Time Will Tell: Recovering Preferences when Choices Are Noisy*

Carlos Alós-Ferrer† Ernst Fehr‡ Nick Netzer§
Department of Economics, University of Zurich

First version: October 2018 This Version: June 2020

Abstract

The ability to uncover preferences from choices is fundamental for both positive economics and welfare analysis. Overwhelming evidence shows that choice is stochastic, which has given rise to random utility models as the dominant paradigm in applied microeconomics. However, as is well known, it is not possible to infer the structure of preferences in the absence of assumptions on the structure of noise. We show that the difficulty can be overcome if data sets are enlarged to include response times. A simple condition on response time distributions (a weaker version of first-order stochastic dominance) ensures that choices reveal preferences without assumptions on the structure of utility noise. Standard random utility models from economics and standard drift-diffusion models from psychology generate data sets fulfilling this condition. Sharper results are obtained if the analysis is restricted to specific classes of noise. Under symmetric noise, response times allow to uncover preferences for choice pairs outside the data set, and if noise is Fechnerian, precise choice probabilities can be forecast out-of-sample. We apply our tools to an experimental data set, illustrating that the application is simple and generates a remarkable prediction accuracy.

JEL Classification: D11 · D81 · D83 · D87
Keywords: revealed preference · random utility models · response times

*We thank the editor and five anonymous referees for very helpful comments. We are particularly grateful to John Clithero for sharing his data. We also thank Drew Fudenberg, Ian Krajbich, Jay Lu, Paola Manzini, Yusufcan Masatlioglu, Paulo Natenzon, Pietro Ortoleva, Antonio Penta, Antonio Rangel, Jakub Steiner, Tomasz Strzalecki, Ryan Webb, Michael Woodford, and seminar audiences at Bocconi University, the University of Bonn, Princeton University, Washington University in St. Louis, the Workshop on Neuroeconomics and the Evolution of Economic Behavior (Vancouver, 2018), the Workshop on Stochastic Choice (Barcelona, 2018), the Sloan-Nomis Workshop on the Cognitive Foundations of Economic Behavior (Vitznau, 2018), the CESifo Area Conference on Behavioral Economics (Munich, 2018), and the 6th Annual Conference of the Society for Economic Measurement (Frankfurt 2019).

†Email: carlos.alos-ferrer@econ.uzh.ch. Department of Economics, University of Zurich. Blümlisalpstrasse 10, 8006 Zurich, Switzerland.

‡Email: ernst.fehr@econ.uzh.ch. Department of Economics, University of Zurich. Blümlisalpstrasse 10, 8006 Zurich, Switzerland.

§Email: nick.netzer@econ.uzh.ch. Department of Economics, University of Zurich. Blümlisalpstrasse 10, 8006 Zurich, Switzerland.

1 Introduction

Revealed-preference arguments lie at the foundation of economics (e.g. Samuelson, 1938; Houthakker, 1950; Arrow, 1959). Preferences revealed by choice are used in positive economics to predict behavior in novel situations, and in normative economics to evaluate the desirability of economic policies. In fact, the very use of a utility function entails the implicit assumption that the preferences represented by that function reflect choice behavior.

The traditional revealed-preference approach assumes that choices are deterministic. This assumption is contradicted by real-world choice behavior, as argued already in the classical work of Fechner (1860) and Luce (1959). Extensive evidence shows that individuals often make different choices when confronted with the same set of options repeatedly (among many others, see Tversky, 1969; Camerer, 1989; Hey and Orme, 1994; Agranov and Ortoleva, 2017). In view of this evidence, the traditional approach has been modified to allow for stochastic choice. The dominant paradigm in applied microeconomics today is to add a random component to cardinal utility (e.g., Thurstone, 1927; Marschak, 1960; McFadden, 1974, 2001). Random utility has several different interpretations, ranging from noise in an individual's perception of the options, over temporary fluctuations of tastes, to unobserved heterogeneity in a population of agents. With assumptions on the distribution of the random utility component, it becomes possible to deduce an underlying deterministic utility function from choice behavior.[1]

A problem with the random utility approach is that the distributional assumptions may actually drive the results. It is a well-known (but rarely stated) fact that, within this approach, nothing can be learned about preferences without making distributional assumptions. The flipside of this result is that anything can be "learned" by making the suitable assumptions on the structure of noise. Unfortunately, these assumptions cannot be verified, because utility is a latent variable that is not directly observed. Hence, the dependence on possibly unwarranted assumptions is not just an abstract theoretical problem but is known to plague empirical research (see Hey and Orme, 1994). For instance, Buschena and Zilberman (2000) showed that, for the same data set, assuming homoskedasticity supports non-expected utility models, while expected utility models cannot be improved upon when heteroskedasticity is allowed.

In this paper, we show that the problem can be overcome by using data on response times. This is because the distribution of response times, which is in principle observable, contains information about the unobservable distribution of utility.[2] We derive,

[1]There are alternative approaches to revealed preference when choice is stochastic, which we do not consider here. One is to work with a distribution over ordinal preferences (Block and Marschak, 1960; Falmagne, 1978; Barberá and Pattanaik, 1986; Gul and Pesendorfer, 2006; Gul et al., 2014; Apesteguía et al., 2017; Apesteguía and Ballester, 2018). Another one is to model stochastic choice as the result of stochastic consideration sets (e.g. Manzini and Mariotti, 2014). See Section 6 for a more detailed literature review.

[2]The time it takes to make a decision ("response time") can always be observed in laboratory experiments, but even outside the laboratory response times are in principle an observable outcome of a

first, a simple and intuitive condition on the distribution of response times that ensures that preferences can be identified from choice data without any assumptions on the structure of noise. Second, we show that under symmetric noise, response times enable the identification of preferences between alternatives for which no previous choice data exist. This would not be possible without response time data unless one is willing to impose stronger (untestable) assumptions than symmetry on the utility noise. Third, we show that if one is willing to assume that utility noise is Fechnerian—an underlying assumption of probit and logit models—response time data enable the calculation of precise choice probabilities for alternatives for which no choice data exist. Again, this would not be possible without response time data under only the Fechnerian assumption.

Our approach is made possible by the fact that, despite being stochastic, choice behavior obeys certain well-known regularities. One of those regularities, often referred to as the *psychometric function*, is the fact that easier choice problems are more likely to elicit correct responses than harder problems. This can be traced back to perceptual discrimination experiments in psychophysics, where an objectively correct response exists (e.g. choosing the brightest or loudest stimulus). It is perhaps one of the most robust facts in all of psychology that the percentage of correct choices increases with the difference in stimuli (Cattell, 1893; Laming, 1985; Klein, 2001; Wichmann and Hill, 2001). Conversely, choice becomes noisier when stimuli are more similar and hence the problem is harder. This finding extends to cases where the correct response is subjective, e.g. favorite colors, and is uncovered by the researcher through ratings (Dashiell, 1937). In economics, the classical work of Mosteller and Nogee (1951) showed that the phenomenon also occurs in decisions under risk. In their data, the alternative with the larger estimated utility was not always chosen, but the percentage of choices in favor of the high-utility option increased in the utility difference between the options. This corresponds exactly to the psychometric function, with the difficulty of the (binary) choice problem being measured by the subjective utility difference between the available options, and easier choices being those with a larger absolute utility difference. In fact, this psychometric relationship is an integral part of standard random utility models, which assume that choice probabilities are monotone in utility differences.

Our approach in this paper rests on integrating a second well-known regularity, often referred to as the *chronometric function*, into the standard random utility framework. The chronometric function describes the fact that easier choice problems take less time to respond than harder problems. As in the case of the psychometric function, there is overwhelming evidence from psychophysics showing this regularity in an astonishing variety of domains, going back to as early as Cattell (1902). For instance, Moyer and Landauer (1967) demonstrated that chronometric effects exist even for the simple question of which of two single-digit numbers is larger (see also Moyer and Bayer, 1976; Dehaene et al., 1990). The finding extends to choice based on subjective preferences,

choice process that can be collected by researchers, firms, or other interested parties. In contrast, the distribution of utility noise is intrinsically unobservable.

as in the work on favorite colors by Dashiell (1937). A growing body of evidence shows that the chronometric phenomenon also applies to economic decisions, that is, response times are decreasing in utility differences. For instance, Chabris et al. (2009) measured response times in binary intertemporal choices and found strong evidence that decisions are faster when utility differences are larger. Krajbich et al. (2015) found the same relation for sharing decisions in a dictator game. Moffatt (2005) and Alós-Ferrer and Garagnani (2018) established the phenomenon for decisions under risk. Krajbich et al. (2012) showed that consumer purchases display the same pattern, and the phenomenon has been repeatedly illustrated in the domain of food choice (Krajbich et al., 2010; Krajbich and Rangel, 2011; Fisher, 2017; Clithero, 2018). These different papers use very different methods to elicit utility differences, such as stated liking ratings (Krajbich and Rangel, 2011), monetary differences (Alós-Ferrer and Garagnani, 2018), and estimation based on a logit or probit specification (Moffatt, 2005; Chabris et al., 2009; Krajbich et al., 2015; Alós-Ferrer and Garagnani, 2018), but they all arrive at the chronometric relationship between utility differences and response times.

To provide an intuition for our results, consider the choice between two options x and y, where x is chosen with probability p and y with probability $1-p$. For the sake of clarity, let us adopt the interpretation that these probabilities describe the choices of a single individual across many repetitions of the problem. We will first show that observing $p > 1/2$ is not sufficient to conclude that the individual prefers x to y, i.e., that the underlying deterministic utility of x is larger than that of y, if no assumptions about the shape of the distribution of utility noise are made. It is possible to rationalize the data by a random utility model (RUM) that has a deterministic utility function with $u(x) < u(y)$ and asymmetric noise with zero mean. The asymmetry is such that the realized utility difference between x and y is often positive, generating $p > 1/2$, but takes large absolute values whenever it is negative. We will argue that asymmetric distributions in fact arise very naturally, even in conventional additive noise models where the utility of an option is given by $\tilde{u}(x) = u(x) + \tilde{\epsilon}(x)$. Wrongfully assuming symmetry then leads to false inferences about preferences.[3]

Now assume we have data on the joint distribution of choices and response times. We will show in Theorem 1 that $p > 1/2$ combined with, informally speaking, a comparatively slow choice of y relative to x is sufficient to conclude that the individual prefers x to y, even without making any assumptions about the shape of the utility distribution. A slow choice of y relative to x reveals that the utility difference cannot be distributed too asymmetrically in the way described above, because negative utility differences with large absolute values would generate quick choices of y, based on the

[3]Our arguments are based on Horan et al. (2019), who describe conditions under which differences of random utilities are symmetrically or asymmetrically distributed. Our example of an asymmetric distribution discussed in Section 3.1 rests on independent, mean-zero Gumbel errors like in any standard logit specification, but with heteroskedasticity across the options. Given our limited understanding of how utility is constructed and how noise affects evaluation in preferential choice problems (which often involve trade-offs between multiple attributes), we see no good reasons for justifying the distributional assumptions required for symmetry, except tractability.

chronometric relationship. This argument does not presume knowledge of the shape of the chronometric function beyond monotonicity (and some technical properties). More formally, let $F(x)(t)$ and $F(y)(t)$ be the cumulative distribution functions of response times conditional on the choice of x and y, respectively. Our criterion states that when we observe

$$F(y)(t) \leq \frac{p}{1-p} F(x)(t) \text{ for all } t \geq 0,$$

then any random utility model with a chronometric function (RUM-CF) which rationalizes the data must satisfy $u(x) \geq u(y)$. A similar statement holds for strict preferences. In the limit as $p \to 1/2$, that is, as choice data alone becomes uninformative, this becomes the condition that choice of y must be slower than choice of x in the first-order stochastic dominance sense. As p grows, hence choice data becomes more indicative of a preference, the condition becomes weaker and only requires that choice of y is not much faster than choice of x.

We then study the case where the analyst has reasons to believe that utility differences are symmetrically distributed, as is often assumed in the literature (e.g. in any application with logit or probit choice). It then follows immediately that $p > 1/2$ implies $u(x) > u(y)$, so preferences are revealed by choices without response times. But now we show that the use of response time data enables the identification of preferences for choice pairs outside the set of available choice data. For the case of deterministic choices and deterministic response times, this has been noted before. Krajbich et al. (2014) argue that a slow choice of z over x combined with a quick choice of the same z over y reveals a preference for x over y, even though the choice between x and y is not directly observed and a transitivity argument is not applicable. The idea is that, based on the chronometric relationship, the positive utility difference $u(z)-u(x)$ must be smaller than the positive utility difference $u(z) - u(y)$, which implies $u(x) > u(y)$. To date, however, it has remained an open question how to implement this idea, since real-world choices and response times are stochastic, and hence it is unclear what "choice of z over x" and "slow versus fast" exactly means. For instance, is "faster than" defined in terms of mean response times, median response times, or some other characteristic of the response time distribution? Our Theorem 2 provides an answer to that question. Suppose z is chosen over x with a probability $p(z, x) > 1/2$, which indeed implies $u(z) - u(x) > 0$ in the symmetric-noise case. Then we define $\theta(z, x)$ as a specific percentile of the response time distribution for z, namely the $0.5/p(z, x)$-percentile. Analogously, if z is chosen over y with a probability $p(z, y) > 1/2$, which implies $u(z)-u(y) > 0$, the corresponding percentile $\theta(z, y)$ can be defined. Our result shows that these observable percentiles are the appropriate measure of preference intensity for the stochastic setting, in the sense that $\theta(z, x) > \theta(z, y)$ implies $u(z) - u(x) < u(z) - u(y)$ and hence a revealed (strict) preference for x over y. That is, inference cannot be based on mean, median, maximum, or minimum response times. The correct measurement is the $0.5/p$-percentile of the distribution of response times, which requires using a different percentile for each choice pair, adjusting for the respective choice frequencies. Since a revealed strict preference for x over y translates into a choice probability $p(x, y) > 1/2$ when the utility distribution is symmetric, this quantification generates out-of-sample predictions that are easy to test empirically.

In the traditional approach without response times, making out-of-sample predictions requires even stronger distributional assumptions than just symmetry. Random utility models like probit or logit are instances of Fechnerian models (Debreu, 1958; Moffatt, 2015), in which the utility difference between the two options follows the exact same distributional form in all binary choice problems. With this Fechnerian assumption (but without assuming a specific functional form for the distribution), already the choice observation $p(z, x) < p(z, y)$ reveals a preference for x over y. Put differently, the Fechnerian assumption enables an exhaustive elicitation of ordinal preferences outside the data set. However, the use of response time data makes it now even possible to move beyond ordinal preferences and make predictions of precise choice probabilities. Theorem 3 provides a closed-form formula to predict $p(x, y)$ based on observables, i.e., choice probabilities and response times, from only the binary choices between z and x and between z and y.

The general pattern that emerges from our results is that response time data allow us to obtain results that would otherwise require an additional distributional assumption, that might be empirically unjustified. Response time data make it possible to get rid of assumptions because the distribution of response times contains information about the distribution of utility noise. This enables the revelation of preferences without any distributional assumptions, makes it possible to extrapolate preferences to cases for which no choice data exist with a symmetry assumption, and even generates precise probability predictions with the Fechnerian assumption.

Our Theorem 1 provides a robust sufficient condition for preference revelation, which essentially goes "from data to models." To investigate how much bite our criterion has, we also look at the converse implication "from models to data," i.e., we study stochastic choice functions with response times (SCF-RTs) that are generated by standard models from the received literature. We take a specific data-generating process as given and apply our agnostic method that does not presume knowledge of the process to the resulting data set. We do this first for the whole class of RUM-CFs that have symmetric distributions, which contains the probit and logit models as special cases but goes far beyond them. We show that our criterion recovers all preferences correctly when any such model generated the data (Proposition 4). In other words, our sufficient condition is also necessary and has maximal bite for the entire class of SCF-RTs generated by symmetric RUM-CFs. Even the analyst who believes in the probit or logit distribution can work with our criterion, because it must always hold in his data. Conversely, a data set where the condition in Theorem 1 is violated cannot be explained through a symmetric RUM-CF. We then show that the result still holds with additional noise in response times, as long as the noise is from an independent source like a stochastic chronometric function or imperfect observation, and does not systematically reverse the chronometric relationship (Proposition 5).

Second, we study the class of drift-diffusion models (DDMs) with constant or collapsing decision boundaries, which are prominent in psychology and neuroscience (e.g. Ratcliff, 1978; Shadlen and Kiani, 2013). These models have recently attracted attention in economics because they can be derived from optimal evidence accumulation mechanisms (Drugowitsch et al., 2012; Tajima et al., 2016; Fudenberg et al., 2018; Baldassi et al., 2019). We show that our criterion again recovers all preferences correctly from data that is generated by a DDM (Proposition 6). Hence our previous statement on believers in probit or logit models also applies to believers in drift-diffusion models. Again, as a consequence of this result, a data set where the condition in Theorem 1 is violated cannot be explained through a DDM.

Finally, we apply our tools to an experimental data set from Clithero (2018). In that experiment, subjects made repeated choices between various snack food items, and response times were recorded. Our results are easy to apply by using non-parametric kernel density estimates of the response time distributions. We show that our condition from Theorem 1 is fulfilled in 61% of all decision problems where choice was random. Hence, a preference is revealed without distributional assumptions in a majority of cases. In addition, the fact that in 39% of the cases the condition from Theorem 1 is violated suggests that, in these cases, models with symmetric noise, including Fechnerian models as the standard logit and probit approaches, are not consistent with the data. These numbers illustrate the empirical relevance of our criterion, but also warn against the unquestioned use of the symmetry assumption.

We then apply Theorem 2 to predict choices in a second phase of the experiment. The accuracy of our out-of-sample predictions, which rest on no distributional assumptions other than symmetry, is remarkable. The prediction is correct in 80.7% of the cases, which is significantly better than for the logit model, and indistinguishable from a fully-parametric drift-diffusion model (Clithero, 2018). We conduct an analogous analysis using Theorem 3 to predict precise choice probabilities, assuming a Fechnerian structure of the utility noise. We again achieve a high prediction accuracy, significantly better than for the logit model. While the drift-diffusion model estimated by Clithero (2018) performs even better in that case, our non-parametric method yields its high accuracy in a straightforward way and does not require computation-intensive structural estimation.

Our paper is related to the vast empirical and theoretical literature on random utility models, that have a long history going back to Thurstone (1927), Luce (1959) and Marschak (1960). These models became more widely used in economics after McFadden (1974) (see also McFadden, 2001). In the meantime, they have become a dominant workhorse in applied economics, as indicated by the more than 19,000 Google Scholar citations of McFadden (1974) (to date). We believe that our paper solves a fundamental issue in this approach by showing how the rigorous and principled use of response times greatly improves the ability to identify preferences.

In addition, our paper contributes to the large literature in psychology, neuroeconomics, and neuroscience on response times and evidence accumulation models (e.g., Ratcliff and Rouder, 1998; Krajbich et al., 2010; Krajbich and Rangel, 2011; Ratcliff et al., 2016). A standard view of human decision-making in neuroscience (e.g., Shadlen and Kiani, 2013) is that the chronometric function reects neural processes of value comparison between the options, and that these processes take longer to differentiate closer values. In particular, Proposition 6 offers a test of whether drift diffusion models with constant or collapsing boundaries can be the data generating process for a given dataset. In addition, our Theorem 1 radically simplifies preference identification relative to structural estimations of drift diffusion models.

Our paper is also related to the emerging literature on response times and process models in economics (Woodford, 2014; Fudenberg et al., 2018; Baldassi et al., 2019; Clithero, 2018; Konovalov and Krajbich, 2019; Schotter and Trevino, 2020). The chronometric relationship between response times and utility differences, in particular, has only recently been an object of interest in economics (e.g., Krajbich et al., 2015; Alós-Ferrer et al., 2016; Echenique and Saito, 2017). It has been formalized in economics by Fudenberg et al. (2018), who show that even in an optimal dynamic process of evidence accumulation, it takes longer to discover that one is close to indifference than to recognize a strong preference (see also Woodford, 2014; for an axiomatization of such evidence accumulation models, see Baldassi et al., 2019). Another possible interpretation of the phenomenon is that it indeed captures an inherent suboptimality of human decision-making. In line with this interpretation, Krajbich et al. (2014) show that choices can sometimes be improved by interventions which force people to spend less time on single decisions. Which of these interpretations one favors is inconsequential for our analysis. We do, however, restrict our attention to decisions which can be conceived of as discovering an internal preference, as opposed to solving a complex problem.[4]

The paper is organized as follows. Section 2 presents the formal setting. Section 3 develops the main results, devoting separate subsections to the unrestricted, symmetric, and Fechnerian cases. Section 4 shows that choice data generated by standard models from economics and psychology fulfills our main criterion for preference revelation. Section 5 contains our empirical application. Section 6 discusses the related literature in more detail, and Section 7 concludes. All proofs omitted from the main text can be found in the Appendix.

[4]Of course, some life-changing decisions, as whether to accept a job, or the parents' choice of the right school for their children, take a long time not because the options are similar, but because they are enormously complex and multidimensional. Likewise, some tasks, like deciding whether a sophisticated mathematical statement is true or false, take time because of their complexity. In contrast, we consider economic decisions where an internal preference exists and has to be discovered and recognized by the decision maker. In our view, these decisions make up a large part of our daily choices.

2 Formal Setting and Definitions

Let X be a finite set of options. Denote by $C = \{(x, y) \mid x, y \in X, x \neq y\}$ the set of all binary choice problems, so (x, y) and (y, x) both represent the problem of choice between x and y. Let $D \subseteq C$ be the set of choice problems on which we have data, assumed to be non-empty and symmetric, that is, $(x, y) \in D$ implies $(y, x) \in D$. To economize notation, we let the set D be fixed throughout.

Definition 1. A *stochastic choice function* (SCF) is a function p assigning to each $(x, y) \in D$ a probability $p(x, y) > 0$, with the property that $p(x, y) + p(y, x) = 1$.

In an SCF, $p(x, y)$ is interpreted as the probability of choosing x in the binary choice between x and y, and $p(y, x)$ is the probability of choosing y. The assumption that $p(x, y) > 0$ for all $(x, y) \in D$ implies that choice is stochastic in a non-degenerate sense, because each alternative is chosen with strictly positive probability.

Since there is no universally agreed-upon definition of random utility models, we will work with a fairly general definition which encompasses several previous ones. In particular, it is convenient for our analysis to directly describe for each $(x, y) \in C$ the distribution of the utility difference between the two options.

Definition 2. A *random utility model* (RUM) is a pair $(u, \tilde{v})$ where $u : X \to \mathbb{R}$ is a utility function and $\tilde{v} = (\tilde{v}(x, y))_{(x,y)\in C}$ is a collection of real-valued random variables, with each $\tilde{v}(x, y)$ having a density function $g(x, y)$ on $\mathbb{R}$, fulfilling the following properties:

(RUM.1) $\mathbb{E}[\tilde{v}(x, y)] = u(x) - u(y)$,

(RUM.2) $\tilde{v}(x, y) = -\tilde{v}(y, x)$, and

(RUM.3) the support of $\tilde{v}(x, y)$ is connected.

In a RUM, the utility function u represents the underlying preferences which the analyst aims to uncover, while the random variables $\tilde{v}(x, y)$ also incorporate the noise, modeled directly as random pairwise utility differences. That is, the density $g(x, y)$ describes the distribution of the random utility difference between x and y. The noise is assumed to have zero mean, so the expected value of the random utility difference $\tilde{v}(x, y)$ must be $u(x) - u(y)$, as required by (RUM.1). This condition can be spelled out in terms of the density $g(x, y)$ as

$$\int_{-\infty}^{+\infty} v g(x, y)(v) dv = u(x) - u(y).$$

We will also use the notation $v(x, y) = u(x) - u(y)$ for $\mathbb{E}[\tilde{v}(x, y)]$. Condition (RUM.2) states that $\tilde{v}(x, y)$ and $\tilde{v}(y, x)$ describe the same random utility difference but with opposite sign, that is, $g(x, y)(v) = g(y, x)(-v)$ for all $v \in \mathbb{R}$. Finally, (RUM.3) is a regularity condition stating that there are no gaps in the distribution of a pair's utility differences.

Our definition reflects the conventional idea that RUMs consist of a deterministic utility function plus mean-zero noise terms, as typically implemented in discrete choice and microeconometrics (see Ben-Akiva and Lerman, 1985; for a historical account, see McFadden, 2001). Those approaches define a RUM by random utilities $\tilde{u}(x) = u(x) + \tilde{\epsilon}(x)$ for each option. Equivalently, a RUM of this kind can be given directly by a vector $\tilde{u} = (\tilde{u}(x))_{x \in X}$ of random variables, with the underlying deterministic utility function u being defined through $u(x) = \mathbb{E}[\tilde{u}(x)]$. Those models are particular cases of ours, taking $\tilde{v}(x, y) = \tilde{u}(x) - \tilde{u}(y)$. Our approach is more general because the utility differences $\tilde{v}(x, y)$ are unrestricted across choice pairs, i.e., the realized utility difference is not constrained to be a difference of realized utilities. This allows us to accommodate pair-specific factors other than utility differences that may affect choice probabilities (and response times), such as dominance relations between some options (see He and Natenzon, 2018, and our discussion in Section 6.3). An additional benefit of this generality is that our positive results on preference revelation become stronger, because they hold within a larger class of models. The added generality does not matter for the impossibility result in Proposition 1, which would also hold for RUMs of the form $\tilde{u}(x) = u(x) + \tilde{\epsilon}(x)$.

A RUM generates choices by assuming that the option chosen is the one with the larger realized utility.

Definition 3. A RUM $(u, \tilde{v})$ *rationalizes* an SCF p if $p(x, y) = \text{Prob}[\tilde{v}(x, y) > 0]$ holds for all $(x, y) \in D$.

For the particular case where $\tilde{v}(x, y) = \tilde{u}(x) - \tilde{u}(y)$, the definition above just specifies $p(x, y) = \text{Prob}[\tilde{u}(x) > \tilde{u}(y)]$. Note that, since $\tilde{v}(x, y)$ is assumed to have a density $g(x, y)$, whether one writes $\text{Prob}[\tilde{v}(x, y) > 0]$ or $\text{Prob}[\tilde{v}(x, y) \geq 0]$ is inconsequential. Denote the cumulative distribution function derived from $g(x, y)$ by $G(x, y)$. Then, the probability that the utility of y exceeds that of x is $G(x, y)(0)$. As a consequence, Definition 3 can be alternatively stated as the condition that

$$G(x, y)(0) = p(y, x)$$

for all $(x, y) \in D$.

We now extend the framework and include response times, by adding conditional response time distributions for each choice. This is the easiest way of describing a joint distribution over choices and response times.

Definition 4. A *stochastic choice function with response times* (SCF-RT) is a pair (p, f) where p is an SCF and f assigns to each $(x, y) \in D$ a strictly positive density function $f(x, y)$ on $\mathbb{R}_+$.

The density $f(x, y)$ describes the distribution of response times conditional on x being chosen in the binary choice between x and y. The corresponding cumulative distribution function is denoted by $F(x, y)$. It would be straightforward to introduce lower

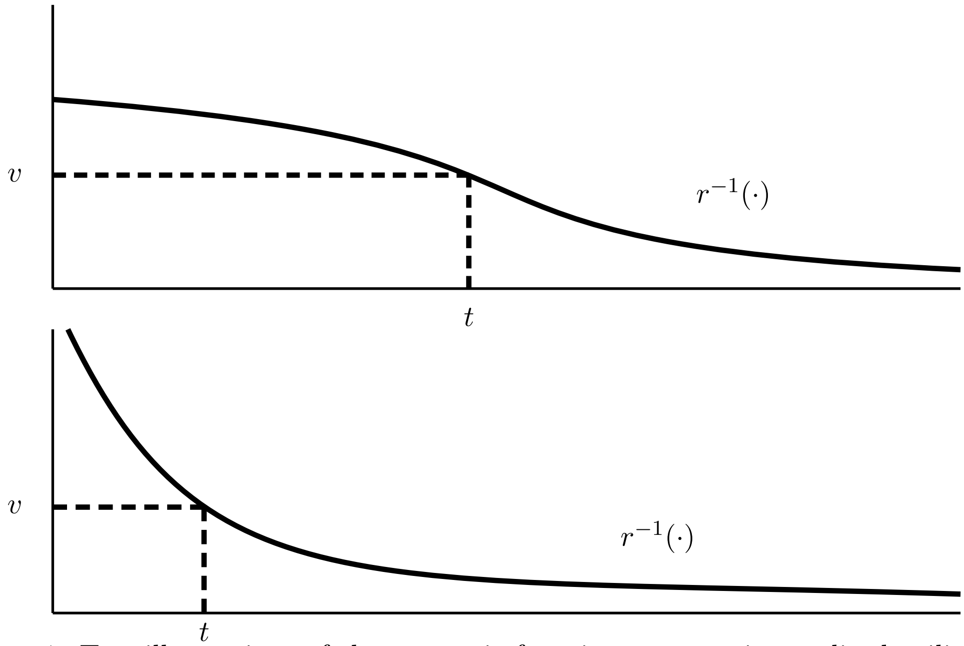

Figure 1: Two illustrations of chronometric functions r, mapping realized utility differences (vertical axis) into response times (horizontal axis).

or upper bounds on response times, for instance due to a non-decision time or a maximal observed response time. We refrain from doing so here for notational convenience and comparability with the literature (e.g. Fudenberg et al., 2018).

Definition 5. A *random utility model with a chronometric function* (RUM-CF) is a triple $(u, \tilde{v}, r)$ where $(u, \tilde{v})$ is a RUM and $r : \mathbb{R}_{++} \to \mathbb{R}_+$ is a continuous function that is strictly decreasing in v whenever $r(v) > 0$, with $\lim_{v \to 0} r(v) = \infty$ and $\lim_{v\to\infty} r(v) = 0$.

In a RUM-CF, r represents the chronometric function. It maps realized utility differences v into response times $r(|v|)$, such that larger absolute utility differences generate shorter response times. The assumption that $\lim_{v\to 0} r(v) = \infty$ and $\lim_{v\to\infty} r(v) = 0$ ensures that the model can encompass all response times observed in an SCF-RT. Our definition allows for functions like $r(v) = 1/v$ that are strictly decreasing throughout, and also for functions that reach $r(v) = 0$ for large enough v. The latter case will arise when we construct chronometric functions from sequential sampling models in Section 4. Figure 1 illustrates both cases, taking advantage of the fact that the inverse $r^{-1}(t)$ is well-defined for the restriction of r to the subset where $r(v) > 0$.

In addition to choices, a RUM-CF generates response times by assuming that the realized response time is related to the realized utility difference through function r. Specifically, given a RUM-CF $(u, \tilde{v}, r)$ and a pair $(x, y) \in C$, the random variable describing the response times when x is chosen over y is given by

$$\tilde{t}(x, y) = r(|\tilde{v}(x, y)|),$$

conditional on $\tilde{v}(x, y) > 0$. This motivates the following definition.

Definition 6. A RUM-CF $(u, \tilde{v}, r)$ *rationalizes* an SCF-RT (p, f) if $(u, \tilde{v})$ rationalizes p and $F(x,y)(t) = \text{Prob}[\tilde{t}(x,y) \leq t \mid \tilde{v}(x,y) > 0]$ holds for all $t > 0$ and all $(x,y) \in D$.

The probability of a response time of at most t, conditional on x being chosen over y, is the probability that the realized utility difference is at least $r^{-1}(t)$, conditional on that difference being positive. Since this probability can be calculated as $\left[1 - G(x,y)(r^{-1}(t))\right] / [1 - G(x,y)(0)]$, the condition in Definition 6 is equivalent to

$$F(x,y)(t) = \frac{1 - G(x,y)\left(r^{-1}(t)\right)}{1 - G(x,y)(0)} \tag{1}$$

for all $t > 0$ and all $(x,y) \in D$.

An alternative approach would have been to assume that response time is a decreasing function of the true absolute utility difference $|v(x,y)|$ between the two options, as opposed to the realized, noisy ones. A first drawback of this approach is that response times would be predicted to be deterministic, in contradiction to all available evidence. Hence a second source of noise would have to be introduced, for instance by making the chronometric function stochastic. Without additional *ad hoc* assumptions, any such model would predict that the conditional distributions of response times for each of the two choices are identical, a further prediction not borne out by the data. Our approach is more parsimonious as it requires only one source of randomness (in utility) to generate both stochastic choices and stochastic response times,[5] and it does not make the implausible prediction of independence between choices and response times. A second drawback of the alternative approach is that the response time for two identical options would be predicted to be infinite, which is unreasonable. In contrast, our model does not make such a prediction, because even with indifference $v(x,y) = 0$, the stochastic utility difference $\tilde{v}(x,y)$ has mean zero but is different from zero with probability one, generating continuous distributions of response times.[6]

When studying stochastic choice data, the analyst might be interested in or willing to make specific assumptions about the distribution of random utility. In doing so, the analyst accepts a restriction to a specific subclass of random utility models. Those might range from symmetry of each density $g(x,y)$ to specific functional forms. We say that an SCF, respectively an SCF-RT, is *rationalizable* within some class of models if there exists a RUM, respectively a RUM-CF, in that class that rationalizes it.

Definition 7. Within a class of models, a rationalizable SCF (SCF-RT) *reveals a preference for x over y* if all RUMs (RUM-CFs) in the class that rationalize it satisfy

[5] That said, a second source of independent noise like a stochastic chronometric function can be introduced without changing our main insights, as we will show in Section 4.

[6] A realized stochastic utility difference of zero would give rise to an infinite response time, but this happens with probability zero. The generated distributions can have finite means, as we would expect in data. Our model does not even predict that these mean response times must be decreasing in the underlying deterministic utility difference. That is an implication with additional assumptions fulfilled by standard models with logit or probit errors, but does not necessarily hold in the general case.

$u(x) \geq u(y)$. It *reveals a strict preference for x over y* if all RUMs (RUM-CFs) in the class that rationalize it satisfy $u(x) > u(y)$.

3 Revealed Preference

In this section, we investigate the use of response times for preference revelation. Specifically, we are interested in preference revelation within difference classes of random utility models, and how the addition of response times improves the results.

3.1 The Unrestricted Case

The first observation is that, without further restrictions on the utility distributions, and without the use of response times, nothing can be learned from choice probabilities. This is well-known among specialists in stochastic choice theory, and hence we do not claim originality.

Proposition 1. *Within the class of all RUMs, a rationalizable SCF reveals no preference between any x and y with $x \neq y$.*

The intuition for the result is simple. Data on the choice between x and y allows us to learn the value $G(x, y)(0)$, but, without distributional assumptions, this does not tell us whether the expected value $v(x, y) = u(x) - u(y)$ is positive or negative, which is what we are interested in.

A solution to the problem would be to impose the seemingly innocuous assumption of symmetry of the distribution. In that case, $G(x, y)(0) \leq 1/2$ indeed implies $v(x, y) \geq 0$, and $G(x, y)(0) \geq 1/2$ implies $v(x, y) \leq 0$. We will investigate the symmetry assumption, and the scope for response times to improve preference revelation under that assumption, in Section 3.2. First, however, we want to illustrate with a simple example that symmetry may sometimes not be an innocuous assumption at all.

Example 1. Consider a simple RUM with additive noise, where $\tilde{u}(x) = u(x) + \tilde{\epsilon}(x)$ and $\tilde{u}(y) = u(y) + \tilde{\epsilon}(y)$. The noise terms $\tilde{\epsilon}(x)$ and $\tilde{\epsilon}(y)$ are independent and have zero mean. It is often assumed in applied work that they follow an identical Gumbel distribution, with cumulative distribution function

$$H(z) = e^{-e^{-((z/\beta)+\gamma)}},$$

where $\beta > 0$ is a parameter and $\gamma \approx 0.5772$ is the Euler-Mascheroni constant. The variance of this distribution is $\beta^2\pi^2/6$. The difference $\tilde{v}(x, y) = \tilde{u}(x) - \tilde{u}(y)$ then follows a logistic distribution, which is easy to describe in closed form and generates the well-known logit choice probabilities (see also our discussion in Section 4 below). This distribution is indeed symmetric around its mean $v(x, y) = u(x) - u(y)$.

Horan et al. (2019) investigate the case where $\tilde{\epsilon}(x)$ and $\tilde{\epsilon}(y)$ are still independent, mean-zero Gumbel random variables, but with different variances. If the noise terms capture mistakes in the evaluation of the options, it is plausible that some options are more difficult to evaluate than others, resulting in option-specific variances. If the noise terms reflect random variation in the tastes of a given individual, or heterogeneity in a population of agents, there is no reason to assume identical distributions for different options either. However, the distribution of $\tilde{v}(x, y)$ is no longer symmetric for option-specific parameters β_x and β_y of the underlying Gumbel distributions. Unfortunately, this distribution cannot be described in closed form when $\beta_x \neq \beta_y$ (presumably one of the reasons why the literature has focussed on the knife-edge case where $\beta_x = \beta_y$).

The following numerical example is inspired by Example 3 in Horan et al. (2019). Suppose that $u(x) = 1$ and $u(y) = 3/4$, so $v(x, y) = 1/4 > 0$ and the individual strictly prefers x over y. Suppose furthermore that $\beta_x = 1$ and $\beta_y = 2$, for instance because y is more difficult to evaluate than x, and thus its utility is more noisy. Then it follows that the probability of choosing x is $p(x, y) = \text{Prob}[\tilde{u}(x) > \tilde{u}(y)] \approx 0.49 < 1/2$.

Consider now an analyst who does not know the data-generating process. Suppose that this analyst erroneously imposes the symmetry assumption when analyzing the SCF with the intention of uncovering the underlying preference. This analyst will correctly deduce $G(x, y)(0) \approx 0.51$, but then, applying symmetry, incorrectly conclude that $v(x, y)$ is strictly negative, i.e., that the SCF reveals a strict preference for y over x.

This example is meant to illustrate (*i*) that symmetry is not necessarily a plausible restriction, and (*ii*) that erroneously making the symmetry assumption can lead to wrong inferences about preferences.[7]

The following result shows that, if response times are available, it may be possible to learn a preference even in the unrestricted class of models. We first introduce the following new concepts. Given two cumulative distribution functions G and H on $\mathbb{R}_+$ and a constant $q \geq 1$, we say that G *q-first-order stochastically dominates* H (also written G q-FSD H) if

$$G(t) \leq q \cdot H(t) \text{ for all } t \geq 0.$$

If, additionally, the inequality is strict for some t, then G *strictly q-first-order stochastically dominates* H (written G q-SFSD H). For $q = 1$, these concepts coincide with the standard notions of first-order stochastic dominance. They are weaker requirements when $q > 1$, and possibly substantially so, because the dominating function G can lie above H to an extent only constrained by the ratio q. In particular, q-FSD implies q'-FSD whenever $q \leq q'$. Furthermore, for any two distributions G and H for which $G(t)/H(t)$ is bounded, we can always find a large enough q such that G q-FSD H.

[7]Other natural examples with asymmetric utility difference distributions can also be constructed in the class of random parameter models, where the decision-maker's preference parameters are stochastic (Apesteguía and Ballester, 2018).

Theorem 1. *Within the class of all RUM-CFs, a rationalizable SCF-RT reveals a preference for x over y if $F(y,x)$ q-FSD $F(x,y)$, and a strict preference if $F(y,x)$ q-SFSD $F(x,y)$, for $q = p(x,y)/p(y,x)$.*

Proof. Let $(u, \tilde{v}, r)$ be any RUM-CF which rationalizes an SCF-RT (p, f), and consider any $(x,y) \in D$. By (1), it holds that

$$1 - G(x,y)(r^{-1}(t)) = p(x,y)F(x,y)(t) \tag{2}$$

for all $t > 0$. Since (RUM.2) implies $1 - G(y,x)(v) = G(x,y)(-v)$, for $(y,x) \in D$ we analogously obtain

$$G(x,y)(-r^{-1}(t)) = p(y,x)F(y,x)(t) \tag{3}$$

for all $t > 0$. With the definition of $Q(x,y)(t) = (p(x,y)F(x,y)(t))/(p(y,x)F(y,x)(t))$ we therefore have

$$Q(x,y)(t) = \frac{1 - G(x,y)(r^{-1}(t))}{G(x,y)(-r^{-1}(t))} \tag{4}$$

for all $t > 0$.

Now suppose $F(y,x)$ q-FSD $F(x,y)$ for $q = p(x,y)/p(y,x)$. This can equivalently be written as $Q(x,y)(t) \geq 1$ for all $t > 0$. Hence it follows from (4) that

$$G(x,y)(-r^{-1}(t)) \leq 1 - G(x,y)(r^{-1}(t))$$

for all $t > 0$. We claim that this implies

$$G(x,y)(-v) \leq 1 - G(x,y)(v) \tag{5}$$

for all $v \geq 0$. The inequality follows immediately for any v for which there exists $t > 0$ such that $r^{-1}(t) = v$. For $v = 0$ it follows from continuity of $G(x,y)(v)$. For any v with $r(v) = 0$ it follows because in that case $G(x,y)(v) = 1$ and $G(x,y)(-v) = 0$, as otherwise the RUM-CF would generate an atom at the response time of zero.

Define a function $H : \mathbb{R} \to [0,1]$ by

$$H(v) = \begin{cases} 1 - G(x,y)(-v) & \text{if } v \geq 0, \\ G(x,y)(v) & \text{if } v < 0. \end{cases}$$

Observe that H is the cumulative distribution function for a distribution that is symmetric around zero and continuous except (possibly) for an atom at zero. Hence

$$\begin{aligned}\int_{-\infty}^{+\infty} v dH(v) &= \int_{(-\infty,0)} v dH(v) + \int_{(0,+\infty)} v dH(v) \\ &= \int_{-\infty}^{0} v g(x,y)(v) dv + \int_{0}^{+\infty} v g(x,y)(-v) dv \\ &= \int_{-\infty}^{0} v g(x,y)(v) dv - \int_{-\infty}^{0} v g(x,y)(v) dv = 0.\end{aligned}$$

Observe furthermore that (5) implies $G(x,y)$ 1-FSD H. Hence we have

$$v(x,y) = \int_{-\infty}^{+\infty} v dG(x,y)(v) \geq \int_{-\infty}^{+\infty} v dH(v) = 0, \tag{6}$$

i.e., a revealed preference for x over y.

If $F(y,x)$ q-SFSD $F(x,y)$ for $q = p(x,y)/p(y,x)$, then (5) is strict for some $v \geq 0$. Hence $G(x,y)$ 1-SFSD H and the inequality in (6) is strict, i.e., a revealed strict preference for x over y. □

The basic idea behind Theorem 1 is that the observable distributions of response times provide information about the unobservable distributions of utilities, based on the chronometric relationship.

To understand the precise condition, assume first that $q = p(x,y)/p(y,x) = 1$ for some $(x,y) \in D$, i.e., both options are equally likely to be chosen. Any RUM-CF that rationalizes this choice must satisfy $G(x,y)(0) = 1/2$. Furthermore, note that the distribution of $\tilde{v}(x,y)$ conditional on $\tilde{v}(x,y) > 0$ generates $F(x,y)$, and the distribution of $\tilde{v}(x,y)$ conditional on $\tilde{v}(x,y) < 0$ generates $F(y,x)$. Thus, if we additionally observe that $F(x,y)(t) = F(y,x)(t)$ for all $t \geq 0$, i.e., identical response time distributions for the two options, then we can conclude that the shape of the utility difference distribution must be identical on the positive and on the negative domain. This requires no knowledge of the properties of r beyond monotonicity. Hence we have *verified* that the distribution is symmetric around zero, so its mean $v(x,y)$ is zero. Our theorem indeed implies a revealed preference for x over y and for y over x in this case, which we also call a *revealed indifference between x and y*. If, by contrast, we observe that $F(y,x)$ 1-SFSD $F(x,y)$, i.e., the choice of y is systematically slower than the choice of x, we can conclude that the utility difference distribution is asymmetric and takes systematically larger absolute values on the positive than on the negative domain. Hence its mean $v(x,y)$ is strictly larger than zero, which translates into a revealed strict preference for x over y. Finally, if we observe $q = p(x,y)/p(y,x) > 1$, then to obtain a revealed preference for x over y it is sufficient that the choice of y is not too much faster than the choice of x, as captured by our concept of q-first-order stochastic dominance. If choice behavior is already indicative of a particular preference, then the response time distributions just need to confirm that the utility difference distribution is not strongly asymmetric in the reverse direction.[8]

Remark. We have stated Theorem 1 as providing a sufficient condition for preference revelation. In Sections 4 and 5 below, we will illustrate that this sufficient condition holds often and is hence useful. However, we could have chosen a different formal approach, casting the condition as an axiom and formulating the analysis in terms of data sets compatible with that axiom. Several results in Section 4 will show that entire subclasses of models (symmetric RUM-CFs and drift-diffusion models) would be covered with this approach. We prefer to state Theorem 1 as we do because we would not see a violation of the q-FSD condition as a fundamental problem for our approach. For instance, in RUM-CFs with asymmetric noise, it is possible that a given strict preference does not generate data fulfilling our sufficient condition. Therefore, a violation of the sufficient condition contradicts specific subclasses of models but not the general approach.

In the remainder of the section, we will discuss several implications of Theorem 1. First, the result can be extended by completing the revealed preferences in a transitive way. A formal statement and proof of this claim can be found in Appendix B.

Second, a long discussion in psychology has examined the question of whether, in tasks where errors can be identified objectively, those are faster or slower than correct responses. Although in many situations errors tend to be slower on average than correct responses, this is not always the case, and no general conclusion can be drawn.[9] Our result might help to clarify this relation for preferential choice. While the definition of error and correct response is not obvious ex-ante for preferential choice, ex-post we can call the choice of y a *revealed error* and the choice of x a *revealed correct response* when x is revealed to be strictly preferred over y. Translated into this language, it follows immediately that slow choices in the first-order stochastic dominance sense indeed reveal an error, provided that choice probabilities are at least minimally informative. However, our actual condition is weaker than first-order stochastic dominance, and hence it might be fulfilled even in cases where errors are faster on average. Thus, our approach is in principle compatible with the seemingly puzzling evidence from psychology.

[8]Theorem 1 and our further results focus on uncovering the sign of the mean of $\tilde{v}(x, y)$, because this mean equals $u(x) - u(y)$ and therefore informs about the ordinal preferences represented by u, for either normative or positive reasons. Potentially, one could be interested in uncovering also other summary statistics of $\tilde{v}(x, y)$, and our tools may be helpful for that purpose, but the relevance of other statistics is not obvious from the point of view of revealed preference theory.

[9]For an overview of classical results, see Luce, 1986, Section 6.4.3. The picture is further complicated if decisions are subject to extraneous impulsive tendencies, as e.g. alternative decision processes reflecting underlying biases. For instance, assumptions linking response times to whether an answer is more intuitive or more deliberative are common in the literature on dual-process thinking (e.g., Kahneman, 2003), but, as pointed out by Krajbich et al. (2015), apparent results in this direction might sometimes hide chronometric effects. Additionally, Achtziger and Alós-Ferrer (2014) show that a simple dual-process model predicts that errors might be either faster or slower than correct responses, depending on whether the underlying processes are in conflict or aligned.

Third, we have described SCF-RTs by unconditional choice probabilities and conditional response time distributions for each choice. This is the natural extension of SCFs and allowed us to work out the intuition for our result. Alternatively, we could have described the joint distribution over choices and response times by an unconditional response time distribution and conditional choice probabilities for each response time. Let $P(x,y)(t)$ denote the probability of a choice of x over y conditional on choice taking place before time t. The ratio of these probabilities can be calculated as

$$Q(x,y)(t) = \frac{P(x,y)(t)}{P(y,x)(t)} = \frac{p(x,y)F(x,y)(t)}{p(y,x)F(y,x)(t)}.$$

Hence, the condition that $F(y,x)$ q-FSD $F(x,y)$ for $q = p(x,y)/p(y,x)$ in Theorem 1 is equivalent to

$$Q(x,y)(t) \geq 1 \text{ for all } t > 0. \tag{7}$$

An analogous formulation holds for the strict case. For a revealed preference without distributional assumptions, we can thus also check if x is more likely to be chosen than y *before all times* t. The simple requirement $p(x,y) \geq p(y,x)$ obtains as a special case of this in the limit as $t \to \infty$.

The formulation based on $Q(x,y)(t)$ suggests a natural but substantially stronger condition. Let $p(x,y)(t)$ denote the probability of a choice of x over y conditional on choice taking place at time t (rather than before t). We obtain the ratio

$$q(x,y)(t) = \frac{p(x,y)(t)}{p(y,x)(t)} = \frac{p(x,y)f(x,y)(t)}{p(y,x)f(y,x)(t)},$$

and can state the following corollary to Theorem 1.

Corollary 1. *Within the class of all RUM-CFs, a rationalizable SCF-RT reveals a preference for x over y if $q(x,y)(t) \geq 1$ for almost all $t \geq 0$, and a strict preference if, additionally, the inequality is strict for a set of t with positive Lebesgue measure.*

The condition that x is more likely to be chosen than y *at almost all times* t can be interpreted as a requirement of stochastic consistency across response times. This is clearly stronger than (7). Appendix C contains an example in which $p(x,y)(t) < p(y,x)(t)$ holds for an interval of response times, so Corollary 1 is not applicable, but a strict preference for x over y is still revealed by Theorem 1. Hence our main criterion arrives at a conclusion even though behavior displays some stochastic inconsistency across response times.[10]

The results in this section are interesting for two main reasons. First, for the analyst who is reluctant to make distributional assumptions in the context of random utility mod-

[10]Notice a similarity to Bernheim and Rangel (2009), who require agreement of choices across choice sets or frames to obtain a revealed preference. A first difference is that we study stochastic choice and contemplate probabilistic agreement of choices across response times. A second difference is that our main criterion can reveal a preference even if there is no such agreement.

els, Theorem 1 provides a robust criterion for preference revelation. Based on observed response times, it is often possible to deduce preferences without such assumptions. The criterion may lead to an incomplete revelation of preferences (we will return to this issue in Sections 4 and 5), but it avoids making mistakes like those illustrated in Example 1. Second, our criterion may be able to arbitrate if choice behavior violates stochastic transitivity (Tversky, 1969; Rieskamp et al., 2006; Tsetsos et al., 2016). For example, assume we observe a stochastic choice cycle with $p(x,y) > 1/2$, $p(y,z) > 1/2$, and $p(z,x) > 1/2$. Such a cycle cannot be rationalized by any model with symmetric utility distributions, but it may be rationalizable by a model with asymmetric utility distributions. In that case, at most two of the three binary choices can reveal a preference, showing which part of the cycle reflects true preferences. A similar argument applies if choices are affected by framing and we observe $p^f(x,y) > p^f(y,x)$ under frame f but $p^{f'}(x,y) < p^{f'}(y,x)$ under frame f'. Again, our response time criterion may be able to detect which frame induces choices that are probabilistically more in line with the true preferences. In other words, since a choice probability above 1/2 is not yet sufficient to reveal a preference, RUM-CFs can explain cyclic or frame-dependent *choices* but are inconsistent with data sets for which our condition reveals cyclic or frame-dependent *preferences*. If the latter were observed, our model would be falsified.

3.2 The Symmetric Case

The assumption of symmetry is often accepted in the literature. Formally, a RUM $(u, \tilde{v})$ or RUM-CF $(u, \tilde{v}, r)$ is *symmetric* if each random variable $\tilde{v}(x,y)$ follows a distribution that is symmetric around its mean $v(x,y)$, that is, if for each $(x,y) \in C$ and all $\delta \geq 0$,

$$g(x,y)(v(x,y) + \delta) = g(x,y)(v(x,y) - \delta).$$

In contrast to Proposition 1, this assumption allows to learn preferences from observed choice probabilities.

Proposition 2. *Within the class of symmetric RUMs, a rationalizable SCF reveals a preference for x over y if $p(x,y) \geq p(y,x)$, and a strict preference if $p(x,y) > p(y,x)$.*

This result is both simple and well-known, and we include a proof in the appendix only for completeness.[11] It allows us to deduce either a strict preference or an indifference for each observed choice pair, and it can be extended by completing the revealed preferences in a transitive way (see Appendix B).

Note that every preference which can be learned with the help of response times without distributional assumptions can also be learned without response times at the price of making the (possibly unwarranted) symmetry assumption.[12] But even if one is

[11] Its first statement in the economics literature that we are aware of is Manski (1977), but an earlier, closely related statement for general stochastic choice can be found already in Edwards (1954).

[12] This is true because $p(x,y) \geq p(y,x)$ is a necessary condition for a revealed preference for x over y according to Theorem 1. The statement holds for weak but not necessarily for strict preferences. We can

willing to make the symmetry assumption, the addition of response times again improves what can be learned about preferences, as the following result will show. It is based on triangulating a preference indirectly through comparisons with a third option. For each $(x, y) \in D$ with $p(x, y) > p(y, x)$, define $\theta(x, y)$ as the $0.5/p(x, y)$-percentile of the response time distribution of x, i.e.,

$$F(x, y)(\theta(x, y)) = \frac{0.5}{p(x, y)}.$$

The percentile $\theta(x, y) > 0$ combines information about choice probabilities and response times. It is larger than the median but converges to the median when $p(x, y) \to 1$. More generally, $\theta(x, y)$ becomes smaller as $p(x, y)$ becomes larger or as the choice of x becomes faster in the usual first-order stochastic dominance sense. Hence a small value of $\theta(x, y)$ is indicative of a strong preference for x over y. Comparison of these percentiles can make it possible to learn preferences for unobserved pairs $(x, y) \in C \setminus D$ even when transitivity is void.

Theorem 2. *Within the class of symmetric RUM-CFs, a rationalizable SCF-RT reveals a preference for x over y, where $(x, y) \in C \setminus D$, if there exists $z \in X$ such that $\theta(x, z) \leq \theta(y, z)$ or $\theta(z, x) \geq \theta(z, y)$, and a strict preference if $\theta(x, z) < \theta(y, z)$ or $\theta(z, x) > \theta(z, y)$.*

Proof. Let $(u, \tilde{v}, r)$ be any symmetric RUM-CF which rationalizes an SCF-RT (p, f). We first claim that, for any $(x, y) \in D$ with $p(x, y) > p(y, x)$, and hence $v(x, y) > 0$ by symmetry, it holds that $\theta(x, y) = r(v(x, y))$. To see the claim, note that from (1) we obtain

$$p(x, y)F(x, y)(t) = 1 - G(x, y)(r^{-1}(t))$$

for all $t > 0$. Evaluated at $t = r(v(x, y))$, which is well-defined because $v(x, y) > 0$ and strictly positive because otherwise the RUM-CF would generate an atom at the response time of zero, this yields

$$p(x, y)F(x, y)(r(v(x, y))) = 1 - G(x, y)(v(x, y)). \tag{8}$$

The RHS of (8) equals 0.5 by symmetry. Hence we obtain

$$F(x, y)(r(v(x, y))) = \frac{0.5}{p(x, y)},$$

and, by definition of $\theta(x, y)$, it follows that $\theta(x, y) = r(v(x, y))$, proving the claim.

Consider now any $(x, y) \in C \setminus D$ for which there exists $z \in X$ with $\theta(x, z) \leq \theta(y, z)$, and hence $0 < r(v(x, z)) \leq r(v(y, z))$ by the above claim. Since r is strictly decreasing

indeed have a revealed strict preference for x over y according to Theorem 1 and a revealed indifference according to Proposition 2, in case $p(x, y) = p(y, x) = 1/2$ and $F(y, x)$ 1-SFSD $F(x, y)$. Any symmetric RUM that rationalizes such an SCF must have $v(x, y) = 0$. However, there is no symmetric RUM-CF that rationalizes the SCF-RT, due to the asymmetric response times. All rationalizing RUM-CFs must be asymmetric and have $v(x, y) > 0$.

in v whenever $r(v) > 0$, it follows that $u(x) - u(z) = v(x, z) \geq v(y, z) = u(y) - u(z)$ and hence $v(x, y) = u(x) - u(y) \geq 0$, i.e., a revealed preference for x over y. If $\theta(x, z) < \theta(y, z)$, all the inequalities must be strict, so the revealed preference is strict. The case where $\theta(z, x) \geq \theta(z, y)$ or $\theta(z, x) > \theta(z, y)$ is analogous. □

It has been observed before that response times can be used to infer preferences for unobserved choices. Krajbich et al. (2014) argue that a slow choice of z over x combined with a quick choice of the same z over y reveals a preference for x over y, even though the choice between x and y is not directly observed and a transitivity argument is not applicable. Based on the chronometric relationship, the positive utility difference $u(z) - u(x)$ must be smaller than the positive utility difference $u(z) - u(y)$, which implies $u(x) > u(y)$. It remained unclear how to generalize the idea to a stochastic framework. Our Theorem 2 answers this question. The condition $\theta(z, x) \geq \theta(z, y)$ is the appropriate formulation of a stochastic choice of z over x being slower than a stochastic choice of z over y. Of course, an analogous argument applies to a quick choice of x over z combined with a slow choice of y over z, as formalized by our alternative condition $\theta(x, z) \leq \theta(y, z)$. Also, not too surprisingly, the preferences revealed by Theorem 2 can further be completed in a transitive way (see Appendix B).

Importantly, we need to compare specific percentiles of the response time distributions that depend on choice probabilities, and not just mean, median or maximum response times. Up to date, empirical applications of response times have overwhelmingly been based on the use of a particular statistic, stating that, say, median response times are shorter or larger for certain types of decisions. Our analysis shows that such an approach is but a first approximation that we can strongly improve upon. Focusing on a particular statistic like the median treats all observations equally and neglects fundamental information, namely choice frequencies. As already in the case of Theorem 1, our analysis again combines information of frequencies and response times. The condition in Theorem 2 might well dictate to compare the 0.85-percentile of the distribution of response times for one choice to the 0.55-percentile of the distribution for a different choice.

The results in this section enable first out-of-sample predictions. Consider an unobserved choice problem $(x, y) \in C \setminus D$. If, based on Theorem 2, the SCF-RT reveals a strict preference for x over y in the class of symmetric RUM-CFs, then we predict that $p(x, y) > p(y, x)$, because each symmetric model with $v(x, y) > 0$ generates such choice probabilities. If the SCF-RT reveals an indifference between x and y, then we can even predict the precise probabilities $p(x, y) = p(y, x) = 1/2$. We will test these predictions in our empirical application in Section 5.

3.3 The Fechnerian Case

Microeconometric models of random utility assume even more structure. For instance, the prominent probit or logit models are special cases of *Fechnerian* models, which go back to the representation result by Debreu (1958). A RUM $(u, \tilde{v})$ or RUM-CF $(u, \tilde{v}, r)$ is Fechnerian if the distribution of each random variable $\tilde{v}(x, y)$ has the same symmetric shape, which is just shifted so that its expected value becomes $v(x, y)$. Formally, there exists a common density g that is symmetric around zero and has full support, i.e., $g(\delta) = g(-\delta) > 0$ for all $\delta \geq 0$, such that, for each $(x, y) \in C$ and all $v \in \mathbb{R}$,

$$g(x, y)(v) = g(v - v(x, y)).$$

This additional structure makes it possible to deduce preferences through comparison with a third option, relying only on choice probabilities.

Proposition 3. *Within the class of Fechnerian RUMs, a rationalizable SCF reveals a preference for x over y, where $(x, y) \in C \setminus D$, if there exists $z \in X$ such that $p(x, z) \geq p(y, z)$, and a strict preference if $p(x, z) > p(y, z)$.*

As in the case of Proposition 2, this result is well-known, and we provide a short proof in the appendix only for completeness.[13] The transitive closure extension can be found in Appendix B.

By Proposition 3 we obtain a revealed preference for an unobserved pair $(x, y) \in C \setminus D$ whenever $(x, z), (y, z) \in D$ for some third option z. Hence, imposing the Fechnerian assumption enables an exhaustive elicitation of ordinal preferences even outside the choice data set, without the use of response times (provided that the assumption is valid). We now show that the use of response times makes it possible to move beyond ordinal preferences and make out-of-sample predictions of precise choice probabilities.

Definition 8. Within a class of models, a rationalizable SCF-RT *predicts choice probability* $\bar{p}(x, y)$ for a non-observed choice $(x, y) \in C \setminus D$ if all RUM-CFs in the class that rationalize it satisfy $\text{Prob}[\tilde{v}(x, y) > 0] = \bar{p}(x, y)$.

Theorem 3. *Within the class of Fechnerian RUM-CFs, a rationalizable SCF-RT predicts a choice probability for each $(x, y) \in C \setminus D$ for which there exists $z \in X$ with $(x, z), (y, z) \in D$. The prediction is*

$$\bar{p}(x, y) = \begin{cases} p(x, z)F(x, z)(\theta(y, z)) & \text{if } p(y, z) > 1/2, \\ p(x, z) & \text{if } p(y, z) = 1/2, \\ 1 - p(z, x)F(z, x)(\theta(z, y)) & \text{if } p(y, z) < 1/2. \end{cases}$$

Proof. Let $(u, \tilde{v}, r)$ be any Fechnerian RUM-CF which rationalizes an SCF-RT (p, f). For any fixed $(x, y) \in C$, this particular RUM-CF predicts

$$p(x, y) = \text{Prob}[\tilde{v}(x, y) > 0] = G(y, x)(0) = G(v(x, y)). \quad (9)$$

[13]The argument can be traced back to Fechner (1860) and Thurstone (1927). Within economics, it has been spelled out e.g. in Ballinger and Wilcox (1997).

Let $(x,y) \in C \setminus D$ and $z \in X$ such that $(x,z),(y,z) \in D$. We distinguish three cases.

Case 1: $p(y,z) > 1/2$. From the proof of Theorem 2 we already know that $\theta(y,z) = r(v(y,z))$. From (1) we obtain

$$p(x,z)F(x,z)(t) = 1 - G(x,z)(r^{-1}(t))$$

for all $t > 0$. Hence, by (RUM.2) and the Fechnerian assumption,

$$p(x,z)F(x,z)(t) = G(z,x)(-r^{-1}(t)) = G(-r^{-1}(t) - v(z,x)) = G(v(x,z) - r^{-1}(t))$$

for all $t > 0$, which for $t = r(v(y,z))$ yields

$$p(x,z)F(x,z)(r(v(y,z))) = G(v(x,z) - v(y,z)) = G(v(x,y)).$$

Combined with (9) and the above expression for $\theta(y,z)$, this implies

$$p(x,y) = p(x,z)F(x,z)(\theta(y,z)),$$

which is the model-independent prediction $\bar{p}(x,y)$ given in the statement.

Case 2: $p(y,z) = 1/2$. It follows from Proposition 2 that $v(y,z) = 0$. We obtain

$$p(x,y) = G(v(x,y)) = G(v(x,z) - v(y,z)) = G(v(x,z)) = p(x,z),$$

which is the model-independent prediction $\bar{p}(x,y)$ given in the statement.

Case 3: $p(y,z) < 1/2$. It follows from Proposition 2 that $v(z,y) > 0$. Following the same steps as in Case 1 but with reversed order for each pair of alternatives yields the model-independent prediction

$$\bar{p}(x,y) = 1 - \bar{p}(y,x) = 1 - p(z,x)F(z,x)(\theta(z,y)),$$

as given in the statement. □

To understand the probability formula in the theorem, just consider the case where $p(x,z) > p(y,z) > 1/2$. Then $u(x) > u(y) > u(z)$ must hold under the Fechnerian assumption, where the first inequality follows from Proposition 3 and the second inequality follows from Proposition 2. Hence we can conclude that the unknown $p(x,y)$ must be strictly smaller than $p(x,z)$, because Fechnerian choice probabilities are strictly monotone in the underlying utility differences $v(\cdot,\cdot)$ across binary choice problems. The theorem now shows that a prediction for $p(x,y)$ can be obtained by multiplying the observed $p(x,z)$ with a discounting factor $F(x,z)(\theta(y,z))$. This factor is an observable, response-time-based indicator of the relative position of $u(y)$ within the interval $[u(x),u(z)]$.

Combined with the Fechnerian assumption, the use of response times allows to predict exact choice probabilities out-of-sample. Without response times, making such a prediction would require assuming a complete and specific functional form for the utility distribution.[14] Hence, in analogy to our earlier results, response times again serve as a substitute for stronger distributional assumptions.[15] Further, as was the case for Theorems 1 and 2, Theorem 3 extracts information from the data by balancing choice frequencies and response times.

4 Behavioral Models from Economics and Psychology

Theorem 1 provided a sufficient condition for preference revelation without distributional assumptions. This analysis left open two questions. First, which SCF-RTs are rationalizable by RUM-CFs? And second, how tight is the sufficient condition? In this section, we try to answer these questions by studying SCF-RTs which are generated by specific behavioral models from the literature. We will do this first for standard RUMs from economics and microeconometrics to which we add chronometric functions, and second for standard sequential sampling models from psychology and neuroscience.

4.1 The View from Economics

Specific RUMs are commonly used for microeconometric estimation using choice data. These models typically start from a utility function $u : X \to \mathbb{R}$ and add an error term with zero mean to each option, such that the overall utility of $x \in X$ is given by a random variable

$$\tilde{u}(x) = u(x) + \tilde{\epsilon}(x).$$

As a next step, even more specific distributional assumptions are imposed. Two popular examples are the probit and the logit model.

In the probit model, the errors are assumed to be normally distributed and i.i.d. across the options. The distribution of the random utility difference $\tilde{v}(x, y) = \tilde{u}(x) - \tilde{u}(y)$ is then also normal and can be described by

$$G(x, y)(v) = \Phi\left(\frac{v - v(x, y)}{\sigma}\right),$$

where σ is a standard deviation parameter and Φ is the cdf of the standard normal distribution. This simple specification gives rise to a Fechnerian model. Generalizations allow for heteroscedasticity or correlation between the error terms of different options,

[14] Some weaker inequality predictions of choice probabilities are already possible under an additive RUM structure without additional distributional assumptions (see e.g. Cohen and Falmagne, 1990).

[15] For the analyst who is willing to make the strong distributional assumptions required, for example, by probit or logit models, response times have no additional value when the available data on choice is rich. However, the literature has shown that the use of response times can be valuable even in the context of logit or probit models when choice data is scarce (e.g. Clithero, 2018; Konovalov and Krajbich, 2019). Our paper differs from these studies by showing that response time data enable the recovery of preferences even when rich choice data cannot recover them, that is, in the absence of untestable assumptions on utility noise.

in which case the parameter σ becomes choice-set-dependent, written $\sigma(x, y)$. Such a generalized model is no longer Fechnerian but still symmetric.

In the logit model, the errors are assumed to follow a Gumbel distribution, again i.i.d. across the options. In that case, the random utility difference follows a logistic distribution described by

$$G(x, y)(v) = \left[1 + e^{-\left(\frac{v - v(x,y)}{\beta}\right)}\right]^{-1},$$

where β is a scale parameter. This model is again Fechnerian, whereas one could think of generalizations where the scale parameter becomes choice-set-dependent, in which case it is no longer Fechnerian but still symmetric.

We now treat an arbitrary symmetric RUM-CF as the real data-generating process and apply our preference revelation method to the resulting data of choices and response times. It is trivial that this data is rationalizable within the class of symmetric models, and hence within the class of all models. More surprising is the fact that our sufficient condition from Theorem 1 always recovers the correct preferences from the data.

Proposition 4. *Consider an SCF-RT (p, f) that is generated by a symmetric RUM-CF $(u, \tilde{v}, r)$. Then, for any $(x, y) \in D$, $u(x) \geq u(y)$ implies $F(y, x)$ q-FSD $F(x, y)$, and $u(x) > u(y)$ implies $F(y, x)$ q-SFSD $F(x, y)$, for $q = p(x, y)/p(y, x)$.*

If we are given a data set that is generated by one of the random utility models commonly employed in the literature, augmented by a chronometric function, then our cautious revealed preference criterion always recovers the correct preferences, despite not using information about the specific data-generating process. Even the analyst who believes in the probit or logit distribution can work with our criterion. It will yield the same revealed preferences as an application of the full-fledged model if his belief is correct, but avoids a mistake if his belief is incorrect.

We add three remarks on this result. First, symmetry of the RUM-CF does not guarantee that also the stronger condition in Corollary 1 is satisfied by the SCF-RT.[16] Second, symmetry of the RUM-CF is not necessary for our condition in Theorem 1 to be applicable. As the proof of Proposition 4 reveals, there are also asymmetric models that generate data from which our criterion recovers preferences correctly. Third, and most importantly, since Proposition 4 shows that any data set generated by a symmetric RUM-CF must fulfill the condition in Theorem 1, the latter becomes an empirical test of the former. That is, if the condition in Theorem 1 is frequently violated in a real data set, this suggests that symmetric RUM-CFs are not an appropriate model for the process generating that dataset.

[16] The example of an SCF-RT presented in Appendix C, which violates this stronger condition, is actually generated by a symmetric RUM-CF. This RUM-CF features a bimodal utility difference distribution. It is possible to show that symmetry and unimodality together imply that the stronger sufficient condition in Corollary 1 is always satisfied. Both the probit and the logit model are unimodal, so working with the stronger condition comes with no loss if either of these models is the data-generating process.

We can go one step further and assume that there is a second source of noise in response times, on top of the noise already generated by random utility. The noise could be part of the behavioral model, e.g. due to a stochastic chronometric function or randomness in the physiological process implementing the response, or it could be due to imperfect observation by the analyst. We assume that this additional noise is independent from the randomness in utility, so that it does not systematically reverse the chronometric relationship. A common approach in the empirical literature (e.g., Chabris et al., 2009; Fischbacher et al., 2013; Alós-Ferrer and Ritschel, 2018) is to add i.i.d. noise to log response times, where taking the log ensures that response times remain non-negative. An equivalent way of modelling this is by means of multiplicative noise, which is technically convenient for our purpose. Formally, a *random utility model with a noisy chronometric function* (RUM-NCF) can be obtained from a RUM-CF by letting the response time when x is chosen over y become the random variable

$$\tilde{t}(x,y) = r(|\tilde{v}(x,y)|) \cdot \tilde{\eta},$$

conditional on $\tilde{v}(x,y) > 0$. Here, $\tilde{\eta}$ is a non-negative random variable with mean one, assumed to be i.i.d. according to a density h on $\mathbb{R}_+$. The probability of a response time of at most t, conditional on x being chosen over y, is now the probability that the realized utility difference is at least $r^{-1}(t/\tilde{\eta})$, conditional on that difference being positive. Hence, for an SCF-RT (p, f) that is generated by a RUM-NCF $(u, \tilde{v}, r, \tilde{\eta})$ we have

$$F(x,y)(t) = \frac{\int_0^\infty \left[1 - G(x,y)(r^{-1}(t/\eta))\right] h(\eta) d\eta}{1 - G(x,y)(0)}$$

for all $t > 0$ and all $(x, y) \in D$, which is analogous to equation (1).

Our preference revelation approach based on RUM-CFs is misspecified when the real data-generating process is a RUM-NCF, because the additional noise is erroneously explained by additional randomness in utility. However, as the next proposition will show, this misspecification is often inconsequential. For the entire class of SCF-RTs that are generated by (and hence are rationalizable in) the class of symmetric RUM-NCFs with full support utility distributions (like probit or logit) and arbitrary response time noise distributions, our previous condition remains the correct criterion for preference revelation.[17]

Proposition 5. *Consider an SCF-RT (p, f) that is generated by a symmetric RUM-NCF $(u, \tilde{v}, r, \tilde{\eta})$ where each $\tilde{v}(x, y)$ has full support. Then, for any $(x, y) \in D$, $u(x) \geq u(y)$ implies $F(y, x)$ q-FSD $F(x, y)$, and $u(x) > u(y)$ implies $F(y, x)$ q-SFSD $F(x, y)$, for $q = p(x, y)/p(y, x)$.*

[17]Without full support of the utility difference distributions, some response times may arise only because of the additional noise but could never be generated by a realized utility difference. The distribution of those response times does not obey the chronometric relationship and would be uninformative of utility.

The proof rests on the insight that q-FSD is invariant to independent perturbations. Whenever an SCF-RT (p, f) satisfies $F(y, x)$ q-FSD $F(x, y)$, then the SCF-RT $(p, \hat{f})$ obtained after perturbing response times by log-additive or multiplicative noise still satisfies $\hat{F}(y, x)$ q-FSD $\hat{F}(x, y)$. The case of RUM-NCFs obtained from symmetric RUM-CFs is a natural application of this insight. However, the robustness of our preference revelation criterion holds more generally for perturbations of any data-generating process for which the criterion has bite. This could be a random utility model that is not symmetric, or it could be one of the sequential sampling models studied in the next section.

4.2 The View from Psychology

A different way of generating stochastic choices and response times is by means of a sequential sampling model as used extensively in psychology and neuroscience. The basic building block for binary choice problems is the *drift-diffusion model* (DDM) of Ratcliff (1978). A *DDM with constant boundaries* is given by a drift rate $\mu \in \mathbb{R}$, a diffusion coefficient $\sigma^2 > 0$, and symmetric barriers B and $-B$ with $B > 0$. A stochastic process starts at $Z(0) = 0$ and evolves over time according to a Brownian motion

$$dZ(t) = \mu dt + \sigma dW(t).$$

The process leads to a choice of x (resp. y) if the upper (resp. lower) barrier is hit first, the response time being the time at which this event occurs.

Although the DDM, as a model anchored in psychology, usually does not make reference to underlying utilities, $Z(t)$ is often interpreted as the difference in spiking rates between neurons computing values for the competing options. Hence it is natural to introduce a link to utility by assuming that the drift rate is determined such that $\mu = \mu(x, y) = -\mu(y, x) \geq 0$ if and only if $v(x, y) \geq 0$. This way, the DDM generates stochastic choices and response times from an underlying deterministic utility function $u : X \to \mathbb{R}$.

The stochastic path of $Z(t)$ can be interpreted as the accumulation of evidence in favor of one or the other option as the brain samples past (episodic) information. Recent research has shown that evidence-accumulation models like the DDM actually represent optimal decision-making procedures under neurologically founded constraints, but optimality requires that the barriers are not constant but rather collapse towards zero as t grows to infinity (Tajima et al., 2016). Similarly, Fudenberg et al. (2018) model optimal sequential sampling when utilities are uncertain and gathering information is costly, and find that the range for which the agent continues to sample should decrease as t grows. A partial intuition for this result is that a value of $Z(t)$ close to zero for small t carries little information, while a value of $Z(t)$ close to zero for large t indicates that the true utilities are most likely close to each other, and hence sampling further evidence has little value. To reflect this idea, a *DDM with collapsing boundaries* works in the exact same way as a DDM with constant boundaries, with the only difference that the barriers are given by a continuous and strictly decreasing function $b : \mathbb{R}_+ \to \mathbb{R}_{++}$ such that $\lim_{t\to\infty} b(t) = 0$. That is, x is chosen if $Z(t)$ hits the upper barrier $b(t)$ before hitting the lower barrier $-b(t)$, and y is chosen if the converse happens, with the response time being the first crossing time (see Figure 2).

We now treat a DDM with an underlying utility function as the real data-generating process and again apply our preference revelation method to the resulting data of choices and response times.[18] The following proposition, the proof of which relies on a result by Fudenberg et al. (2018), shows that our sufficient condition from Theorem 1 is again tight and always recovers the correct preferences.

Proposition 6. *Consider an SCF-RT (p, f) that is generated by a DDM with constant or collapsing boundaries and underlying utility function u. Then, $u(x) \geq u(y)$ implies $F(y,x)$ q-FSD $F(x,y)$, and $u(x) > u(y)$ implies $F(y,x)$ q-SFSD $F(x,y)$, for $q = p(x,y)/p(y,x)$.*

To understand the intuition behind the proof, recall from our discussion of Theorem 1 that the condition ensuring a revealed preference can be reformulated as $P(x,y)(t) \geq P(y,x)(t)$ for all t, that is, the probability of choosing x over y before any pre-specified response time t should be larger than the probability of choosing y over x before t. In a DDM, $u(x) > u(y)$ means that the drift rate $\mu(x,y) > 0$ favors x, since the upper barrier reflects a choice of x. Hence the probability of hitting the upper barrier first before any pre-specified response time t is indeed larger than the probability of hitting the lower barrier first. The proof actually establishes that the stronger condition in Corollary 1 is always satisfied by an SCF-RT that is generated by a DDM. Furthermore, the result would continue to hold for DDMs with more general boundary functions that are not necessarily constant or collapsing, but these models do not always generate well-behaved choices and have received less attention in the literature.

We remark that, as in the case of Proposition 4 above, Proposition 6 is of independent interest. Since it states that any data set generated by a DDM must fulfill the condition in Theorem 1, the latter becomes an empirical test of the former. That is, if the condition in Theorem 1 is frequently violated in a real data set, this suggests that drift-diffusion models are not an appropriate model for the process generating that data set.

We conclude the section with remarks about rationalizability. An SCF-RT that is generated by a DDM is always rationalizable within the class of all RUM-CFs when we consider just two options, $D = \{(x,y),(y,x)\}$, which is the setting in which DDMs are typically applied. This is a corollary of the following more general result.

[18] For the case of constant boundaries, closed-form solutions for choice probabilities and response time distributions generated by the DDM are known, see e.g. Palmer et al. (2005). Closed-form solutions are not available for the case of collapsing boundaries. Webb (2019) explores the link between bounded accumulation models as the DDM and random utility models and shows how to derive distributional assumptions for realized utilities of the latter if the true data-generating process is of the DDM form.

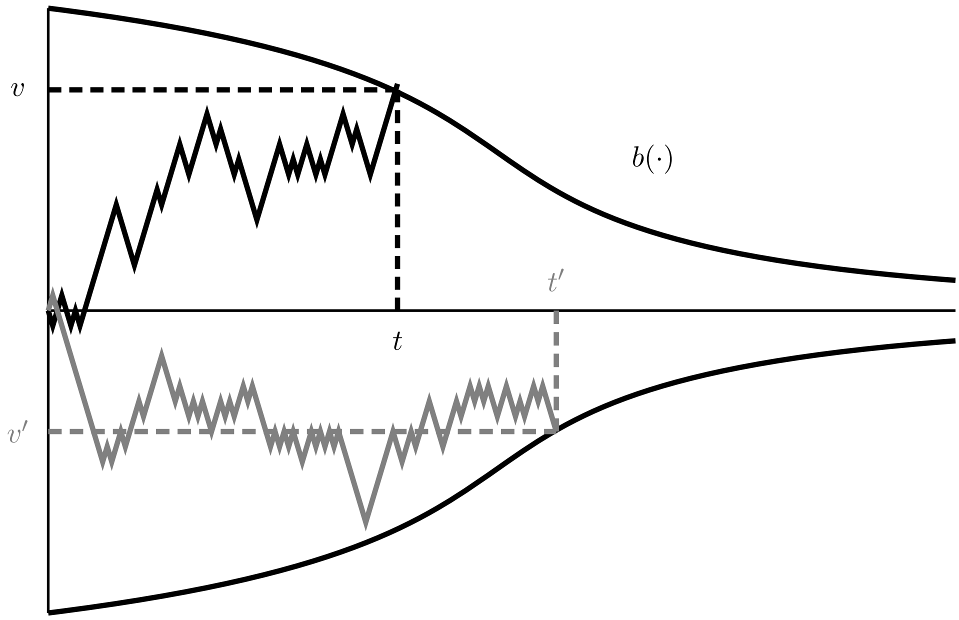

Figure 2: Illustration of a DDM with collapsing decision boundaries.

Proposition 7. *Suppose $D = \{(x, y), (y, x)\}$. Then, any SCF-RT (p, f) is rationalizable within the class of all RUM-CFs.*

In the proof, we fix an arbitrary chronometric function r that satisfies Definition 5 (and has $r(v) = 0$ for large v) and construct an associated density $g(x, y)$ such that the data is rationalized. The construction is particularly illuminating for an SCF-RT that is generated by a DDM with collapsing boundaries. In that case, it is very natural to choose the chronometric function of the RUM-CF as the inverse of the boundary function of the DDM. The associated density $g(x, y)$ then describes the distribution of the value of $Z(t)$ at the endogenous decision point (see again Figure 2). This interpretation is in line with Fudenberg et al. (2018), who argue that $Z(t)$ is a signal about the true utility difference. In other words, if we think of the chronometric function as the inverse of the collapsing boundary function, then realized utility differences can be interpreted as realized signals about the underlying deterministic utility difference.

With more than two options, the question of rationalizability is less straightforward. It is not clear that the above construction yields utility difference distributions that are consistent with a single utility function u. For SCF-RTs generated by DDMs, this problem is further complicated by the fact that there is no agreed upon discipline on how utility differences $v(x, y)$ map into drift rates $\mu(x, y)$, beyond the basic ordinal requirement that $\mu(x, y) \geq 0$ if and only if $v(x, y) \geq 0$. Hence, rationalizability of an arbitrary DDM-generated data set in terms of a RUM-CF is not guaranteed. Our approach, however, can be generalized naturally to solve this problem. Define a *generalized random utility model with a chronometric function* (GRUM-CF) by replacing (RUM.1) in Definition 2 by the weaker requirement

(GRUM.1) $\mathbb{E}[\tilde{v}(x, y)] \geq 0 \;\Leftrightarrow\; u(x) - u(y) \geq 0,$

keeping everything else unchanged. That is, realized utility differences (or "decision values" as in the recent neuroeconomics literature, see e.g. Glimcher and Fehr, 2013, Chapters 8–10) are interpreted as signals about ordinal preferences rather than cardinal utility. It is easy to see that Theorem 1 remains valid for preference revelation within the class of all GRUM-CFs. Furthermore, any SCF-RT generated by a DDM with constant or collapsing boundaries and underlying utility function u is rationalizable in the class of all GRUM-CFs.[19]

5 Empirical Application

In this section, we apply our results to an experiment by Clithero (2018), whom we thank for sharing the data. In the first phase of this experiment, 31 subjects made repeated binary choices between various snack food items and one fixed reference item. Each subject made decisions for 17 such items, and each decision was repeated 10 times. Response times were recorded. In a second phase of the experiment, each subject was confronted once with each of the possible binary choices between the 17 non-reference items from the first phase, resulting in $17 \times 16/2 = 136$ choice problems per subject.

Our strategy of analysis is as follows. First, we use the choice and response time data from the first phase of the experiment to illustrate Theorem 1. We show that the q-FSD condition is fulfilled in substantially more than half of the choice problems. That is, Theorem 1 has significant empirical bite, allowing for preference identification in the absence of any assumptions on utility noise. At the same time, there remains a non-negligible fraction of choice problems in which q-FSD is violated, indicating that these choices are not typical realizations of RUM-CFs with symmetric noise distributions or of drift-diffusion models. Second, we take full advantage of the experiment's design, where choices in the first phase are always made against the same reference item, and use Theorem 2 to derive a complete preference ordering for each subject, relying only on the choice and response time data from the first phase. We then use these preferences to predict choices in the second phase of the experiment. That is, we check the accuracy of the out-of-sample predictions embodied in Theorem 2, which rest on no distributional assumptions other than symmetry. The prediction accuracy is remarkable. It clearly outperforms the accuracy of conventional logit models, and it is comparable to fully parametric, computation-intensive drift-diffusion models. Third, we conduct an analogous analysis relying on Theorem 3 to predict choice probabilities for the second phase of the experiment, assuming a Fechnerian structure of the utility noise. We again achieve a remarkable prediction accuracy.

[19]This approach could be extended even further, building a connection to the economic literature on consumer theory without transitive preferences. Shafer (1974) showed that every complete, continuous, and strongly convex binary relation R (not necessarily transitive) on a Euclidean space can be represented by a continuous, real-valued, two-variable function v such that xRy if and only if $v(x,y) \geq 0$, with $v(x,y) = -v(y,x)$. Hence, one could replace u in the definition of RUM-CFs and DDMs by a complete but not necessarily transitive binary relation R on X. The appropriate reformulation of (RUM.1) would be $\mathbb{E}[\tilde{v}(x,y)] \geq 0 \Leftrightarrow xRy$, and DDMs could be linked to R by $\mu(x,y) \geq 0 \Leftrightarrow xRy$.

The purpose of this section is twofold. On the one hand, we illustrate how our methods can be applied in practice. On the other hand, we demonstrate that our theoretical results are empirically relevant, reveal the limits of assumptions made in conventional models, and outperform conventional models that do not use response times in out-of-sample prediction.

5.1 Applying Theorem 1

We start with an application of Theorem 1 to the first phase of the experiment. Our result says that a revealed preference for x over y can be deduced if $F(y,x)$ q-FSD $F(x,y)$ for $q = p(x,y)/p(y,x)$, which can be rewritten as

$$p(x,y)F(x,y)(t) - p(y,x)F(y,x)(t) \geq 0 \text{ for all } t \geq 0.$$

This statement makes use of exact choice probabilities and the full distributions of response times. Since our dataset is finite, we need to translate the statement into its empirical counterpart. To do so, we replace the choice probabilities by empirical choice frequencies, and the response time distributions by non-parametric kernel density estimates.[20]

Fix a subject and a snack food item z, and call the reference item r. Suppose choice was actually stochastic, so the subject did not make the same choice in all 10 repetitions of the decision problem. The dataset contains 77 such decision problems involving 770 choices. We then plot the function

$$\hat{D}(t) = \hat{p}(z,r)\hat{F}(z,r)(t) - \hat{p}(r,z)\hat{F}(r,z)(t),$$

where $\hat{p}(z,r)$ and $\hat{p}(r,z)$ are the empirical choice frequencies across the 10 repetitions, and $\hat{F}(z,r)$ and $\hat{F}(r,z)$ are cumulative distribution functions obtained from non-parametric kernel density estimates of the associated response times.[21] If $\hat{D}(t) \geq 0$ for all $t \geq 0$, we can deduce a revealed preference for item z over the reference r. Similarly, if $\hat{D}(t) \leq 0$ for all $t \geq 0$, we deduce a revealed preference for r over z. If $\hat{D}(t)$ intersects the x-axis at some t, a revealed preference cannot be deduced without distributional assumptions.

Figure 3 shows the function $\hat{D}(t)$ for all stochastic choices of four selected subjects in the data set, normalized so that $\hat{D}(t) \geq 0$ for large t and plotting response times on a logarithmic scale.

[20] An alternative approach is to replace the response time distributions by the empirical cumulative distribution functions, i.e., step functions with steps at the observed response times. This alternative approach yields results similar to the ones based on kernel estimates, and we report them in the footnotes.

[21] The kernel density estimates were performed in Stata using the -akdensity- module, which delivers cumulative distribution functions as output. We estimate the distribution of log response times to avoid boundary problems arising from $t \geq 0$. The estimates use an Epanechnikov kernel with optimally chosen non-adaptive bandwidth. For the case where some choice is made only one out of 10 times, and hence only a single response time is available, an optimal bandwidth cannot be determined endogenously, so we set it to 0.1, yielding a distribution function close to a step function at the observed response time.

There is, of course, a large heterogeneity among subjects. For instance, the top-left subject only exhibited stochastic choice for two of the 17 snack food items. Our q-FSD condition is fulfilled in one of these cases (blue line) but not in the other (red line). The top-right subject exhibited stochastic choice for 13 items, with q-FSD being satisfied in 10 of the cases. Hence we obtain a distribution-free revealed preference for more than three quarters of the stochastic choices of this subject.

The bottom-left subject displays stochastic choice for three snacks, of which two fulfill our revealed preference condition. The third case is interesting and represents a case where q-FSD is clearly violated. The blue line initially falls below zero because the quickest choices favor one of the options (the reference r in that case), but it eventually increases above zero because slower choices favor the other option (the snack item z). Overall the snack z is chosen in 6 out of 10 cases, but the countervailing pattern of response times makes it impossible to conclude that the distribution of the noise is symmetric enough to obtain a revealed preference. It is worth remarking again that patterns like this one are precluded, except as noisy realizations in a finite data set, both by RUM-CFs with symmetric distributions (as shown by our Proposition 4) and by drift-diffusion models (as shown by our Proposition 6).

The bottom-right subject is particularly interesting. Of the four snacks for which choice was stochastic, three violate q-FSD. The fourth one, however, is such that both the item z and the reference r were chosen exactly 5 times, hence the blue line becomes zero for large t. Yet it lies weakly above zero for all t, with many inequalities being strict, meaning that a strict preference (for the reference r in that case) is revealed according to our criterion, while an approach based on the symmetry assumption would conclude an indifference (see our discussion in footnote 12).

Across all subjects and decision problems in which choice was stochastic, our q-FSD condition is fulfilled in 61.0% of the cases.[22] Put differently, in substantially more than half of the stochastic choice problems a preference is revealed without distributional assumptions. We take this as evidence of the empirical relevance of the q-FSD criterion. We also note that classical first-order stochastic dominance obtains only in 35.1% of the cases, and in more than half of these cases (15 out of 27) the less frequently chosen option is actually the faster one, so that there is no systematic relationship between choice frequencies and classical first-order stochastic dominance. Finally, the 39.0% of the cases where q-FSD is violated suggest that neither symmetric RUM-CFs nor drift-diffusion models are appropriate models for the data-generating process.

[22]With cumulative distributions functions estimated by their empirical step function counterparts, this number is 64.9%. If we do not count as a violation the cases where $\hat{D}(t)$ intersects the x-axis right after the quickest choice, which can happen by chance whenever the quickest choice is in the "wrong" direction, the number increases to 87.0%.

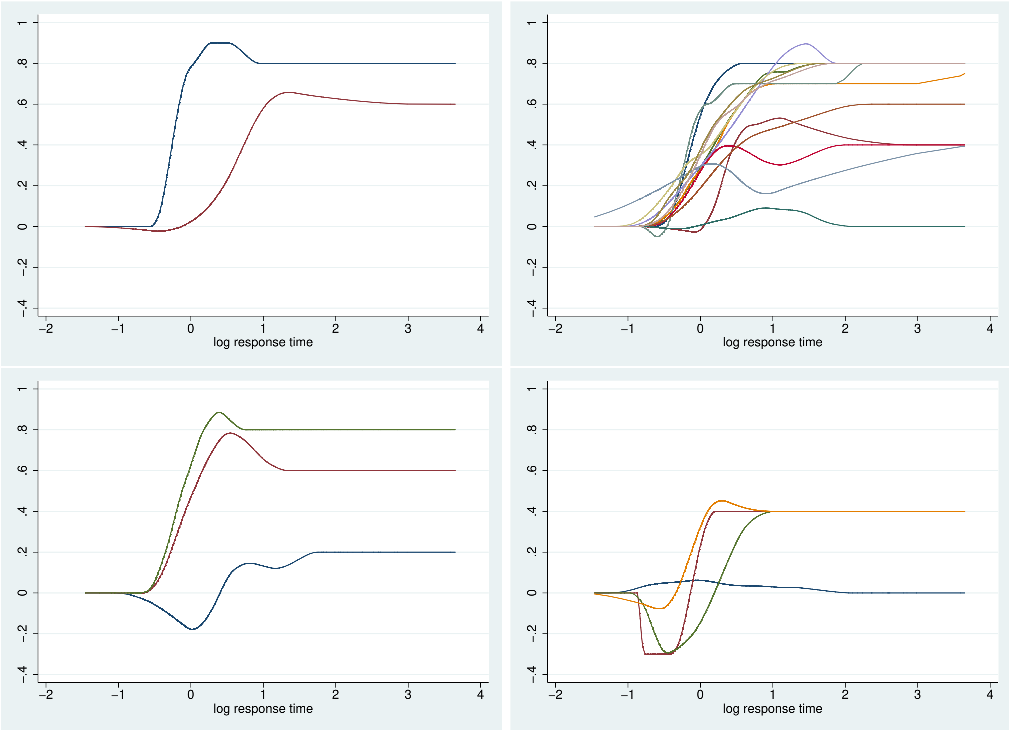

Figure 3: Functions $\hat{D}(t)$ for four selected subjects.

5.2 Applying Theorem 2

Theorem 2 allows to make out-of-sample predictions. The experiment of Clithero (2018) is particularly well-suited for this purpose, since all first-phase choices were made between different snack food items and a fixed reference item, while in the second phase each possible pair of non-reference items was presented once as a binary decision problem. In a setting like this, we can use the first phase to derive complete preferences and use them to predict choices in the second phase, making no distributional assumptions other than symmetry.

The key to the application is realizing that Theorem 2 (plus transitivity) reveals the entire preference relation if one has choice and response time data for the choice between each item z and one fixed reference item r. By Proposition 2, all items z with $p(z,r) > 1/2$ are revealed to be strictly better than the reference r. Theorem 2, in turn, implies that items better than r are ordered inversely by $\theta(z,r)$. That is, for any z and z' with $p(z,r) > 1/2$ and $p(z',r) > 1/2$, z is revealed to be strictly preferred over z' if $\theta(z,r) < \theta(z',r)$. Similarly, all items z with $p(z,r) < 1/2$ are strictly worse than r and are ordered by $\theta(r,z)$. Last, any item z with $p(z,r) = 1/2$ is revealed indifferent to r.

To deduce the entire preference relation from first-phase choices, we need to identify the empirical counterparts $\hat{\theta}(z,r)$ of the percentiles $\theta(z,r)$, and analogously for $\hat{\theta}(r,z)$.

Again, we rely on the non-parametric kernel density estimates of the response time distributions and define $\hat{\theta}(z,r)$ such that

$$\hat{F}(z,r)(\hat{\theta}(z,r)) = \frac{0.5}{\hat{p}(z,r)},$$

and analogously for $\hat{\theta}(r,z)$.[23] This empirical percentile is also well-defined if choice was deterministic, in which case it becomes the median response time of the option that was chosen in all 10 repetitions. While our theoretical analysis does not cover the case of truly deterministic choice, it predicts that the percentile converges to the median as $p(z,r) \to 1$. We implement this prediction but ranking also the deterministically chosen options based on $\hat{\theta}(z,r)$ or $\hat{\theta}(r,z)$.

We use the so-derived revealed preference relations to predict choices in the second phase, that is, strictly out-of-sample. With symmetric utility noise, z is chosen over z' with a probability larger than 1/2 if and only if z is preferred over z'. Hence we code a choice of z over z' in the second phase as correctly predicted if z was revealed to be preferred over z' in the first phase.[24] Despite the fact that second-phase choices are noisy themselves, and the data set contains only one decision for each pair of items, the accuracy of our prediction is remarkable. Across all subjects and binary decision problems, our predictions are correct in 80.7% of the cases.[25]

Clithero (2018) conducted the same prediction exercise based on additional structural assumptions. First, he fitted a conventional logit model to the first-phase choices and used the individual parameters to predict second-phase choices.[26] This method yielded a prediction accuracy of only 73.8%. This accuracy of the logit model is significantly lower than for our approach based on Theorem 2 (paired t-test based on subject-level prediction accuracy that we obtained from John Clithero, $t = 5.4046$, $p < 0.0001$). Second, Clithero (2018) fitted a fully parametric drift-diffusion model, using first-phase choices and response times. This method yielded a prediction accuracy of 81.2%, which is statistically indistinguishable from our approach (paired t-test, $t = -0.7781$, $p = 0.4426$). Hence we achieve the same accuracy as a drift-diffusion model, but without any parametric estimation and with no structural assumptions beyond symmetry of noise and the chronometric relationship.

[23]More precisely, we estimate $\hat{F}(z,r)(t)$ at all response times t that were ever observed in the entire data set. Then we define $\hat{\theta}(z,r)$ as the average between the largest of the t for which $\hat{F}(z,r)(t) \leq 0.5/\hat{p}(z,r)$ and the smallest of the t for which $\hat{F}(z,r)(t) \geq 0.5/\hat{p}(z,r)$. With the alternative approach based on empirical step function cdfs, the condition $\hat{F}(z,r)(t) = 0.5/\hat{p}(z,r)$ is always satisfied for an entire interval of t, corresponding to one step of the cdf. We then define $\hat{\theta}(z,r)$ as the midpoint of this interval.

[24]The revealed preference relations that we obtain are always strict, except for one subject and one pair of items. We ignore this single case of indifference in our prediction evaluation.

[25]With empirical step function cdfs, the prediction is correct in 80.0% of the cases.

[26]The fitting approach relies on a Markov Chain Monte Carlo method assuming comparability of choices across individuals. Comparability across individuals is necessary because, if not, $\hat{p}(z,r) \in \{0,1\}$ would deliver no information. The Monte Carlo approach generates information from the fact that $\hat{p}(z,r) \notin \{0,1\}$ for some other individual.

Our prediction accuracy is even higher for a subset of decisions. The second phase of the experiment collected only one decision for each pair of snacks. Since choices are stochastic, prediction accuracy must be limited simply because of the inherent noise. Previous research has documented that noisier decisions, which are thus more difficult to predict accurately, have longer response times on average (e.g. Chabris et al., 2009; Alós-Ferrer and Garagnani, 2018). We do have the response times of the second-phase decisions, which we did not use so far. As a simple exercise, we divided the data in four response time quartiles for each subject and re-examined the corresponding subsets of decisions. For the slowest decisions, prediction accuracy drops to 68.5%. For the second-slowest quartile, it rises to 79.0%. For the second-fastest quartile, it is already at 86.7%. Finally, for the quartile containing the fastest decisions, prediction accuracy is at a remarkable 88.7%. This shows the actual potential of our method.

5.3 Applying Theorem 3

If one is willing to assume that utility noise has a Fechnerian structure (but still not commit to a specific form such as logit or probit), Theorem 3 allows an even more fine-tuned out-of-sample prediction, in the form of choice probabilities $\bar{p}(x, y)$ for unobserved choices. We apply this result to the same data set as above.

The procedure is analogous to that illustrated in the last subsection. For each pair of snacks (z, z'), we obtain a probability prediction from the formula in Theorem 3 by replacing all expressions by their empirically estimated counterparts, just like before. However, we can use Theorem 3 to independently predict either $\bar{p}(z, z')$ or $\bar{p}(z', z)$, and the two predictions are not guaranteed to satisfy $\bar{p}(z, z') = 1 - \bar{p}(z', z)$. Put differently, we may obtain two different predictions for the same choice probability by swapping the roles of z and z' in Theorem 3. The issue does not arise with a large and rationalizable data set like in our theoretical analysis, but it will arise in finite noisy data. It turns out that differences are small in our data set. On average across all second-phase decisions, the difference between $\bar{p}(z, z')$ and $1 - \bar{p}(z', z)$ is 0.04 and thus less than five percentage points.[27] For the following evaluation, we take the average of the two predictions.

To measure the accuracy of our prediction, we compute the mean absolute error across all decisions in the second phase of the experiment. If z was chosen over z', the absolute error of the prediction in this decision is defined as $1 - \bar{p}(z, z')$, which implies a smaller error if the observed choice was predicted to be more likely. If z' was chosen over z, the absolute error is $\bar{p}(z, z')$. Proceeding in this way, we obtain a mean absolute error of 0.237.[28]

The logit model of Clithero (2018) yielded a mean absolute error of 0.263. Compared with our method, this error is significantly larger (paired t-test, $t = -3.7208$, $p = 0.0008$), so that our response-time based approach again outperforms the conventional random

[27] With empirical step function cdfs, the average difference is 0.06. Since probability predictions are on a much coarser grid in that case, there is actually no difference at all in 57.2% of the cases.

[28] With empirical step function cdfs, the number is 0.235.

utility model. The drift-diffusion model of Clithero (2018) achieved a mean absolute error of 0.209, which is significantly lower than for our method (paired t-test, $t = 7.3948$, $p < 0.0001$). We emphasize that our non-parametric method yields its high accuracy in a computationally straightforward way and does not require structural estimation like a drift-diffusion model.

Our prediction is again more accurate for faster decisions. We predict probabilities closer to 0 or 1 for second-phase decisions that turn out to happen faster, in line with the literature cited earlier.[29] For the slowest decisions, the mean absolute prediction error is 0.354. For the second-slowest quartile, it falls to 0.253. For the second-fastest quartile, it is already at 0.188. Finally, for the quartile containing the fastest decisions, the error is at only 0.152.

6 Relation to the Literature

6.1 Stochastic Choice and Response Times: History

The realization that choice is inherently stochastic has permeated economics for decades,[30] and psychology even longer. Models of stochastic choice in economics view stochastic choice functions – the most direct theoretical counterpart of observable choice frequencies – as a primitive. Block and Marschak (1960) posed the question of when a stochastic choice function can be represented as a distribution over preferences, with the interpretation that either choices are generated by randomly sampled preferences at the individual level, or reflect the choices of a population of randomly sampled agents. Not every stochastic choice function can be viewed as such a distribution, but Falmagne (1978), Barberá and Pattanaik (1986), and McFadden and Richter (1990) derived necessary and sufficient conditions under which this is possible. Those *random preference models* are behaviorally equivalent to *random utility models*, which assume the existence of a utility function that is perturbed by an error term. For finite sets of alternatives, Block and Marschak (1960) showed that choice probabilities can be rationalized by a random preference model if and only if they can be rationalized by a random utility model. The applied literature overwhelmingly relies on the second approach, because it allows for parametric estimation once one fixes functional forms for the utility function and the noise term. Random utility models have a long history, with contributions including Marschak (1960) and going back to Thurstone (1927), but they first became widespread in economics after being used in transportation science (Ben-Akiva and Lerman, 1985), as McFadden (2001) explains in his Nobel Prize Lecture. It was Thurstone (1927), in

[29] For the slowest quartile of decisions, the average predicted probability of the more likely choice is 0.74. It becomes 0.81 for the second-slowest quartile, 0.85 for the second-fastest quartile, and 0.89 for the fastest quartile.

[30] The extent to which this is true seems to be often underestimated. For instance, Samuelson (1938), which is widely considered to be one of the seminal contributions to have started the modern revealed preference approach, was, as its title indicated, a note on the earlier paper of Georgescu-Roegen (1936), which explicitly viewed consumer choice as stochastic.

psychology, who first introduced what is known today as the probit model, where utility errors are assumed to be normally distributed. This is the first example of a Fechnerian model as described in our Section 3.3. In these models, choice probabilities can be derived from a fixed cumulative distribution function of utility differences (see also Debreu, 1958). A different example is the logit model, where one assumes a logistic distribution for utility differences. The logit model is particularly interesting because choice probabilities can be written in a closed form that is equivalent to the celebrated choice model of Luce (1959), where the choice probability of an option is proportional to a normalized value which can be interpreted as the utility of the option (see Anderson et al., 1992, Chapter 1 for a detailed treatment). Generating choice probabilities from decision values by assuming a logit function (or "softmax") is also a common approach in decision neuroscience.

Parallel to developments in economics and related sciences, cognitive psychology – inspired by the same earlier observations and models – followed a different path. The motivation was to explain both choice frequencies and data associated with the choice process such as response times. A key development in this literature was the drift-diffusion model of Ratcliff (1978) described in Section 4.2 (see also Ratcliff and Rouder, 1998). The DDM predicts logit probabilities in binary choice and, as argued by Woodford (2014), could be seen as providing a foundation for the latter. The model essentially evolved from a chain of contributions adapting the sequential ratio probability test of Wald (1945) as a descriptive model of choice, and has grown to become a veritable standard in cognitive psychology and neuroscience. Although the reasons for its original success are probably linked to its flexibility in fitting data, it has been argued to reflect neural processes implementing choices (Shadlen and Kiani, 2013; Shadlen and Shohamy, 2016). Recent developments have integrated attentional processes in the form of eye-tracking data into the drift-diffusion model (Krajbich et al., 2010; Krajbich and Rangel, 2011). Empirical estimates of these attentional drift-diffusion models indicate that non-attended options tend to be discounted in the process of evidence accumulation.

As discussed in the main text, models as the one of Ratcliff (1978) can also be interpreted as models where latent utility differences are discovered by the decision maker, and that discovery takes time (Fudenberg et al., 2018). It is precisely this reinterpretation that provides a powerful intuition for the chronometric function: it takes time to tell apart two close values, while differentiating two values which are far apart is quickly done. Our work constitutes a bridge between the disparate branches of the literature, by providing a framework where response times can be understood as an integral part of random utility models.

6.2 Response Times: Recent Developments

Our work is related to a recent strand of the literature which makes the empirical point that response time data can help with structural estimation of preferences by using the chronometric function. Schotter and Trevino (2020) and Konovalov and Krajbich (2019) propose to estimate indifference points by using the longest response times in a data set and then deduce ordinal preference relations from the indifference points. Other studies have shown how response times are indicative of effort allocation (Moffatt, 2005) or can be used to improve out-of-sample predictions of choices (Clithero, 2018). All those works, however, consider fully-specified structural models and add response times to improve the estimation, an approach which can be useful when choice data is scarce or not reliable. In contrast, our paper solves a different problem. As Proposition 1 shows, in the absence of unverifiable assumptions about the structure of utility noise, it is impossible to uncover preferences from choices. We show that this fundamental problem can be overcome with response time data. We provide, in particular, a simple and intuitive condition which ensures that preferences can be recovered non-parametrically in the absence of any assumptions on the distribution of utility noise, i.e., under conditions where the recovery of preferences fails even with rich choice data.

Echenique and Saito (2017) provide an axiomatization of the chronometric relationship, viewed as a mapping from utility differences to response times as in our model. They consider deterministic choices and deterministic response times only. Their main interest is a characterization of finite and incomplete data sets that can be rationalized by a deterministic utility function together with a chronometric function. That is, they do not consider stochastic choice or the problems that arise when utility is noisy.

While response times are generally receiving increased attention as a tool to improve economic analysis, detailed studies are still scarce (a review and discussion can be found in Spiliopoulos and Ortmann, 2018). Examples include the studies of risky decision-making by Wilcox (1993, 1994), the web-based studies of Rubinstein (2007, 2013), and the study of belief updating by Achtziger and Alós-Ferrer (2014). The study of Alós-Ferrer et al. (2016) uses the chronometric relationship to understand the preference reversal phenomenon (Grether and Plott, 1979), where decision-makers typically make lottery choices which contradict their elicited certainty equivalents when one of the lotteries has a salient, large outcome. Alós-Ferrer et al. (2016) show that, if reversals are due to a bias in the elicitation process rather than in the choice process, choices associated with reversals should take longer than comparable non-reversals. The reason is simply that reversals (where noisy valuations "flip") are more likely when the actual utilities are close, and hence, by the chronometric relationship, response times must be longer. The prediction is readily found in the data, providing insights into the origin and nature of reversals.

6.3 Stochastic Choice: Recent Developments

Our work is also related to the recent literature on stochastic choice theory using extended data sets. Caplin and Martin (2015) and Caplin and Dean (2015) consider state-dependent data sets, which specify choice frequencies as functions of observable states. Caplin and Martin (2015) study rationalizability by maximization of expected utility when the decision-maker has a prior on the state and updates it through Bayes' rule after receiving signals on the state. Caplin and Dean (2015) study rationalizability when the decision-maker additionally decides how much effort to invest in obtaining costly signals (through attention strategies). While their results focus on rationalizability, state-dependent choice data adds an additional dimension which potentially could help with preference revelation. Extended data also play a role in the model of Eliaz and Rubinstein (2014), where decision makers observe and can imitate other agents, but can also observe correlates of their decisions, and in particular whether those decisions were hasty or not.

Our paper also contributes to the theory of behavioral welfare economics, which aims at eliciting preferences from inconsistent choice data (among others, see Bernheim and Rangel, 2009; Rubinstein and Salant, 2012; Masatlioglu et al., 2012; Benkert and Netzer, 2018). Most of this literature considers deterministic choices. Two exceptions are Manzini and Mariotti (2014) and Apesteguía and Ballester (2015). Manzini and Mariotti (2014) show that underlying preferences can be identified when stochastic choice is due to stochastic consideration sets. Apesteguía and Ballester (2015) propose as a welfare measure the preference relation which is closest (in a certain, well-defined sense) to the observed stochastic choices. Similarly to our results in Section 4, they show that this procedure recovers the true underlying preference if the data is generated by random utility models fulfilling a monotonicity condition. To the best of our knowledge, this literature has not yet discovered the value of response time data for preference revelation.

The difficulty of a choice problem can be influenced by additional factors, on top of the utility difference between the options. For instance, if the options are multidimensional, then a choice problem involving a dominant alternative may be very simple, generating accurate and quick responses even if the underlying utility difference is small (see e.g. He and Natenzon, 2018). This is a well-known problem for conventional RUMs, because if the utility difference is small, error rates are predicted to be high. However, our definition of RUMs sidesteps this problem. The reason is that the added generality in Definition 2 allows for pair-specific random utility differences $\tilde{v}(x, y)$ that are not necessarily simple differences $\tilde{u}(x) - \tilde{u}(y)$ between option-specific random variables. For instance, He and Natenzon (2018) consider "moderate utility models" relating choice probabilities $p(x, y)$ to utility differences $u(x) - u(y)$ and an additional distance $d(x, y)$, which reflects factors beyond utility differences that may influence the pair-specific choice difficulty. The assumption is that $p(w, x) \geq p(y, z)$ if and only if $(u(w) - u(x))/d(w, x) \geq (u(y) - u(z))/d(y, z)$, which retains the basic regularities of conventional models while allowing for violations of weak stochastic transitivity. Such considerations can be reflected in our setting by working with a formulation like $\tilde{v}(x, y) = u(x) - u(y) + d(x, y)(\tilde{\epsilon}(x) - \tilde{\epsilon}(y))$. A similar point applies if decisions are subject to frames (e.g. Salant and Rubinstein, 2008; Benkert and Netzer, 2018) or to pair-specific impulsive tendencies or intuitive processes (Achtziger and Alós-Ferrer, 2014) which might influence response times independently of the underlying utilities.

7 Conclusion

Choice theory has traditionally focused on choice outcomes and has ignored auxiliary data such as response times. This neglect comes at a cost even for traditional choice-theoretic questions. In the context of stochastic choice, ignoring response time data means discarding information about the distribution of random utility, which then has to be compensated by making distributional assumptions. In this paper, we have developed a suite of tools that utilize response time data for a recovery of preferences without or with fewer distributional assumptions. Table 1 summarizes our main results on what can be learned from a data set. The table illustrates that using response times is always a substitute for making the next, stronger distributional assumption. We have also shown that our tools are easy to apply and generate remarkably accurate results.

Throughout most of the paper, we have interpreted SCF-RTs as describing the choices of a single individual who is confronted with the same set of options repeatedly. Random utility then reflects fluctuating tastes or noisy perception of the options. However, our tools also work when the data is generated by a heterogeneous population of individuals, each of whom makes a deterministic choice at a deterministic response time (as in Echenique and Saito, 2017). Random utility then reflects a distribution of deterministic utility functions within the population, and response times vary because the difficulty of the choice problem varies with the subjective utility difference. At first glance, a one-to-one translation of our results seems to require the assumption that the same chronometric function applies to all individuals.[31] However, what we called a noisy chronometric function in Section 4 can readily be interpreted as a distribution of chronometric functions within the population. In this population interpretation, a revealed preference for x over y means that utilitarian welfare with x is larger than with y. The use of response times is a novel way to approach the long-standing problem of how to measure the cardinal properties of utility that utilitarianism relies on.[32]

There is a range of interesting questions that we leave for future research. First, our results lend themselves to empirical application. A first natural step was to work with experimental choice data from the lab, where response times are easy to measure. Future work could study real-world data e.g. from online marketplaces, where the time a consumer spends contemplating the options could be (and presumably is already) recorded. A challenge will be to differentiate response time in our sense from other concepts such as

[31] Empirical results by Chabris et al. (2009) and Konovalov and Krajbich (2019) indicate that response times (even as little as one observation per individual) can indeed be used to track down parametric differences in utilities across individuals.

[32] See d'Aspremont and Gevers (2002) for a discussion. The requirement that each individual's chronometric function is drawn independently from a restricted class, as described in Section 4, mirrors the interpersonal comparability of utility units that utilitarianism requires.

Assumptions	Choice data only	Choice and response time data
None	Cannot learn anything about preferences (Proposition 1).	Learn preferences within sample (Theorem 1).
Symmetric noise	Learn preferences within sample (Proposition 2).	Also learn preferences for option pairs outside the data set (Theorem 2).
Fechnerian noise	Learn preferences for option pairs outside the data set (Proposition 3).	Also learn choice probabilities for option pairs outside the data set (Theorem 3).

Table 1: Summary of main results.

the time required to read information or to deliberate on the consequences of an action, which may have other qualitative predictions (as in Rubinstein, 2007, 2013).

Second, we have not attempted a full characterization of rationalizability for arbitrary SCF-RTs beyond those studied in Section 4. For the case without response times, characterizations are relatively simple and have been given in the literature.[33] The problem is substantially more involved when response time distributions have to be rationalized, too. However, some useful necessary conditions are easy to obtain. For instance, an SCF-RT can be rationalizable in the class of all RUM-CFs only if the preferences revealed according to Theorem 1 have no cycles, as otherwise there cannot exist a utility function that is consistent with the revealed preferences. Analogous conditions hold for rationalizability in the classes of symmetric or Fechnerian RUM-CFs. These simple conditions provide a specific test of our response-time-based model and allow it to be falsified by the data.

Finally, response times are a particularly simple measure with a well-established relation to underlying preferences, but they may not be the only auxiliary data with that property. Physiological measures such as pupil dilation, blood pressure, or brain activation may also carry systematic information about preferences. It is worth exploring to what extent these measures can improve the classical revealed preference approach and should therefore be added to the economics toolbox.

References

Achtziger, A. and Alós-Ferrer, C. (2014). Fast or Rational? A Response-Times Study of Bayesian Updating. *Management Science*, 60(4):923–938.

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

Alós-Ferrer, C. and Garagnani, M. (2018). Strength of Preference and Decisions Under Risk. Working Paper, University of Zurich.

[33]For instance, it can be shown that an SCF is rationalizable in our class of symmetric RUMs if and only if the binary relation R^s defined in Appendix B has no cycles, in the sense of Suzumura (1976).

Alós-Ferrer, C., Granić, D.-G., Kern, J., and Wagner, A. K. (2016). Preference Reversals: Time and Again. *Journal of Risk and Uncertainty*, 52(1):65–97.

Alós-Ferrer, C. and Ritschel, A. (2018). The Reinforcement Heuristic in Normal Form Games. *Journal of Economic Behavior and Organization*, 152:224–234.

Anderson, S. P., Thisse, J.-F., and De Palma, A. (1992). *Discrete Choice Theory of Product Differentiation*. MIT Press, Cambridge, MA.

Apesteguía, J. and Ballester (2015). A Measure of Rationality and Welfare. *Journal of Political Economy*, 123:1278–1310.

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

Apesteguía, J., Ballester, M. A., and Lu, J. (2017). Single-Crossing Random Utility Models. *Econometrica*, 85(2):661–674.

Arrow, K. J. (1959). Rational Choice Functions and Orders. *Economica*, 26:121–127.

Baldassi, C., Cerreia-Vioglio, S., Maccheroni, F., and Marinacci, M. (2019). A Behavioral Characterization of the Drift Diffusion Model and its Multi-Alternative Extension to Choice under Time Pressure. *Management Science*, forthcoming.

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

Barberá, S. and Pattanaik, P. K. (1986). Falmagne and the Rationalizability of Stochastic Choices in Terms of Random Orderings. *Econometrica*, 54(3):707–715.

Ben-Akiva, M. E. and Lerman, S. R. (1985). *Discrete Choice Analysis: Theory and Applications to Travel Demand*. MIT Press, Cambridge, MA.

Benkert, J.-M. and Netzer, N. (2018). Informational Requirements of Nudging. *Journal of Political Economy*, forthcoming.

Bernheim, B. D. and Rangel, A. (2009). Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics. *Quarterly Journal of Economics*, 124:51–104.

Block, H. D. and Marschak, J. (1960). Random Orderings and Stochastic Theories of Responses. In Olkin, I., editor, *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*, pages 97–132. Stanford University Press, Stanford.

Buschena, D. and Zilberman, D. (2000). Generalized Expected Utility, Heteroscedastic Error, and Path Dependence in Risky Choice. *Journal of Risk and Uncertainty*, 20:67–88.

Camerer, C. F. (1989). Does the Basketball Market Believe in the 'Hot Hand'. *American Economic Review*, 79:1257–1261.

Caplin, A. and Dean, M. (2015). Revealed Preference, Rational Inattention, and Costly Information Acquisition. *American Economic Review*, 105(7):2183–2203.

Caplin, A. and Martin, D. (2015). A Testable Theory of Imperfect Perception. *Economic Journal*, 125:18–202.

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

Cattell, J. M. (1902). The Time of Perception as a Measure of Differences in Intensity. *Philosophische Studien*, 19:63–68.

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

Clithero, J. A. (2018). Improving Out-of-Sample Predictions Using Response Times and a Model of the Decision Process. *Journal of Economic Behavior and Organization*, 148:344–375.

Cohen, M. and Falmagne, J.-C. (1990). Random Utility Representation of Binary Choice Probabilities: A New Class of Necessary Conditions. *Journal of Mathematical Psychology*, 34:88–94.

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

d'Aspremont, C. and Gevers, L. (2002). Social Welfare Functionals and Interpersonal Comparability. In Arrow, K. J., Sen, A. K., and Suzumura, K., editors, *Handbook of Social Choice and Welfare*, pages 450–541. Elsevier.

Debreu, G. (1958). Stochastic Choice and Cardinal Utility. *Econometrica*, 26(3):440–444.

Dehaene, S., Dupoux, E., and Mehler, J. (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.

Drugowitsch, J., Moreno-Bote, R., Churchland, A. K., Shadlen, M. N., and Pouget, A. (2012). The Cost Of Accumulating Evidence In Perceptual Decision Making. *Journal of Neuroscience*, 32:3612–3628.

Echenique, F. and Saito, K. (2017). Response Time and Utility. *Journal of Economic Behavior and Organization*, 139(15):49–59.

Edwards, W. (1954). The Theory of Decision Making. *Psychological Bulletin*, 51(4):380–417.

Eliaz, K. and Rubinstein, A. (2014). A Model of Boundedly Rational "Neuro" Agents. *Economic Theory*, 57:515–528.

Falmagne, J.-C. (1978). A Representation Theorem for Finite Random Scale Systems. *Journal of Mathematical Psychology*, 18(1):52–72.

Fechner, T. G. (1860). *Elemente der Psychophysik [Elements of Psychophysics]*. Breitkopf & Härtel, Leipzig.

Fischbacher, U., Hertwig, R., and Bruhin, A. (2013). How to Model Heterogeneity in Costly Punishment: Insights from Responders' Response Times. *Journal of Behavioral Decision Making*, 26(5):462–476.

Fisher, G. (2017). An Attentional Drift-Diffusion Model Over Binary-Attribute Choice. *Cognition*, 168:34–45.

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

Georgescu-Roegen, N. (1936). The Pure Theory of Consumer's Behaviour. *Quarterly Journal of Economics*, 50(4):545–593.

Glimcher, P. W. and Fehr, E. (2013). *Neuroeconomics: Decision Making and the Brain*. Elsevier, Amsterdam, 2 edition.

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

Gul, F., Natenzon, P., and Pesendorfer, W. (2014). Random Choice as Behavioral Optimization. *Econometrica*, 82(5):1873–1912.

Gul, F. and Pesendorfer, W. (2006). Random Expected Utility. *Econometrica*, 74(1):121–146.

He, J. and Natenzon, P. (2018). Moderate Expected Utility. Working Paper, Washington University in Saint Louis.

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

Horan, S., Manzini, P., and Mariotti, M. (2019). When Is Coarseness Not a Curse? Comparative Statics of the Coarse Random Utility Model. Discussion Paper.

Houthakker, H. S. (1950). Revealed Preference and the Utility Function. *Economica*, 17(66):159–174.

Kahneman, D. (2003). Maps of Bounded Rationality: Psychology for Behavioral Economics. *American Economic Review*, 93(5):1449–1475.

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

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

Krajbich, I., Armel, C., and Rangel, A. (2010). Visual Fixations and the Computation and Comparison of Value in Simple Choice. *Nature Neuroscience*, 13(10):1292–1298.

Krajbich, I., Bartling, B., Hare, T., and Fehr, E. (2015). Rethinking Fast and Slow Based on a Critique of Reaction-Time Reverse Inference. *Nature Communications*, 6(7455):1–9.

Krajbich, I., Lu, D., Camerer, C., and Rangel, A. (2012). The Attentional Drift-Diffusion Model Extends to Simple Purchasing Decisions. *Frontiers in Psychology*, 3(Article 193):1–18.

Krajbich, I., Oud, B., and Fehr, E. (2014). Benefits of Neuroeconomic Modeling: New Policy Interventions and Predictors of Preference. *American Economic Review (Papers and Proceedings)*, 104(5):501–506.

Krajbich, I. and Rangel, A. (2011). Multialternative Drift-Diffusion Model Predicts the Relationship Between Visual Fixations and Choice in Value-Based Decisions. *Proceedings of the National Academy of Sciences*, 108(33):13852–13857.

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

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

Luce, R. D. (1986). *Response Times: Their Role in Inferring Elementary Mental Organization.* Oxford University Press, New York.

Manski, C. F. (1977). The Structure of Random Utility Models. *Theory and Decision*, 8(3):229–254.

Manzini, P. and Mariotti, M. (2014). Stochastic Choice and Consideration Sets. *Econometrica*, 82(3):1153–1176.

Marschak, J. (1960). Binary Choice Constraints on Random Utility Indicators. In Arrow, K. J., editor, *Stanford Symposium on Mathematical Methods in the Social Sciences*, pages 312–329. Stanford University Press, Stanford, CA.

Masatlioglu, Y., Nakajima, D., and Ozbay, E. Y. (2012). Revealed Attention. *American Economic Review*, 102(5):2183–2205.

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

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

McFadden, D. L. and Richter, M. K. (1990). Stochastic Rationality and Revealed Preference. In Chipman, J. S., McFadden, D. L., and Richter, M. K., editors, *Preferences, Uncertainty, and Optimality: Essays in Honor of Leonid Hurwicz*, pages 163–186. Westview Press, Boulder, Colorado.

Moffatt, G. P. (2015). *Experimetrics: Econometrics for Experimental Economics.* Palgrave Macmillan, London.

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

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

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

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

Palmer, J., Huk, A. C., and Shadlen, M. N. (2005). The Effect of Stimulus Strength on the Speed and Accuracy of a Perceptual Decision. *Journal of Vision*, 5:376–404.

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

Ratcliff, R. and Rouder, J. N. (1998). Modeling Response Times for Two-Choice Decisions. *Psychological Science*, 9(5):347–356.

Ratcliff, R., Smith, P. L., Brown, S. D., and McKoon, G. (2016). Diffusion Decision Model: Current Issues and History. *Trends in Cognitive Science*, 20:260–281.

Rieskamp, J., Busemeyer, J. R., and Mellers, B. A. (2006). Extending the Bounds of Rationality: Evidence and Theories of Preferential Choice. *Journal of Economic Literature*, 44(3):631–661.

Rubinstein, A. (2007). Instinctive and Cognitive Reasoning: A Study of Response Times. *Economic Journal*, 117(523):1243–1259.

Rubinstein, A. (2013). Response Time and Decision Making: An Experimental Study. *Judgment and Decision Making*, 8(5):540–551.

Rubinstein, A. and Salant, Y. (2012). Eliciting Welfare Preferences from Behavioural Data Sets. *Review of Economic Studies*, 79:375–387.

Salant, Y. and Rubinstein, A. (2008). $(\{A\}, f)$: Choice with frames. *Review of Economic Studies*, 75:1287–1296.

Samuelson, P. A. (1938). A Note on the Pure Theory of Consumer's Behavior. *Economica*, 56(17):61–71.

Schotter, A. and Trevino, I. (2020). Is Response Time Predictive of Choice? An Experimental Study of Threshold Strategies. *Experimental Economics*, forthcoming.

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

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

Shafer, W. J. (1974). The Nontransitive Consumer. *Econometrica*, 42:913–919.

Spiliopoulos, L. and Ortmann, A. (2018). The BCD of Response Time Analysis in Experimental Economics. *Experimental Economics*, 21(2):383–433.

Suzumura, K. (1976). Remarks on the Theory of Collective Choice. *Economica*, 43:381–390.

Tajima, S., Drugowitsch, J., and Pouget, A. (2016). Optimal Policy for Value-Based Decision-Making. *Nature Communications*, 7:12400.

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

Tsetsos, K., Moran, R., Moreland, J., Chater, N., Usher, M., and Summerfield, C. (2016). Economic Irrationality is Optimal During Noisy Decision Making. *Proceedings of the National Academy of Sciences*, 113(11):3102–3107.

Tversky, A. (1969). Intransitivity of Preferences. *Psychological Review*, 76:31–48.

Wald, A. (1945). Sequential tests of statistical hypotheses. *Annals of Mathematical Statistics*, 16(2):117–186.

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

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

Wilcox, N. T. (1993). Lottery Choice: Incentives, Complexity, and Decision Time. *Economic Journal*, 103(421):1397–1417.

Wilcox, N. T. (1994). On a Lottery Pricing Anomaly: Time Tells the Tale. *Journal of Risk and Uncertainty*, 8(7):311–324.

Woodford, M. (2014). Stochastic Choice: An Optimizing Neuroeconomic Model. *American Economic Review: Papers & Proceedings*, 104(5):495–500.

Appendices – For Online Publication

A Omitted Proofs

A.1 Proof of Proposition 1

Consider any SCF and fix an arbitrary utility function $u : X \to \mathbb{R}$. We will construct a RUM with utility function u that rationalizes the SCF. Since u is arbitrary, it follows that no preference between any x and y with $x \neq y$ is revealed.

For all $(x, y) \in C \setminus D$, choose arbitrary random variables $\tilde{v}(x, y)$ so that (RUM.1-3) are satisfied. For $(x, y) \in D$, where w.l.o.g. $v(x, y) \geq 0$, define the densities

$$g(x,y)(v) = \begin{cases} 0 & \text{if } \delta(x,y) < v \\ d(x,y) & \text{if } 0 \leq v \leq \delta(x,y) \\ p(y,x) & \text{if } -1 \leq v < 0 \\ 0 & \text{if } v < -1 \end{cases}$$

and $g(y, x)(v) = g(x, y)(-v)$ for all $v \in \mathbb{R}$, with

$$d(x,y) = \frac{p(x,y)^2}{p(y,x) + 2v(x,y)} > 0 \text{ and } \delta(x,y) = \frac{p(y,x) + 2v(x,y)}{p(x,y)} > 0.$$

We then obtain that

$$\begin{aligned} \int_{-\infty}^{+\infty} vg(x,y)(v)dv &= \int_{-1}^{0} vp(y,x)dv + \int_{0}^{\delta(x,y)} vd(x,y)dv \\ &= -\frac{1}{2}p(y,x) + \frac{1}{2}\delta(x,y)^2 d(x,y) = v(x,y), \end{aligned}$$

so that (RUM.1-3) are satisfied. It also follows that $G(x, y)(0) = p(y, x)$, so this RUM indeed rationalizes the SCF. □

A.2 Proof of Corollary 1

The condition that $q(x, y)(t) \geq 1$ for almost all $t \geq 0$ can be rewritten as

$$p(x,y)f(x,y)(t) \geq p(y,x)f(y,x)(t)$$

for almost all $t \geq 0$. It implies

$$p(x,y)F(x,y)(t) = \int_0^t p(x,y)f(x,y)(\tau)d\tau \geq \int_0^t p(y,x)f(y,x)(\tau)d\tau = p(y,x)F(y,x)(t)$$

for all $t \geq 0$. Hence $Q(x, y)(t) \geq 1$ holds for all $t > 0$, which implies a revealed preference for x over y by (7) and Theorem 1. The argument for strict preferences is analogous. □

A.3 Proof of Proposition 2

Let $(u, \tilde{v})$ be a symmetric RUM which rationalizes an SCF p. By symmetry, we have that $G(x, y)(v(x, y)) = 1/2$ for all $(x, y) \in C$. Hence $G(x, y)(0) < 1/2$ implies $v(x, y) > 0$, because $G(x, y)(v)$ is increasing in v. Furthermore, $G(x, y)(0) = 1/2$ implies $v(x, y) = 0$, because $G(x,y)(v)$ is strictly increasing in v in the connected support of $\tilde{v}(x,y)$, by (RUM.3).

Suppose $p(x,y) \geq p(y,x)$ for some $(x,y) \in D$. From Definition 3 it then follows that $G(x,y)(0) = p(y,x) \leq 1/2$ and hence $v(x,y) \geq 0$, i.e., a revealed preference for x over y. If $p(x,y) > p(y,x)$, then analogously $G(x,y)(0) = p(y,x) < 1/2$ and hence $v(x,y) > 0$, i.e., a revealed strict preference for x over y. □

A.4 Proof of Proposition 3

Let $(u, \tilde{v})$ be any Fechnerian RUM which rationalizes an SCF p. For each $(x,y) \in C$, the Fechnerian assumption implies

$$G(x,y)(0) = \int_{-\infty}^{0} g(x,y)(v)dv = \int_{-\infty}^{0} g(v - v(x,y))dv = \int_{-\infty}^{-v(x,y)} g(v)dv = G(v(y,x)),$$

where G is the (strictly increasing) cumulative distribution function for g. Hence $G(z,x)(0) \geq G(z,y)(0)$ implies $G(v(x,z)) \geq G(v(y,z))$ and therefore $v(x,z) \geq v(y,z)$, which in turn implies $v(x,y) = v(x,z) - v(y,z) \geq 0$. Furthermore, $G(z,x)(0) > G(z,y)(0)$ implies $v(x,y) > 0$.

Consider any $(x,y) \in C \setminus D$ such that there exists $z \in X$ with $p(x,z) \geq p(y,z)$. Hence $G(z,x)(0) \geq G(z,y)(0)$ and $v(x,y) \geq 0$, i.e., a revealed preference for x over y. If $p(x,z) > p(y,z)$ then analogously $v(x,y) > 0$, i.e., a revealed strict preference for x over y. □

A.5 Proof of Proposition 4

Consider any symmetric RUM-CF $(u, \tilde{v}, r)$. Suppose $v(x,y) \geq 0$. Fix any $t > 0$. If $r^{-1}(t) \geq v(x,y)$, let $\delta(t) = r^{-1}(t) - v(x,y) \geq 0$. We obtain

$$\begin{aligned}
1 - G(x,y)(r^{-1}(t)) &= 1 - G(x,y)(v(x,y) + \delta(t)) \\
&= G(x,y)(v(x,y) - \delta(t)) \\
&= G(x,y)(-r^{-1}(t) + 2v(x,y)) \\
&\geq G(x,y)(-r^{-1}(t)),
\end{aligned}$$

where the second equality follows from symmetry. If $r^{-1}(t) < v(x,y)$, which requires that $v(x,y) > 0$, let $\delta(t) = v(x,y) - r^{-1}(t) > 0$, so that

$$\begin{aligned}
1 - G(x,y)(r^{-1}(t)) &= 1 - G(x,y)(v(x,y) - \delta(t)) \\
&= G(x,y)(v(x,y) + \delta(t)) \\
&= G(x,y)(-r^{-1}(t) + 2v(x,y)) \\
&> G(x,y)(-r^{-1}(t)),
\end{aligned}$$

where the second equality again follows from symmetry, and the inequality is strict since $G(x,y)(-r^{-1}(t)) < G(x,y)(v(x,y)) = 1/2 < G(x,y)(-r^{-1}(t) + 2v(x,y))$, because $G(x,y)(v)$ is strictly increasing in v in the connected support of $\tilde{v}(x,y)$, by (RUM.3).

Suppose an SCF-RT (p, f) is generated by $(u, \tilde{v}, r)$. It follows that equation (4) derived in the proof of Theorem 1 holds for any $(x,y) \in D$ and all $t > 0$. Combined with the above inequalities, whenever $v(x,y) \geq 0$ we obtain that $Q(x,y)(t) \geq 1$ for all $t > 0$, or $F(y,x)$ q-FSD $F(x,y)$ for $q = p(x,y)/p(y,x)$. If $v(x,y) > 0$, then the above case where $r^{-1}(t) < v(x,y)$ indeed arises for large enough t, which additionally implies that $Q(x,y)(t) > 1$ for some t, or $F(y,x)$ q-SFSD $F(x,y)$ for $q = p(x,y)/p(y,x)$. □

A.6 Proof of Proposition 5

Let (p,f) be an SCF-RT that is generated by a symmetric RUM-NCF $(u,\tilde{v},r,\tilde{\eta})$ in which each density $g(x,y)$ is strictly positive on $\mathbb{R}$. Consider the underlying symmetric RUM-CF $(u,\tilde{v},r)$ and note that it generates an SCF-RT $(p,\hat{f})$, i.e., response time densities $\hat{f}(x,y)$ with full support. This holds because each $g(x,y)$ is strictly positive on $\mathbb{R}$ by assumption, and r must be strictly positive on $\mathbb{R}^{++}$ as otherwise $(u,\tilde{v},r,\tilde{\eta})$ would have generated an atom at the response time of zero.

For any $(x,y) \in D$, it then follows from Proposition 4 that $u(x) \geq u(y)$ implies $\hat{F}(y,x)$ q-FSD $\hat{F}(x,y)$ and $u(x) > u(y)$ implies $\hat{F}(y,x)$ q-SFSD $\hat{F}(x,y)$, for $q = p(x,y)/p(y,x)$.

Fix any $(x,y) \in D$. The fact that $(u,\tilde{v},r,\tilde{\eta})$ generates (p,f) implies

$$F(x,y)(t) = \int_0^\infty \left[\frac{1 - G(x,y)(r^{-1}(t/\eta))}{1 - G(x,y)(0)}\right] h(\eta)d\eta \tag{10}$$

for all $t > 0$. Since (u,g,r) generates $(p,\hat{f})$, we obtain

$$\frac{1 - G(x,y)(r^{-1}(z))}{1 - G(x,y)(0)} = \hat{F}(x,y)(z)$$

for all $z > 0$. Evaluated at $z = t/\eta$ and substituted into (10), this yields

$$F(x,y)(t) = \int_0^\infty \hat{F}(x,y)(t/\eta)h(\eta)d\eta$$

for all $t > 0$, and analogously

$$F(y,x)(t) = \int_0^\infty \hat{F}(y,x)(t/\eta)h(\eta)d\eta.$$

Now $\hat{F}(y,x)$ q-FSD $\hat{F}(x,y)$ implies

$$F(y,x)(t) = \int_0^\infty \hat{F}(y,x)(t/\eta)h(\eta)d\eta \leq \int_0^\infty q \cdot \hat{F}(x,y)(t/\eta)h(\eta)d\eta = q \cdot F(x,y)(t),$$

for all $t \geq 0$, i.e., $F(y,x)$ q-FSD $F(x,y)$. Furthermore, $\hat{F}(y,x)$ q-SFSD $\hat{F}(x,y)$ implies that the inequality is strict for some t, i.e., $F(y,x)$ q-SFSD $F(x,y)$. □

A.7 Proof of Proposition 6

Suppose an SCF-RT (p,f) is generated by a DDM with constant or collapsing boundaries and underlying utility function u. Consider any $(x,y) \in D$. We need to show that $v(x,y) \geq 0$ and hence $\mu(x,y) \geq 0$ implies $F(y,x)$ q-FSD $F(x,y)$, and $v(x,y) > 0$ and hence $\mu(x,y) > 0$ implies $F(y,x)$ q-SFSD $F(x,y)$, for $q = p(x,y)/p(y,x)$.

Consider first the case of constant boundaries. It is well-known (see e.g. Palmer et al., 2005) that $\mu(x,y) \geq 0$ implies $p(x,y) \geq p(y,x)$ and $\mu(x,y) > 0$ implies $p(x,y) > p(y,x)$. Furthermore, the distributions of response times conditional on either choice are identical, i.e., $F(x,y)(t) = F(y,x)(t)$ for all $t \geq 0$ (see again Palmer et al., 2005). The conclusion follows then immediately.

Consider now the case of collapsing boundaries. For this case, Fudenberg et al. (2018, proof of Theorem 1) show that (adapting their notation to ours)

$$q(x,y)(t) = \frac{p(x,y)f(x,y)(t)}{p(y,x)f(y,x)(t)} = \exp\left(\frac{\mu(x,y)b(t)}{\sigma^2/2}\right)$$

for all $t \geq 0$. We obtain that $q(x,y)(t) \geq 0$ for all $t \geq 0$ when $\mu(x,y) \geq 0$, and $q(x,y)(t) > 0$ for all $t \geq 0$ when $\mu(x,y) > 0$. The conclusion now follows as in the proof of Corollary 1. □

A.8 Proof of Proposition 7

Suppose $D = \{(x,y),(y,x)\}$ and consider an arbitrary SCF-RT (p,f). Fix any function $b : \mathbb{R}_+ \to \mathbb{R}_{++}$ that is continuous and strictly decreasing with $\lim_{t\to\infty} b(t) = 0$. Let its inverse be $r = b^{-1}$, with the understanding that $r(v) = 0$ if $v > b(0)$. Note that $r : \mathbb{R}_{++} \to \mathbb{R}_+$ satisfies the requirements of a chronometric function in Definition 5. Now define the cumulative distribution function $G(x,y)$ by

$$G(x,y)(v) = \begin{cases} 0 & \text{if } v < -b(0), \\ p(y,x)F(y,x)(t) & \text{if } v = -b(t) \text{ for some } t \geq 0, \\ p(y,x) & \text{if } v = 0, \\ 1 - p(x,y)F(x,y)(t) & \text{if } v = b(t) \text{ for some } t \geq 0, \\ 1 & \text{if } v > b(0), \end{cases}$$

which, due to the assumed properties of b, is well-defined and describes a random variable $\tilde{v}(x,y)$ with connected support $[-b(0),+b(0)]$ that admits a density $g(x,y)$. Let $g(y,x)(v) = g(x,y)(-v)$ for all $v \in \mathbb{R}$. Finally, choose any function $u : X \to \mathbb{R}$ such that

$$u(x) - u(y) = \int_{-b(0)}^{+b(0)} vg(x,y)(v)dv,$$

and let $\tilde{v}(w,z)$ for all $(w,z) \in C \backslash D$ be arbitrary so that (RUM.1-3) are satisfied. It now follows immediately that the RUM-CF $(u,\tilde{v},r)$ rationalizes the SCF-RT (p,f). □

B Transitive Closure

Collect in a binary relation R^{rt} on X all the preferences that are directly revealed according to Theorem 1, by defining

$$(x,y) \in R^{rt} \ \Leftrightarrow \ F(y,x) \ q\text{-FSD} \ F(x,y) \text{ for } q = \frac{p(x,y)}{p(y,x)}, \text{ or } x = y.$$

For any binary relation R, denote by $T(R)$ the transitive closure of R, i.e., $(x,y) \in T(R)$ if and only if there exists a sequence $x_1, x_2, \ldots, x_n$ of any length $n \geq 2$ with $x_1 = x$, $x_n = y$ and $(x_k, x_{k+1}) \in R$ for all $k = 1, \ldots, n-1$. Let P denote the asymmetric part of R, and $T_P(R)$ the asymmetric part of $T(R)$. With this notation, we obtain the following result.

Corollary 2. *Within the class of all RUM-CFs, a rationalizable SCF-RT reveals a preference for x over y if $(x,y) \in T(R^{rt})$, and a strict preference if $(x,y) \in T_P(R^{rt})$.*

Proof. Let $(u, \tilde{v}, r)$ be any RUM-CF which rationalizes an SCF-RT (p, f). For any $x, y \in X$ with $(x,y) \in T(R^{rt})$, it follows that there exists a sequence $x_1, x_2, \ldots, x_n$ with $x_1 = x$ and $x_n = y$ such that, for each $k = 1, \ldots, n-1$, we have $(x_k, x_{k+1}) \in R^{rt}$ and hence $v(x_k, x_{k+1}) \geq 0$ by definition of R^{rt} and Theorem 1 (or trivially, if $x_k = x_{k+1}$). This implies $v(x,y) = \sum_{k=1}^{n-1} v(x_k, x_{k+1}) \geq 0$, i.e., a revealed preference for x over y.

If $(x,y) \in T_P(R^{rt})$, the above sequence cannot at the same time satisfy $(x_{k+1}, x_k) \in R^{rt}$ for each $k = 1, \ldots, n-1$. Hence $(x_{k^*}, x_{k^*+1}) \in P^{rt}$ for some $k^* = 1, \ldots, n-1$. We claim that this implies $v(x_{k^*}, x_{k^*+1}) > 0$, and therefore $v(x,y) > 0$, i.e., a revealed strict preference for x over y. The fact that $(x_{k^*+1}, x_{k^*}) \notin R^{rt}$ implies that $x_{k^*} \neq x_{k^*+1}$ and that there exists $t^* \geq 0$ such that

$$F(x_{k^*}, x_{k^*+1})(t^*) > \frac{p(x_{k^*+1}, x_{k^*})}{p(x_{k^*}, x_{k^*+1})} F(x_{k^*+1}, x_{k^*})(t^*).$$

Together with $(x_{k^*}, x_{k^*+1}) \in R^{rt}$ this implies $F(x_{k^*+1}, x_{k^*})$ q-SFSD $F(x_{k^*}, x_{k^*+1})$ for $q = p(x_{k^*}, x_{k^*+1})/p(x_{k^*+1}, x_{k^*})$, so that our claim follows from Theorem 1. ☐

Next, collect in a binary relation R^s all the preferences that are directly revealed by symmetry according to Proposition 2, i.e.,

$$(x,y) \in R^s \ \Leftrightarrow \ p(x,y) \geq p(y,x), \text{ or } x = y.$$

Corollary 3. *Within the class of symmetric RUMs, a rationalizable SCF reveals a preference for x over y if $(x,y) \in T(R^s)$, and a strict preference if $(x,y) \in T_P(R^s)$.*

Proof. The proof is similar to the proof of Corollary 2 and therefore omitted. ☐

Proceeding analogously, collect in binary relation R^{srt} all the preferences that are directly revealed according Theorem 2, using both the symmetry assumption and response time data, i.e.,

$$(x,y) \in R^{srt} \ \Leftrightarrow \ (x,y) \in C \backslash D \text{ and } \exists z \in X \text{ with } \theta(x,z) \leq \theta(y,z) \text{ or } \theta(z,x) \geq \theta(z,y).$$

Corollary 4. *Within the class of symmetric RUM-CFs, a rationalizable SCF-RT reveals a preference for x over y if $(x,y) \in T(R^s \cup R^{srt})$, and a strict preference if $(x,y) \in T_P(R^s \cup R^{srt})$.*

Proof. Let $(u, \tilde{v}, r)$ be any symmetric RUM-CF which rationalizes an SCF-RT (p, f). For any $x, y \in X$ with $(x, y) \in T(R^s \cup R^{srt})$, it follows that there exists a sequence $x_1, x_2, \ldots, x_n$ with $x_1 = x$ and $x_n = y$ such that, for each $k = 1, \ldots, n-1$, either $(x_k, x_{k+1}) \in R^s$ or $(x_k, x_{k+1}) \in R^{srt}$. It follows by definition of R^s and Proposition 2, or by definition of R^{srt} and Theorem 2, that $v(x_k, x_{k+1}) \geq 0$ in either case. This implies $v(x, y) = \sum_{k=1}^{n-1} v(x_k, x_{k+1}) \geq 0$, i.e., a revealed preference for x over y.

If $(x, y) \in T_P(R^s \cup R^{srt})$, there must exist $k^* = 1, \ldots, n-1$ such that, in the above sequence, neither $(x_{k^*+1}, x_{k^*}) \in R^s$ nor $(x_{k^*+1}, x_{k^*}) \in R^{srt}$, hence either $(x_{k^*}, x_{k^*+1}) \in P^s$ or $(x_{k^*}, x_{k^*+1}) \in P^{srt}$. If $(x_{k^*}, x_{k^*+1}) \in P^s$, then $v(x_{k^*}, x_{k^*+1}) > 0$ by definition of R^s and Proposition 2. If $(x_{k^*}, x_{k^*+1}) \in P^{srt}$, then $(x_{k^*}, x_{k^*+1}) \in C \setminus D$ and $\exists z \in X$ such that $\theta(x_{k^*}, z) < \theta(x_{k^*+1}, z)$ or $\theta(z, x_{k^*}) > \theta(z, x_{k^*+1})$ by definition of R^{srt}, hence $v(x_{k^*}, x_{k^*+1}) > 0$ by Theorem 2. This implies $v(x, y) > 0$ in either case, i.e., a revealed strict preference for x over y. □

Finally, collect in R^f all the preferences that are directly revealed with the Fechnerian assumption by Proposition 3, so

$$(x, y) \in R^f \;\Leftrightarrow\; (x, y) \in C \setminus D \text{ and } \exists z \in X \text{ with } p(x, z) \geq p(y, z).$$

Corollary 5. *Within the class of Fechnerian RUMs, a rationalizable SCF reveals a preference for x over y if $(x, y) \in T(R^s \cup R^f)$, and a strict preference if $(x, y) \in T_P(R^s \cup R^f)$.*

Proof. The proof is similar to the proof of Corollary 4 and therefore omitted. □

C Additional Material

C.1 Theorem 1 versus Corollary 1

Assume we have data on a single pair (x, y) such that $p(x, y) = 3/4$, $p(y, x) = 1/4$,

$$f(x,y)(t) = \begin{cases} \frac{4}{3}t^3 & \text{if } 0 \leq t < 1, \\ \frac{4}{3}\left(\frac{2-t}{t^3}\right) & \text{if } 1 \leq t < 2, \\ \frac{4}{3}\left(\frac{t-2}{t^3}\right) & \text{if } t \geq 2, \end{cases}$$

and

$$f(y,x)(t) = \frac{4t^3}{(1+t)^5} \quad \text{for all } t \geq 0.$$

Note that $f(x, y)$ and $f(y, x)$ are continuous densities,[34] with corresponding cdfs

$$F(x,y)(t) = \begin{cases} \frac{1}{3}t^4 & \text{if } 0 \leq t < 1, \\ \frac{1}{3} + \frac{4}{3}\left(\frac{t-1}{t^2}\right) & \text{if } 1 \leq t < 2, \\ 1 - \frac{4}{3}\left(\frac{t-1}{t^2}\right) & \text{if } t \geq 2, \end{cases}$$

and

$$F(y,x)(t) = \frac{t^4}{(1+t)^4} \quad \text{for all } t \geq 0.$$

It is easy to see that $F(y, x)$ q-SFSD $F(x, y)$ holds for $q = p(x, y)/p(y, x) = 3$. For $0 \leq t < 1$, the condition $F(y, x)(t) \leq 3F(x, y)(t)$ reduces to $(1+t)^4 \geq 1$, which is satisfied. For $t \geq 1$, we have $F(y, x)(t) < 1 \leq 3F(x, y)(t)$. Hence a strict preference for x over y is revealed according to Theorem 1, provided that the SCF-RT is rationalizable (which we will show below). At the same time, $f(y, x)(t) > 3f(x, y)(t)$ holds for an open interval of response times around $t = 2$. This follows immediately from $f(y, x)(2) > 0 = 3f(x, y)(2)$ and continuity of $f(x, y)$ and $f(y, x)$. Hence Corollary 1 is not applicable.

The SCF-RT is rationalizable because it is generated by the RUM-CF $(u, \tilde{v}, r)$ with $u(x) - u(y) = 1/2$, $r(v) = 1/v$, and a symmetric bimodal distribution given by

$$g(x,y)(v) = \begin{cases} \frac{1}{(1-v)^5} & \text{if } v \leq 0, \\ 1 - 2v & \text{if } 0 < v \leq \frac{1}{2}, \\ 2v - 1 & \text{if } \frac{1}{2} < v \leq 1, \\ \frac{1}{v^5} & \text{if } v > 1, \end{cases}$$

and

$$G(x,y)(v) = \begin{cases} \frac{1}{4(1-v)^4} & \text{if } v \leq 0, \\ \frac{1}{4} + v(1-v) & \text{if } 0 < v \leq \frac{1}{2}, \\ \frac{3}{4} - v(1-v) & \text{if } \frac{1}{2} < v \leq 1, \\ 1 - \frac{1}{4v^4} & \text{if } v > 1. \end{cases}$$

[34] The fact that $f(x, y)(t) = 0$ for $t = 0, 2$ and $f(y, x)(t) = 0$ for $t = 0$ is not a problem for our SCF-RT definition which requires strictly positive densities, because only isolated points are affected.