978). Confidence in judgment: Persistence in the illusion of validity. *Psychological Review, 85,* 396–416.

Estes, W. (1976). The cognitive side of probability learning. *Psychological Review, 83,* 37–64.

Fischoff, B. (1991). Value elicitation: Is there anything in there? *American Psychologist, 46,* 835–847.

Glöckner, A., & Pachur, T. (2012). Cognitive models of risky choice: Parameter stability and predictive accuracy of prospect theory. *Cognition, 123,* 21–32.

Goldstein, W., & Einhorn, H. (1987). Expression theory and the preference reversal phenomena. *Psychological Review, 94,* 236–254.

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

Hertwig, R. (2012). The psychology and rationality of decisions from experience. *Synthese, 187,* 269–292.

Hey, J., & Orme, C. (1994). Investigating generalizations of expected utility theory using experimental data. *Econometrica, 62,* 1291–1326.

Humphrey, S. (1995). Regret aversion or event-splitting effects? More evidence under risk and uncertainty. *Journal of Risk and Uncertainty, 11,* 263–274.

Humphrey, S. (1996). Do anchoring effects underlie event-splitting effects? An experimental test. *Economics Letters, 51,* 303–308.

Humphrey, S. (1999). Probability learning, event-splitting effects and the economic theory of choice. *Theory and Decision, 46,* 51–78.

Humphrey, S. (2001). Non-transitive choice: Event-splitting effects or framing effects? *Economica, 68,* 77–96.

Humphrey, S. (2006). Does learning diminish violations of independence, coalescing and monotonicity? *Theory and Decision, 61,* 93–128.

Jullien, B., & Salanie, B. (2000). Estimating preferences under risk: The case of racetrack bettors. *Journal of Political Economy, 108,* 503–530.

Kahneman, D. (2003). Experiences of collaborative research. *American Psychologist*, 58(9), 723–730.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47,* 263–292.

Kolmogoroff, A. (1933). *Grundbegriffe der Wahrscheinlichkeitsrechnung.* Berlin: Springer.

Luce, R. (1959). *Individual choice behavior.* New York: Wiley.

Luce, R. (1998). Coalescing, event commutativity, and theories of utility. *Journal of Risk and Uncertainty, 16,* 87–113.

McCord, M., & de Neufville, R. (1986). Lottery equivalents: Reduction of the certainty effect problem in utility assessment. *Management Science, 32,* 56–60.

Nillson, H., Rieskamp, J., & Wage, E. (2011). Hierarchical Bayesian parameter estimation for cumulative prospect theory. *Journal of Mathematical Psychology, 55,* 84–93.

Payne, J. (2005). It is whether you win or lose: The importance of the overall probabilities of winning or losing in risky choice. *Journal of Risk and Uncertainty, 30,* 5–19.

Prelec, D. (1998). The probability weighting function. *Econometrica, 66,* 497–528.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior & Organization, 3,* 323–343.

Rieskamp, J. (2008). The probabilistic nature of preferential choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 1446–1465.

Schmidt, U., & Seidl, C. (2014). Reconsidering the common ratio effect: The roles of compound independence, reduction, and coalescing. *Theory and Decision, 77,* 323–339.

Sneddon, R., & Luce, R. (2001). Empirical comparisons of bilinear and nonbilinear utility theories. *Organizational Behavior and Human Decision Processes, 84,* 71–94.

Starmer, C., & Sugden, R. (1993). Testing for juxtaposition and event-splitting effects. *Journal of Risk and Uncertainty, 6,* 235–254.

Stott, H. (2006). Cumulative prospect theory's functional menagerie. *Journal of Risk and Uncertainty, 32,* 101–130.

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

Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *The Quarterly Journal of Economics, 106,* 1039–1061.

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

von Neumann, J., & Morgenstern, O. (1947). *The theory of games and economic behavior*. Princeton: Princeton University Press.

von Winterfeldt, D., & Edwards, W. (1986). *Decision analysis and behavioral research*. New York: Cambridge University Press.

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

Wu, G., & Gonzalez, R. (1996). Curvature of the probability weighting function. *Management Science, 42,* 1676–1690.

Affiliations

Andreas Glöckner[1,2] · Baiba Renerte[3,4] · Ulrich Schmidt[5,6,7]

✉ Baiba Renerte
  baiba.renerte@uni-konstanz.de

[1] Department of Social Psychology, University of Cologne, Richard-Strauss-Str. 2, 50931 Cologne, Germany

[2] Max Planck Institute for Research on Collective Goods, Kurt-Schumacher-Str. 10, 53115 Bonn, Germany

[3] Department of Economics, University of Konstanz, Universitätsstr. 10, 78464 Constance, Germany

[4] Thurgau Institute of Economics, Hauptstr. 90, 8280 Kreuzlingen, Switzerland

[5] Kiel Institute for the World Economy, Kiellinie 66, 24105 Kiel, Germany

[6] Department of Economics, University of Kiel, Olshausenstr. 40, 24098 Kiel, Germany

[7] Department of Economics and Econometrics, University of Johannesburg, Auckland Park, Johannesburg 2092, South Africa