Common Ratio and Common Consequence Effects Arise from True Preferences

Carlos Alós-Ferrer, Ernst Fehr, Helga Fehr-Duda, and Michele Garagnani*

This Version: August 2024

Abstract

Recent contributions suggest that the empirical evidence for the common ratio effect could be explained as noise instead of underlying preferences under "common assumptions." We revisit this argument using a more general method which allows to unambiguously distinguish noise from preferences nonparametrically and with less stringent assumptions. The results are independent of the assumed behavioral model or how noise affects choices. Applying this method to new experimental data we show that there is a systematic preference for the common ratio and the common consequence effects which cannot be explained by noise.

JEL Classification: C91 · D81 · D91

1 Introduction

One of the most prominent violations of Expected Utility Theory (EUT) is the empirical observation that, when choosing between a sure but small amount and a larger but risky amount, most participants choose the safe alternative, but when all probabilities are scaled down by a common ratio, most people tend to choose the riskier option (Kahneman and Tversky, 1979). This phenomenon, called the "common ratio effect" (CR), is closely-related to the common consequence effect shown by Allais (1953), and has been widely replicated (see Ruggeri et al., 2020 for a large-scale replication and Blavatskyy et al., 2023 for a re-analysis of 143 previous studies). Together, both effects are often informally referred to as the *certainty effect*.

In a recent contribution, McGranaghan et al. (2024) provided evidence that the empirical support for the CR effect could amount to noise instead of a systematic deviation from EUT. In particular, what the authors show is that when paired valuations (i.e., elicited monetary valuations for each risky option) are used to study this phenomenon, there is scarce evidence for a CR preference in the data. However, when paired choices

*Carlos Alós-Ferrer: Lancaster University Management School (c.alosferrer@lancaster.ac.uk); Ernst Fehr: Department of Economics, University of Zurich (ernst.fehr@econ.uzh.ch); Helga Fehr-Duda: Department of Banking and Finance, University of Zurich (helga.fehr@bf.uzh.ch); Michele Garagnani: Centre for Brain, Mind, and Markets, Department of Finance, University of Melbourne (michele.garagnani@unimelb.edu.au)

are used, as in the majority of previous studies, the observed patterns of behavior do reproduce the CR effect, but they are also compatible with stochastic choice. The authors then suggest that the CR effect could be explained by participants having EUT preferences but noisy behavior, in line with random utility models (McFadden, 2001).

These results are in agreement with the long-standing literature which finds that choices are stochastic rather than deterministic (Hey and Orme, 1994; Agranov and Ortoleva, 2017; Alós-Ferrer and Garagnani, 2021, 2022a,b). This limits the possible inferences from observed behavior in experiments (see Alós-Ferrer et al., 2021, for a recent discussion on inference from stochastic choices). In particular, several previous papers have shown that random utility models based on EUT could generate a CR effect (Ballinger and Wilcox, 1997; Loomes, 2005; Hey, 2005; Blavatskyy, 2007, 2010; Bhatia and Loomes, 2017). Contributing to this literature, McGranaghan et al. (2024) show that, under certain specific assumptions on the underlying behavioral model and the structure of noise, the CR effect could appear at the aggregate level using paired choices even when subjects have EUT preferences but choice is noisy. They then reproduce standard experiments which involved paired choices using paired valuations instead, which, assuming either additive or symmetric noise in the determination of indifference points, should lead to unbiased tests for the CR effect. Specifically, assuming that reported valuations arise from additive noise centered around the true valuations yields an unbiased means test (McGranaghan et al., 2024, Proposition 2(i)). However, as McGranaghan et al. (2024) discuss, standard models of additive *utility noise* (additive random utility models; RUMs) do not generate additive noise for elicited valuations. Assuming symmetric valuation noise (but allowing for a non-additive structure) yields an unbiased sign test (McGranaghan et al., 2024, Proposition 2(ii)).[1] Using those tests, McGranaghan et al. (2024) conclude that there is scarce evidence of a systematic preference for the CR effect at the aggregate level, although they document substantial heterogeneity in behavior.

In this work, we point out that a different way to tackle this problem is already available. Essentially, McGranaghan et al. (2024) argue that previous evidence for the CR effect might be due to noise, and we agree with their conclusions. However, the picture drawn by those authors is incomplete because it is now possible to *distinguish preferences from noise non-parametrically and at the individual level*, hence isolating anomalies as the CR effect independently of noise. In the analysis below, we do precisely that. We apply the method proposed by Alós-Ferrer et al. (2021) (TWT; see also Alós-Ferrer and Garagnani, 2024) to new experimental evidence and reveal genuine, individual preferences compatible with the common ratio and the common consequence effects, which cannot be rationalized as resulting from noise or the specific model of behavior

[1] While the authors make a good case for symmetric valuation noise, it should be noted that in the related literature on the preference reversal phenomenon (e.g., Lichtenstein and Slovic, 1971; Grether and Plott, 1979), valuations are typically found to be biased compared to choices (e.g., Bateman et al., 2007). Hence, an argument could be made as to whether valuation noise is actually symmetric and whether valuations can actually yield unbiased tests. This is, however, not our point.

one assumes. We hence conclude that, while the critique of McGranaghan et al. (2024) regarding previous evidence is correct, new evidence and techniques are already available which show that CR is an actual behavioral anomaly contradicting EUT which cannot be explained away by noise alone.

The reason we are able to go beyond previous analyses, and in particular the conclusions of McGranaghan et al. (2024), is that the TWT method put forward in Alós-Ferrer et al. (2021), and recently expanded in Alós-Ferrer et al. (2024), reveals preferences not for any particular combination of assumed utilities and distribution of noise, but simultaneously for all such combinations in a wide family of models. The results we rely on are not based on preference estimation, but rather on true, nonparametric preference revelation techniques independently of any assumptions on utility functions or noise distributions. The price to pay for this generality is twofold. First, one needs repeated observations, including response times, for each binary choice of interest (a, b) and every decision maker. Second, the results are based on sufficient conditions (see Section 2 below). In words, if the data for a choice pair (a, b) from a particular decision maker fulfills a certain condition, the formal-analytical results imply that *any* model of behavior and *any* distribution of behavioral noise which rationalize (reproduce) the actual data must be such that the utility of a exceeds the utility of b. Hence, if the sufficient conditions hold, a preference for a over b is truly revealed. Preferences are hence disentangled from noise. We refer the reader to Alós-Ferrer et al. (2021) and Alós-Ferrer et al. (2024) for a more detailed discussion of the method.

In other words, the TWT method is agnostic on the shape of the utility and the properties of the noise, and reveals preferences not for a particular shape of utility and distribution of noise, but simultaneously for all combinations of utility and noise that fit the data. Since the approach relies on sufficient conditions, however, no conclusion is warranted when those fail. Thus, the approach is well-suited to show the existence of behavioral preference anomalies as the CR effect,[2] but it delivers *lower bounds* on the extent of those violations in a given population.

For our purposes, what the TWT method allows us to do in the context of the CR effect is to identify when inference from the data can unequivocally tell us whether CR patterns can be attributed to systematic preferences, instead of just explained by noise. This is the key difference to previous methods (e.g., McGranaghan et al., 2024; Ballinger and Wilcox, 1997; Loomes, 2005; Hey, 2005; Blavatskyy, 2007, 2010; Bhatia and Loomes, 2017; Vieider, 2018), as we do not need to impose particular assumptions on behavior or the distribution of noise. Our preference revelation results hold for a very general class of models comprehending, but not limited to, random utility models (McFadden, 2001), random parameter models (Loomes and Sugden, 1998; Apesteguía and Ballester, 2018), and evidence accumulation models as the Drift-Diffusion Model (Ratcliff, 1978; Fudenberg et al., 2018; Baldassi et al., 2020). Alós-Ferrer et al. (2024) has shown that

[2] Alós-Ferrer et al. (2024) uses this approach to show transitivity violations in risky choices, addressing a log debate where such violations have frequently been argued to be due to noise.

the results also cover non-transitive models, as those based on (Generalized) Regret theory (Loomes and Sugden, 1982, 1987), Skew-Symmetric-Bilinear utility (Fishburn, 1984a,b), and Salience theory (Bordalo et al., 2012), among others.

We collected data on binary, risky choices allowing for both the common ratio and the common consequence effects, including a substantial number of repetitions for each choice, and measuring response times. We then applied the TWT method the reveal preferences, disentangling them from any assumptions on noise. The sufficient conditions bite very often, revealing preferences in almost 90% of the cases. This allows us to show that the common ratio and common consequence effects are present quite often in the data, while at the same time showing that no model of noise would be able to explain those anomalies while maintaining the EUT assumption. Hence, while we agree with McGranaghan et al. (2024) that previous data did not satisfactorily establish the CR (and common consequence) effects, we conclude that more recent techniques, coupled with our new data and analyses, are able to do so. Hence, the challenge to the EUT represented by the certainty effect remains.

2 Nonparametric Preference Revelation

The general framework we work with was proposed by Alós-Ferrer et al. (2021) (TWT) and further expanded in Alós-Ferrer et al. (2024). An application of related results to predictions out of sample has been presented in Alós-Ferrer and Garagnani (2024). The framework is based on a generalization of RUMs (McFadden, 1974, 2001), where an agent is assumed to have an underlying utility function u, but to be affected by random utility shocks. In this generalization, error terms are modeled directly for utility differences, i.e. the realized utility difference given a binary choice $\{x, y\}$ is $u(x) - u(y) + \varepsilon_{x,y}$ for a mean-zero random variable $\varepsilon_{x,y}$ and hence the probability of x being chosen when y is also available is

$$p(x, y) = \text{Prob}(\varepsilon_{x,y} > u(y) - u(x))$$

where tie-breaking conventions are irrelevant for continuously-distributed errors. In classical, additive RUMs, error terms are alternative-specific, i.e. $\varepsilon_{x,y} = \varepsilon_x - \varepsilon_y$ for appropriate ε_x, ε_y. Under specific assumptions on the distributions of the error terms, one obtains particular models, as the celebrated logit choice (Luce, 1959) or the classical probit choice (Thurstone, 1927). Pair-specific error terms substantially generalize the class of models, e.g. allowing to encompass trembling-hand models (Loomes et al., 2002), where error rates are inherently pair-specific, or random parameter models (Loomes and Sugden, 1998; Apesteguía and Ballester, 2018), where noise affects a parameter in a class of utility functions, one of which is randomly sampled for each choice (see Alós-Ferrer et al., 2024, for details).

The previous literature (Ballinger and Wilcox, 1997; Loomes, 2005; Hey, 2005; Blavatskyy, 2007, 2010; Bhatia and Loomes, 2017) has argued that classical, additive RUMs can generate the common ratio effect in paired choice tasks simply due to the structure of noise. The argument is summarized in (McGranaghan et al., 2024, Section IB). Essentially, when participants choose between a sure but small amount and a larger but risky amount, even if they prefer the former, there is a probability that they will choose the latter due to noise. If the difference in utilities ("strength of preference") is large enough relative to the noise, the probability of a mistake will be positive but not large. When the winning probabilities are scaled down to create the second pair in a common ratio experiment, utility differences become much smaller, and the probability of an error becomes much larger (assuming that the distribution of noise remains largely unchanged across choice pairs). This makes reversals of the type observed in the CR effect more likely than the opposite ones.

This argument rests on a specific relation between errors and strength of preference which reflects a robust empirical regularity. Specifically, widespread evidence from psychology and neuroscience shows that error rates are smaller for easier choice problems, e.g. discrimination tasks where stimuli are more similar, than for harder ones (e.g. Cattell, 1893; Dashiell, 1937; Mosteller and Nogee, 1951; Laming, 1985; Klein, 2001; Wichmann and Hill, 2001). Random utility models reflect this empirical regularity.[3] McGranaghan et al. (2024, Section IVB) show that strength of preference (in the form of a proxy they call "distance to indifference") influences whether a CR effect appears in paired choice tasks.

The TWT approach that we use builds upon the same and related regularities. In addition to the effects on error rates just described, the TWT results are made possible by using response times and incorporating a further, robust empirical regularity from psychology and neuroscience. Specifically, response times are shorter for easier choice problems when stimuli are similar to each other) than for harder ones (e.g. Dashiell, 1937; Moyer and Landauer, 1967; Moyer and Bayer, 1976; Laming, 1985; Dehaene et al., 1990; Klein, 2001; Wichmann and Hill, 2001). This effect has also been demonstrated in economic choices (e.g., Moffatt, 2005; Chabris et al., 2009; Konovalov and Krajbich, 2019; Alós-Ferrer and Garagnani, 2022a,b). The intuition for these regularities (both for error rates and for response times) is that computational processes in the human brain are inherently noisy and hence more likely to produce errors when the differences in decision values are small (Glimcher et al., 2005; Shadlen and Kiani, 2013; Shadlen and Shohamy, 2016). This is captured in sequential-sampling models as the Drift-Diffusion Model (Ratcliff, 1978; Fudenberg et al., 2018; Webb, 2019; Baldassi et al., 2020), but the properties and the empirical regularities they reflect go beyond any specific model. We refer the reader to Alós-Ferrer et al. (2021) for a more detailed discussion of these well-known properties.

Theorem 1 in Alós-Ferrer et al. (2021), expanded by Theorem 1 of Alós-Ferrer et al. (2024), identifies a sufficient conditions on the distributions of response times conditional

[3]Indeed, RUMs can be traced back to psychological work motivated by this empirical regularity, starting with Thurstone (1927).

on each possible choice (x or y for a given pair (x, y)) which ensure that any model within the class described above that fits the data (in terms of choice frequencies and distributions of response times) reveals a preference for x over y, in the sense that $u(x) > u(y)$ for the any underlying u. Specifically, the condition weights choice frequencies and response times. For a given decision maker, let $p(x, y)$ be the proportion of x choices out of all repetitions of the choice pair $\{x, y\}$, and let $F(x, y)$ be the cumulative distribution of response times for those choices, conditional on x being chosen (and analogously for $p(y, x)$ and $F(y, x)$). A preference for x over y is revealed if

$$p(x, y)F(x, y)(t) - p(y, x)F(y, x)(t) \geq 0 \qquad \text{(TWT)}$$

for all $t \geq 0$, and a strict preference is revealed if, additionally, the inequality is strict for some t.

Thus, if condition (TWT) is fulfilled, it revels that x is preferred to y another for *any* utility function and *any* distribution of the error term that the analyst might consider, and hence the results are completely non-parametric and independent of functional forms or assumptions on the noise. If the condition is not fulfilled, preferences are not revealed, and no conclusion can be drawn from the data; in particular, the data might be explained with models such that $u(x) > u(y)$ and also by different models such that $u(y) > u(x)$, differing on their assumptions on the noise terms.

In the next section, we describe an experiment which collected data on the common ratio and common consequence effects, including choice repetitions and measuring response times, with the explicit purpose of allowing for an application of the TWT method. In Section 4 below, we report the analysis of this data. That is, the objective is to reveal preferences for particular combinations of risky choice pairs, where some pairs are built from others to detect the behavioral anomalies. For the CR effect, this means constructing "normal pairs" with two risky alternatives and associated "certainty pairs" where all probabilities of non-zero outcomes are scaled up, making one of the alternatives a sure option. The CR effect amounts to a preference reversal among the two related pairs. In our terms, we aim to show reversals in the *revealed* preferences, which are hence independent of any assumptions on the noise.

3 The Experiment

We conducted a new experiment to test for the presence of the CR effect in revealed preferences. We further included the common consequence (CC) effect, since it is a closely-related, prominent violation of EUT frequently studied in the same context (Kahneman and Tversky, 1979). Together, CR and CC effects constitute the standard demonstrations of the certainty effect.

A total of 115 subjects (70 females; median age 24, ranging from 20 to 30) participated in 20 experimental sessions with up to 7 subjects each. Participants were recruited from the student population of the University of Zurich, excluding students majoring in psychology and economics, as well as subjects who had already participated in similar experiments involving lottery choices. The experiment was conducted at the laboratory for experimental and behavioral economics of the University of Zurich and was programmed in PsychoPy.

The experiment included 18 pairs of risky options (lotteries), each of them repeated 11 times for a total of 198 binary choices per participant. Repetitions of the same choice were non-consecutive, with at least two other choices between any two repetitions of the same lottery pair. For each pair, participants were asked to choose one of the two lotteries. All lotteries were presented in as tables (see Appendix D for an sample decision screen). The screen position of the lotteries (left or right) was counterbalanced across repetitions and subjects. To control for order effects, each subject was randomly assigned to one of four different sequences of lottery pairs. There was no feedback during the course of the experiment, that is, subjects did not receive any information regarding their earnings until the very end of the experiment. All decisions were made independently and at a subject's individual pace. Following the main part, subjects were asked to complete a short questionnaire eliciting gender and age.

Lottery choice was incentivized using the random incentive mechanism (Cubitt et al., 1998). To determine a participant's payoff, one of the 198 lottery pairs was randomly selected, the lottery actually chosen by the participant was played out, and the participant received the realized outcome as payoff. In addition, participants received a show-up fee of 5 CHF for an average total remuneration of 21.78 CHF. Sessions lasted between 45 and 60 minutes including instructions and payment.

The set of lottery pairs used in the experiment included three CR and three CC double pairs, each repeated eleven times. Three further lottery pairs were included and will be described further below. For CR, the construction is as follows.

Given monetary outcomes X, Y with $X > Y > 0$, a probability p with $0 < p < 1$, and an integer $k \geq 1$, define the pair of lotteries $S_{CR}(k)$ and $R_{CR}(k)$ by

$$S_{CR}(k) = \begin{cases} Y \text{ with prob.} & 1/k \\ 0 & \text{otherwise,} \end{cases} \qquad R_{CR}(k) = \begin{cases} X \text{ with prob.} & p/k \\ 0 & \text{otherwise.} \end{cases}$$

Note that $R_{CR}(k)$ is always riskier than $S_{CR}(k)$, hence the latter is "safer."

The CR effect involves a preference reversal between two related pairs of lotteries. The *certainty pair* is given by setting $k = 1$ above. Then $S_C = S_{CR}(1)$ is a sure gamble, while $R_C = R_{CR}(1)$ has probability p of a non-zero outcome. The *normal pair* is derived from the certainty pair by scaling down the probabilities of non-zero outcomes by the ratio $1/k$. That is, given some $k > 1$, one defines $S_N = S_{CR}(k)$ and $R_N = R_{CR}(k)$, none of which involve a certain outcome. For instance, letting $X = 21$, $Y = 16$, and $p = 0.8$, we obtain the certainty pair with $S_C = (1, 16)$ and $R_C = (0.80, 21)$. Dividing the winning probabilities by $k = 4$, we obtain the normal pair with $S_N = (0.25, 16)$ and

$R_N = (0.20, 21)$. Under EUT, any decision maker who chooses S_C in the certainty pair must also choose S_N in the normal pair (and analogously for R_C, R_N). Empirically, a large number of people choose R_N and S_C, hence contradicting EUT.

For the CC effect, the construction is as follows. Given X, Y, p, q with $X > Y > 0$ and $0 < p < 1$, $0 < r \leq q < 1$, $p + q < 1$, define the pair of lotteries $S_{CC}(q)$ and $R_{CC}(q)$ by

$$S_{CC}(r) = \begin{cases} Y \text{ with prob.} & 1 - r \\ 0 & \text{otherwise,} \end{cases} \qquad R_{CC}(r) = \begin{cases} X \text{ with prob.} & p \\ Y \text{ with prob.} & q - r \\ 0 & \text{otherwise.} \end{cases}$$

Analogously to the CR construction above, $R_{CC}(r)$ is always riskier than $S_{CC}(r)$. The CC effect also amounts to a preference reversal between two related pairs of lotteries. The *certainty pair* is given by setting $r = 0$ above. Then $S_C = S_{CC}(0)$ offers a sure gain, while $R_C = R_{CC}(0)$ has a probability $p + q$ of non-zero outcomes. The *normal pair* is derived from the certainty pair by transferring a probability q from the common consequence Y to the consequence zero. That is, one sets $r = q$ to obtain $S_N = S_{CC}(q)$ and $R_N = R_{CC}(q)$, none of which involves a certain outcome. For example, letting $X = 23$, $Y = 17$, $p = 0.15$, and $q = 0.8$, we obtain the certainty pair with $S_C = (1, 17)$ and $R_C = (0.15, 23; 0.8, 17)$. Transferring a probability $r = 0.8$ from 17 to zero, we obtain the normal pair $S_N(0.8) = (0.2, 23)$ and $R_N = (0.15, 23)$. As in the case of the CR effect, any decision maker who chooses S_C in the certainty pair must also choose S_N in the normal pair (and analogously for R_C, R_N). Empirically, however, many people choose R_N and S_C, hence contradicting EUT. This is the well-known *Allais paradox* (Allais, 1953).

The experiment further included *middle pairs* bridging the gap between certainty and normal pairs. For CR, the middle pairs were constructed setting k' with $1 < k' < k$. For CC, the middle pairs were constructed using r with $0 < r < q$. For the sake of brevity, the analysis of the middle pairs is relegated to Appendix C.

4 Results

4.1 Choice proportions

Previous analyses of the certainty effect use data where each participant makes each choice once. In contrast, each of our participants made each choice 11 times, which enables application of the TWT preference revelation result. In line with previous results in the stochastic choice literature, participants were often inconsistent in their choices during the experiment. In particular, decisions over the 12 lottery pairs used to study the CR and CC effects were deterministic (choosing the same option across the 11 repetitions of the same pair) only in 42.17% of the cases. Every participant in the experiment was inconsistent at least once across those pairs.

Before we proceed to our main analysis, we briefly comment on choice frequencies mirroring previous analyses in the literature. Say that a participant chose an option for a given pair (S, R) if that participant chose that option more than 50% of the time across all the 11 repetitions of the pair. Thus, we consider that a participant displayed a specific pattern in a double pair, e.g., (S_C, R_N), if across the repetitions, she chose option R for more than half of the repetitions of the normal pair and option S for more than half of the repetitions of the certainty pair. For the sake of brevity, we report here average results across the three CR double pairs and the three CC double pairs. Appendix B provide separate analyses and figures for each double pair in the experiment. The conclusions remain unchanged.

The choice data from our experiment replicates standard patterns from the literature. That is, both in the CR and in the CC pairs the majority of subjects chose the risky alternative in normal pairs (CR 58.26%, CC 51.30%), but then were more likely to choose the sure alternative in the certainty pairs (CR 83.77%, CC 65.51%). The differences are significant according to McNemar's tests (which compare paired proportions) for all three CR double pairs and two of the three CC pairs (see Appendix B for the tests).[4] These results are similar in magnitude to those reported in a large reanalysis of existing data (Blavatskyy et al., 2023).

However, as argued by McGranaghan et al. (2024), these choice patterns could be due to behavioral noise, and not reflect a systematic preference anomaly. Hence, in the following subsection we apply the TWT method to reveal preferences independently of noise.

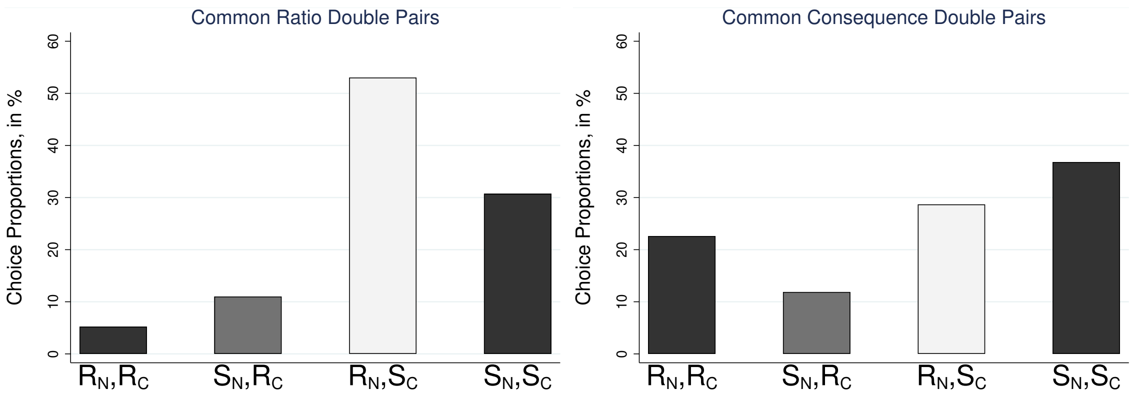

Figure 1: Proportions of participants favoring each possible choice pattern, separately for common ratio (left) and common consequence (right) double pairs.

[4]Results are unchanged if we compare the individual choice proportions (out of the 11 repetitions) across normal and certainty pairs with Wilcoxon Signed-Rank tests.

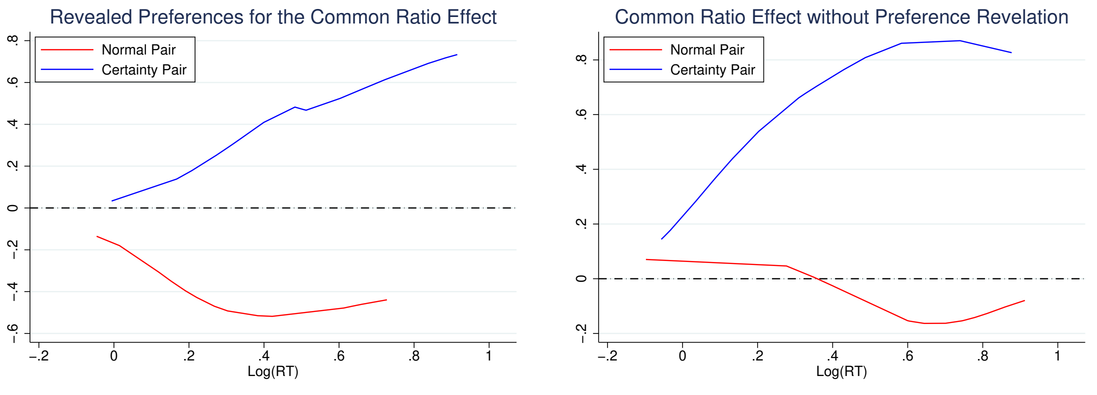

Figure 2: Revealed preferences according to the TWT method which reflect a genuine common ration effect (left). Common ratio effect without preference revelation (right).

4.2 Genuine Certainty Effects

We now use the TWT method to reveal preferences independently from noise or the assumed behavioral model. To do so, we closely follow the procedure described in Alós-Ferrer et al. (2021). In particular, the kernel density estimates $F(x, y)$, which are needed to check condition (TWT), use an Epanechnikov kernel with optimally chosen non-adaptive bandwidth and were performed in *Stata* using the *akdensity* function. We estimated the distribution of log-transformed response times to avoid boundary problems. We refer the reader to Alós-Ferrer et al. (2021), Alós-Ferrer and Garagnani (2024), or Alós-Ferrer et al. (2024) for further details.

Figure 2 gives an illustration of how the TWT method is applied to data in our experiment. For each double pair $\{(S_N, R_N), (S_C, R_C)\}$, we aim to compute how many individuals show a preference reversal (reflecting either the CR or the CC effect) in their *revealed* preferences. Thus, for each individual and each choice pair $\{S, R\}$, we compute the empirical choice frequencies $p(S, R)$ and $p(R, S)$ and obtain the conditional response time distribution functions $F(S, R)$ and $F(R, S)$. Figure 2 plots the function in condition (TWT), i.e. $p(S, R) \cdot F(S, R)(t) - p(R, S) \cdot F(R, S)(t)$ as a function of (log-transformed) response times. The figure is hence equivalent to a graphical test that inequality (TWT) holds for all response times. Independently for each choice pair, if the line obtained in this way does not cross the 0, then the condition for revealing a preference is fulfilled.

Figure 2 depicts two actual examples corresponding to the two most-interesting cases. In each panel, the participant's data for a given double pair is represented with one line referring to the normal pair and the other line representing the certainty pair. Hence, lines above the 0 indicate a revealed preference for the safer alternative S.

The left-hand panel illustrates the data from participant nr. 9 in the experiment and the CR double-pair in Set I (see Appendix A). Neither line crosses the 0. The one for the certainty pair (S_C, R_C) is fully above the horizontal line, hence revealing a (strict) preference for S_C. The one for the normal pair (S_N, R_N) is fully below the horizontal line, hence revealing a (strict) preference for R_N for this participant. Taken together, both results constitute an example of the common ratio effect (a preference for S_C in the certainty pair but a preference for R_N in the normal pair) *in terms of revealed preferences*. This preference pattern cannot be explained by noise, and hence a genuine common ratio effect is revealed for this participant.

The right-hand panel of Figure 2 illustrates the data from a different participant (nr. 31 in the experiment) and the same double pair. In this case, the data for the certainty pair also reveals a preference for S_C. For the normal pair, this participant chose R_N most of the time (6 out of 11 times). However, this is *not* evidence of the CR effect, because the function given in (TWT) crosses the horizontal line and hence preference revelation does not obtain. Hence, this data is not evidence for the CR effect, and might correspond precisely to the kind of noisy evidence that McGranaghan et al. (2024) warn about.

We conducted the analysis described above for each of the 115 experimental participants and each of the three CR double pairs and the three CC double pairs. In each case, we determined whether the preferences were revealed or not. Instances where preferences in one of the two choice pairs were not revealed cannot contradict EUT, simply because preferences are not available. This might include cases where an analysis of choice frequencies suggests an effect, but might indeed simply hide noisy choices.

In our dataset, the TWT method overwhelmingly revealed preferences. For each individual, we computed how many times preferences were revealed across the 12 pairs used to identify CR and CC effects. The average proportion of revealed preferences across individuals was 89.49% (median 91.67%, SD 10.88, min 58.33%, max 100.00%). The proportion of revealed preferences was slightly higher for certainty pairs (92.90%) compared to normal pairs (86.09%; WSR, $N = 115$, $z = 3.555$, $p = 0.0003$). Although this is beyond the scope of the present contribution, this difference might suggest that there was less noise in the certainty pairs compared to the normal pairs. One possible interpretation of this finding is that certainty pairs are less complex, as they involve a lower number of outcomes, and hence their subjective value is easier to compute, which leads to less stochastic choices (Murawski and Bossaerts, 2016; Oprea, 2020).

Whenever preferences were revealed for both pairs in a double pair (certainty and associated normal pair), we recorded what was revealed. A revelation of preferences for (S_N, S_C) or (R_N, R_C) does not contradict EUT, while a revelation of preferences for (S_N, R_C) or (R_N, S_C) would. The last one corresponds to the *certainty effect*, in the form of the CR effect or the CC effect, depending on the nature of the double pair. By the very nature of the TWT method, we then know that this pattern of revealed preferences cannot be explained by noise. Following the terminology of McGranaghan et al. (2024), if we observe a revealed (R_N, S_C) pattern, as predicted by CR or the CC effects, we then reveal a genuine certainty effect preference (either CR or CC).

For the CR effect, the results of our revealed preference analysis show that the majority of participants genuinely preferred the risky alternative in normal pairs (63.06%, averaged across the three normal pairs; conditional on preferences being revealed), but a large

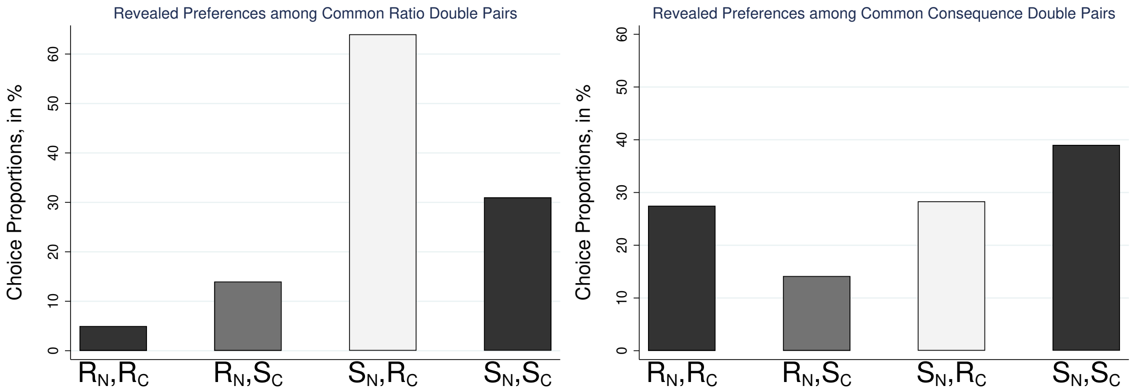

Figure 3: Proportion of participants with each possible pattern of revealed preferences, separately for common ratio (left) and common consequence (right) double pairs. The data is conditional on preference revelation and then averaged across the three double pairs used for each effect.

majority also genuinely preferred the sure alternative in the certainty pairs (84.68%). For the CC effect, 50.00% of participants preferred the risky alternative more frequently in normal pairs (actually a majority for two of the three CC normal pairs), but 60.85% preferred the sure alternative more often in the certainty pairs.

We hence restrict the analysis to double pairs where preferences are revealed for both pairs according to the TWT method. Consider the CR effect first. For all three double pairs, the TWT method uncovered a larger proportion of participants with a revealed preference for R in the normal pair compared to the certainty pair. The differences are highly significant according to McNemar's tests comparing paired observations (Set I, normal pair 60.64% vs. certainty pair 8.51%, $N = 94$, $\chi^2 = 39.36$, $p < 0.001$; Set II, 65.26% vs. 11.58%, $N = 95$, $\chi^2 = 38.82$, $p < 0.001$; Set III, 51.55% vs. 14.43%, $N = 97$, $\chi^2 = 24.00$, $p < 0.001$). Hence, we conclude that the CR effect is present in our data in terms of revealed preferences and cannot be accounted for by noise in the choice process.

We obtain a similar conclusion for the CC effect. For two of the three double pairs used to study this effect, again the TWT method uncovered a significantly larger proportion of participants with a revealed preference for R in the normal pair compared to the certainty pair (Set IV, 52.38% vs. 36.90%, $N = 84$, $\chi^2 = 39.36$, $p = 0.015$; Set VI, 52.69% vs. 34.41%, $N = 93$, $\chi^2 = 7.81$, $p = 0.008$). For the third pair, the CC effect was already absent when using choice data only, and there were also no significant differences in terms of revealed preferences (Set V, 31.11% vs. 33.33%, $N = 90$, $\chi^2 = 0.15$, $p = 0.845$).

Figure 3 illustrates these results. This figure plots the proportion of people such that a particular combination of preferences were revealed, over all the participants such that preferences were revealed for both pairs in the double pair. In particular, we insist that the proportions illustrated in the figure, in contrast to Figure 1, are *not* choice frequencies. The Figure illustrates that revealed preference reversals of the type

(R_N, S_C) are substantial and much more frequent than reversals of the opposite nature, (S_N, R_C). For ease of presentation, data is averaged across the three double pairs used in the experiment for each effect; Figure B.2 reports separate figures for each double pair.

We hence conclude that choices in our experiment reveal patterns of genuine preferences (independently of any assumptions on possible behavioral noise or underlying utilities) in alignment with the certainty effect, both in the form of common ratio and common consequence anomalies. Hence, while McGranaghan et al. (2024) were right to point out that previous evidence (based on choice frequencies only) did not conclusively demonstrate those effects, the techniques made available in Alós-Ferrer et al. (2021), Alós-Ferrer and Garagnani (2024), and Alós-Ferrer et al. (2024) do provide precisely such a demonstration.

5 Conclusions

Recent contributions as McGranaghan et al. (2024) point out that previous experimental evidence based on choice frequencies fails to convincingly demonstrate the existence of the common ratio effect as a systematic pattern of behavior. Using paired valuation and paired choice tasks, they provide evidence that the stochastic nature of choices alone could in principle explain this phenomenon.

In this work we add to this evidence using a recently-available method (TWT) which allows to unambiguously and nonparametrically disentangle preferences from inherent choice noise, independently of assumptions on underlying behavioral models or noise distributions. Applying this method to new experimental data, we show definite evidence for the presence of a systematic "preference" for the common ratio effect and the related common consequence effect (Allais paradox), which cannot be explained by noise. As a consequence, we show that, in spite of the reasonable doubts raised by McGranaghan et al. (2024) and others, violations of EUT are systematic and do not depend on the specific assumptions on utilities or noise.

Our evidence does not contradict the results of McGranaghan et al. (2024). We both start from the same observation, namely that noise could explain previous evidence for the CR effect in paired choice tasks. McGranaghan et al. (2024) move on to propose an alternative analysis through paired valuation tasks and find no evidence for the CR effect. We propose an alternative method relying on preference revelation using response times, and find clear evidence for the CR effect (and also for the common consequence effect). The main result of McGranaghan et al. (2024) (Result 1) reports absence of evidence using tests for paired valuations, which is however not evidence of absence. Additionally, the sign test that substantiates their analysis is only unbiased if the underlying, unobservable noise distribution is symmetric. As observed in Alós-Ferrer et al. (2021), this assumption might be unwarranted in general. For example, Alós-Ferrer et al. (2021, Proposition 4) shows that condition (TWT) must be fulfilled if the data generating process displays symmetric noise. In the experiment described here, the condition failed in 10.5% of the cases. However, an empirical application in Alós-Ferrer et al. (2021) finds the condition to fail in 39% of the choices. In related work, we found the condition to fail on 43.3% and 23% of choices in two different datasets for risky choice (Alós-Ferrer et al., 2024). Additionally, the literature has documented a large number of discrepancies between choice and valuation tasks and suggested that noise might be very different across tasks (Lichtenstein and Slovic, 1971; Schmidt and Hey, 2004; Bateman et al., 2007). Taken together, this evidence suggests that the assumption of symmetric (valuation) noise might be unwarranted in general, which might explain why the sign test of McGranaghan et al. (2024) detects no systematic evidence.

References

Agranov, M. and P. Ortoleva (2017). Stochastic Choice and Preferences for Randomization. *Journal of Political Economy 125*(1), 40–68.

Allais, M. (1953). Le Comportement de l'Homme Rationnel Devant le Risque, Critique des Postulates et Axiomes de l'Ecole Americaine. *Econometrica 21*, 503–546.

Alós-Ferrer, C., E. Fehr, and M. Garagnani (2024). Identifying Nontransitive Preferences. Working Paper, University of Zurich.

Alós-Ferrer, C., E. Fehr, and N. Netzer (2021). Time Will Tell: Recovering Preferences when Choices are Noisy. *Journal of Political Economy 129*(6), 1828–1877.

Alós-Ferrer, C. and M. Garagnani (2021). Choice Consistency and Strength of Preference. *Economics Letters 198*, 109672.

Alós-Ferrer, C. and M. Garagnani (2022a). Strength of Preference and Decisions Under Risk. *Journal of Risk and Uncertainty 64*, 309–329.

Alós-Ferrer, C. and M. Garagnani (2022b). The Gradual Nature of Economic Errors. *Journal of Economic Behavior & Organization 200*, 55–66.

Alós-Ferrer, C. and M. Garagnani (2024). Improving Risky-Choice Predictions Using Response Times. *Journal of Political Economy: Microeconomics 2(2)*, 335–354.

Apesteguía, J. and M. A. Ballester (2018). Monotone Stochastic Choice Models: The Case of Risk and Time Preferences. *Journal of Political Economy 126*(1), 74–106.

Baldassi, C., S. Cerreia-Vioglio, F. Maccheroni, and M. Marinacci (2020). A Behavioral Characterization of the Drift Diffusion Model and its Multi-Alternative Extension to Choice under Time Pressure. *Management Science 66*(11), 5075–5093.

Ballinger, T. P. and N. T. Wilcox (1997). Decisions, Error and Heterogeneity. *Economic Journal 107*(443), 1090–1105.

Bateman, I., B. Day, G. Loomes, and R. Sugden (2007). Can Ranking Techniques Elicit Robust Values? *Journal of Risk and Uncertainty 34*(1), 49–66.

Bhatia, S. and G. Loomes (2017). Noisy Preferences in Risky Choice: A Cautionary Note. *Psychological Review 124*(5), 678–687.

Blavatskyy, P., V. Panchenko, and A. Ortmann (2023). How Common is the Common-Ratio Effect? *Experimental Economics 26*(2), 253–272.

Blavatskyy, P. R. (2007). Stochastic Expected Utility Theory. *Journal of Risk and Uncertainty 34*, 259–286.

Blavatskyy, P. R. (2010). Reverse Common Ratio Effect. *Journal of Risk and Uncertainty 40*(3), 219–241.

Bordalo, P., N. Gennaioli, and A. Shleifer (2012). Salience Theory of Choice under Risk. *Quarterly Journal of Economics 127*(3), 1243–1285.

Cattell, J. M. (1893). On Errors of Observation. *The American Journal of Psychology 5*(3), 285–293.

Chabris, C. F., C. L. Morris, D. Taubinsky, D. Laibson, and J. P. Schuldt (2009). The Allocation of Time in Decision-Making. *Journal of the European Economic Association 7*(2-3), 628–637.

Cubitt, R. P., C. Starmer, and R. Sugden (1998). On the Validity of the Random Lottery Incentive System. *Experimental Economics 1*, 115–131.

Dashiell, J. F. (1937). Affective Value-Distances as a Determinant of Aesthetic Judgment-Times. *American Journal of Psychology 50*, 57–67.

Dehaene, S., E. Dupoux, and J. Mehler (1990). Is Numerical Comparison Digital? Analogical and Symbolic Effects in Two-Digit Number Comparison. *Journal of Experimental Psychology: Human Perception and Performance 16*(3), 626–641.

Fishburn, P. C. (1984a). SSB Utility Theory: An Economic Perspective. *Mathematical Social Sciences 8*(1), 63–94.

Fishburn, P. C. (1984b). SSB Utility Theory and Decision-Making Under Uncertainty. *Mathematical Social Sciences 8*(3), 253–285.

Fudenberg, D., P. Strack, and T. Strzalecki (2018). Speed, Accuracy, and the Optimal Timing of Choices. *American Economic Review 108*(12), 3651–3684.

Glimcher, P. W., M. C. Dorris, and H. M. Bayer (2005). Physiological Utility Theory and the Neuroeconomics of Choice. *Games and Economic Behavior 52*(2), 213–256.

Grether, D. M. and C. R. Plott (1979). Theory of Choice and the Preference Reversal Phenomenon. *American Economic Review 69*(4), 623–638.

Hey, J. D. (2005). Why We Should Not Be Silent About Noise. *Experimental Economics 8*, 325–345.

Hey, J. D. and C. Orme (1994). Investigating Generalizations of Expected Utility Theory Using Experimental Data. *Econometrica 62*(6), 1291–1326.

Kahneman, D. and A. Tversky (1979). Prospect Theory: An Analysis of Decision Under Risk. *Econometrica 47*(2), 263–291.

Klein, A. S. (2001). Measuring, Estimating, and Understanding the Psychometric Function: A Commentary. *Attention, Perception, & Psychophysics 63*(8), 1421–1455.

Konovalov, A. and I. Krajbich (2019). Revealed Strength of Preference: Inference from Response Times. *Judgment and Decision Making 14*(4), 381–394.

Laming, D. (1985). Some Principles of Sensory Analysis. *Psychological Review 92*(4), 462–485.

Lichtenstein, S. and P. Slovic (1971). Reversals of Preference Between Bids and Choices in Gambling Decisions. *Journal of Experimental Psychology 89*(1), 46–55.

Loomes, G. (2005). Modelling the Stochastic Component of Behaviour in Experiments: Some Issues for the Interpretation of Data. *Experimental Economics 8*(4), 301–323.

Loomes, G., P. G. Moffatt, and R. Sugden (2002). A Microeconometric Test of Alternative Stochastic Theories of Risky Choice. *Journal of Risk and Uncertainty 24*(2), 103–130.

Loomes, G. and R. Sugden (1982). Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty. *Economic Journal 92*(368), 805–824.

Loomes, G. and R. Sugden (1987). Some Implications of a More General Form of Regret Theory. *Journal of Economic Theory 41*(2), 270–287.

Loomes, G. and R. Sugden (1998). Testing Different Stochastic Specifications of Risky Choice. *Economica 65*(260), 581–598.

Luce, R. D. (1959). *Individual Choice Behavior: A Theoretical Analysis*. New York: Wiley.

McFadden, D. L. (1974). Conditional Logit Analysis of Qualitative Choice Behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics*, pp. 105–142. New York: Academic Press.

McFadden, D. L. (2001). Economic Choices. *American Economic Review 91*(3), 351–378.

McGranaghan, C., K. Nielsen, T. O'Donoghue, J. Somerville, and C. D. Sprenger (2024). Distinguishing Common Ratio Preferences from Common Ratio Effects Using Paired Valuation Tasks. *American Economic Review 114*, 307–347.

Moffatt, P. G. (2005). Stochastic Choice and the Allocation of Cognitive Effort. *Experimental Economics 8*(4), 369–388.

Mosteller, F. and P. Nogee (1951). An Experimental Measurement of Utility. *Journal of Political Economy 59*, 371–404.

Moyer, R. S. and R. H. Bayer (1976). Mental Comparison and the Symbolic Distance Effect. *Cognitive Psychology 8*(2), 228–246.

Moyer, R. S. and T. K. Landauer (1967). Time Required for Judgements of Numerical Inequality. *Nature 215*(5109), 1519–1520.

Murawski, C. and P. Bossaerts (2016). How Humans Solve Complex Problems: The Case of the Knapsack Problem. *Scientific Reports 6*(1), 34851.

Oprea, R. (2020). What Makes a Rule Complex? *American Economic Review 110*(12), 3913–3951.

Ratcliff, R. (1978). A Theory of Memory Retrieval. *Psychological Review 85*, 59–108.

Ruggeri, K., S. Alí, M. L. Berge, G. Bertoldo, L. D. Bjørndal, A. Cortijos-Bernabeu, C. Davison, E. Demić, C. Esteban-Serna, M. Friedemann, et al. (2020). Replicating Patterns of Prospect Theory for Decision Under Risk. *Nature Human Behaviour 4*(6), 622–633.

Schmidt, U. and J. D. Hey (2004). Are Preference Reversals Errors? An Experimental Investigation. *Journal of Risk and Uncertainty 29*(3), 207–218.

Shadlen, M. N. and R. Kiani (2013). Decision Making as a Window on Cognition. *Neuron 80*, 791–806.

Shadlen, M. N. and D. Shohamy (2016). Decision Making and Sequential Sampling from Memory. *Neuron 90*, 927–939.

Thurstone, L. L. (1927). A Law of Comparative Judgement. *Psychological Review 34*, 273–286.

Vieider, F. M. (2018). Violence and Risk Preference: Experimental Evidence from Afghanistan, Comment. *American Economic Review 108*(8), 2366–2382.

Webb, R. (2019). The (Neural) Dynamics of Stochastic Choice. *Management Science 64*(1), 230–255.

Wichmann, A. F. and N. J. Hill (2001). The Psychometric Function: I. Fitting, Sampling, and Goodness of Fit. *Attention, Perception, & Psychophysics 63*(8), 1293–1313.

APPENDICES.

Appendix A Lottery pairs used in the experiment

Set	Pair	OutR1	ProbR1	OutS1	ProbS1	Diff.EV
	Cert	21	80%	16	**100%**	0.80
I	Middle	21	60%	16	75%	0.60
	Normal	21	20%	16	25%	0.20
	Cert	27	72%	19	**100%**	0.40
II	Middle	27	58%	19	75%	0.22
	Normal	27	18%	19	25%	0.22
	Cert	14	85%	11	**100%**	0.90
III	Middle	14	68%	11	80%	0.72
	Normal	14	17%	11	20%	0.21

Table A.1: List of the common ratio pairs used in the experiment. The second outcome is zero for all lotteries.

Set	Pair	OutR1	ProbR1	OutR2	ProbR2	OutR3	ProbR3	OutS1	ProbS1	OutS2	ProbS2	Diff.EV
	Cert	23	15%	17	80%	0	5%	17	**100%**			0.05
IV	Middle	23	15%	17	50%	0	35%	17	70%	0	30%	0.05
	Normal	23	15%			0	85%	17	20%	0	80%	0.05
	Cert	31	12%	26	86%	0	2%	26	**100%**			0.08
V	Middle	31	12%	26	56%	0	32%	26	70%	0	30%	0.08
	Normal	31	12%			0	88%	26	14%	0	86%	0.08
	Cert	18	10%	12	85%	0	5%	12	**100%**			0.00
VI	Middle	18	10%	12	60%	0	30%	12	75%	0	25%	0.00
	Normal	18	10%			0	90%	12	15%	0	85%	0.00

Table A.2: List of the common consequence pairs used in the experiment.

Appendix B Analysis for each lottery double-pair

For the sake of brevity, the main text (including figures) reports averages across all CC or CR double pairs. Here we provide analyses and figures separately for each double pair.

Appendix B.1 Choice proportions for each lottery double-pair

This subsection provides the statistical tests using only choice frequencies for each of the double pairs, as well as the disaggregated figures corresponding to Figure 1.

First, recall that we say that a participant chose an option for a fixed (S, R) pair is she chose it more than half of the time across all 11 repetitions of that pair. We provide McNemar's tests comparing the proportion of participants that chose the riskier option R (in this sense) in the normal pair vs. the corresponding certainty pair. Identifying the double pairs (sets) as given in Appendix A, the proportion was significantly larger for normal pairs for all three CR double pairs (McNemar's tests; Set I, normal pair 61.73% vs. certainty pair 15.65%, $\chi^2 = 35.56$, $p < 0.001$; Set II, normal pair 62.61% vs. certainty pair 16.52%, $\chi^2 = 36.48$, $p < 0.001$; Set III, normal pair 50.43% vs. certainty pair 16.52%, $\chi^2 = 23.40$, $p < 0.001$). For CC double pairs, the proportion was significantly larger for two of the three double pairs (McNemar's tests; Set IV, normal pair 59.13% vs. certainty pair 34.78%, $\chi^2 = 17.82$, $p < 0.001$; Set V, normal pair 33.91% vs. certainty pair 33.04%, $\chi^2 = 0.02$, $p = 1.000$; Set VI, normal pair 60.87% vs. certainty pair 35.65%, $\chi^2 = 15.29$, $p < 0.001$).

The conclusions are unchanged if we compare the individual proportion of R choices (across the 11 repetitions of each pair) across normal and certainty pairs with Wilcoxon Signed-Rank tests (Set I, average 62.61% for normal pairs vs. 19.13% for certainty pairs, $z = 6.785$, $p < 0.001$; Set II, 62.77% vs. 19.92%, $z = 6.826$, $p < 0.001$; Set III, 48.46% vs. 19.76%, $z = 5.284$, $p < 0.001$, Set IV, 59.37% vs. 38.34%, $z = 4.780$, $p < 0.001$; Set V, 36.05% vs. 36.21%, $z = 0.622$, $p = 0.536$; Set VI, 58.81% vs. 38.10%, $z = 4.436$, $p < 0.001$).

Figure B.1 illustrates the choice frequencies separately for all six double pairs in the experiment (Figure 1 depicts the averages across the panels in this figure). For this figure, a participant is classified as displayed a specific pattern in a double pair, e.g., (S_C, R_N), if she chose option R for more than half of the repetitions of the normal pair and option S for more than half of the repetitions of the certainty pair.

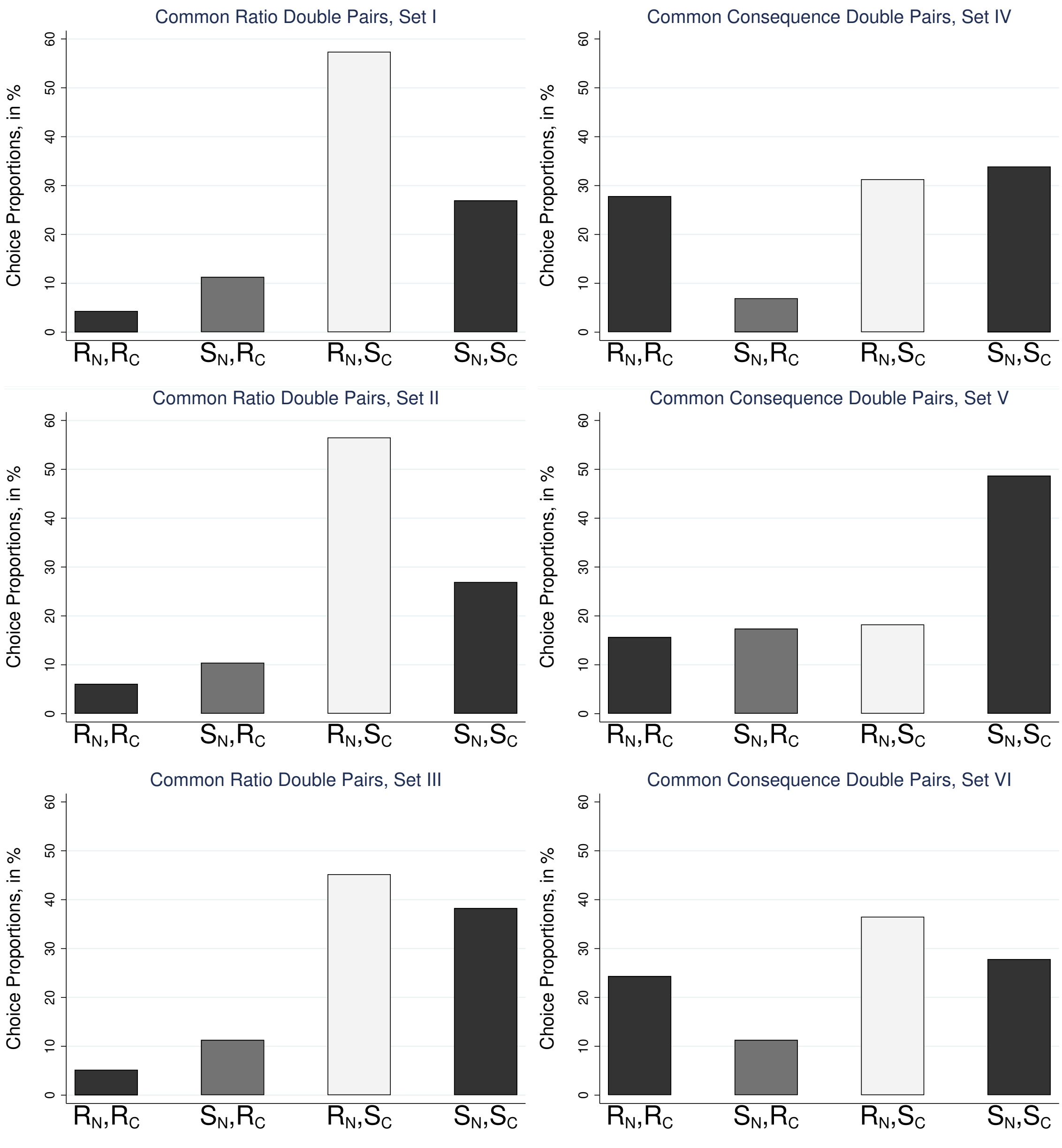

Figure B.1: Proportions of participants favoring each possible choice pattern, for each of the six double pairs in the experiment. Left column: common ratio double pairs. Right column: common consequence double pairs.

Appendix B.2 Genuine certainty effects for each lottery double-pair

Figure B.2 illustrates the proportion of participants with each possible pattern of revealed preferences, separately for each of the six double pairs in the experiment (Figure 3 depicts the averages across the panels in this figure). For this figure, a participant is classified as displayed a specific pattern in a double pair, e.g., (S_C, R_N), if a preference for R was revealed by the TWT method in the normal pair and preference for S was revealed in the corresponding certainty pair.

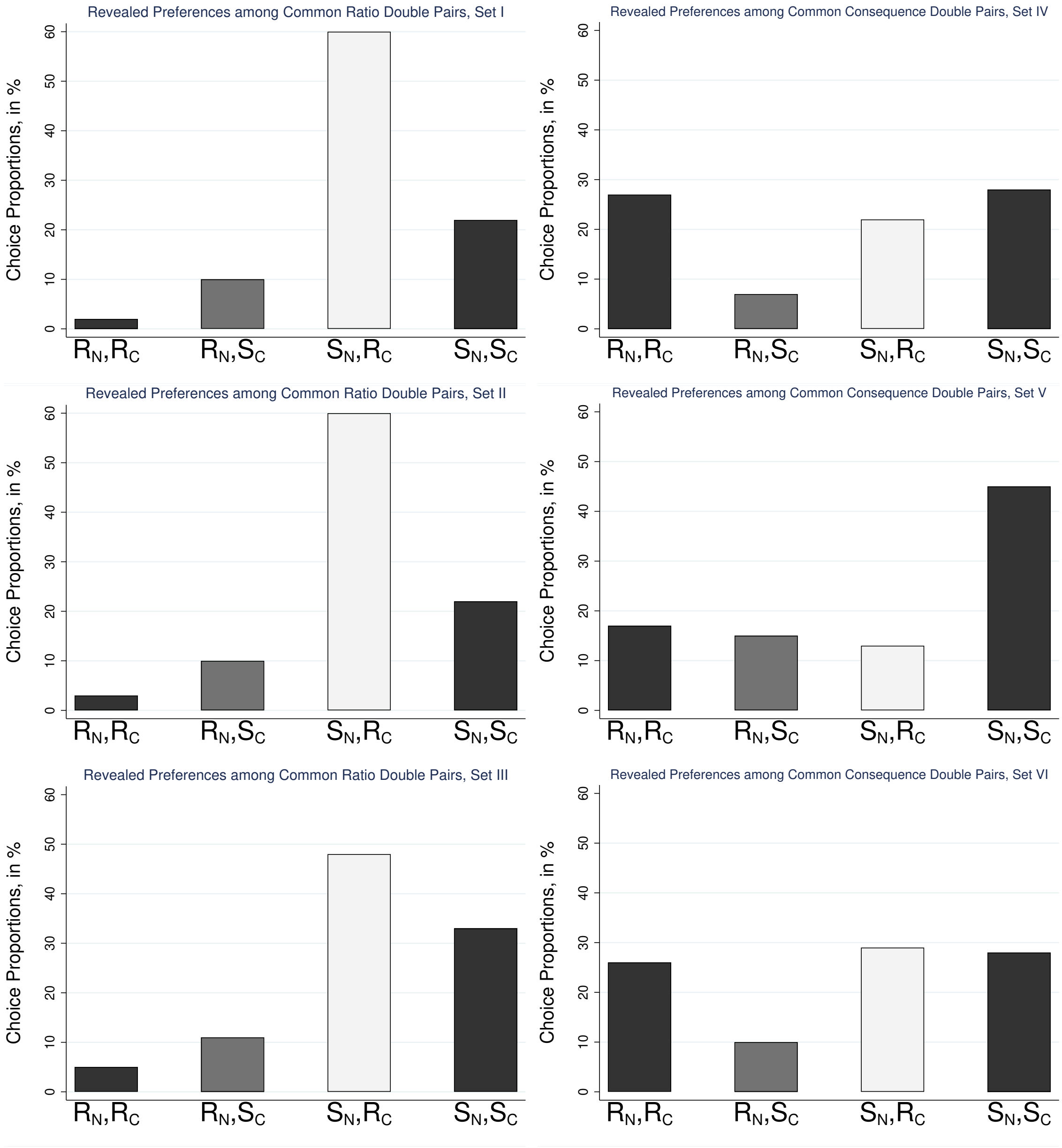

Figure B.2: Proportion of participants with each possible pattern of revealed preferences for each of the six double pairs in the experiment (conditional on preferences being revealed for the double pair). Left column: common ratio double pairs. Right column: common consequence double pairs.

Appendix C Additional Analysis, Middle Pairs

In the experiment participants also made choices for *middle pairs* constructed to bridge certainty and normal pairs (see the actual pairs in Appendix A). For the common ratio effect, these pairs were constructed to have the same outcomes of Normal and certainty pairs, but an intermediate winning probability. For the common consequence effect, the middle pairs were such that the common-consequence probability for a non-zero outcome was intermediate with respect to the normal and certainty pairs. That is, it was reduced compared to the certainty pair, but not taken all the way to zero as in the normal pair.

The analysis of the normal pairs qualitatively reproduces the results presented in the main text, and hence we only briefly summarize them here. Figures C.1 and C.2 show that choice patterns when considering most-frequent choices across repetitions are similar to actual revealed preference patterns. That is, observed choice patterns are mostly attributable to systematic preferences and not to noise.

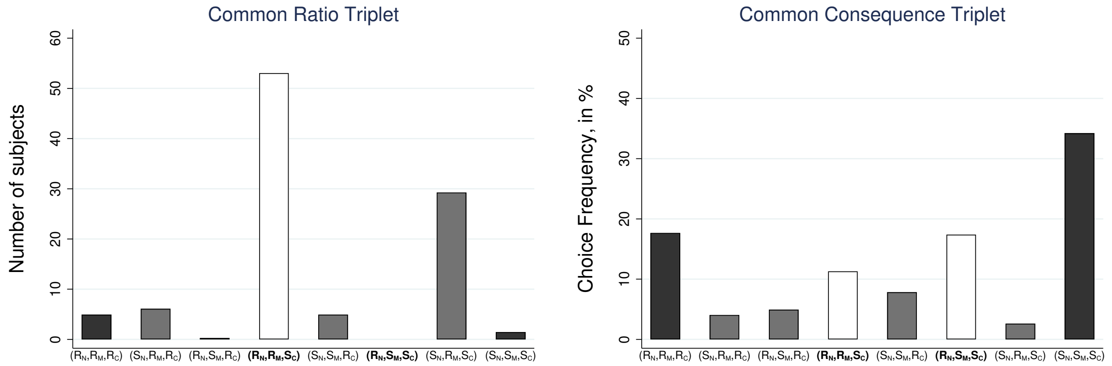

Figure C.1: Proportion of participants favoring each possible choice pattern (most frequent choice across repetitions), separately for common ratio (left) and common consequence (right) sets, including middle pairs.

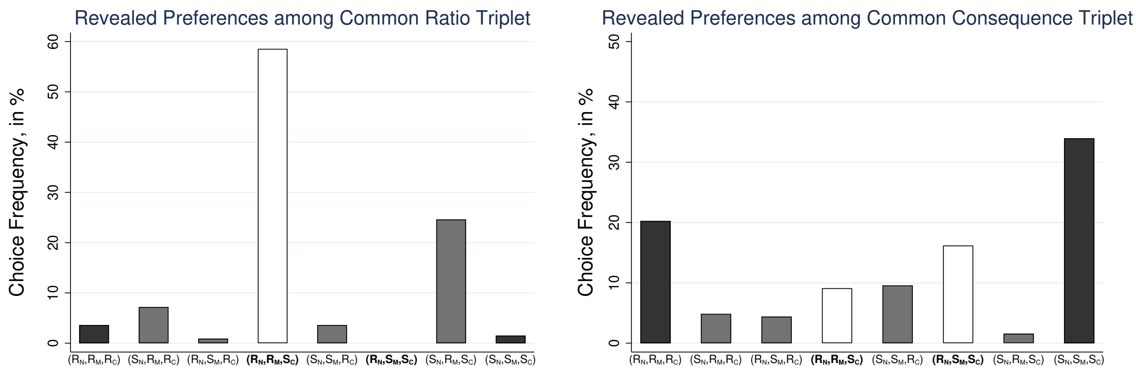

Figure C.2: Proportion of participants displaying each possible pattern of revealed preferences for common ratio (left) and common consequence (right) sets, including middle pairs.

Appendix D Experimental Instructions

[Instructions were in English and only presented on the screen of the participant's computer. No written instructions were provided.]

General Instructions

In the following you will be presented with numerous choices between two options, so-called "TRIALS."

These options involve different amounts of money that may be uncertain. Your choices are expressions of your personal preferences, therefore there are no objectively right or wrong answers.

You can work at your own speed. However, you will have to wait until all the other participants have completed their responses before you can leave your seat.

After you have made all your choices ONE TRIAL will be picked at random. Your PAYMENT depends on the choice you made in that trial. Please think carefully about every single choice because each one could be relevant for your payment.

Independently of your choices you will receive an additional CHF 10 for participating in the study.

Options

The box on the right-hand side of the screen represents an example of an option ("A").

On the left you see the possible amounts of money, on the right you see the probabilities associated with each amount of money.

If this option were relevant for your payment you would receive CHF 70 with 60% probability or CHF 0 (zero) with 40% probability.

The computer will determine your actual payment based on the given probabilities: If a randomly drawn number lies between 1 and 60, you would receive CHF 70, if it lies between 61 and 100 you would receive zero.

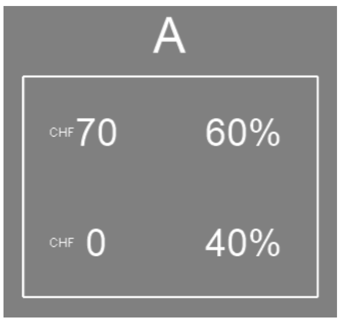

Figure D.1: Example of a lottery representation displayed on the "Options" and "Options Continued" screens.

Options, Continued

All options in this study will be presented in this way.

Some of them will only have one amount of money, hence only one line.

Some of them will have three amounts of money, hence three lines.

If you have a question, please raise your hand.

If you are ready to start the study, press the space bar.

Choice Screen

Figure D.2: Decision screen for lottery choices.