Narrow Bracketing and Dominated Choices

Matthew Rabin

and

Georg Weizsäcker[1]

Abstract

We consider a decisionmaker who "narrowly brackets", i.e. evaluates her decisions separately. Generalizing an example by Tversky and Kahneman (1981) we show that if the decisionmaker does not have constant-absolute-risk-averse preferences, there exists a simple pair of independent binary decisions where she will make a first-order stochastically dominated combination of choices. We also characterize, as a function of preferences, a lower bound on the monetary cost that can be incurred due to a single mistake of this kind. Empirically, we conduct a real-stakes laboratory

[1]Rabin: University of California — Berkeley, 549 Evans Hall, Berkeley, CA 94720-3880, USA, rabin@econ.berkeley.edu. Weizsäcker: London School of Economics, Houghton Street, London, WC2A 2AE, U.K., g.weizsacker@lse.ac.uk. We are grateful to Dan Benjamin, Syngjoo Choi, Erik Eyster, Thorsten Hens, Michele Piccione, Peter Wakker, Heinrich Weizsäcker, seminar participants at Amsterdam, Caltech, Cambridge, Helsinki, IIES Stockholm, IZA Bonn, LSE, Zurich, Harvard, Nottingham, NYU, Oxford, Pompeu Fabra, and the LEaF 2006, FUR 2008 and ESSET 2008 conferences, and especially to Vince Crawford and three anonymous referees for helpful comments, and to Zack Grossman and Paige Marta Skiba for research assistance. The survey experiment was made possible by the generous support of TESS (Time-Sharing Experiments in the Social Sciences) and the efforts by the staff of Knowledge Networks. We also thank the ELSE Centre at University College London for the generous support of the laboratory experiment, and Rabin thanks the National Science Foundation (Grants SES-0518758 and SES-0648659) for financial support.

experiment replicating Tversky and Kahneman's original experiment, finding that 28% of the participants violate dominance. In addition, we conduct a representative survey among the general U.S. population that asks for hypothetical large-stakes choices. There we find higher proportions of dominated choice combinations. A statistical model suggests that the average preferences are close to prospect-theory preferences and that about 89% of people bracket narrowly. Results do not vary much with the personal characteristics of participants.

Keywords: Lottery choice, narrow framing, representative-sample experiments

JEL Classification: B49

Contact: rabin@econ.berkeley.edu, g.weizsacker@lse.ac.uk

A mass of evidence, and the ineluctable logic of choice in a complicated world, suggests that people "narrowly bracket": a decisionmaker who faces multiple decisions tends to choose an option in each case without full regard to the other decisions and circumstances that she faces. In the context of monetary risk, Tversky and Kahneman (1981) present an experiment that demonstrates both how powerful this propensity is, and its clear welfare cost. In their experiment, people narrowly bracket even when faced with only a pair of independent simple binary decisions that are presented on the same sheet of paper, and as a result make a combination of choices that is inconsistent with any reasonable preferences. In our slight reformulation, we present subjects with the following:

> You face the following pair of concurrent decisions. First examine both decisions, then indicate your choices, by circling the corresponding letter. Both choices will be payoff relevant, i.e. the gains and losses will be added to your overall payment.
>
> Decision (i): Choose between
>
> A. a sure gain of £2.40
>
> B. a 25% chance to gain £10.00 and a 75% chance to gain £0.00.
>
> Decision (ii): Choose between
>
> C. a sure loss of £7.50
>
> D. a 75% chance to lose £10.00, and a 25% chance to lose £0.00.

If, as predicted by Kahneman and Tversky's (1979) prospect theory, the decision-maker is risk-averting in gains and risk-seeking in losses *and if* she applies these preferences separately to the decisions, then she will tend to choose A and D. This prediction was confirmed: 60% of Tversky and Kahneman's (1981) participants chose A and D with small real stakes, and 73% did so for large hypothetical stakes. But A and D is first-order stochastically dominated: the joint distribution resulting from the combination of B and C is a $\frac{1}{4}$ chance of gaining £2.50 and a $\frac{3}{4}$ chance of losing £7.50; the joint distribution of A and D is a $\frac{1}{4}$ chance of gaining £2.40 and a $\frac{3}{4}$ chance of losing £7.60. The BC combination is equal to the AD combination plus a sure payoff of £0.10.

In this paper, we explore both the empirical and theoretical generality of this experiment. Our experiments — laboratory experiments both with real and hypothetical payments, and a hypothetical-payment survey with a representative sample from the general U.S. population — confirm the pattern of frequent AD choices, although at a somewhat lower level. With large hypothetical payoffs (£1 being replaced by £100 in the laboratory, and by \$100 in the survey) about 60% of participants choose AD; for small real and hypothetical stakes, 28% and 34% of the subjects do so. We also introduce three other hypothetical large-stakes sets of decisions where between 40% and 50% of subjects make dominated choices, despite giving up amounts of \$50 and \$75 rather than the \$10 in the large-payoff version of the original example.

These dominance violations demonstrate that subjects are narrowly bracketing, because the choice of AD is clearly due to the separate presentation.[2] Theoretically,

[2] A related example of narrow bracketing is Redelmeier and Tversky's (1992) demonstration that the investment choice in a risky asset can depend on whether the asset is framed as part of a portfolio of other assets or as a stand-alone investment. See also the replication and variation in Langer and Weber (2001), and the literature cited there. Other evidence on narrow bracketing in lottery choice include Gneezy and Potters (1997) and Thaler et al (1997) who test whether mypoic loss aversion — a form of narrow bracketing — may serve as a possible explanation of the equity-premium puzzle, and by Camerer (1989) and Battalio et al (1990), both of whom present treatment

we are not aware of any broad-bracketed utility theory ever proposed that would allow for the choice of AD over BC.[3] Empirically, we find in a "broad presentation" treatment, offering an explicit choice between the combinations AC, AD, BC and BD, that violation rates are reduced to 0% and 6%, respectively, in the laboratory and the survey.[4]

The theoretical part of the paper contributes to understanding the generality with which narrow bracketing can lead to dominated choices. For a given preference relation $\succeq$ and a given set of choice sets, we define narrow bracketing as the application of $\succeq$ separately to each choice set. In contrast, a broad bracketer applies $\succeq$ to the set that comprises all possible choice combinations. This definition is widely applicable and, in particular, a natural connection can be made to arbitrage possibilities. Building on Diecidue and Wakker (2002), one can use the definition of narrow bracketing to re-phrase de Finetti's (1974) Dutch-book theorem: for a narrow bracketer who is

variations that suggest narrow bracketing of standard lottery choices. Papers that have explored the principles of what we call narrow bracketing include Kahneman and Lovallo (1993), Benartzi and Thaler (1995) and Read, Loewenstein, and Rabin (1999). Other research has found evidence of violation of dominance due to errors besides narrow bracketing; see, e.g., Birnbaum *et al* (1992) and Mellers, Weiss, and Birnbaum (1992).

[3] Even models that allow for dominance violations — such as the disappointment-theory models of Bell (1985) and Loomes and Sugden (1986), the related "choice-acclimating personal equilibrium" concept in Kőszegi and Rabin (2007), and the gambling preferences in Diecidue, Schmidt, and Wakker (2004) — do not permit the preference for AD over BC, since BC is simply AD plus a sure amount of money.

[4] We examine the broad presentation in three examples in the survey experiment, and the violation rates are reduced from 66% to 6%, from 40% to 3%, and from 50% to 29%. The surprisingly high violations of dominance in the last example even under broad presentation are inconsistent with any hypothesis about choice behavior that we are aware of and we do not understand what motives or errors were induced by the design of this example.

not risk neutral, there exists a series of choices between correlated gambles such that her combined choice will be dominated by another feasible combination of choice.[5] Being a narrow bracketer, she ignores all correlations between gambles that appear in different choices and thus she can be tricked into a suboptimal combination. In light of this, Tversky and Kahneman's example is is striking in that it works even if the gambles are *uncorrelated*. Hence, even if a decisionmaker merely ignores other uncorrelated choices, prospect-theory preferences may induce her to choose a first-order stochastically dominated portfolio.[6] This extension is not trivial: most prospects in life are relatively independent of each other and it is a considerably weaker assumption that people ignore only independent background choices. Moreover, the example induces dominance in only two binary decisions.

Yet we show below that Tversky and Kahneman's example can itself be generalized considerably. The main result in Section 1 establishes that the logic of their example extends broadly beyond prospect-theoretic preferences: if a narrow bracketer's risk attitudes are not identical at all possible ranges of outcomes — essentially, if she does not have constant-absolute-risk-aversion (CARA) preferences — then there exists a pair of independent binary lottery problems where she chooses a dominated combination.

The logic behind this simple result is itself simple. In the Tversky and Kahneman example, narrow bracketing means that a prospect-theoretic chooser takes a less-than-expected-value certain amount over the lottery in the gain domain due to risk aversion, but takes the lottery over its expected value in the loss domain due to risk-lovingness. Since her payoff is the sum of the two gambles, she'd be better off doing

[5] The result is the single-person analogue of the fundamental theorem of asset pricing, which equates the market's freedom from arbitrage with as-if risk neutrality of prices.

[6] Our instructions made clear that all draws are independent.

the opposite. But the potential for dominance does not depend on where or how her risk attitudes differ: if a person's absolute risk aversion is not identical over all possible ranges, then there exists a non-risky alternative payment that a person would prefer to a given lottery in one range that involves sacrificing more money than a non-risky payment that she would reject in favor of the same lottery in another range. If offered both decisions between such non-risky payments and the corresponding lotteries, a narrow bracketer will therefore choose the non-risky alternative only in the case where it involves more sacrifice. But then reversing her choices leaves her with the same risk but a higher distribution of outcomes.

This result relies solely on monotonicity and completeness of preferences as well as the existence of certainty equivalents. In particular, it does not assume that the decisionmaker is an expected-utility maximizer in the sense of weighting prospects linearly in probabilities. Section 1 also establishes two stronger results for the case of linear-in-probability evaluations. First, we characterize as a function of preferences a lower bound on the maximum amount that the decisionmaker can leave on the table in only two choices, establishing that a narrow bracketer with preferences significantly different from CARA can be made to give up substantial amounts of money in such cases. Second, we show that a pair of independent binary choices can induce dominance even for a decisionmaker who has an arbitrarily small propensity to narrowly bracket.

While our results establish the existence of situations that generate dominance violations, we do not address the empirical prevalence of such situations or scale of welfare loss from dominated choices.[7] However, a narrow bracketer who faces a large

[7] Also, we take as given the perception of what is in each bracket and do not discuss the origin of narrow bracketing.

enough set of varied decisions will make a dominated choice overall if only one pair of those decisions generates a dominated choice. And people will typically lose utility from bracketing narrowly even when they do not violate dominance; we focus on dominance violations because they are suboptimal for all monotonic preferences.

The high rates of dominance violations in the experiments indicate directly that many subjects are not broad-bracketing utility maximizers. But to demonstrate that narrow-bracketing utility maximization has greater explanatory power, we use a simple statistical model to jointly estimate the subjects' utility and the extent of narrow bracketing. Agents are all assumed to maximize a common utility function, but to differ as to whether they bracket narrowly or broadly. We estimate that about 89% of decisions are made with narrow brackets, and that average preferences accord to the prospect-theory value function, with risk aversion in gains and around the status quo point, and a preference for risk in losses. While the estimation strategy comes with strong assumptions — most notably, that preferences are homogeneous across the population — Section 3 provides additional statistics that robustly indicate narrow bracketing.

The data on personal characteristics in our survey sample also allows us to ask *who* brackets narrowly. We find few strong correlations of bracketing propensity with observable background characteristics. In each of the subgroups that we examined, between 0% and 22% of people are broad bracketers, with few significant deviations from the average level of 11%. There is more variation, however, in estimated preferences. "Non-white" respondents are more risk-neutral with respect to lotteries around zero and in the gains domain, which makes them less likely to violate dominance. Although less pronounced, estimates also suggest that men are more risk neutral than women and that the math-skilled are more risk neutral than the less math-skilled respondents. Perhaps surprisingly, we find no significant effect of education on the violation rates.

Although this paper concentrates on the "positive" questions of when and how narrow bracketing leads people to dominated choice, we elaborate in Section 4 on the implications our analysis has for the normative status of various models of risky choice. We then conclude the paper with a brief discussion of whether such violations may be observed in markets where agents interact, and with some methodological implications for assessments of risk preferences. Our Appendix presents proofs, and more detail on our experimental procedures and results are on the journal's Web site.

1 Theory

Assume that a person simultaeneously faces I different choice sets $M_1, ..., M_I$, where every possible $m_i \in M_i$ induces a lottery, or probability distribution, $L_i(x_i|m_i)$ over changes in wealth $x_i \in \mathbb{R}$. A possible vector of choices $m = (m_1, ..., m_I)$ induces a probability distribution over the sum of wealth changes $x^I = \sum_i x_i$, denoted by $F(x^I|m)$. We restrict attention to the case that the lotteries L_i are independent across the "brackets" $i = 1, ..., I$, and will state conditions for which there exists a set of choice sets $\{M_1, ..., M_I\}$ such that the chosen distribution $F(x^I|m)$ is dominated.

As described in the introduction, we denote the decisionmaker's preferences by $\succeq$ and define her as a *narrow bracketer* if she applies $\succeq$ separately to each of the I choice sets, without consideration to the fact that the relevant outcome is the combined outcome from all her choices. That is, she chooses from each M_i by evaluating the lotteries $L_i(x_i|m_i)$, but not the summed distribution $F(x^I|m)$. We assume further that $\succeq$ is complete and strictly monotonic over the set of all possible lotteries, and that according to $\succeq$, certainty equivalents exist for all available lotteries. Because preferences are monotonic, the agent will never choose a dominated lottery within any single bracket — but the resulting distribution $F(x^I|m)$ may be dominated. To define the size of a first-order stochastic dominance (FOSD) violation, we say that F_1 *dominates* F_2 *by an amount* δ if it holds for all x in their support that $F_1(x+\delta) \leq F_2(x)$. This measure has a straightforward interpretation: if F_1 dominates F_2 by an amount δ, any decisionmaker with monotonic preferences will find F_1 at least as desirable as receiving F_2 plus a sure payment of δ. We say that the decisionmaker violates FOSD by an amount δ if she chooses a distribution F_2 that is FOS-dominated by amount δ by another available distribution F_1.

The propositions below establish that lower bounds for the largest possible δ are linked to the decisionmaker's variability of risk attitudes, which can be captured by describing certainty equivalents and their variability. Let CE_L be the decisionmaker's certainty equivalent for L. Denote by $\widetilde{L} \equiv L+\triangle x$ a lottery that is generated by adding $\triangle x$ to all payoffs in L, keeping the probabilities constant: $\widetilde{L}$ is a shifted version of L. Finally, define the decisionmaker's risk premium for lottery L, π_L, as the difference between the lottery's expected value, μ_L, and its certainty equivalent: $\pi_L = \mu_L - CE_L$. A larger π_L corresponds to more risk aversion towards L, and the agent is risk neutral if $\pi_L = 0$.

Proposition 1 shows that a decisionmaker will violate dominance in joint decisions to the degree that a shift can induce a change in the risk premium:

> **Proposition 1:** Suppose that the decisionmaker is a narrow bracketer and there exist a lottery L and a shifted version thereof, $\widetilde{L} = L + \triangle x$, such that $|\pi_L - \pi_{\widetilde{L}}| > \delta$. Then there exists a pair of independent binary choices such that the decisionmaker violates FOSD by the amount δ.

The proof of Proposition 1 follows the logic behind Tversky and Kahneman's example. But whereas they applied prospect theory to generate a shift in lotteries (from D to B) that moved the decisionmaker from risk seeking to risk averse, the proposition clarifies that neither a sign change in risk aversion nor the type of lottery that Tversky and Kahneman used is necessary to generate dominance. The construction works whenever the risk premium changes by any amount, and for any shift of any lottery. Therefore, dominance can result from narrow bracketing for all but a very restricted class of preferences. In particular, among all expected-utility preferences, the proposition shows that a dominance violation is possible for all utility functions v outside the constant-absolute-risk-aversion family, i.e. any preferences that cannot be represented by $v(x) = CARA(x; \beta, \alpha, r) \equiv \beta - \alpha \exp(-rx)$ for any $(\beta, \alpha, r) \in \mathbb{R}^3$. This is because under expected utility (and indeed more generally) the CARA family encompasses exactly those utility functions where the risk premium is constant for all shifts of all lotteries L. But the proof of Proposition 1 does not rely on preferences being EU-representable, and hence the violations can occur even for a large class of non-EU preferences — in particular, preferences that are representable under probability weighting formulations.

For any given δ, however, Proposition 1 is silent about the set of preferences for which there is a pair of lotteries L and $\widetilde{L}$ with the property $|\pi_L - \pi_{\widetilde{L}}| > \delta$. To investigate when the decisionmaker is in danger of making a *large* mistake, we now consider preferences that are EU-representable by a (possibly reference-dependent) strictly increasing and continuously differentiable function v whose expected value she maximizes. This allows us to characterize a lower bound for the size of possible dominance violations by comparing preferences to the CARA family using the following metric:

Definition: For an interval $[\underline{x}, \overline{x}] \subset \mathbb{R}$ of changes in wealth,

$$K(v, \underline{x}, \overline{x}) \equiv \inf_{(\beta,\alpha,r)\in\mathbb{R}^3} \max_{y\in[v(\underline{x}),v(\overline{x})]} |v^{-1}(y) - CARA^{-1}(y; \beta, \alpha, r)|$$

is the *horizontal distance* between v and the family of CARA functions.

That is, for an interval $[\underline{x}, \overline{x}]$, K is the smallest distance in horizontal direction such that all CARA functions reach at least this distance from v, somewhere on the interval. K is a monetary amount that indicates the change in risk attitudes across different ranges within $[\underline{x}, \overline{x}]$, as CARA represents a constant risk attitude and K measures the distance between v and CARA. (K's arguments $v, \underline{x}, \overline{x}$ are suppressed from here onwards.) We can now state another simple proposition (although with a long proof):

Proposition 2: Suppose that the decisionmaker is a narrow bracketer and that preferences are EU-representable by a function v that is strictly increasing and continuously differentiable and has a horizontal distance of K from the CARA family on the interval $[\underline{x}, \overline{x}]$. Then for all $\epsilon > 0$ there exists a pair of independent binary choices — each between a binary lottery and a sure payment, and using only payoffs in $[\underline{x}, \overline{x}]$ — such that the decisionmaker violates FOSD by an amount greater than $K - \epsilon$.

Proposition 2 shows that one can find an example where narrow bracketing causes the decisionmaker to leave K on the table. The proof provides a construction of two candidate binary lotteries L^A and L^B, where at least one of them can always be shifted in a way that yields a variation of the risk premium by K. Hence, Proposition 1 can be applied to generate the violation.

K is defined conditional not only on v, but also on the interval $[\underline{x}, \overline{x}]$. If the interval is expanded, K increases. Indeed, with almost all functional forms of v that are commonly used, such as a two-part linear function or a constant relative risk aversion function, K becomes infinitely large as the interval increases to infinite size. This is a strong limiting result, but we note that for larger and larger payoff sizes the assumption of narrow bracketing is arguably less and less plausible.

Our final theoretical result shows that it is not only fully narrow bracketers who make dominated choices. To formulate the sense in which "partial narrow bracketing" causes problems, we abandon our simple definition of narrow bracketing, and impose some additional structure on the preferences. A convenient formulation is the global-plus-local functional form of Barberis and Huang (2004) and Barberis, Huang and Thaler (2006): assume that the agent's choices are determined by maximizing, over possible choice vectors m, the expression

$$U(m) = \kappa \int u(x^I) dF(x^I|m) + (1-\kappa) \sum_i \int u(x_i) dL_i(x_i|m_i).$$

Here u is a valuation function for money, which the decisionmaker applies both globally to total earnings — as captured in the first term — and locally to each choice set M_i — as captured in the second term. Notice that each element m_i of m enters U in two ways, by contributing to the distribution F of total wealth changes and through the narrow evaluation of payoffs in bracket i alone. The parameter $\kappa \in [0,1)$ is the weight of the global part, so that $1-\kappa$ is the degree of narrow bracketing. When $\kappa \to 1$, choices correspond to fully broad bracketing, and when $\kappa = 0$, there is fully narrow bracketing. The proposition shows that if u is different from CARA, then an arbitrarily mild degree of narrow bracketing puts the decisionmaker in danger of FOSD violations: κ could be arbitrarily close to 1.[8]

[8] The proof in the appendix covers a somewhat stronger statement, allowing for different valuation functions in the broad versus narrow parts of the valuations. That is, the proposition holds even if the function representing the broad valuation (here, κu) is CARA. It suffices if the narrow valuation

Proposition 3: Suppose that the decisionmaker maximizes $U(\cdot)$, where u is strictly increasing, twice continuously differentiable and not a member of the CARA family of functions. Then there is a pair of independent binary choices, each between a 50/50 lottery and a sure payment, where the decisionmaker violates FOSD.

All three propositions highlight that departures from constant absolute risk aversion lead to dominance violations. In fact, the converse of all propositions is also true: a person with CARA preferences will never make dominated choices, as even under broad bracketing her choice within each bracket is independent of the background risk generated in other brackets. Among economists, there is widespread agreement that CARA is not the best-fitting class of preferences. To the extent that this agreement is based on data analyses, however, it is important to note that all estimates of risk attitudes will crucially depend on the maintained assumptions about bracketing.[9] We are not aware of a study that simultaneously describes risk attitudes and narrowness of bracketing, and will provide such an estimation in the following sections.

2 Experimental Design and Procedure

We conducted two experiments in different formats: one laboratory experiment that replicates and systematically varies the Tversky and Kahneman experiment ("Example 1", hereafter), and one survey experiment with a large and representative subject

function (here, $(1-\kappa)u$) differs from CARA.

[9] The evidence on lottery choice behavior points at a decreasing degree of absolute risk aversion, for the average decisionmaker — see e.g. Holt and Laury (2002) for laboratory evidence. As in most related studies, this stylized result implicitly assumes narrow bracketing in the sense that all income from outside the experiment is ignored in the analysis. Dohmen et al (2005), in contrast, measure risk aversion also under the assumption that people integrate other assets.

pool, where we introduce additional tasks. We describe the procedures of both experiments before describing the additional choice tasks and the data.

2.1 Procedure of the Laboratory Experiment

For the laboratory experiment, 190 individuals (mostly students) were recruited from the subject pool of the ELSE laboratory at University College London. We held 15 sessions of sizes ranging between 7 and 18 participants, in four different treatments. Each participant faced one treatment only, consisting of one particular variant of the $A/B/C/D$ choices of Example 1. The wording was as given in the introduction.

In the first treatment, "Incentives-Small Scale", which was conducted in four sessions with $N = 53$ participants in total, we used the payoffs that were given in the introduction, and these payments were made for real. In a "Flat Fee-Small Scale" treatment (three sessions, $N = 44$), participants made the same two choices A/B and C/D, but only the show-up fee was paid, as explained below. In the third treatment, "Incentives-Small Scale-Broad Presentation" (four sessions, $N = 45$), they made only one four-way decision, choosing between the distributions of the sum of earnings that would result from the four possible combinations of A and C, A and D, B and C, and B and D. That is, in this treatment we imposed a broad view by adding up the payoffs from the two decisions. For example, the combination of A and D would be presented as "a 25% chance to gain £2.40 and a 75% chance to lose £7.60."[10] Finally, in a "Flat Fee-Large Scale" treatment (three sessions, $N = 48$), the participants made the two hypothetical choices of the second treatment, but we multiplied all payoff numbers by a factor of 100. Hence, they could make hypothetical gains and

[10] In this treatment, the order of the four choice options was randomly changed between the participants. In the three treatments with two binary choices, we maintained the same order as in Tversky and Kahneman (1981).

losses of up to £1000 in this treatment.

On the first sheet of the experimental instructions, it was clarified that all random draws in the course of the experiment would be determined by independent coin flips. All choices were made by paper and pencil, with only very few oral announcements that followed a fixed protocol for all treatments, and with the same experimenter present in all sessions. After the choices on Example 1, the experiments moved on to a second part. This second part is not analyzed in the paper; the tasks and data are described in Online Appendix 1. The tasks of the second part differed between the 15 sessions, but the participants were not made aware of the contents of the second part before making their choices in the first part, so that the Example 1 choices cannot have been affected by the differences in the second part. The participants also had to fill in a questionnaire and a sheet with five mathematical problems. Finally, the relevant random draws were made and the participants were paid in cash. The entire procedure, including payments, took about 40-50 minutes in each session.

An email was sent to the participants 24 hours before the session in which they participated, and made them aware that (i) they would receive a show-up fee of £22, (ii) that they "may" make gains and losses relative to their show-up fee, and (iii) that overall, they would be "about equally likely to make gains as losses (on top of the £22)."[11] Upon arrival at the laboratory, the participants learned whether the experiment used monetary incentives or not, i.e. whether the outcome amounts were added to/subtracted from their show-up fee. This procedure aims at minimizing possible effects of earnings differences between treatments with hypothetical and real payments, by ruling out both ex-ante differences and anticipated ex-post differences in average earnings. In those sessions where we used real monetary incentives, the

[11] The email text and complete instructions of all experiments are in Online Appendix 3.

second part of the experiment was designed such that the expected average of total earnings would indeed be at £22. (On average, the subjects received £21.85 in these sessions, with a standard deviation of £7.70.) A further role of the 24-hour advance notice about the show-up fee was to make the losses more akin to real losses, as the participants may have "banked" the show-up fee. The amount £22 was not mentioned on the day of the experiment before the subjects had made their Example 1 choices, and all gains and losses were presented using the words "gain" and "lose".

Treatment	# of obs.	Sessions
Incentives-Small Scale	53	1-4
Flat Fee-Small Scale	44	5-7
Incentives-Small Scale-Broad Presentation	45	8-11
Flat Fee-Large Scale	48	12-15

Table 1: Overview of laboratory treatments.

2.2 Procedure of the Survey Experiment

The survey experiment used the survey tool of TESS (Time-Sharing Experiments in the Social Sciences), which regularly conducts questionnaire surveys with a stratified sample of American households, those on the Knowledge Networks panel. The panel members were recruited based on their telephone directory entries and are used to answering questions via special TV-connected terminals at their homes. For each new study, they are contacted by email. In the case of our questionnaire, a total of 1910 panel members were contacted, of whom 1292 fully completed the study. A further 30 respondents participated but left at least one question unanswered. (We included their responses in the analysis, wherever possible.) Each participant was presented with one or several decision tasks, plus a short questionnaire that asked for information on mathematics education and gave the participants three mathematical problems to solve. The data set also contains information on each participant's personal background characteristics such as gender, employment status, income and obtained level of education. None of the lotteries was paid out, i.e. all choices were hypothetical. The amounts used in the decision tasks ranged from $-\$1550$ to $+\$2500$.

In addition to the binary lottery choices that we report here, subjects were also asked to state certainty equivalents for 11 different lotteries. In Online Appendix 2, we describe the procedure and data, and discuss why we feel that these certainty equivalence data are unreliable, as many participants cannot plausibly have understood the procedure. We therefore do not include these data in the analysis.

Participants were randomly assigned to 10 different treatment groups, and each treatment contained a different set of one, two or six decision tasks (including lottery choices and certainty equivalent statements). Within each decision, the order in which the choice options appeared was randomized. After excluding two treatments where only certainty-equivalent statements were collected, the sample contains 2543 choices made by 1130 participants in 8 treatments. Table 2 summarizes the lottery choice (LC) tasks in each of the 8 treatments and lists the number of certainty equivalent (CE) tasks in the same treatments. Further details on the lottery choice tasks (Example 1 etc.) are given in the next subsection.

In each treatment, the participants' interfaces forced them to read through all their decisions before they could make their choices. Importantly, the instructions stated clearly on the first screen that the participants should make their choices as if all of their outcomes were paid. Hence, it is unlikely that choices were made under a misunderstanding that only subsets of the decisions were relevant. Also, as in the laboratory experiment, the instructions made clear that all random draws were independent.

Treatment	# of obs.	# of LC tasks	Description	# of CE tasks
1	88	2	Example 1	0
2	86	1	Example 1 — broad presentation	0
3	107	2	Example 2	0
4	108	1	Example 2 — broad presentation	0
5	168	3	Example 2 Example 4 — broad presentation	3
6	185	3	Example 2 — separate screens Example 3	3
7	174	2	Example 4	4
8	184	3	Example 2 — broad presentation Example 4 — separate screens	3

Table 2: Overview of survey experiment treatments

2.3 The Lottery Choice Problems

The lottery choice tasks of the survey experiment are similar to those in Tversky and Kahneman's example, but with a slightly different wording. Our first set of decisions is parallel to the original example, but using U.S. dollars instead of pounds:

Example 1:

Decision 1: Choose between:

A. winning $240

B. a 25% chance of winning $1000 and a 75% chance of not winning or losing any money

Before answering, read the next decision.

Decision 2: Choose between:

C. losing $750

D. a 75% chance of losing $1000, and a 25% chance of not winning or losing any money

This example was conducted in two treatments, once as described above (Treatment 1) and once in the broad four-way presentation of the four combined choices AC, AD, BC and BD (Treatment 2), analogous to the third laboratory treatment. In both of these treatments, the participants made no other choices.

In all other treatments, we used only 50/50 gambles. The following are the new examples that we designed to generate dominance violations (the labels of choice options were changed from the instructions, for the sake of the exposition):

Example 2:

Decision 1: Choose between:

A. not winning or losing any money

B. a 50% chance of losing \$500 and a 50% chance of winning \$600

Before answering, read the next decision.

Decision 2: Choose between:

C. losing \$500

D. a 50% chance of losing \$1000, and a 50% chance of not winning or losing any money

Example 2 was designed to bring loss aversion into play: if participants weigh losses heavier than gains, they will tend to choose A over B, and if they are risk seeking in losses, they will tend to choose D over C. Such a combination is dominated with an expected loss of \$50 relative to the reversed choices. This example, too, was conducted in isolation — i.e. with no other decisions for the participants — and presented as stated here (Treatment 3) and presented as a broad four-way choice (Treatment 4). In addition, the example was presented together with other decisions, in three different ways. In Treatment 5, the two decisions appeared on the same screen. In Treatment 6, they appeared on separate screens, with four other tasks appearing in between. This variation was included to detect potential effects (e.g. distractions) caused by other choices. In Treatment 8, the example was presented as a broad four-way choice alongside with other decisions.

Similar to Example 1, Example 3 uses possible risk aversion in gains and risk lovingness in losses.

Example 3:

Decision 1: Choose between:

A. winning \$1500

B. a 50% chance of winning \$1000, and a 50% chance of winning \$2100

Before answering, read the next decision. [...]

Decision 2: Choose between:

C. losing \$500

D. a 50% chance of losing \$1000, and a 50% chance of not winning or losing any money

The choice of A and D is dominated with a loss of \$50 on average. The example was only conducted as stated here, in Treatment 6.[12] An important new feature of the example is that all possible combined outcomes involve positive amounts. Therefore, although a narrow evaluation of the second decision would consider negative payoffs, a broad-bracketing decisionmaker's choices can only be influenced by preferences over gains. In particular, under the assumption of broad-bracketed choice, a decision for D over C would be evidence of risk-lovingness in gains.[13]

[12] In treatment 6, the second decision of Example 3 is also the second decision of Example 2, so that the occurences of dominance violations are correlated between the two examples.

[13] This is true only if Example 3 is viewed separately from the third lottery choice decision in Treatment 6 (Decision 1 in Example 2). Considering all three decisions, the treatment involves eight possible choice combinations, only one of which involves a possible loss as only one of its eight possible outcomes. Hence, the choice of D in Example 3 would still be indicative of a preference for risk with almost all payoffs being positive.

The final example uses a more difficult spread between the payoffs and it involves some payoffs that are not multiples of \$100:

Example 4:

Decision 1: Choose between:

A. winning \$850

B. a 50% chance of winning \$100 and a 50% chance of winning \$1600

Before answering, read the second decision.

Decision 2: Choose between:

C. losing \$650

D. a 50% chance of losing \$1550, and a 50% chance of winning \$100

As before, a decisionmaker who rejects the risk in the first decision but accepts it in the second decision (A and D) would violate dominance, here with an expected loss of \$75 relative to B and C. An new feature is that these choices sacrifice expected value in the second decision, not in the first. This implies that for all broad-bracketing risk averters the combined choice of A and C would be optimal: it generates the highest available expected value at no variance. Different from the other examples, the prediction for a broad-bracketed risk averter is therefore independent of the exact nature of her preferences. A further property of the example is that A and C would be predicted even for some narrow bracketers who have preferences like in prospect theory, with diminishing sensitivity for larger gains and losses, loss aversion, and risk aversion/lovingness in the gain/loss domains. This is because the risky choice D involves a possible gain of \$100 so that a prospect-theoretic decisionmaker would only accept the gamble D if the preference for risk in the loss domain is strong relative to the effect of loss aversion (which makes her averse to lotteries with payoffs on both sides of zero). In particular, the preference for risk in the loss domain needs to be slightly stronger than in the often-used parameterization of Tversky and Kahneman (1992) – see footnote 17. Under the assumption of narrow bracketing, the example therefore helps to discriminate between different plausible degrees of risk lovingness in the loss domain. The example was conducted in Treatments 5, 7, and 8, with differences between broad versus narrow presentation, and with and without other decisions appearing in between the two decisions.

3 Experimental Results

3.1 Results of the Laboratory Experiment

Table 3 lists the frequencies of observing each of the four possible choice combinations in Example 1, in the four different laboratory treatments.

Treatment	A and C	A and D	B and C	B and D
Incentives-Small Scale	0.21	0.28	0.11	0.40
Flat Fee-Small Scale	0.16	0.34	0.09	0.41
Incentives-Small Scale-Broad Presentation	0.11	0.00	0.38	0.51
Flat Fee-Large Scale	0.15	0.54	0.08	0.23

Table 3: Laboratory choice frequencies in Example 1.

There is little difference between the observed behavior in the first two treatments. No matter whether the outcomes are actually paid or not, about half of the respondents choose the sure gain of A over the lottery B, and slightly more than two thirds choose the uncertain loss of lottery D over the sure loss in C. This confirms prospect theory's prediction of risk-seeking behavior in the losses domain, but with less clear evidence of risk aversion in the gain domain.[14] The dominance-violating combina-

[14]The choices in the sessions' second parts also show much risk taking behavior.

tion of A and D was chosen by 28 percent and 34 percent of the two treatments' participants, respectively. The difference in these two frequencies is insignificant at any conventional level ($p = 0.346$, one-tailed Fisher exact test), i.e. we find little indication that the frequency of dominance violations decreases if the decisions are paid for real. Similarly, testing for differences in the entire distribution of choices, not just in the frequency of A and D, we cannot reject the null hypothesis of no difference ($p = 0.866$, two-tailed Pearson chi-square test). Overall, there is no statistically significant effect of actually paying the (small-scale) decisions in this data set.

Comparing the Flat Fee-Small Scale treatment with the Flat Fee-Large Scale treatment, the frequencies of the dominated AD combination increases from 34 percent to 54 percent, which is statistically significant ($p = 0.042$, one-tailed Fisher exact test).[15]

Finally, a comparison between the Incentives-Small Scale treatment and the Incentives-Small Scale-Broad Presentation treatment suggests a strong effect of narrow bracketing. If the participants have to view the decision problem from a broad perspective, the number of combined A and D choices goes from 28% to 0% ($p < 0.001$, one-tailed Fisher exact test), and also the overall distributions of choices are significantly different ($p < 0.001$, two-tailed Pearson chi-square test). This clearly indicates that the subjects did not view the two decisions in the Incentives-Small Scale treatment as a combined problem.

[15] But a Pearson chi-square test still supports the hypothesis of identical four-way distributions of choices between the two treatments ($p = 0.221$, two-tailed). In light of the significant result of the Fisher exact test we attribute this failure to reject to the low numbers of observations. In any case, the results do not indicate that a large (hypothetical) payoff scale makes people more likely to bracket broadly.

3.2 Results of the Survey Experiment

3.2.1 Data Summary

Treatment	Description	A and C	A and D	B and C	B and D
1	Example 1	0.16	0.66	0.03	0.15
2	Example 1 — broad presentation	0.22	0.06	0.24	0.48
3	Example 2	0.24	0.53	0.05	0.19
4	Example 2 — broad presentation	0.09	0.38	0.12	0.41
5	Example 2	0.22	0.53	0.05	0.20
	Example 4 — broad presentation	0.72	0.04	0.13	0.12
6	Example 2 — separate screens,	0.26	0.44	0.07	0.23
	Example 3	0.23	0.50	0.10	0.17
7	Example 4	0.35	0.36	0.16	0.13
8	Example 2 — broad presentation,	0.15	0.24	0.24	0.36
	Example 4 — separate screens	0.33	0.43	0.10	0.14

Table 4: Survey choice frequencies

Table 4 summarizes the frequencies of the representative-survey responses, where A and D is the dominated combination of choices in each of the four examples. The table shows that the frequencies of dominance violations is large when the examples are presented as two choices. A and D is the most frequently chosen choice combination in all four examples, with average frequencies in Examples 1, 2, 3, and 4 of 0.66, 0.50, 0.50 and 0.40, respectively. In treatments 5, 6 and 8 of the survey, the participants made two binary choices and one four-way choice, which allows for a total of 16 choice combinations. Of these, dominance can be violated by any of 7 combinations that involve A and D in either example. One of these 7 combinations was chosen at frequencies of 0.56, 0.59 and 0.55, respectively.

All of these observed violations are inconsistent with any model where rational agents with monotonic preferences make broad-bracketed choices. But the table also contains more direct evidence against broad bracketing. In the three examples that we presented broadly, Examples 1, 2 and 4, the frequencies of dominated choice are significantly reduced. In Examples 1 and 4, violations are reduced drastically, to 0.06 and 0.03, respectively. But in the case of Example 2, there remains a large proportion of respondents (0.29, summing across treatments) who choose the dominated A and D even when the choice is presented to them in a broad way. This is puzzling, and clearly is not due to narrow bracketing. We have no good explanation for the high violation rate in this task.[16]

Evidence that participants are not broadly bracketing also also comesfrom thinking about the nature of preferences that broad bracketers would have to have in order to make the observed constellation of choices. The designs of Examples 3 and 4 help for this: in Example 3, 67% of the participants choose D over C, which broad bracketers would do only if they are risk loving in the gains domain. Beyond massive evidence for risk aversion over gains from previous experiments, however, the high frequencies of A and C in Example 4, particularly when the example is presented broadly (72%), suggest the opposite — because the choice of A and C is predicted for a risk-averse broad bracketer in that example.

3.2.2 Further Summary Statistics that Support Narrow Bracketing

To make such a comparison of hypothetical underlying preferences more rigorous, it is convenient to look at treatments 5 and 8, which contain the same set of 16 avail-

[16] Perhaps (to give a couple of *bad* explanations) the fact that AD has fewer nonzero outcomes led some participants to choose it, or BC is unattractive due to the large-looming loss of \$1000 that may appear even larger when contrasted with the small gain of \$100.

able choice combinations, only bracketed in different ways. One can ask whether any distribution of broadly bracketed preferences in the population would predict the choices in these two treatments. Obviously, no monotonic preferences would allow a dominated choice, so (considering the violation rates reported above) a distribution of broad monotonic preferences can at most generate 44% of the choices in these treatments. But because the available composite lotteries are identical and the allocation of participants into treatments was random, all behavioral differences between the two treatments are further evidence against broad bracketing. The largest possible proportion of choices that can be generated by a model with a stable distribution of broad-bracketing agents is therefore given by adding up the smaller of the two observed frequencies of all 9 undominated choice combinations. For brevity, we do not report the full distribution of choice combinations, but only the result of the addition: at most 33% of all choices could be generated by broad bracketers with monotonic preferences.

This upper bound permits arbitrary heterogeneity in the preferences of broad bracketers. Restricting preferences further in various ways yields some insights into the plausibility of more specific models. For example, assume an expected-utility model where all agents have two-part CRRA preferences with a kink at 0, so that

$$v(x) = \begin{Bmatrix} x^{1-\gamma} \text{ for } x \geq 0 \\ -\delta(-x^{1-\gamma}) \text{ for } x < 0 \end{Bmatrix}, \tag{3.1}$$

and allow the two parameters γ and δ to vary arbitrarily across the population. Under braod bracketing, this model rules out few additional choice combinations and explains up to 31% of the choices in treatments 7 and 10. In contrast, a distribution of CARA agents with different risk attitudes (a much less flexible model with one parameter per agent) could only explain up to 17%. Restricting preferences to meet the most standard model of economic decisionmaking — expected utility over total earnings — would require near-risk-neutrality and could not explain virtually any of the participants' choices: this model would predict B and C in Example 2 and A and C in Example 4 but only 2% of the choices in treatment 5 follow this prediction.

We also briefly summarize the evidence in Table 4 that speaks for or against models where agents have narrow brackets. As mentioned earlier, the surprisingly high frequency of A and D choices in the broad presentation of Example 2 represents a failure of all reasonable models, including those with narrow brackets. But in the remaining choices, the results are consistent with narrowly-bracketed preferences. In particular, it is straightforward to find prospect-theoretic preferences, e.g. of the form (3.1), that have a very good fit in the binary choices: preferences that exhibit a sufficient degree of loss aversion and a sufficiently fast decrease in the sensitivity to gains and losses would predict the modal choice in each of the survey's 13 binary choice problems — and a forteriori it would correctly predict the modal choice of A and D in all four examples.[17]

The table also shows that the experimental variation of presenting the two tasks of

[17] For a narrowly bracketed model with no free parameters, consider Tversky and Kahneman's (1992) estimated utility function, which is given by expression (3.1) with parameters $\gamma = 0.12$ and $\delta = 2.25$. Among the three four-way choices with broad presentation, this model would correctly predict the modal choice in Examples 1 and 4, but not in Example 2. Among the binary choices, it would predict A and D in Examples 1, 2 and 3, but not in Example 4. In treatments 5 and 8, the model would only be partially successful, correctly predicting the choice combinations in 44% and 8%, respectively. Partly, the poor performance in treatment 8 is driven by the strange behavior in Example 2 under broad presentation, and partly by the fact that the model would not predict D in the narrow presentation of Example 4, because the preference for risks in the negative domain is too small to offset the effect of the kink that discounts the high payoff of \$100 relative to the other payoffs. To predict D, the sensitivity parameter γ would have to be at least 0.15.

Examples 2 and 4 on separate screens, and hence including other choices in between, yielded no strong effect.[18] One can also ask more generally whether the inclusion of other tasks in a treatment appears to influence the choice frequencies. With the exception of the broadly presented Example 2 — which shows an effect towards fewer violations when other choices are included — there appears to be no such effect. Indeed, all of the binary choices can be quite reliably predicted independent of other choices, but strongly dependent on the framing of the choice itself: in each case where a binary choice problem was framed as a risky choice with positive payoffs (the A-versus-B problems in Examples 1, 3 and 4), at least 67% participants rejected the risk. In contrast, in each of the four cases where a binary choice was presented as a risky choice with negative payoffs (the C-versus-D choices in Examples 1, 2 and 3), the risk was *accepted* by at least 67% of the participants.

Summing up, we find that broad-bracketing models can explain only a small minority of choices, whereas most choices can be fairly well organized by assuming narrow brackets, with risk aversion over gains and strong risk lovingness over loss. To provide a fuller and more systematic statistical test of this claim, we now estimate simultaneously the preferences and brackets that fit the behavior best.

3.2.3 Simultaneous Estimation of Preferences and Degree of Bracketing

We analyze the lottery-choice data from the survey experiment under the assumption that there exist two types of decisionmakers: one broad type who integrates her lottery-choice decisions into a joint decision problem, and a narrow type who makes

[18] The differences in the frequencies of A and D are statistically significant but small, and have opposite directions between the two examples. In Example 2, the frequency decreases from 0.53 to 0.44 when other choices are included ($p = 0.091$, two-tailed Fisher exact test) and in Example 4, the frequency increases from 0.36 to 0.43 ($p = 0.026$).

all decisions one at a time. Apart from their different bracketing, we assume that the decisionmakers have identical preferences that are EU-representable with a utility function $v(\cdot)$. In order to generate a positive likelihood of observing any feasible choice vector, we assume logistic choice: the broad type calculates the expected utility $E[v(m)]$ from each available choice vector m in the set of choice combinations $M \equiv M_1 \times ... \times M_I$, and probabilistically makes her choice according to

$$\Pr(m|\lambda, \text{broad}) = \frac{\exp(\lambda E[v(m)]}{\sum_{m' \in M} \exp(\lambda E[v(m')]},$$

where $\lambda \in \mathbb{R}_0^+$ is the precision parameter that governs how well the choice probabilities approximate best responses. The narrow type calculates the expected utitlity for each choice m_i within each bracket i, and chooses m_i with probability

$$\Pr(m_i|\lambda, \text{narrow}) = \frac{\exp(\lambda E[v(m_i)]}{\sum_{m_i' \in M_i} \exp(\lambda E[v(m_i')]}.$$

Hence, her choices in bracket i are independent of the choices in other brackets. Letting θ be the proportion of broad bracketers, the overall likelihood of observing choice vector m is

$$\Pr(m|\theta, \lambda, v) = \theta \Pr(m|\lambda, \text{broad}) + (1 - \theta) \Pr(m|\lambda, \text{narrow}), \tag{3.2}$$

where the narrow type's likelihood of choosing the vector m is calculated as the product $\Pr(m|\lambda,\text{narrow}) = \Pi_i \Pr(m_i|\lambda,\text{narrow})$.

For the preferences v, we allow for a flexible hybrid CRRA-CARA utility function both above and below the status-quo point of $x = 0$, which we take as the agent's reference point. The hybrid CRRA-CARA function is given by[19]

$$v(x) = \left\{ \begin{array}{ll} \frac{1-\exp(-r_+ x^{1-\gamma_+})}{r_+} & \text{if } x \geq 0 \\ -\frac{1-\exp(-r_-(-x)^{1-\gamma_-})}{r_-} & \text{otherwise} \end{array} \right\},$$

[19] See Abdellaoui, Barrios and Wakker (2007) and Holt and Laury (2002) for related analyses with this hybrid function.

where $r_+, r_-, \gamma_+, \gamma_- \in (0,1)$. The parameters r_+ and γ_+ govern the shape of the function for positive x-values, and r_- and γ_- for negative x-values. This separation into two separate domains introduces a kink at 0 and makes v flexible in terms of allowing for changes in the degree of risk aversion. For $r_+ \to 0$ or $r_- \to 0$, the respective parts above or below the reference point exhibit constant relative risk aversion, and for $\gamma_+ \to 0$ or $\gamma_- \to 0$ they exhibit constant absolute risk aversion.

Simultaneously to estimating the four parameters of v via maximum likelihood, we estimate the proportion of broad types θ and the noise parameter λ. θ and λ are estimated as $\widehat{\theta} = 0.1119$ (std. dev. 0.0491) and $\widehat{\lambda} = 0.0133$ (0.0012). The obtained log likelihood is $ll^* = -1926.4$. The estimate $\widehat{\theta} = 0.1119$ indicates the degree of broad bracketing: only one out of nine choice vectors is estimated to be made by a broad-bracketing decisionmaker. Hence, our statistical model supports even more strongly the arguments above for the main empirical claim of the paper — that narrow bracketing is ubiquitous.

Figure 3.1 shows the estimated v function, with the parameter estimates for v given in the caption of the figure.

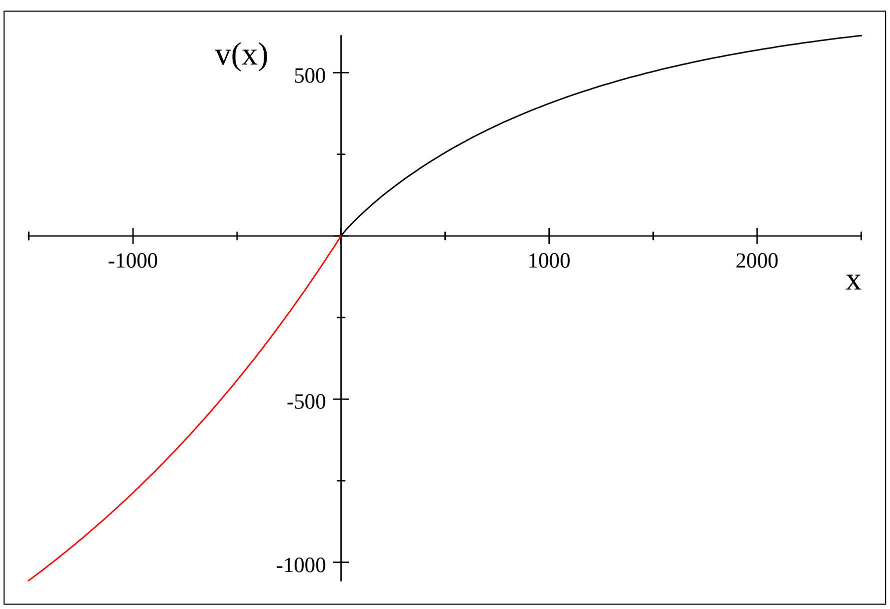

Figure 3.1: Estimated preferences v. Parameter estimates (and estimated standard deviations in parentheses) are $\widehat{r}_+ = 0.0014$ (0.0004), $\widehat{\gamma}_+ = 0.0740$ (0.0109), $\widehat{r}_- = 0.0005$ (0.0001), and $\widehat{\gamma}_+ = 0.0000$ (0.0000).

The estimates of the preferences are reminiscent of prospect theory's value function, with risk aversion around zero and in the positive domain, and a preference for risk in the negative domain. In part, this is because of the restrictions that we impose on the parameters: we require that $r_+, r_-, \gamma_+, \gamma_-$ all lie in $(0, 1)$, so that the function is necessarily concave above 0 and convex below 0. However, the reader is referred to our working paper (Rabin and Weizsäcker, 2007) for an analogous estimation with a flexible reference point, where the degree of risk aversion is unrestricted at any given x-value and the function is generally more flexible. The estimates in the working paper essentially confirm the depicted shape of the utility function. They also allow to reject the hypothesis that v has the CARA form, so that the propositions of Section 1 apply. Using a numerical approximation, we can also calculate the minimal horizontal distance K between the estimated preferences and the family of CARA functions, on the interval $[-\$1550, \$2500]$. The result is $K = \$183.4$, indicating (by Proposition 2) that the typical decisionmaker can be made to leave \$183.4 on the table in a single pair of choices.

We can also examine how well each of the extreme cases of the degree of bracketing can organize the data. Suppose we restrict $\theta = 0$, so that all decisionmakers are assumed to be narrow bracketers. The resulting model has a log likelihood of $ll^*_{\theta=0} = -1928.7$. While this implies that the restriction is rejected at statistical significance of $p = 0.032$, the log likelihood is still fairly close to that of the unrestricted model. In particular, the fit of the fully narrow model is hugely better than the fit when we restrict $\theta = 1$, the fully broad model. This latter model yields $ll^*_{\theta=1} = -2128.8$, i.e. it performs not much better than a uniformly random model of choice, with log likelihood $ll_{\text{rand}} = -2158.5$.[20]

3.2.4 Allowing for Heterogeneous Preferences

The above tests depend, of course, on the maintained assumptions about the preferences. In particular, the simplifying assumption that all agents have the same preferences is very strong. But in light of the data summary in Section 3.2.2, it seems impossible that allowing for heterogeneity would rescue the broad-bracketing model. There, we had found that even allowing for an arbitrary degree of heterogeneity, only small parts of the data can be accounted for within a broad-bracketing model.

A different kind of heterogeneity can arise if broad and narrow types have different preferences. We address this by estimating the parameters of (3.2) but with separate functions $v_{\text{narrow}}(\cdot)$ and $v_{\text{broad}}(\cdot)$ for the two types. To save on the number of parameters, we assume that both $v_{\text{narrow}}(\cdot)$ and $v_{\text{broad}}(\cdot)$ follow the two-part CRRA form given in (3.1). Hence, we estimate six parameters, $\gamma_{\text{narrow}}, \delta_{\text{narrow}}, \gamma_{\text{broad}}, \delta_{\text{broad}}, \theta$ and λ. This estimation also has the advantage that the kink parameters δ_{narrow} and δ_{broad} can be interpreted as the factors by which losses weigh heavier than gains. The resulting parameter estimates are $\widehat{\gamma}_{\text{narrow}} = 0.220$ (0.024), $\widehat{\delta}_{\text{narrow}} = 1.772$ (0.138), $\widehat{\gamma}_{\text{broad}} = 0.296$ (0.094), $\widehat{\delta}_{\text{broad}} = 2.752$ (1.589), $\widehat{\theta} = 0.110$ (0.058) and $\widehat{\lambda} = 0.026$ (0.003), and the model has a maximum log-likelihood of -1951.3. While there is

[20] As another goodness-of-fit measure we counted how often each model has its modal prediction on the choice that actually occurred. The best-fitting model among those with $\theta = 0$ correctly predicts a total of 63.3% of all the lottery choices. In contrast, the best-fitting model with $\theta = 1$ only has a rate of 48.6% correct predictions. Given that most choices are binary choices, this statistic further illustrates the weakness of the broad-bracketing model. A fully random model would correctly predict 44.6% of the choices. The value function of Kahneman and Tversky (1992), with their estimated parameters, correctly predicts 59.1% of the choices.

a hint of difference in preference between the types, it is statistically insignificant. The results also show that the high estimated frequency of narrow bracketing does not depend on assuming homogeneity of preferences between the two types, since the estimated proportion of broad bracketers remains at 11%. Under the restrictions that $\gamma_{\text{narrow}} = \gamma_{\text{broad}} = \gamma$ and $\delta_{\text{narrow}} = \delta_{\text{broad}} = \delta$, we estimate the parameters as $\widehat{\gamma} = 0.223$ (0.023), $\widehat{\delta} = 1.806$ (0.144), $\widehat{\theta} = 0.115$ (0.049) and $\widehat{\lambda} = 0.026$ (0.003), with a log-likelihood of -1951.5. We see that the kink parameter δ that governs loss aversion lies close to 2, confirming previous studies.

3.2.5 Testing for Demographic Differences

In Online Appendices 4 and 5, we consider the whether preferences and bracketing may differ by background characteristics of the decisionmakers in the survey sample. The panel is designed to be representative of the general U.S. population and we have a variety of potentially relevant characteristics in the data set. Before analyzing results, we chose to separate the data set into pairs of subsamples according to the following variables: gender, age, racial/ethnic background, household income, revealed mathematical skills, attendance of a mathematics course in college, and educational degree. In Online Appendix 4, we re-estimate the statistical model (with homogeneity in preferences between broad and narrow types) for each of the subsamples. In Online Appendix 5, we conduct a series of behavioral regressions with the personal characteristics as explanatory variables.

Between the subgroups of respondents, differences in the dominance violation rates may appear for two reasons: differences in the groups' preferences, and differences in the groups' propensities to narrowly bracket. It is therefore instructive to consider separately the two comparisons of preference parameters $(r_+, r_-, \gamma_+, \gamma_-)$ and the bracketing parameter θ. Online Appendix 4 shows that although preferences are sometimes quite different between the subgroups, the frequency of broad bracketers is below 22% in all subgroups. Although there is some variation, the only notable difference in θ is between men and women, where 21% of men are broad bracketers, compared to 0% of women. However, the men in our sample are also especially risk loving in the domain of losses, which makes them choose dominated combinations with about equal likelihood as women. A stronger behavioral difference appears between white and non-white respondents. Non-whites are much more risk neutral towards lotteries around zero and in the domain above zero. This difference in risk attitudes translates into a rate of dominance violations that is 28% higher for whites than for non-whites.

We find very little effects with regard to education and math skills.[21] Our results therefore only partially confirm recent studies by Benjamin, Brown and Shapiro (2006), Frederick (2006) and Dohmen et al (2007) who find that risk preferences change systematically with measures of IQ or mathematics skills. Like these studies, we find more risk neutrality among the math-skilled respondents. But (perhaps due to the different pools of participants and/or to the different behavioral outcome variables) we do not find robust differences in most choice rates and no significant difference in the bracketing parameter θ. These results suggest that it is not a question of numerical complexity that determines whether or not decisionmakers integrate broadly. Even math-skilled respondents are susceptible to narrow bracketing, and therefore to making dominated choices.

[21] The respondents who correctly answered three mathematical questions have a 9% lower violation rate than the remaining respondents. Respondents with a bachelor's degree have a 5% higher violation rate than those with lower level of schooling. Respondents who report to have attended a math course in college have a 8% higher violation rate. None of these differences is significant at 5 percent in logistic regressions.

4 Conclusion

We believe that the analysis in this paper can be usefully related to several recent strands of the literature on risk preferences. Rabin (2000a) formalizes a common intuition among researchers that the conventional diminishing-marginal-utility-of-wealth expected-utility model not only fails in systematic ways to empirically describe risk behaviors, but cannot even in theory provide a calibrationally plausible account of modest-scale risk aversion. As suggested in Rabin (2000a, 2000b) and Rabin and Thaler (2001), and brought into further focus by Cox and Sadiraj (2006) and Rubinstein (2006), the culprit in the failure of the conventional model to account for departures from modest-scale risk neutrality is the premise that choice is determined by final wealth: maintaining the assumption that people's weighting of prospects are linear in probabilities, the reality of widespread modest-scale risk aversion could be accounted for by assuming as in prospect theory that *changes* in wealth are the carriers of value.[22] Yet we note that the existing evidence of preferences over changes in wealth is essentially evidence of preferences over *choice-by-choice isolated changes* in wealth. As our new evidence reconfirms, most decisionmakers do not integrate their experimental choices with other choices that occur even within the same experimental session, and certainly not with simultaneous risks outside the experiment.

[22] Indeed, papers such as Benartzi and Thaler (1995), Bowman, Minehart, and Rabin (1999), Kőszegi and Rabin (2006, 2007, forthcoming) and many others have over the years emphasized that much of the insight of prospect theory as an alternative to the expected-utility-of-wealth model can be gleaned even assuming linear-probability preferences. Safra and Segal (2006) make the case even more clearly that the assumption that final wealth is the carrier of utility is the culprit: they show that a wide range of models that allow non-linear probability weighting but assume final wealth as the carrier of utility cannot provide plausible accounts of modest-scale risk aversion. Barberis, Huang and Thaler (2006) give a closely related discussion in a different model.

Because the typical decisionmaker narrowly brackets preferences that do not have a constant degree of absolute risk aversion, one can conclude that she is apt to make combinations of choices that are dominated. Our analysis thus reinforces Wakker's (2005) observation that choice-by-choice consistency with von Neumann's and Morgenstern's axioms does not per se yield the features of rationality generally associated with "expected-utility theory".

More generally, our experimental and theoretical results illustrate the problems inherent in trying to find a model of risky choice that is both positively and normatively compelling: narrow bracketing is highly prevalent in decisions under risk, but will lead to decisions that are bad by every normative standard we are aware of having been proposed by researchers. And the normative error associated with narrow bracketing here is especially striking because of the presentation of two simple independent choices side by side. Indeed, the finding of narrow bracketing in response to an in-your-face invitation to combine independent gambles is likely to be a lower bound on the propensity of people to narrowly bracket in risky choice. The huge number of risky choices people face in their lives are largely independent of each other, and are presented in stronger isolation. In this sense, the natural framing of life is likely to be more separated than any of the conditions in our or other experiments demonstrating narrow bracketing.

The finding of narrow bracketing in simultaneous side-by-side choice is relevant for a second reason. Narrow bracketing has, under various names, been interpreted as an error by many researchers — such as Kahneman and Lovallo (1993), Benartzi and Thaler (1995), Read, Loewenstein and Rabin (1999), and others. Recently researchers have argued — see, e.g., Palacios-Huerta (1999) for an earlier articulation and example, and Kőszegi and Rabin (forthcoming) for a general model that formalizes some of the issues — that *given correctly identified preferences* some instances of what has been thought of as narrow bracketing may in fact not be an error. Whereas sacrificing near-certain long-term gains because of over-attentiveness to short-term losses is a mistake from the perspective of previous preference models, Kőszegi and Rabin (forthcoming) argue, by defining "gain-loss utility" over changes in beliefs, that people may care about small changes in wealth even if they recognize that the changes contribute negligible risk to the consumption ultimately determined by their wealth. The reason is that gains and losses in money are news about future consumption, and this news generates immediate "prospective gain-loss utility". Isolated treatment of risks whose resolution is spread out over time can be rational with such preferences. However, the type of narrow bracketing we observe here is irrational even under this broader conception, since the realizations of the lotteries occur simultaneously and "news utility" cannot give an account of why people might knowingly separate the gambles out. Unless one believes in preferences that are even nuttier than proposed in this recent research, narrowly bracketed choices like those in this paper *are* errors.

The paper establishes that, under a very wide set of preferences, the failure to combine decisions can lead a decisionmaker to make such errors. We have not explored at all whether such situations are likely to arise naturally, or whether economic actors out to make money will have the wish and the ability to induce dominance by narrowly-bracketing agents. However, our results showing that dominance violations can occur without either the inclusion of choices that involve only losses, as does the second decision in the original example, or correlation among the choices, may add reason to believe that such violations might be prevalent.

Against all this, we note that the frequency of dominated choices is not itself likely to be of fundamental interest for welfare analysis because the prevalence of such choices does not tell us how much utility the decisionmakers forgo. A complete welfare analysis of the losses due to narrow bracketing would measure the utility loss occuring with the preferences that people seem to have, given the array of choices they face. We focus on stochastically dominated choices because this allows establishing that narrow bracketing induces mistakes independent of what preferences prevail (in the economy or in economic theory), and as such provide a rather stark and uncontroversial indication that prevalent behavior is inconsistent with normative models of risky choice.

A final methodological note concerns the question of how to devise empirical estimates of risk preferences. Narrow bracketing implies that empirical estimates of risk attitudes will vary widely with the assumptions about the scope of the decision problem that the agents face, and how well those assumptions match the way agents themselves isolate choices in their minds. The currently prevalent approach of measuring, for instance, a coefficient of relative risk aversion over wealth gives the researcher the freedom to choose from a range of possible definitions of wealth (from one-hour experimental earnings to lifetime wealth). This has the undesirable property that the choice of definition changes the measured coefficient by several orders of magnitude. The method might be made more coherent and justified if we linked the definition of wealth to the decisions or sets of decisions that people are focusing on. As a first step in the direction of adding more such discipline, our statistical analysis demonstrates that it is possible to include a simultaneous estimation of the agents' degrees of bracketing.

References

Abdellaoui, Mohammed, Carolina Barrios, and Peter P. Wakker.

2007. "Reconciling revealed utiltiy with introspective preference: Experimental arguments based on Prospect Theory." *Journal of Econometrics*, 138(1): 356-378.

Barberis, Nicholas, and Ming Huang. 2004. "Preferences with frames: A new utility specification that allows for the framing of risks." http://badger.som.yale.edu/faculty/ncb25/tc18b.pdf.

Barberis, Nicholas, Ming Huang, and Richard H. Thaler. 2006. "Individual preferences, monetary gambles, and stock market participation: A case for narrow framing." *American Economic Review,* 96(4): 1069-1090.

Battalio, Raymond C., John H. Kagel, and Komain Jiranyakul. 1990. "Testing between alternative models of choice under uncertainty: Some initial results." *Journal of Risk and Uncertainty,* 3(1): 25-50.

Bell, David E. 1985. "Disappointment in decision making under uncertainty." *Operations Research,* 33(1): 1-27.

Benartzi, Shlomo, and Richard H. Thaler. 1995. "Myopic loss aversion and the equity premium puzzle." *Quarterly Journal of Economics,* 110(1): 73-92.

Benjamin, Daniel J., Sebastian A. Brown, and Jesse M. Shapiro. 2005. "Who is 'behavioral'? Cognitive ability and anomalous preferences." http://home.uchicago.edu/~jmshapir/iq050506.pdf.

Birnbaum, Michael H., Gregory Coffey, Barbara A. Mellers, and Robin Weiss. 1992. "Utility measurement: Configural-Weight Theory and the judge's point of view." *Journal of Experimental Psychology: Hu-*

man Perception and Performance, 18(2): 331-346.

Bowman, David, Deborah Minehart, and Matthew Rabin. 1999. "Loss aversion in a consumption-savings model." *Journal of Economic Behavior and Organization,* 38(2): 155-178.

Camerer, Colin F. 1989. "An experimental test of several generalized utility theories." *Journal of Risk and Uncertainty,* 2(1): 61-104.

Cox, James C., and Vjollca Sadiraj. 2006. "Small- and large-stakes risk aversion: Implications of concavity calibrations for decision theory." *Games and Economic Behavior*, 56(1): 45-60.

de Finetti, Bruno. 1974. *Theory of Probability.* Vol. 1, New York: Wiley.

Diecidue, Enrico, Ulrich Schmidt, and Peter P. Wakker. 2004. "The utility of gambling reconsidered." *Journal of Risk and Uncertainty,* 29(3): 241-259.

Diecidue, Enrico, and Peter P. Wakker. 2002. "Dutch books: Avoiding strategic and dynamic complications, and a comonotonic extension." *Mathematical Social Sciences* 43(2): 135-149.

Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G. Wagner. 2005. "Individual risk attitudes: New evidence from a large, representative, experimentally-validated survey." IZA Discussion Paper 1730.

Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde. 2007. "Are risk aversion and impatience related to cognitive ability?" IZA Discussion Paper 2735.

Frederick, Shane. 2005. "Cognitive reflection and decision making." *Journal of Economic Perspectives,* 19(4): 25-42.

Gneezy, Uri, and Jan Potters. "An experiment on risk taking and evaluation periods." *Quarterly Journal of Economics,* 112(2): 631-645.

Holt, Charles A., and Susan K. Laury. 2002. "Risk aversion and incentive effects." *American Economic Review,* 92(5): 1644-1655.

Kahneman, Daniel, and Dan Lovallo. 1993. "Timid choices and bold forecasts: A cognitive perspective on risk-taking." *Management Science,* 39(1): 17-31.

Kőszegi, Botond, and Matthew Rabin. 2006. "A model of reference-dependent preferences." *Quarterly Journal of Economics,* 121(4): 1133-1166.

Kőszegi, Botond, and Matthew Rabin. 2007. "Reference-dependent risk attitudes." *American Economic Review,* 97(4): 1047-1073.

Kőszegi, Botond, and Matthew Rabin. Forthcoming. "Reference-Dependent Consumption Plans," *American Economic Review.*

Langer, Thomas, and Martin Weber. 2001. "Prospect theory, mental accounting, and differences in aggregated and segregated evaluation of lottery portfolios." Management Science 47(5): 716-733.

Loomes, Graham, and Robert Sugden. 1986. "Disappointment and dynamic consistency in choice under uncertainty." *Review of Economic Studies,* 53(2): 271-282.

Mellers, Barbara A., Robin Weiss, and Michael H. Birnbaum. 1992. "Violations of Dominance in Pricing Judgments." *Journal of Risk* *and Uncertainty,* 5(1): 73-90.

Palacios-Huerta, Ignacio. 1999. "The Aversion to the Sequential Resolution of Uncertainty." *Journal of Risk and Uncertainty,* 18(3): 249-269.

Rabin, Matthew. 2000a. "Risk aversion and expected utility theory: A calibration theorem." *Econometrica,* 68(5): 1281-1292.

Rabin, Matthew. 2000b. "Diminishing marginal utility of wealth cannot explain risk aversion." in *Choices, Values, and Frames*, ed. Daniel Kahneman and Amos Tversky, 202-208. New York: Cambridge University Press.

Rabin, Matthew, and Richard H. Thaler. 2001. "Anomalies: Risk aversion." *Journal of Economic Perspectives,* 15(1): 219-232.

Rabin, Matthew, and Georg Weizsäcker. 2007. "Narrow bracketing and dominated choices." IZA Discussion Paper 3040.

Read, Daniel, Loewenstein, George, and Matthew Rabin. 1999. "Choice bracketing." *Journal of Risk and Uncertainty,* 19(1-3): 171-197.

Redelmeier, Donald A., and Amos Tversky. 1992. "On the framing of multiple prospects." *Psychological Science* 3(3): 191-193.

Rubinstein, Ariel. 2006. "Dilemmas of an economic theorist." *Econometrica,* 74(4): 865-883.

Safra, Zvi, and Uzi Segal. 2006. "Calibration results for non-expected utility theories." Boston College Working Papers in Economics 645.

Thaler, Richard H., Amos Tversky, Daniel Kahneman, and Alan-Schwartz. 1997. "The effect of myopia and loss aversion on risk taking:

An experimental test." *Quarterly Journal of Economics,* 112(2): 647-661.

Tversky, Amos, and Daniel Kahneman. 1981. "The framing of decisions and the psychology of choice." *Science,* 211(4481): 453-458.

Tversky, Amos, and Daniel Kahneman. 1992. "Advances in prospect theory: Cumulative representation of uncertainty." *Journal of Risk and Uncertainty,* 5(4): 297-323.

Wakker, Peter P. 2005. "Formalizing Reference Dependence and Initial Wealth in Rabin's Calibration Theorem." http://people.few.eur.nl/wakker/pdf/calibcsocty05.pdf.

5 Appendix: Proofs

Proof of Proposition 1: The result is shown by a simple construction of two decision problems, each of which leaves the agent approximately indifferent. Suppose w.l.o.g. that $\widetilde{L}$ has the larger risk premium, i.e. $\pi_{\widetilde{L}} - \pi_L > \delta$. Therefore, to make the the agent indifferent between accepting $\widetilde{L}$ and a sure outcome, a relatively smaller sure outcome suffices. Offer her the following pair of choices, for small $\epsilon > 0$:

"Choose between lottery $\widetilde{L}$ and a sure payment of $CE_{\widetilde{L}} + \epsilon$".

"Choose between lottery L and a sure payment of $CE_L - \epsilon$".

She will take the sure payment in the first choice, and the lottery in the second. In sum, she will own the joint lottery given by $CE_{\widetilde{L}} + \epsilon + L$. But she could have made the reverse choices, which would have resulted in the joint lottery $CE_L - \epsilon + \widetilde{L}$. From the stated assumptions ($\mu_{\widetilde{L}} = \mu_L + \triangle x$, $\widetilde{L} = L + \triangle x$ and $\pi_{\widetilde{L}} - \pi_L > \delta$), it follows that $CE_L - \epsilon + \widetilde{L}$ is identical to $CE_{\widetilde{L}} + \epsilon + L$ plus a sure payment that exceeds $\delta - 2\epsilon$:

$$
\begin{aligned}
CE_L - \epsilon + \widetilde{L} &= CE_L - \epsilon + L + \triangle x \\
&= \mu_L - \pi_L - \epsilon + L + \triangle x \\
&> \mu_{\widetilde{L}} + \delta - \pi_{\widetilde{L}} - \epsilon + L \\
&= CE_{\widetilde{L}} + \delta - \epsilon + L
\end{aligned}
$$

Hence, a dominance of size δ can be approached. ■

Proof of Proposition 2: Let $CARA^*(\cdot)$ denote the CARA function that is closest to v (or the limiting function if the minimum does not exist), i.e. reaches the distance K at a "K-distance value" $x \in [\underline{x}, \overline{x}]$.[23] The proof proceeds in four steps and one lemma. The first three steps show the result for the case that $CARA^*$ is concave, and step 4 covers the case that $CARA^*$ is convex. Several additional notations will be used repeatedly in the proof: For a given binary lottery L with possible outcomes x'_L and x''_L and expected value $\mu_L \in (x'_L, x''_L)$, let $L(\cdot)$ denote the function that describes the straight line connecting $(x'_L, v(x'_L))$ and $(x''_L, v(x''_L))$ (the "lottery line"). Let $C_L(\cdot)$ be the function describing a straight line through $(x'_L, CARA^*(x'_L))$ and $(x''_L, CARA^*(x''_L))$. Similarly, let $H_L(\cdot)$ describe the straight line through $(x'_L, CARA^*(x'_L - K))$ and $(x''_L, CARA^*(x''_L + K))$, and let $J_L(\cdot)$ describe the straight line through $(x'_L, CARA^*(x'_L + K))$ and $(x''_L, CARA^*(x''_L - K))$. The two

[23]Since K is defined as an infimum over an open set of parameters $(\beta, \alpha, r) \in \mathbb{R}^3$ it may be that K can only be approached but not reached with equality at some or all K-distance values — this may occur for one or more of the CARA parameters growing to infinity. To cover this case, the precise definition of a "K-distance value" is a value such that for all $\epsilon > 0$ the horizontal distance lies within an ϵ-neighborhood of K for a sequence of CARA functions that converges to the limiting function. In the following, we will deal with the case that K can be reached with equality at all K-distance points, but all arguments extend to the case where K can only be approached.

lines C_L and H_L intersect at a point denoted by $x^*_{H_L}$, and the two lines C_L and J_L intersect at a point denoted by $x^*_{J_L}$. The lemma stated after the proof contains some properties of $x^*_{H_L}$ and $x^*_{J_L}$, as well as other properties of CARA functions.

From here on, we define risk premia with respect to the functions that represent underlying preferences. For example, for utility function v and binary lottery L, the risk premium π^v_L is defined as $\mu_L - CE^v_L$, where CE^v_L satisfies the indifference condition $v(CE^v_L) = \Pr(x')v(x') + (1 - \Pr(x'))v(x'')$. Finally, for any given point $(x, v(x))$ on the graph of v and for any given CARA function $\widetilde{CARA}(\cdot)$, we describe the horizontal distance between $(x, v(x))$ and $\widetilde{CARA}$ by the function $\eta(x, v, \widetilde{CARA}) = x - \widetilde{CARA}^{-1}(v(x))$. Under the assumptions of the proposition, it holds that $|\eta(x, v, CARA^*)| \leq K$ for all $x \in [\underline{x}, \overline{x}]$.

Step 1: Existence of four K-distance points on the graph of v. Assume that v has a horizontal distance of K from the CARA family on $[\underline{x}, \overline{x}]$. Then $[\underline{x}, \overline{x}]$ contains (at least) four distinct x-values $\{x_1, x_2, x_3, x_4\}$ with $|\eta(x_i, v, CARA^*)| = K$ for $i = 1...4$ and with three sign changes of $\{\eta(x_i, v, CARA^*)\}^4_{i=1}$ from one value to the next: there exist $x_1 < x_2 < x_3 < x_4$ such that either $[\eta(x_1, v, CARA^*) = -K$, $\eta(x_2, v, CARA^*) = K$, $\eta(x_3, v, CARA^*) = -K$ and $\eta(x_4, v, CARA^*) = K]$ or $[\eta(x_1, v, CARA^*) = K$, $\eta(x_2, v, CARA^*) = -K$, $\eta(x_3, v, CARA^*) = K$ and $\eta(x_4, v, CARA^*) = -K]$.

This statement holds because if there are fewer sign changes between K-distance values then $CARA^*$ cannot be the closest CARA function but there exists a CARA function $CARA''_\epsilon$ with the property that $|\eta(x, v, CARA''_\epsilon)| < K - \iota$ for all $x \in [\underline{x}, \overline{x}]$ and some $\iota > 0$. Towards a contradiction, assume that there are no more than two sign changes of $\eta(\cdot, v, CARA^*)$ between K-distance values in $[\underline{x}, \overline{x}]$. With no more than two such sign changes, we can assume without loss of generality that the sign

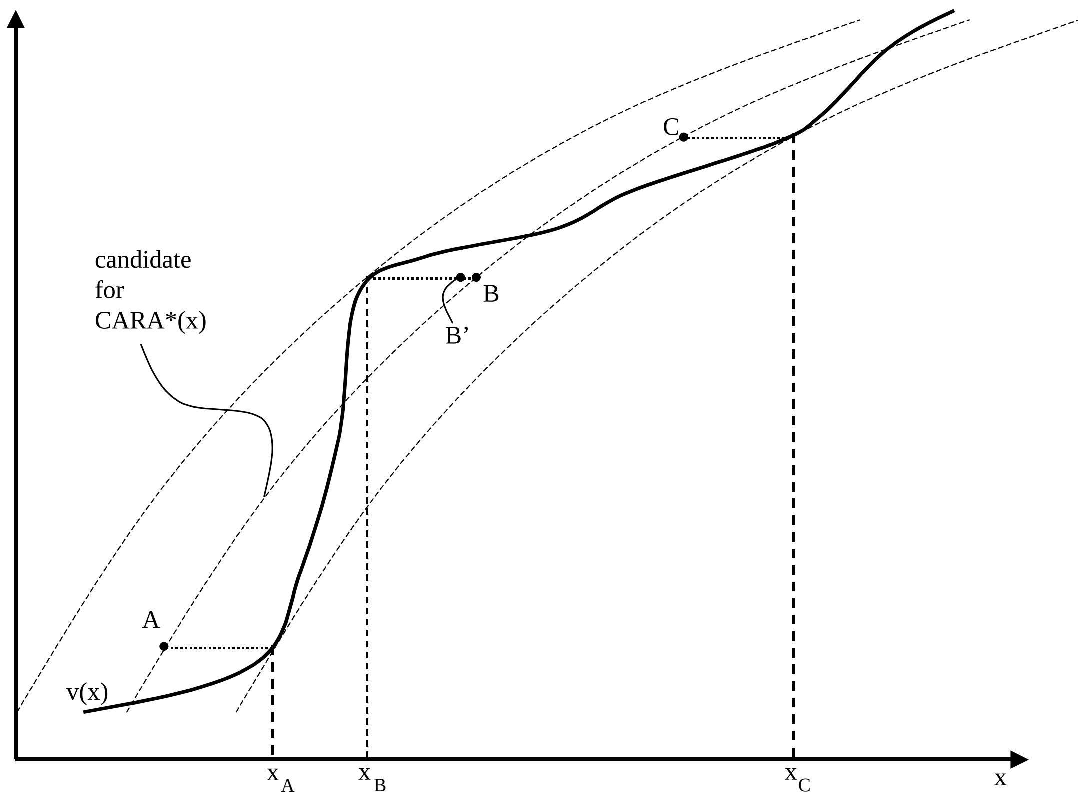

Figure 5.1: Three K-distance values

of $\eta(\cdot, v, CARA^*)$ does not change from negative to positive and back to negative within the set of K-distance values – so we have a sequence like in Figure 5.1.[24] We will discuss several cases, with the main argument given for Case 1. For each case, denote by X_B the (possibly empty) set of K-distance values that lie to the left of $CARA^*$ so that $X_B = \{x \in [\underline{x}, \overline{x}] | \eta(x, v, CARA^*) = -K\}$. If X_B is nonempty, let X_A and X_C be the two (possibly empty) sets of K-distance values below X_B and above X_B, respectively: $X_A = \{x \in [\underline{x}, \overline{x}] | x < X_B \text{ and } \eta(x, v, CARA^*) = K\}$ and $X_C = \{x \in [\underline{x}, \overline{x}] | x > X_B \text{ and } \eta(x, v, CARA^*) = K\}$. Then all K-distance values lie in the set $X_A \cup X_B \cup X_C$. If X_A and X_C are nonempty, select x_A to be the maximum of X_A and x_C to be the minimum of X_C.

Case 1: The sets X_A, X_B, X_C are all nonempty and X_B is a singleton with a unique element x_B. This case is depicted in Figure 5.1. (In the figure X_A and X_C are also singletons, but the arguments below cover the more general case of non-singleton sets.) By assumption, the minimum-distance CARA function is $CARA^*$, depicted as the middle dashed line through points A, B, C. Consider the point B', with coordinates $(x_B + K - \epsilon, v(x_B))$ for some small $\epsilon > 0$. By the lemma, property (i), there exists a CARA function $CARA'_\epsilon$ that connects the three points A, B', C.

Consider the following two-step manipulation of CARA functions: first, replace $CARA^*$ by $CARA'_\epsilon$; second, replace $CARA'_\epsilon$ by the function $CARA''_\epsilon : x \to CARA'_\epsilon(x - \epsilon/2)$ – i.e. shift $CARA'_\epsilon$ horizontally to the right by $\epsilon/2$. For sufficiently small ϵ, this manipulation will always be feasible and will result in a smaller horizontal distance to v: let $\widetilde{X}_\epsilon = \{x \in [\underline{x}, \overline{x}] || \eta(x, v, CARA''_\epsilon)| > K - \epsilon/4\}$ be the set of x-values that

[24] This is equivalent to assuming that if the sign changes twice, then it is from positive to negative to positive, like in Figure 5.1. If the sequence of sign switches is from negative to positive to negative, then construct X_A, X_B, X_C analogously, starting with the definition of $X_B = \{x \in [\underline{x}, \overline{x}] | \eta(x, v, CARA^*) = K\}$. All ensuing arguments are analogous.

have a horizontal distance that is strictly larger than $K - \epsilon/4$ if $CARA^*$ is replaced by $CARA''_\epsilon$. We will show that this set is empty for sufficiently small ϵ, by checking all possible $x \in [\underline{x}, \overline{x}]$.

Checking $x \in X_A \cup X_B \cup X_C$: for $x = x_B$ the two-step manipulation yields $|\eta(x, v, CARA''_\epsilon)| = |\eta(x, v, CARA'_\epsilon)| + \epsilon/2 = |\eta(x, v, CARA^*)| - \epsilon/2$ by construction of $CARA'_\epsilon$ and $CARA''_\epsilon$. For $x \in \{x_A, x_C\}$, the first manipulation from $CARA^*$ to $CARA'_\epsilon$ does not change the horizontal distance $|\eta|$, and the second manipulation reduces it by $\epsilon/2$ so that $|\eta(x, v, CARA''_\epsilon)| = |\eta(x, v, CARA^*)| - \epsilon/2$. For $x \in X_A \cup X_C \backslash \{x_A, x_C\}$, i.e. for K-distance values that lie below x_A and above x_C, the initial manipulation yields a strict reduction in $|\eta|$ for sufficiently small ϵ because $CARA'_\epsilon$ is more concave than $CARA^*$, and the second manipulation yields a further reduction by $\epsilon/2$. Hence, $|\eta(x, v, CARA''_\epsilon)| > K - \epsilon/4$ cannot hold at any initial K-distance value.

Checking $x \notin X_A \cup X_B \cup X_C$: first observe that only in the vicinity of the initial K-distance values in $X_A \cup X_B \cup X_C$ can v's distance from $CARA^*$ be arbitrarily close to K. Since $CARA^*$ and $CARA''_\epsilon$ converge for $\epsilon \to 0$, this observation implies that $\widetilde{X}_\epsilon$ converges to $X_A \cup X_B \cup X_C$ as $\epsilon \to 0$. Hence, any sequence of pairs $(\epsilon, \widetilde{x}_\epsilon)$ such that $\epsilon \to 0$ and $\widetilde{x}_\epsilon \in \widetilde{X}_\epsilon$ has the property that $\widetilde{x}_\epsilon$ lies arbitrarily close to $X_A \cup X_B \cup X_C$ for sufficiently small ϵ. But $\widetilde{x}_\epsilon$ cannot lie to the left of x_A or to the right of x_C, by the same argument that we gave above for $x \in X_A \cup X_C \backslash \{x_A, x_C\}$: for any such values the two manipulations will yield a reduction in horizontal distance by strictly more than $\epsilon/2$ so that $|\eta(\widetilde{x}_\epsilon, v, CARA''_\epsilon)| > K - \epsilon/4$ can hold only if $\widetilde{x}_\epsilon \in (x_A, x_C)$. It follows that $\widetilde{x}_\epsilon$ must become arbitrarily close to the set $\{x_A, x_B, x_C\}$ so that there exists a subsequence of $\widetilde{x}_\epsilon$, $\widetilde{\widetilde{x}}_\epsilon$, that converges to x_A, x_B or x_C.

Could $\widetilde{\widetilde{x}}_\epsilon$ converge to x_A? $\widetilde{\widetilde{x}}_\epsilon$ needs to satisfy $|\eta(\widetilde{\widetilde{x}}_\epsilon, v, CARA''_\epsilon)| > K - \epsilon/4$, which implies $|\eta(\widetilde{\widetilde{x}}_\epsilon, v, CARA'_\epsilon)| > K + \epsilon/4$, since $CARA''_\epsilon$ results from a right shift of $CARA'_\epsilon$ by $\epsilon/2$. But this means that as $CARA^*$ is replaced by $CARA'_\epsilon$ in the first manipulation, the horizontal distance $\eta(x, v, \cdot)$ increases at $x = \widetilde{\widetilde{x}}_\epsilon$ by more than $1/4$ of the decrease at $x = x_B$, or $\frac{|\eta(\widetilde{\widetilde{x}}_\epsilon,v,CARA^*)-\eta(\widetilde{\widetilde{x}}_\epsilon,v,CARA'_\epsilon)|}{|\eta(x_B,v,CARA^*)-\eta(x_B,v,CARA'_\epsilon)|} > 1/4$. This is impossible if $\widetilde{\widetilde{x}}_\epsilon$ converges to x_A because by L'Hospital's rule we have

$$\begin{aligned}
&\lim_{\epsilon\to 0} \frac{|\eta(\widetilde{\widetilde{x}}_\epsilon, v, CARA^*) - \eta(\widetilde{\widetilde{x}}_\epsilon, v, CARA'_\epsilon)|}{|\eta(x_B, v, CARA^*) - \eta(x_B, v, CARA'_\epsilon)|} \\
=&\lim_{\epsilon\to 0} \frac{\frac{d}{d\epsilon}|\eta(\widetilde{\widetilde{x}}_\epsilon, v, CARA^*) - \eta(\widetilde{\widetilde{x}}_\epsilon, v, CARA'_\epsilon)|}{\frac{d}{d\epsilon}|\eta(x_B, v, CARA^*) - \eta(x_B, v, CARA'_\epsilon)|} \\
=&\frac{\frac{d}{d\epsilon}|\eta(\lim_{\epsilon\to 0}\widetilde{\widetilde{x}}_\epsilon, v, CARA^*) - \eta(\lim_{\epsilon\to 0}\widetilde{\widetilde{x}}_\epsilon, v, CARA'_\epsilon)|}{1} \\
=&\frac{d}{d\epsilon}|\eta(x_A, v, CARA^*) - \eta(x_A, v, CARA'_\epsilon)| = 0,
\end{aligned}$$

a contradiction. The same argument rules out that $\widetilde{\widetilde{x}}_\epsilon$ converges to x_C.

Could $\widetilde{\widetilde{x}}_\epsilon$ converge to x_B? Again, $\widetilde{\widetilde{x}}_\epsilon$ would need to satisfy $|\eta(\widetilde{\widetilde{x}}_\epsilon, v, CARA''_\epsilon)| > K - \epsilon/4$, which implies that as $CARA^*$ is replaced by $CARA''_\epsilon$ the absolute horizontal distance $|\eta(x, v, \cdot)|$ is reduced at $x = \widetilde{\widetilde{x}}_\epsilon$ by less than $1/2$ of the reduction at $x = x_B$. But by L'Hospital's rule we have

$$\begin{aligned}
&\lim_{\epsilon\to 0} \frac{|\eta(\widetilde{\widetilde{x}}_\epsilon, v, CARA^*) - \eta(\widetilde{\widetilde{x}}_\epsilon, v, CARA''_\epsilon)|}{|\eta(x_B, v, CARA^*) - \eta(x_B, v, CARA''_\epsilon)|} \\
=&\lim_{\epsilon\to 0} \frac{\frac{d}{d\epsilon}|\eta(\widetilde{\widetilde{x}}_\epsilon, v, CARA^*) - \eta(\widetilde{\widetilde{x}}_\epsilon, v, CARA''_\epsilon)|}{\frac{d}{d\epsilon}|\eta(x_B, v, CARA^*) - \eta(x_B, v, CARA''_\epsilon)|} \\
=&\frac{\frac{d}{d\epsilon}|\eta(\lim_{\epsilon\to 0}\widetilde{\widetilde{x}}_\epsilon, v, CARA^*) - \eta(\lim_{\epsilon\to 0}\widetilde{\widetilde{x}}_\epsilon, v, CARA''_\epsilon)|}{\frac{d}{d\epsilon}|\eta(x_B, v, CARA^*) - \eta(x_B, v, CARA''_\epsilon)|} \\
=&\frac{\frac{d}{d\epsilon}|\eta(x_B, v, CARA^*) - \eta(x_B, v, CARA''_\epsilon)|}{\frac{d}{d\epsilon}|\eta(x_B, v, CARA^*) - \eta(x_B, v, CARA''_\epsilon)|} = 1,
\end{aligned}$$

in contradiction to the preceding sentence, so that $\widetilde{\widetilde{x}}_\epsilon$ cannot converge to x_B either. Therefore, $\widetilde{X}_\epsilon$ must be empty under the assumptions of Case 1, for small enough ϵ.

Case 2: The sets X_A, X_B, X_C are all nonempty and X_B is has multiple elements. In this case, define $x_B^{\min} = \min X_B$ and $x_B^{\max} = \max X_B$. The argument proceeds analogously to Case 1, but instead of using the value x_B in the construction, we will appropriately choose for each ϵ a value $x_{B,\epsilon} \in \{x_B^{\min}, x_B^{\max}\}$ which will determine the functions $CARA'_\epsilon$ and $CARA''_\epsilon$.

For any given (small) $\epsilon > 0$ consider the point $B'^{\min}$ with coordinates $(x_B^{\min} + K - \epsilon, v(x_B^{\min}))$ and construct CARA functions $CARA'^{\min}_\epsilon$ and $CARA''^{\min}_\epsilon$ just as in Case 1 but using $x_B^{\min}$ instead of x_B: $CARA'^{\min}_\epsilon$ is the CARA function through A, $B'^{\min}$ and C, and $CARA''^{\min}_\epsilon$ is a right shift of $CARA'^{\min}_\epsilon$ by $\epsilon/2$. Among the x-values in X_B, let $x_{B,\epsilon}$ be the value with the smallest reduction in horizontal distance that is achieved by replacing $CARA^*$ by $CARA'^{\min}_\epsilon$, given by

$$x_{B,\epsilon} = \arg\min_{x \in X_B} |\eta(x, v, CARA^*)| - |\eta(x, v, CARA'^{\min}_\epsilon)|.$$

Note that $|\eta(x, v, CARA^*)| - |\eta(x, v, CARA'^{\min}_\epsilon)|$ is simply the horizontal distance between $CARA^*$ and $CARA'^{\min}_\epsilon$ at the level $v(x)$ – it can be rewritten as $|\eta(\widetilde{x}, CARA'^{\min}_\epsilon, CARA^*)|$ for $\widetilde{x} = (CARA^*)^{-1}(v(x))$.[25] This yields that $x_{B,\epsilon} \in \{x_B^{\min}, x_B^{\max}\}$ because the two functions cross in A and C: for such a pair of functions, property (v) of the lemma implies that their horizontal distance cannot have a local minimum between A and C.

If $x_{B,\epsilon} = x_B^{\min}$, then we set $CARA'_\epsilon = CARA'^{\min}_\epsilon$ and $CARA''_\epsilon = CARA''^{\min}_\epsilon$. If $x_{B,\epsilon} = x_B^{\max}$, then we repeat the construction of the two CARA functions, but using $x_B^{\max}$ instead of $x_B^{\min}$: we construct $CARA'^{\max}_\epsilon$ through A, C and the point with coordinates $(x_B^{\max} + K - \epsilon, v(x_B^{\max}))$, and we construct $CARA''^{\max}_\epsilon$ by shifting $CARA'^{\max}_\epsilon$ to the right by $\epsilon/2$. Then we set $CARA'_\epsilon = CARA'^{\max}_\epsilon$ and $CARA''_\epsilon = CARA''^{\max}_\epsilon$.

Notice that if $x_{B,\epsilon} = x_B^{\max}$, $CARA'^{\max}_\epsilon$ lies above $CARA'^{\min}_\epsilon$ at all $x \in X_B$: $x_{B,\epsilon} =$

[25] The reformulation is $|\eta(x, v, CARA^*)| - |\eta(x, v, CARA'^{\min}_\epsilon)| = (CARA^*)^{-1}(v(x)) - (CARA'^{\min}_\epsilon)^{-1}(v(x)) = \widetilde{x} - (CARA'^{\min}_\epsilon)^{-1}(CARA^*(\widetilde{x})) = \eta(\widetilde{x}, CARA^*, CARA'^{\min}_\epsilon)$.

$x_B^{\max}$ implies that the horizontal distance decreases less at $x_B^{\max}$ than at $x_B^{\min}$ when $CARA^*$ is replaced by $CARA_\epsilon'^{\min}$ – hence, the decrease at $x_B^{\max}$ is less than ϵ. But when replacing $CARA^*$ by $CARA_\epsilon'^{\max}$, the distance at $x_B^{\max}$ is decreased by ϵ, so that $CARA_\epsilon'^{\max}$ lies above $CARA_\epsilon'^{\min}$ at $x = x_B^{\max}$. Since two CARA functions can only cross each other twice (by property (v) of the lemma), and $CARA_\epsilon'^{\min}$ and $CARA_\epsilon'^{\max}$ cross in A and C, $CARA_\epsilon'^{\max}$ lies above $CARA_\epsilon'^{\min}$ at all $x \in [x_A - K, x_C - K]$. By an analogous argument, if $x_{B,\epsilon} = x_B^{\min}$ then $CARA_\epsilon'^{\min}$ lies above $CARA_\epsilon'^{\max}$ at all $x \in [x_A - K, x_C - K]$. Therefore, we have established that the construction of $x_{B,\epsilon}$, $CARA_\epsilon'$ and $CARA_\epsilon''$ has the property that $|\eta(x, v, CARA^*)| - |\eta(x, v, CARA_\epsilon')| \geq \epsilon$ for all $x \in X_B$, and hence $|\eta(x, v, CARA^*)| - |\eta(x, v, CARA_\epsilon'')| \geq \epsilon/2$ for all $x \in X_B$.

The arguments of Case 1 now apply to this construction, ruling out that $|\eta(x, v, CARA_\epsilon'')| > K - \epsilon/4$ for small enough ϵ at all x that are bounded away from X_B. In addition, we need to rule out two possibilities: First, could $|\eta(x, v, CARA_\epsilon'')| > K - \epsilon/4$ be true for $x \in X_B$? This is impossible due to the observation in the previous paragraph, which implies that $|\eta(x, v, CARA_\epsilon'')| \leq K - \epsilon/2$ for $x \in X_B$. Second, could there exist a sequence $(\epsilon, \widetilde{x}_\epsilon)$ such that $\epsilon \to 0$, $\widetilde{x}_\epsilon \in \widetilde{X}_\epsilon$ and $\widetilde{x}_\epsilon$ becomes arbitrarily close to X_B? No: $\widetilde{x}_\epsilon$ would have a subsequence $\widetilde{\widetilde{x}}_\epsilon$ converging to some value $\widetilde{\widetilde{x}} \in X_B$, and due to the observation in the previous paragraph we could use the same argument as in Case 1, where convergence of $\widetilde{\widetilde{x}}_\epsilon$ to x_B was ruled out: it would have to hold that replacing $CARA^*$ by $CARA_\epsilon''$ reduces $|\eta(x, v, \cdot)|$ at $x = \widetilde{\widetilde{x}}_\epsilon$ by less than $1/2$ of the reduction at $x = \widetilde{\widetilde{x}}$, which is disproven by L'Hospital's rule for small enough ϵ.

Case 3: X_A is empty but X_B and X_C are nonempty, so that there is only one sign change between K-distance values. The argument follows exactly the cases above, except that x_A is chosen to be the smallest possible value, $x_A = \underline{x}$.[26]

[26] The case that X_C is empty is analogous, with $x_C = \overline{x}$.

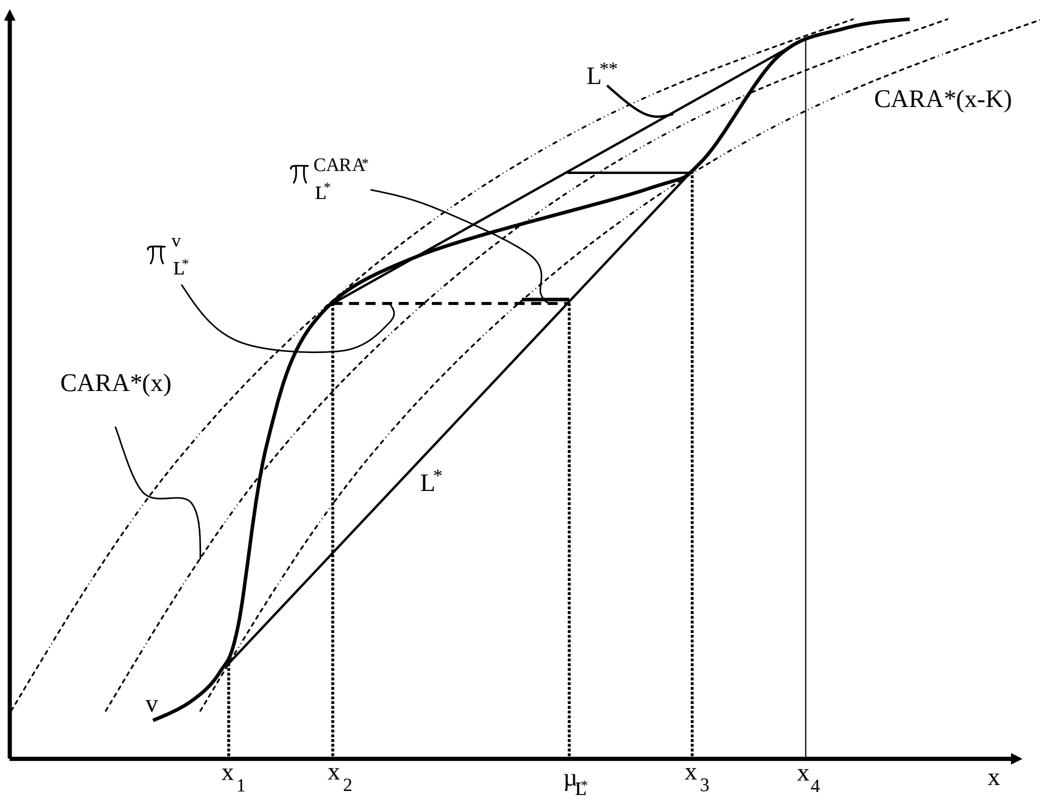

Figure 5.2: Lotteries L^* and L^{**}

Case 4: X_B is empty or both X_A and X_C are empty, so that no sign change between K-distance points exists. In this case $CARA^*$ cannot be the best-fitting CARA function, as it can obviously be improved by a horizontal shift that moves it closer to v.

Step 2: Constructing the candidate lotteries L^A and L^B: As seen in Proposition 1, it suffices to find a lottery that can be shifted within $[\underline{x}, \overline{x}]$ such that its risk premium varies by at least K. Step 1 guarantees the existence of four K-distance values $x_1 < x_2 < x_3 < x_4$ where v lies on alternating sides of $CARA^*$. We start by considering the lottery L^* in Figure 5.2, which has as its two outcomes x_1 and x_3, and has an expectation μ_{L^*} that is chosen such that the decisionmaker is indifferent between receiving lottery L^* or a sure payment of x_2. This can be expressed as

$L^*(\mu_{L^*}) = v(x_2)$, where $L^*(\cdot)$ is the function describing the lottery line. This lottery L^* has the property that $\pi^v_{L^*} = \pi^{CARA^*}_{L^*} + 2K$, as can be seen in Figure 5.2: $\pi^{CARA^*}_{L^*}$ is the horizontal distance between $(\mu_{L^*}, L^*(\mu_{L^*}))$ and the graph of $CARA^*(x-K)$ at v-value $v(x_2)$ — this is because $CARA^*$'s risk premium is constant across shifts of a lottery (property (ii) of the lemma), so that $\pi^{CARA^*(x-K)}_{L^*}$ and $\pi^{CARA^*(x)}_{L^*}$ are identical. Hence, it holds that $\pi^v_{L^*} = \pi^{CARA^*}_{L^*} + 2K$, as the point $(x_2, v(x_2))$ lies $2K$ to the left of the graph of $CARA^*(x-K)$.

Analogously, consider the lottery L^{**}, with outcomes x_2 and x_4 and the property that $L^{**}(\mu_{L^{**}}) = v(x_3)$ (see Figure 5.2). Here, it holds that $\pi^v_{L^{**}} = \pi^{CARA^*}_{L^{**}} - 2K$.

Now consider changing the payoffs of lottery L^*, while holding its expected value μ_L constant, at the level where $L(\mu_L) = v(x_2)$. It holds that any such lottery L — with payoffs x' and x'' that lie in $[\underline{x}, \overline{x}]$, and with $L(\mu_L) = v(x_2)$ — has the property that $\pi^v_L \geq \pi^{CARA^*}_L$. This is because $v(x') \leq CARA^*(x'+K)$ and $v(x'') \leq CARA^*(x''+K)$: if v reaches both of these upper bounds, then $\pi^v_L = \pi^{CARA^*}_L$. All lower values of $v(x')$ and $v(x'')$ result in a lower location of the lottery line $L(\cdot)$, and hence in a larger π.

Analogously, it holds that any binary lottery L with $L(\mu_L) = v(x_3)$ has the property that $\pi^v_L \leq \pi^{CARA^*}_L$ if both payoffs lie in $[\underline{x}, \overline{x}]$.

Notice that the above properties suggest a construction that would produce the desired result: if we can shift lottery L^* far enough to the right so that this shift results in a lottery $\widetilde{L}^*$ with an expected utility of $L(\mu_{\widetilde{L}^*}) = v(x_3)$, then we are done: due to CARA's constant π (see property (ii) of the lemma), we would have $\pi^{CARA^*}_{L^*} = \pi^{CARA^*}_{\widetilde{L}^*}$, so the two properties $\pi^v_{L^*} = \pi^{CARA^*}_{L^*} + 2K$ and $\pi^v_{\widetilde{L}^*} \leq \pi^{CARA^*}_{\widetilde{L}^*}$ would imply that the risk premium varies by at least $2K$, i.e. more than we need. We could then apply Proposition 1 to conclude that an FOSD-violation by $2K$ can be generated. The trouble is that we cannot always shift L^* far enough. It may be that the upper payoff

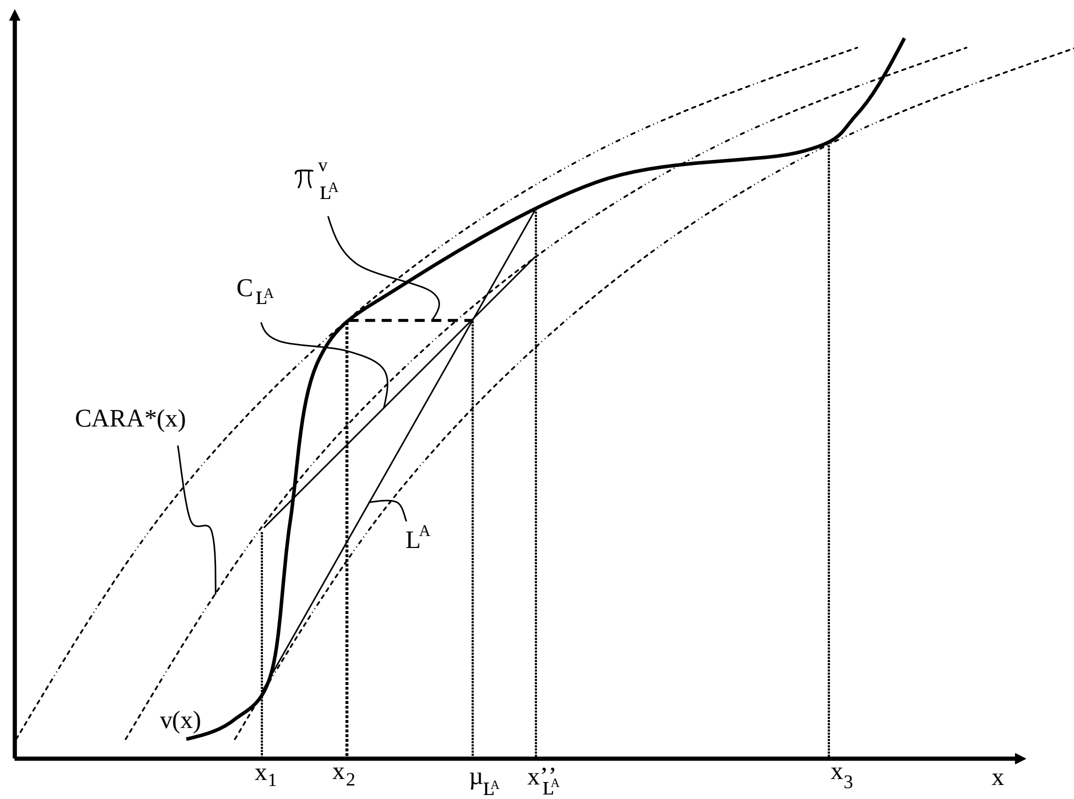

Figure 5.3: Lottery L^A , in the case $v(x_1) < CARA(x_1)$

bound $\overline{x}$ is close or equal to x_4, and that x_3 is close to x_4 , such that $x_3 - \mu_{L^*}$ exceeds $\overline{x} - x_3$. Similarly, it may be that the analogous shift of L^{**} to the left is not possible either.

However, we can use the lotteries L^* and L^{**} to generate two "shorter" lotteries that can be shifted far enough for the risk premium to vary by at least K (as will be shown in the next steps).

First, there exists a binary lottery L^A with the properties that $\pi^v_{L^A} = \pi^{CARA^*}_{L^A} + K$, that the low outcome is x_1, and that its expected utility is $L(\mu_{L^A}) = v(x_2)$. This lottery is depicted in Figure 5.3. L^A must exist due to continuity: consider the lottery L^*, and decrease its high payoff outcome x'', but keep the low lottery outcome x_1 constant and keep the expected utility constant at $L(\mu) = v(x_2)$. At the starting

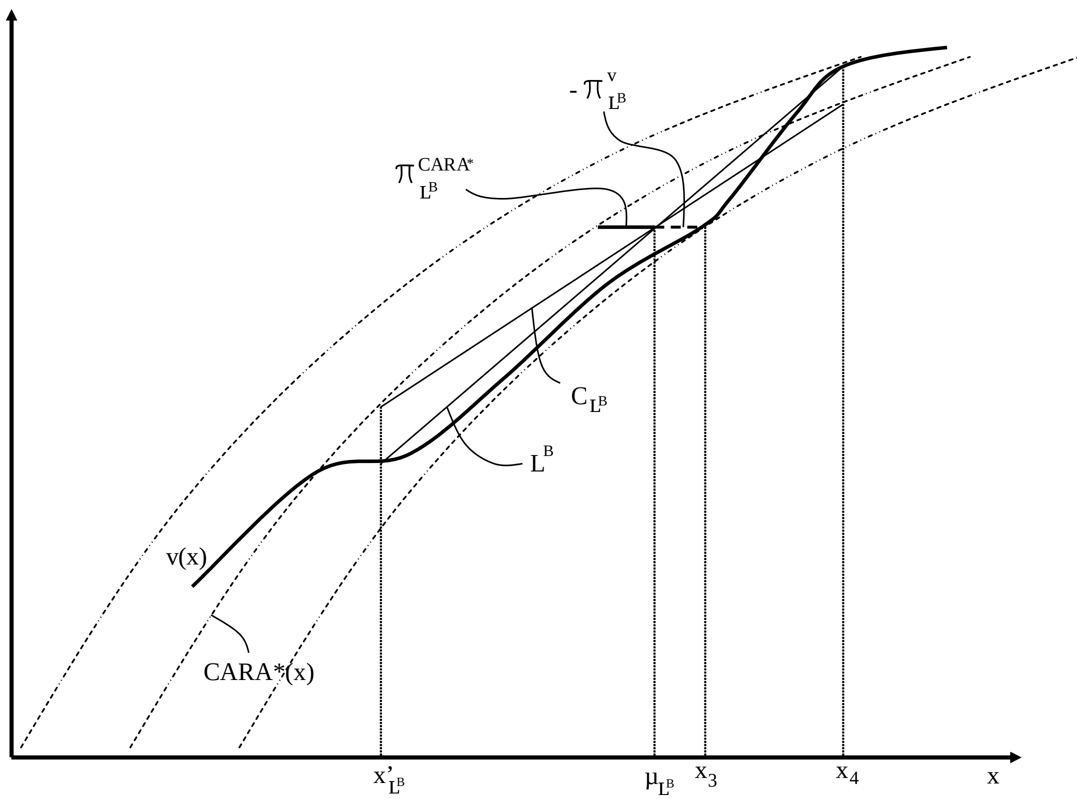

Figure 5.4: Lottery L^B , in the case $v(x_1) < CARA(x_1)$

point of this variation, i.e. with a high outcome of x_3, it holds that $\pi^v_{L^*} = \pi^{CARA^*}_{L^*} - 2K$. As the high outcome approaches x_2, it holds that the risk premia of v and $CARA^*$ of the resulting lotteries converge — they both approach 0. Hence, by continuity of π^v and π^{CARA^*}, the variation will generate a lottery L^A with high outcome x''_{L^A} such that $\pi^v_{L^A} = \pi^{CARA^*}_{L^A} + K$. Let L^A be the shortest lottery that has this property.

Second, and analogously, there exists a lottery L^B with the property $\pi^v_{L^B} = \pi^{CARA^*}_{L^B} - K$, with a high outcome of x_4, and with expected value of $L(\mu_{L^B}) = v(x_3)$. Again, let L^B be the shortest such lottery, i.e. the one with the highest low outcome. This lottery is depicted in Figure 5.4.

Step 3: If $CARA^*$ is concave, then either L^A or L^B can be shifted far enough. In the following we will show that either L^A can be shifted to the right, resulting in a lottery $L^{\widetilde{A}}$ with expected utility $L^{\widetilde{A}}(\mu_{L^{\widetilde{A}}}) = v(x_3)$ or L^B can be shifted to the left, resulting in a lottery $L^{\widetilde{B}}$ with expected utility $L^{\widetilde{B}}(\mu_{L^{\widetilde{B}}}) = v(x_2)$. By the observations in the previous step, this suffices to prove the result (for the case that $CARA^*$ is concave). We consider 4 cases — two with $v(x_1) < CARA^*(x_1)$ (as in Figures 5.3 and 5.4), and two with $v(x_1) > CARA^*(x_1)$.

Case 1: $v(x_1) < CARA^*(x_1)$, and L^A is longer than L^B, i.e. $x''_{L^A} - x_1 > x_4 - x'_{L^B}$. First notice that the line H_{L^A} lies above the lottery line of L^A. (See Figure 5.3: H_{L^A} is a straight line through $(x_1, v(x_1))$ and $(x''_{L^A}, CARA^*(x''_{L^A} + K))$ — it is not drawn in the figure, to avoid having too many lines.) Hence, the value $x^*_{H^A}$, where H_{L^A} intersects C_{L^A}, lies to the left of μ_{L^A}, i.e. $x^*_{H^A} \leq \mu_{L^A}$. Similarly, for lottery L^B, consider the auxiliary straight line H_{L^B} through $(x_4, v(x_4))$ and $(x'_{L^B}, CARA^*(x'_{L^B} - K))$. H_{L^B} lies below C_{L^B}, so it holds that $\mu_{L^B} \leq x^*_{H^B}$.

Now shift L^B to the left, such that the low outcome of the resulting lottery $L^{B'}$ is equal to x_1 — which is clearly possible without leaving $[\underline{x}, \overline{x}]$. Let $x''_{L^{B'}}$ be the high outcome of $L^{B'}$, and $\mu_{L^{B'}}$ be its expected value. We will show that for this shifted lottery we have $L^{B'}(\mu_{L^{B'}}) \leq v(x_2)$, which implies the desired result, because by continuity of v there must then exist another shifted version of L^B, $L^{\widetilde{B}}$, with $L^{\widetilde{B}}(\mu_{L^{\widetilde{B}}}) = v(x_2)$.

For the shifted lottery $L^{B'}$, consider the auxiliary lines $H_{L^{B'}}$ and $C_{L^{B'}}$, which intersect at a point $x^*_{H^{B'}}$. Due to property (iii) of the lemma, the relative x-location of this intersection is constant, i.e. it holds that $x^*_{H^{B'}} = x''_{L^{B'}} - (x_4 - x^*_{H^B})$. This implies that $\mu_{L^{B'}} \leq x^*_{H^{B'}}$, because both $x_{L^{B'}}$ and $x^*_{H^{B'}}$ were shifted by the same $\triangle x$ (from $x^*_{L^B}$ and $x^*_{H^B}$).

It holds that $C_{L^A}(\mu_{L^A}) = v(x_2)$ by construction. $x^*_{H^A} \leq \mu_{L^A}$ then implies that $C_{L^A}(x^*_{H^A}) \leq v(x_2)$, because C_{L^A} is increasing. From property (iv) of the lemma, it holds that $C_{L^{B'}}(x^*_{H^{B'}}) < C_{L^A}(x^*_{H^A})$: L^A is longer than $L^{B'}$ and both have the same low outcome x_1, so the property applies here.

From above, we know that $\mu_{L^{B'}} \leq x^*_{H^{B'}}$. But for all values $x \leq x^*_{H^{B'}}$ it holds that $H_{L^{B'}}(x) \leq C_{L^{B'}}(x^*_{H^{B'}})$, because $H_{L^{B'}}$ and $C_{L^{B'}}$ intersect in $x^*_{H^{B'}}$. Hence, $H_{L^{B'}}(\mu_{L^{B'}}) \leq C_{L^{B'}}(x^*_{H^{B'}})$, and since the lottery line $L^{B'}$ lies below $H_{L^{B'}}$, it holds that $L^{B'}(\mu_{L^{B'}}) \leq H_{L^{B'}}(\mu_{L^{B'}})$. Collecting the above inequalities, it holds that $L^{B'}(\mu_{L^{B'}}) < v(x_2)$.

Case 2: $v(x_1) < CARA^*(x_1)$, and L^A is shorter than L^B, i.e. $x''_{L^A} - x_1 \leq x_4 - x'_{L^B}$: Now shift lottery L^A to the right, resulting in a lottery $L^{A'}$ with high outcome x_4 and expected value $\mu_{L^{A'}}$. Construct the auxiliary lines $H_{L^{A'}}$ and $C_{L^{A'}}$, which intersect at $x^*_{H^{A'}}$. Analogous to the arguments in Case 1, it holds that $x^*_{H^{A'}} \leq \mu_{L^{A'}}$ and $x^*_{H^B} \geq \mu_{L^B}$. Also analogous to the above arguments, one can show that $C_{L^B}(x^*_{H^B}) \geq v(x_3)$, $C_{L^{A'}}(x^*_{H^{A'}}) \geq C_{L^B}(x^*_{H^B})$, $H_{L^{A'}}(\mu_{L^{A'}}) \geq C_{L^{A'}}(x^*_{H^{A'}})$, and $L^{A'}(x^*_{L^{A'}}) \geq H_{L^{A'}}(\mu_{L^{A'}})$.[27] Hence, $L^{A'}(\mu_{L^{A'}}) \geq v(x_3)$, and it follows that there exists a lottery $L^{\widetilde{A}}$ with $L^{\widetilde{A}}(\mu_{L^{\widetilde{A}}}) = v(x_3)$.

Case 3: $CARA^*(x_1) < v(x_1)$, and L^A is shorter than L^B, i.e. $x''_{L^A} - x_1 \leq x_4 - x'_{L^B}$: Again, shift L^A to the right, resulting in a lottery $L^{A'}$ with high outcome x_4 and expected value $\mu_{L^{A'}}$. Consider Figure 5.5, which depicts the auxiliary lines $C_{L^{A'}}$ and C_{L^B}. Since both lotteries L^B and $L^{A'}$ have the same high outcome, $C_{L^{A'}}$ lies above C_{L^B}. As in Case 2, we will argue that $L^{A'}(\mu_{L^{A'}}) \geq v(x_3)$. To this end, we first ask where the lottery line $L^{A'}$ lies, relative to $C_{L^{A'}}$. Observe that if the entire line $L^{A'}$ lies

[27] The following repeats the arguments, for this case. It holds that $C_{L^B}(\mu_{L^B}) = v(x_3)$ by construction. $x^*_{H^B} \geq \mu_{L^B}$ then implies that $C_{L^B}(x^*_{H^B}) \geq v(x_3)$, because C_{L^B} is increasing. Property (iv) of the lemma yields $C_{L^B}(x^*_{H^B}) \leq C_{L^{A'}}(x^*_{H^{A'}})$: L^B is longer than $L^{A'}$ and both have the same high outcome x_4. From above, we know that $x^*_{H^{A'}} \leq \mu_{L^{A'}}$, so that $H_{L^{A'}}(x) \geq C_{L^{A'}}(x^*_{H^{A'}})$. Hence, $H_{L^{A'}}(\mu_{L^{A'}}) \geq C_{L^{A'}}(x^*_{H^{A'}})$, and since the lottery line $L^{A'}$ lies above $H_{L^{A'}}$, it holds that $L^{A'}(\mu_{L^{A'}}) \geq H_{L^{A'}}(\mu_{L^{A'}})$.

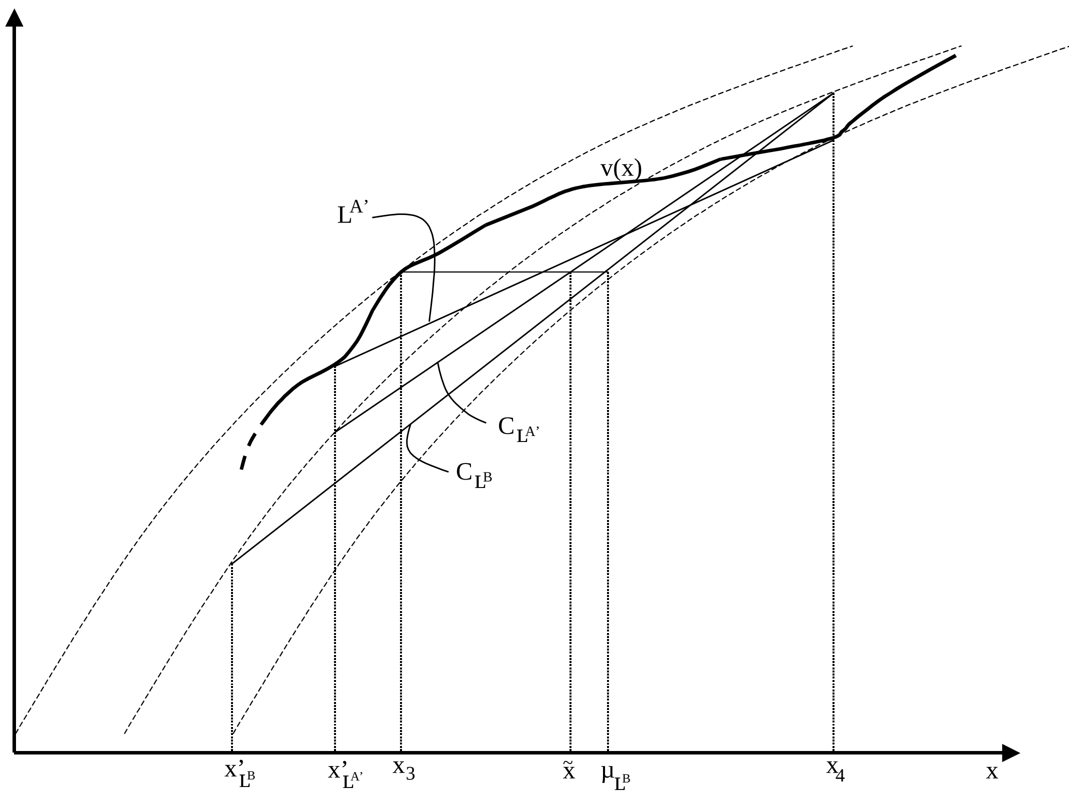

Figure 5.5: Location of C_{L^B}, $C_{L^{A'}}$ and $L^{A'}$ in Case 3 of Step 3.

above $v(x_3)$, then we are done — $L^{A'}(\mu_{L^{A'}}) \geq v(x_3)$ must be true in this case. Hence, assume to the contrary that the low payoff of $L^{A'}$ lies below $v(x_3)$, $L^{A'}(x'_{L^{A'}}) < v(x_3)$, as in the figure. $L^{A'}$ crosses the horizontal line at level $v(x_3)$ somewhere, because its high payoff is x_4, and v is increasing. Furthermore, we will show in the next paragraph that at v-level $v(x_3)$, $L^{A'}$ necessarily lies to the left of $C_{L^{A'}}$. Otherwise, there would exist a lottery $\widehat{L}$ that is shorter than L^B but has all of the properties of L^B: it would hold that $\pi^v_{\widehat{L}} = \pi^{CARA*}_{\widehat{L}} + K$, $x''(\widehat{L}) = x_4$, and $\widehat{L}(\mu_{\widehat{L}}) = v(x_3)$. This would contradict the definition of L^B as the shortest lottery with these properties.

The following shows existence of $\widehat{L}$, for this case. Consider the lottery $\widehat{\widehat{L}}$ with outcomes $x'_{L^{A'}}$ and x_4, and an expected value $\mu_{\widehat{\widehat{L}}}$ that is chosen such that $\widehat{\widehat{L}}(\mu_{\widehat{\widehat{L}}}) = v(x_3)$. If $L^{A'}$ lies to the right of $C_{L^{A'}}$ at v-level $v(x_3)$, then it follows that $\pi^v_{\widehat{\widehat{L}}} > \pi^{CARA*}_{\widehat{\widehat{L}}} + K$. Hence, we can increase the lower outcome of this lottery, keeping the expected value $\mu_{\widehat{\widehat{L}}}$ constant, until (by continuity) there results a lottery $\widehat{L}$ with $\pi^v_{\widehat{L}} = \pi^{CARA*}_{\widehat{L}} + K$, as claimed in the previous paragraph.

The observation that $L^{A'}$ lies to the left of $C_{L^{A'}}$ at v-level $v(x_3)$ implies that there exists a value $\widetilde{x}$ where $C_{L^{A'}}$ crosses the horizontal line at level $v(x_3)$, see Figure 5.5. (This is true because otherwise $C_{L^{A'}}$ would lie entirely above the horizontal line at $v(x_3)$ and could therefore not lie to the right of $L^{A'}$ at $v(x_3)$.) Also, it holds that $L^{A'}$ lies above $C_{L^{A'}}$ at $\widetilde{x}$, i.e. $L^{A'}(\widetilde{x}) \geq C_{L^{A'}}(\widetilde{x}) = v(x_3)$.

Next, construct the auxiliary lines J_{L^B} and $J_{L^{A'}}$, which intersect with C_{L^B} and $C_{L^{A'}}$, respectively, at $x^*_{J^B}$ and $x^*_{J^{A'}}$. Analogous to Case 1, it holds that $x^*_{J^B} \geq \mu_{L^B}$ and $x^*_{J^{A'}} \leq \mu_{L^{A'}}$. The first of these inequalities implies that $C_{L^B}(x^*_{J^B}) \geq C_{L^B}(\mu_{L^B})$ because C_{L^B} is increasing, and hence $C_{L^B}(x^*_{J^B}) \geq v(x_3)$.

Now, we have

$$C_{L^{A'}}(x^*_{J^{A'}}) \geq C_{L^B}(x^*_{J^B}) \geq v(x_3),$$

where the first inequality is implied by property (iv) of the lemma. Since $C_{L^{A'}}$ is increasing and reaches the level $v(x_3)$ at $\widetilde{x}$, this implies that $x^*_{J^{A'}} \geq \widetilde{x}$. Since $L^{A'}$ is increasing and lies above $C_{L^{A'}}$ at $\widetilde{x}$, it follows that $L^{A'}(x^*_{J^{A'}}) \geq L^{A'}(\widetilde{x}) \geq v(x_3)$. Finally, since $x^*_{J^{A'}} \leq \mu_{L^{A'}}$ (and $L^{A'}$ is increasing) it follows that $L^{A'}(\mu_{L^{A'}}) \geq v(x_3)$, the desired result.

Case 4: $CARA^*(x_1) < v(x_1)$, and L^A is longer than L^B, i.e. $x''_{L^A} - x_1 > x_4 - x'_{L^B}$. This is analogous to Case 3. Shift L^B to the left, so that the resulting lottery $L^{B'}$ has the lower outcome x_1. Then, analogous versions of all arguments for Case 3 apply, yielding $L^{B'}(\mu_{L^{B'}}) \leq v(x_2)$.[28]

Step 4: The previous steps show the result for the case that $CARA^*$ is concave. This can be used to show that if $CARA^*$ is convex, the result holds as well.

Suppose that $CARA^*$ is convex. Consider the functions that describe v and $CARA^*$, from upside-down: $\widehat{v} = -v(-x)$ and $\widehat{CARA}^*(x) = -CARA^*(-x)$. Both of these mirrored functions are increasing in x, on the interval $[-\overline{x}, -\underline{x}]$, and $\widehat{CARA}^*$ is concave. Also, for any CARA function $CARA : x \to CARA(x)$, it holds that $\widehat{CARA} : x \to -CARA(-x)$ is also a CARA function, and that the horizontal distance between v and $CARA$ on $[\underline{x}, \overline{x}]$ is equal to the horizontal distance between $\widehat{v}$ and

[28] The following sketches the steps: $C_{L^{B'}}$ lies above C_{L^A}. If the entire line $L^{B'}$ lies below $v(x_2)$, then $L^{B'}(\mu_{L^{B'}}) \leq v(x_2)$ holds. Assume to the contrary that $L^{B'}(x''_{L^{B'}}) > v(x_2)$. $L^{B'}$ crosses the horizontal line at level $v(x_2)$ to the right of $C_{L^{B'}}$, because otherwise the definition of L^A would be violated. Hence, there exists a value $\widetilde{\widetilde{x}}$ where $C_{L^{B'}}$ crosses the horizontal line at level $v(x_2)$, and $L^{B'}(\widetilde{\widetilde{x}}) \leq C_{L^{B'}}(\widetilde{\widetilde{x}}) = v(x_2)$. For the locations where the auxiliary lines J_{L^A} and $J_{L^{B'}}$ intersect with C_{L^A} and $C_{L^{B'}}$, respectively, it holds that $x^*_{J^A} \leq \mu_{L^A}$ and $x^*_{J^{B'}} \geq \mu_{L^{B'}}$. Therefore, and from property (iv) of the lemma, we have $C_{L^{B'}}(x^*_{J^{B'}}) \leq C_{L^A}(x^*_{J^A}) \leq v(x_2)$. Since $C_{L^{B'}}$ is increasing and reaches $v(x_2)$ at $\widetilde{\widetilde{x}}$, this implies that $x^*_{J^{B'}} \leq \widetilde{\widetilde{x}}$. Since $L^{B'}$ is increasing and $x^*_{J^{B'}} \geq \mu_{L^{B'}}$, it follows that $L^{B'}(\mu_{L^{B'}}) \leq L^{B'}(x^*_{J^{B'}}) \leq L^{B'}(\widetilde{\widetilde{x}}) \leq v(x_2)$.

$\widehat{CARA}$ on $[-\overline{x}, -\underline{x}]$. Hence, $\widehat{CARA}^*$ is the best-fitting CARA function for $\widehat{v}$ on the interval $[-\overline{x}, -\underline{x}]$, with distance K (because if there were a CARA function $\widetilde{CARA}(\cdot)$ with a lower horizontal distance, then it would also be true that $-\widetilde{CARA}(-x)$ is a CARA function with a lower horizontal distance from v on the interval $[\underline{x}, \overline{x}]$).

Hence, from the result of step 3, we know that there is a pair of shifted lotteries, with payoffs in $[-\overline{x}, -\underline{x}]$, such that $\widehat{v}$'s risk premium varies by at least K between the two lotteries. This pair of lottery choice tasks can be reflected around zero, by reversing the sign of all payoffs. Since $\widehat{v}$ is the mirror image of v, the resulting pair of lotteries has the property that v's risk premium varies by at least K between the two lotteries. Also, all payoffs of the reflected lotteries lie in $[\underline{x}, \overline{x}]$. Hence, we can apply Proposition 1 to obtain the result. ■

Lemma*:* Properties of CARA:

(For notation, see the first two paragraphs of the proof of Proposition 2.)

(i) Consider a set of three points $\{(x_1, v_1), (x_2, v_2), (x_3, v_3)\}$ that is strictly monotonic: $x_1 < x_2 < x_3$ and $v_1 < v_2 < v_3$. There exists a CARA function that connects the three points.

(ii) For any CARA function $CARA(\cdot)$ and any binary lottery L, π_L^{CARA} is constant across shifted versions of L, where $\triangle x$ is added to both payoffs as in Section 1. That is, $\pi_{\widetilde{L}}^{CARA} - \pi_L^{CARA} = 0$.

(iii) Let $(L, \widetilde{L})$ be a pair of shifted lotteries. Then $x^*_{H_{\widetilde{L}}} - x'_{\widetilde{L}}$ and $x^*_{J_{\widetilde{L}}} - x'_{\widetilde{L}}$ are both independent of the shift $\triangle x$. That is, the relative location of $x^*_{H_L}$ and $x^*_{J_L}$ does not change if L is shifted.

(iv) Let $\alpha, r > 0$, so that $CARA$ is strictly concave. Consider a pair of lotteries $(\widetilde{L}, \widetilde{\widetilde{L}})$ such that $\widetilde{\widetilde{L}}$ that is longer than $\widetilde{L}$, i.e. $x''_{\widetilde{\widetilde{L}}} - x'_{\widetilde{\widetilde{L}}} \geq x''_{\widetilde{L}} - x'_{\widetilde{L}}$. If both $\widetilde{\widetilde{L}}$ and $\widetilde{L}$ have the same low payoff ($x'_{\widetilde{\widetilde{L}}} = x'_{\widetilde{L}}$), then $C_L(x^*_{H_{\widetilde{\widetilde{L}}}}) \geq C_L(x^*_{H_{\widetilde{L}}})$ and $C_L(x^*_{J_{\widetilde{\widetilde{L}}}}) \geq C_L(x^*_{J_{\widetilde{L}}})$, i.e.

for the longer lottery, both values $C_L(x^*_{H_L})$ and $C_L(x^*_{J_L})$ are larger than for the shorter lottery. If both have the same high payoff ($x''_{\tilde{\tilde{L}}} = x''_{\tilde{L}}$), then the reverse inequalities are true, $C_L(x^*_{H_{\tilde{\tilde{L}}}}) \leq C_L(x^*_{H_{\tilde{L}}})$ and $C_L(x^*_{J_{\tilde{\tilde{L}}}}) \leq C_L(x^*_{J_{\tilde{L}}})$.

(v) Consider two CARA functions $CARA_1(x) = \beta_1 - \alpha_1 \exp(-r_1 x)$ and $CARA_2(x) = \beta_2 - \alpha_2 \exp(-r_2 x)$. Their horizontal distance $\eta(x, CARA_1, CARA_2)$ is unimodal: its derivative w.r.t. x changes its sign no more than once.

Proof: (i) For a CARA function with parameters α, β, r to connect (x_1, v_1) and (x_3, v_3), we need

$$v_1 = \beta - \alpha \exp(-r x_1)$$
$$v_3 = \beta - \alpha \exp(-r x_3).$$

First restrict attention to the case that the CARA function is strictly concave (i.e. $\alpha, r > 0$) and that $\beta > v_3$. Under these conditions, the above equations imply that α can be expressed as a function of β,

$$\begin{aligned} \alpha &= \exp\left(\frac{x_3}{x_3 - x_1} \ln(\beta - v_1) - \frac{x_1}{x_3 - x_1} \ln(\beta - v_3)\right) \\ &= \frac{(\beta - v_1)^{\frac{x_3}{x_3 - x_1}}}{(\beta - v_3)^{\frac{x_1}{x_3 - x_1}}}. \end{aligned}$$

Plugging the first of the expressions for α into the equation for v_1 yields an expression for r, as a function of β:

$$r = \frac{\ln(\beta - v_1) - \ln(\beta - v_3)}{x_3 - x_1}$$

Now vary β, and ask what values can the CARA function have at x_2, if it connects (x_1, v_1) and (x_3, v_3). Using the above expressions for α and r, it holds that

$$\begin{aligned} CARA(x_2) &= \beta - \alpha \exp(-r x_2) \\ &= \beta - (\beta - v_1)^{\frac{x_3 - x_2}{x_3 - x_1}} (\beta - v_3)^{\frac{x_2 - x_1}{x_3 - x_1}}. \end{aligned}$$

As β approaches v_3 from above, this expression approaches v_3. Also, as β approaches ∞, r approaches 0, and therefore the function becomes approximately linear between x_1 and x_3. Hence, by varying β, all points (x_2, v_2) that lie in the triangle between (x_1, v_3), (x_1, v_1) and (x_3, v_3) can be reached by a concave CARA function that also connects (x_1, v_1) and (x_3, v_3).

Now, consider the points that lie below the straight line between (x_1, v_1) and (x_3, v_3). The upside-down images of convex CARA functions are concave CARA functions ($-CARA(-x)$, like in step 4 of the proof of Proposition 2), so that the above argument applies: we can find a convex CARA function connecting $(x_1, v_1), (x_2, v_2)$, and (x_3, v_3) iff we can find a concave CARA function connecting $(-x_3, -v_3), (-x_2, -v_2)$, and $(-x_1, -v_1)$. This can be achieved by the above construction.

(ii) For a binary lottery L with lower payoff x'_L, express the higher payoff as $x''_L = x'_L + \delta''$ and the expected value as $\mu_L = x'_L + \widehat{\delta}$. By the definition of the certainty equivalent CE_L^{CARA},

$$\begin{aligned} CARA(CE_L^{CARA}) &= \Pr(x'_L)CARA(x'_L) + (1 - \Pr(x'_L))CARA(x''_L) \\ &= \frac{\delta'' - \widehat{\delta}}{\delta''}CARA(x'_L) + \frac{\widehat{\delta}}{\delta''}CARA(x'_L + \delta''). \end{aligned}$$

Using the functional form of $CARA$, and $\pi_L^{CARA} = \mu_L - CE_L^{CARA}$, this is equivalent to

$$\exp(-r(\widehat{\delta} - \pi_L^{CARA})) = \frac{\delta'' - \widehat{\delta}}{\delta''} + \frac{\widehat{\delta}}{\delta''}\exp(-r\delta''),$$

which implicitly defines π_L^{CARA} independent of the lottery's lower payoff x'_L. Since shifting the lottery to $\widetilde{L}$ is equivalent to a shift in x'_L while keeping δ'' and $\widehat{\delta}$ constant, it follows that $\pi_L^{CARA} = \pi_{\widetilde{L}}^{CARA}$.

(iii) It is straightforward to check that the relative location of $x^*_{H_L}$ is

$$\frac{x^*_{H_L} - x'_L}{x''_L - x'_L} = \frac{e^{Kr}}{e^{-r(x''_L - x'_L)} + e^{Kr}}.$$

Since $\triangle x$ is added to both x''_L and x'_L, the term $(x''_L - x'_L)$ is constant. Hence, $\frac{x^*_{H_L} - x'_L}{x''_L - x'_L}$ is constant.

Similarly, one can check that $x^*_{J_L}$ has the relative location

$$\frac{x^*_{J_L} - x'_L}{x''_L - x'_L} = -\frac{(1 - e^{-rK})}{(e^{-rK} - 1 - e^{rK - r(x''_L - x'_L)} + e^{-r(x''_L - x'_L)})}.$$

This expression, too, depends only on $(x''_L - x'_L)$, so it is constant across shifts of the lottery.

(iv) To calculate $C_L(x^*_{H_L})$, first solve $\frac{x^*_{H_L} - x'_L}{x''_L - x'_L} = \frac{e^{Kr}}{e^{-r(x''_L - x'_L)} + e^{Kr}}$ for $x^*_{H_L}$:

$$x^*_{H_L} = \frac{1}{e^{-r(x''_L - x'_L)} + e^{Kr}} \left(x'_L e^{-r(x''_L - x'_L)} + x''_L e^{Kr} \right)$$

The functional form of C_L is given by

$$C_L(x) = \beta - (\alpha \exp(-r x'_L))(1 + \frac{x - x'_L}{x''_L - x'_L}(\exp(-r(x''_L - x'_L)) - 1)).$$

Plugging $x^*_{H_L}$ into this expression yields

$$C_L(x^*_{H_L}) = \beta - \alpha \frac{e^{Kr - r x''_L} + e^{-r x''_L}}{e^{-r(x''_L - x'_L)} + e^{Kr}}.$$

Taking partial derivatives, one can see that the expression is increasing in both variables x'_L and x''_L, if $\alpha > 0$ and $r > 0$. Now, observe that if $\widetilde{\widetilde{L}}$ is longer than $\widetilde{L}$ and the two lotteries have the same lower payoff, then $\widetilde{\widetilde{L}}$ can be generated from $\widetilde{L}$ by increasing $x''_{\widetilde{L}}$, so the inequality $C_L(x^*_{H_{\widetilde{\widetilde{L}}}}) \geq C_L(x^*_{H_{\widetilde{L}}})$ follows from the monotonicity of $C_L(x^*_{H_L})$ in x''_L. If they have the same higher payoff, then $\widetilde{\widetilde{L}}$ can be generated from $\widetilde{L}$ by decreasing $x'_{\widetilde{L}}$, and $C_L(x^*_{H_{\widetilde{\widetilde{L}}}}) \leq C_L(x^*_{H_{\widetilde{L}}})$ follows from the monotonicity in x'_L.

Similarly, one can check that the intersection of C_L and J_L occurs at v-level

$$C_L(x^*_{J_L}) = \beta - \alpha e^{-r x'_L} + \alpha e^{-rK} \frac{e^{-r x'_L} - e^{-r x''_L}}{e^{-rK} + e^{-r(x''_L - x'_L)}},$$

and again, the partial derivatives with respect to x'_L and x''_L show that the expression is increasing in both x'_L and x''_L, if $\alpha > 0$ and $r > 0$. The same reasoning as in the previous paragraph then yields the results for $C_L(x^*_{J_{\tilde{\tilde{L}}}})$ and $C_L(x^*_{J_{\tilde{L}}})$.

(v) The horizontal distance between $CARA_1$ and $CARA_2$ is given by $\eta(x, CARA_1, CARA_2) = x - CARA_2^{-1}(CARA_1(x))$. The inverse of $CARA_2$ is given by $CARA_2^{-1}(v) = -\frac{1}{r_2}\ln(\frac{\beta_2 - v}{\alpha_2})$. Hence, $\eta(x, CARA_1, CARA_2) = x + \frac{1}{r_2}\ln(\frac{\beta_2 - \beta_1 + \alpha_1 \exp(-r_1 x)}{\alpha_2})$.

The statement is true because the second derivative of η w.r.t. x is

$$\frac{d^2}{dx^2}\eta(x, CARA_1, CARA_2) = \frac{(\beta_2 - \beta_1)(-r_1)^2 \alpha_1 \exp(-r_1 x)}{((\beta_2 - \beta_1)\alpha_1 \exp(-r_1 x))^2},$$

which has a constant sign. ■

Proof of Proposition 3. We show a more general version of the proposition, which covers the case that choices are representable by maximizing over m the expression

$$U(m) = \int u(x^I) dF(x^I|m) + \sum_i \int v(x_i) dL_i(x_i|m_i).$$

Here, u is the global evaluation of the distribution of total earnings, and v is the local or narrow evaluation of wealth changes in each bracket. Both are assumed to be increasing and twice continuously differentiable. This generalizes the assumptions made in Section 1, where $v(x) = (1 - \kappa)u(x)$. In the following version of the proposition, there are only two choices to be made ($I = 2$) but this is only a simplification. As we will see, the gambles offered in brackets 1 and 2 (to be constructed in the proof) can be made arbitrarily small for the result to hold, so that by the continuity of u and v these gambles would generally not affect the choices in any other existing brackets, unless these choices were knife-edge cases. Hence, fixing $I = 2$ is not restrictive — for larger I, we can construct the $I - 2$ other choices to be non-knife-edge.

Proposition 3′: Let u and v be strictly increasing and twice continuously differentiable, and suppose that v is not a member of the CARA family of functions. Then there is a pair of binary choices, each between a 50/50 lottery and a sure payment, where a decisionmaker who maximizes $U(\cdot)$ violates FOSD.

Proof: We will show that there exists a pair of 50/50 lottery choices of the following kind, violating FOSD:

Reject gamble 1: A 50/50 gamble between $x_1 - y_1$ and $x_1 + z$, with $y_1, z > 0$, versus x_1 for sure.

Accept gamble 2: A 50/50 gamble between $x_2 - y_2$ and $x_2 + z$, with $y_2, z > 0$, versus x_1 for sure. Importantly, $y_2 > y_1$.

Due to $y_2 > y_1$, the choice combination $(reject_1, accept_2)$ is FOS-dominated by $(reject_2, accept_1)$. We find appropriate values of x_1, x_2, y_1, y_2, z in two steps.

(i) If v is differentiable but not CARA, then there exist payoffs $x_1^* \neq x_2^*$ such that the Arrow-Pratt degree of risk aversion $r_{AP}(x) = -\frac{v''(x)}{v'(x)}$ is different at the two locations. This can be shown by showing the reverse statement: suppose that $r_{AP}(x)$ is constant and equal to some value r, at all x. Then we have a first-order linear differential equation for the function v': $v''(x) + rv'(x) = 0$, which has the general solution $v'(x) = \alpha \exp(-rx)$. Integrating yields the general solution $v(x) = \beta - \alpha \exp(-rx)$, i.e. that v is CARA.

(ii) Consider a fair 50/50 gamble in an arbitrary bracket i: the decisionmaker can either receive a certain payoff x_i or a 50/50 gamble between $x_i + \epsilon$ and $x_i - \epsilon$. (We will show that as ϵ gets small, the expression $\frac{v''(x)}{v'(x)}$ will be important, and that we would need different certainty equivalents to offset the gamble at x_1^* and x_2^*.) Fix the choices in the other bracket at m_{-i}, which will generate a distribution $L_{-i}(\cdot|m_{-i})$ over the sum of earnings in bracket $-i$. For any m_{-i}, there exists a payment $c_i(x_i, \epsilon, m_{-i})$ that would need to be added to x_i in order to make $x_i + c_i$ into a certainty equivalent, i.e. for the decisionmaker to be indifferent between accepting and rejecting the gamble. This payment $c_i(x_i, \epsilon, m_{-i})$ is implicitly defined by $U(x + c_i(x_i, \epsilon, m_{-i})|reject_i, m_{-i}) = U(x|accept_i, m_{-i})$. Using the global-plus-local form, this is

$$\begin{aligned}
&\int u(x + x_i + c_i(x_i, \epsilon, m_{-i}))dL_{-i}(x|m_{-i}) + v(x_i + c_i(x_i, \epsilon, m_{-i})) \\
&- \int \frac{1}{2}[u(x + x_i + \epsilon) + u(x + x_i - \epsilon)]dL_{-i}(x|m_{-i}) - \frac{1}{2}[v(x_i + \epsilon) + v(x_i - \epsilon)] \\
=& \, 0.
\end{aligned}$$

With fix x_i and m_{-i}, we can view this as $h(\epsilon, c_i(\epsilon)) = 0$, and apply the implicit function theorem to find that

$$\begin{aligned}
\frac{dc_i}{d\epsilon} &= -\frac{(\partial h/\partial \epsilon)}{(\partial h/\partial c_i)} \\
&= \frac{\int \frac{1}{2}u'(x + x_i + \epsilon) - \frac{1}{2}u'(x + x_i - \epsilon)dL_{-i}(x|m_{-i}) + \frac{1}{2}[v'(x_i + \epsilon) - v'(x_i - \epsilon)]}{\int u'(x + x_i + c_i(\epsilon))dL_{-i}(x|m_{-i}) + v'(x_i + c_i(\epsilon))}.
\end{aligned}$$

Now consider the pair of brackets labeled 1 and 2, and *simultaneously* offer such a 50/50 gamble to win or lose ϵ in both of them, starting from sure payoffs of x_1^* and x_2^* (taken from part (i) above). Now, since the decisionmaker can either accept or reject the gamble in the other bracket j (with $j \in \{1, 2\}$ and $j \neq i$), we have to consider two cases for m_{-i} in the expression for $\frac{dc_i}{d\epsilon}$. But differentiating both of these expressions once again with respect to ϵ, and taking the limit $\epsilon \to 0$, it is straightforward to see that for both possible choices $m_j \in \{accept_j, reject_j\}$, the second derivative of $c_i(\epsilon)$ has the same limit:

$$\lim_{\epsilon \to 0} \frac{d}{d\epsilon}\frac{dc_i}{d\epsilon}(accept_j) = \lim_{\epsilon \to 0} \frac{d}{d\epsilon}\frac{dc_i}{d\epsilon}(reject_j) = \frac{u''(x_i^* + x_j^*) + v''(x_i^*)}{u'(x_i^* + x_j^*) + v'(x_i^*)}$$

(The complete expressions for $\frac{d}{d\epsilon}(\frac{dc}{d\epsilon})$, evaluated separately at $m_j \in \{accept_j, reject_j\}$, are available upon request.)

Now compare this expression between $i = 1$ and $i = 2$. In the denominator and in the numerator, respectively, the same expressions $u''(x_i^* + x_j^*)$ and $u'(x_i^* + x_j^*)$ are added in both brackets. But $\frac{v''(x_1^*)}{v'(x_1^*)} \neq \frac{v''(x_2^*)}{v'(x_2^*)}$ by the result of (i). Hence, there is an $\bar{\epsilon}$ such that the second derivative of $c_i(x_i, \epsilon, m_{-i})$ is different between $i = 1$ and $i = 2$, for all $\epsilon < \bar{\epsilon}$. This holds irrespective of the behavior in bracket $j \neq i$, because the limit of $\frac{d}{d\epsilon}\frac{dc_i}{d\epsilon}$ is the same for both possible values of m_j. That is, even if the choice in bracket j changes arbitrarily often as ϵ approaches 0, it is still true that $\frac{d}{d\epsilon}\frac{dc_i}{d\epsilon}$ approaches a different limit in the two brackets. Hence, there is an $\bar{\bar{\epsilon}} < \bar{\epsilon}$ such that the certainty equivalent itself is different, i.e. $c_1(x_1^*, \bar{\bar{\epsilon}}) \neq c_2(x_2^*, \bar{\bar{\epsilon}})$, because if two functions have different continuous second derivatives over an interval, then they have a different value somewhere in that interval. W.l.o.g., assume $c(x_1^*, \bar{\bar{\epsilon}}) < c(x_2^*, \bar{\bar{\epsilon}})$. Then, by the definition of a certainty equivalent and by monotonicity of u and v, there exists a value c with $c(x_1^*, \bar{\bar{\epsilon}}) < c < c(x_2^*, \bar{\bar{\epsilon}})$, such that the decisionmaker strictly prefers to get $x_1^* + c$ for sure over the 50/50 gamble between $x_1^* + \bar{\bar{\epsilon}}$ and $x_1^* - \bar{\bar{\epsilon}}$, and he strictly prefers the 50/50 gamble between $x_2^* + \bar{\bar{\epsilon}}$ and $x_2^* - \bar{\bar{\epsilon}}$ over $x_2^* + c$ for sure. Hence, the desired payoffs exist: pick a δ that is small enough to not change these choices, and set $x_1 = x_1^* + c, x_2 = x_2^* + c, y_1 = \bar{\bar{\epsilon}} + c, y_2 = \bar{\bar{\epsilon}} + c + \delta, z = \bar{\bar{\epsilon}} - c$. ■