Time is Knowledge: What Response Times Reveal

Jean-Michel Benkert, Shuo Liu and Nick Netzer*

This version: February 2025
First version: August 2024

Abstract

Response times contain information about economically relevant but unobserved variables like willingness to pay, preference intensity, quality, or happiness. Here, we provide a general characterization of the properties of latent variables that can be detected using response time data. Our characterization unifies and generalizes results in the literature, solves identification problems of binary response models, and has many new applications. We apply the result to test and support the hypothesis that marginal happiness is decreasing in income, a principle that is commonly accepted but so far not established empirically.

Keywords: response times, chronometric effect, binary response model, non-parametric identification, decreasing marginal happiness

JEL Classification: C14, D60, D91, I31

*Benkert: Department of Economics, University of Bern, jean-michel.benkert@unibe.ch. Liu: Guanghua School of Management, Peking University, shuo.liu@gsm.pku.edu.cn. Netzer: Department of Economics, University of Zurich, nick.netzer@econ.uzh.ch. We are grateful for very helpful comments by Ernst Fehr, Lukas Hensel, Blaise Melly, Costanza Naguib, Andrew Oswald, Martin Vaeth and seminar participants at CUHK-Shenzhen, Hong Kong Baptist University, HU Berlin, NYU Abu Dhabi, Peking University, Shanghai Jiao Tong University, Xiamen University, the Universities of Bamberg, Basel, Bern, Edinburgh, Hamburg, and Hong Kong. Shuo Liu acknowledges financial support from the National Natural Science Foundation of China (grants 72322006 and 72192844).

1 Introduction

Traditionally, economists have ignored choice process data like response times. Only recently has the literature realized that response times contain valuable information about unobserved variables like willingness to pay or accept (Krajbich et al., 2012; Cotet, Zhao and Krajbich, 2025), preference intensity (Chabris et al., 2009; Alós-Ferrer, Fehr and Netzer, 2021; Alós-Ferrer and Garagnani, 2022), quality of alternatives (Card et al., 2024), or happiness (Liu and Netzer, 2023).

In this paper, we take a systematic approach to understanding what information response times contain. We study the identification of distributional properties of latent variables in canonical binary response models, a framework extensively used in economics and applicable to all the settings described above. Our approach builds on the observation in the aforementioned empirical literature that decisions are faster when the absolute value of the latent variable is larger. The observable response times are therefore informative about the unobservable latent variable. As our main result, we provide a full characterization of the distributional properties of latent variables that can be detected using response time data. Our characterization highlights how different assumptions about the exact relationship between the latent variable and response time determine what properties can be identified. The result then allows us to propose solutions to various identification problems noted in earlier work (e.g., Haile, Hortaçsu and Kosenok, 2008; Bond and Lang, 2019).

Due to its generality, our approach lends itself to a wide range of applications. Here, we briefly highlight three, each of which will be discussed in the theoretical part of the paper. First, our approach can be used to detect polarization of political attitudes (Lelkes, 2016) from ordinal survey questions. This task would be difficult, if not impossible, without response times (Vaeth, 2023). Second, it becomes possible to infer properties of demand functions that are important for the pricing of firms (Johnson and Myatt, 2006) from purchase decisions at a single price. Third, we show how to uncover correlations that are not directly observable to an analyst, for example because subjects' responses may be uninformative when they conflict with social norms (Coffman, Coffman and Ericson, 2017). On a more abstract level, our result enables detection of properties of a single latent distribution such as the sign of the mean, inequality, or unimodality. These properties play important roles in the context of revealed preference theory, pricing, and political analysis. For the case of multiple latent distributions, the result enables detection of properties such as first-order stochastic dominance, the ranking of means, likelihood-ratio dominance, single-crossing relationships, and correlations with observable variables. These properties matter for the analysis of survey data, out-of-sample prediction of behavior, and monotone comparative statics.

Our theoretical results make it possible to tackle long-standing empirical challenges and debates in the economics literature. We showcase this potential by applying the results to test the hypothesis of decreasing marginal happiness of income, a principle that is central to redistributive policies. Oswald (2008) and Kaiser and Oswald (2022) question the empirical foundation of this principle by arguing that an observed concave relationship between income and self-reported happiness may result from a concave reporting function rather than from decreasing marginal happiness. Conventional approaches used in the happiness literature are insufficient to establish the principle (Bond and Lang, 2019). In the empirical part of the paper, we show that the principle becomes testable with our response time-based method. Our analysis of existing survey data reveals that the hypothesis of decreasing marginal happiness cannot be rejected.

The central assumption underlying our analysis is the so-called *chronometric function* that associates each choice with a response time. This function is monotone, in the sense that a larger absolute value of the latent variable generates a faster decision, possibly after controlling for individual heterogeneity. From a theoretical perspective, such a relationship emerges naturally in evidence-accumulation models (see Chabris et al., 2009; Fudenberg, Strack and Strzalecki, 2018; Card et al., 2024), where a stronger stimulus generates faster decisions. The empirical evidence for a monotone chronometric function in the laboratory is vast. Among many others, Kellogg (1931), Moyer and Bayer (1976), and Palmer, Huk and Shadlen (2005) document the effect for choice situations with an objective stimulus, and Moffatt (2005), Chabris et al. (2009), Konovalov and Krajbich (2019), and Alós-Ferrer and Garagnani (2022) for subjective value-based environments. Field evidence is also emerging. Card et al. (2024) document that editorial decisions take longer when the submitted paper's quality implies a closer decision. Using eBay data on bargaining behavior, Cotet, Zhao and Krajbich (2025) show that sellers' response times to an offer systematically depend on its perceived value. In the context of an online survey, Liu and Netzer (2023) demonstrate that faster responses are associated with a stronger sense of approval for the selected answer. The consistent support across diverse settings and studies underscores the validity and robustness of the chronometric effect as modeled in our work.

To understand our main insight, consider a decision-making environment where one or multiple agents choose between two options. Choice is determined by the realization x of an unobservable latent variable, with $x \leq 0$ generating choice of option a and $x > 0$ generating choice of option b. An analyst observing the choices wants to learn about the underlying binary response model, and in particular about the cumulative distribution function G of the latent variable, from those data. Unfortunately, the only property of G that is identified without additional assumptions or data is its value at zero, $G(0)$, which is given by the observed probability or frequency of choosing a. This information is extremely limited and does, for example, not imply anything about the mean of G without additional distributional assumptions. Now suppose that the speed of the decision is given by $c(|x|)$ for a strictly decreasing chronometric function c, assumed here to be identical for both choice options and all subjects just for ease of exposition. Since a choice of a arises at time t or earlier if x is sufficiently far below 0, where "sufficiently far" is determined by the chronometric function, the observed probability or frequency of choosing a at time t or earlier pins down the value of $G(-c^{-1}(t))$, and analogous for choices of b. Observing the joint distribution of responses and response times therefore allows the analyst to identify a composition of the distribution G and the (inverse of the) chronometric function c.

Consequently, if the analyst had perfect knowledge of the chronometric function linking values to response times, she could recover the entire latent distribution from the data. Such detailed knowledge is not necessary for inferring only specific distributional properties. Our main result characterizes which properties (or their violations) can be detected under which assumptions on the chronometric function. For example, detecting properties that are preserved under monotone transformations requires only knowledge of monotonicity of the chronometric function. If we further restrict attention to chronometric functions that are identical for both choice options, as in the above illustration, then we can detect properties that are preserved under symmetric monotone transformations, and analogously for other classes of transformations. Our result also provides a simple recipe how to detect or reject any property of interest. It involves constructing a candidate distribution based on the empirical data using a representative chronometric function that the analyst deems possible. If the property of interest holds for this candidate distribution, it must hold for all chronometric functions that can be obtained from the representative one using any transformation under which the property is preserved. We also discuss an extension that combines this approach with direct distributional assumptions. Additionally, we show how to incorporate individual heterogeneity and noise into the analysis, and we provide necessary and sufficient conditions for the rationalizability of response time data.

The existing econometrics literature has studied identification of binary response models through assumptions on the distribution of the latent variable and exogenous variation of observables (e.g., Manski, 1988; Matzkin, 1992). Our approach based on response times is complementary and allows us to avoid some controversial assumptions in that literature. Our general characterization yields as special cases several existing results on the use of response time data. In the context of stochastic choice, Alós-Ferrer, Fehr and Netzer (2021) have shown that response time data can be used to obtain revealed preferences without making distributional assumptions about the random utility component and to improve out-of-sample predictions. Several of their results, such as a sufficient condition for the mean of a distribution to be positive, follow as corollaries from our characterization here. We then generalize these results, for example allowing for additional heterogeneity and dispensing with unnecessary assumptions. In the context of happiness surveys, Liu and Netzer (2023) have shown how to solve identification problems of ordered response models by detecting first-order stochastic dominance between distributions from response time data. Once more, our result yields theirs as a corollary and allows for further generalization. Our unified approach also clarifies the mathematical structure underpinning the existing results.

The paper is organized as follows. Section 2 introduces the formal framework and presents the main result, along with three extensions. Section 3 applies the result and derives theoretical conditions for detecting the various properties of interest that we discussed above. Section 4 uses several of these results to empirically test the hypothesis that marginal happiness is decreasing in income. Section 5 concludes. Additional details and supporting materials are provided in the appendices.

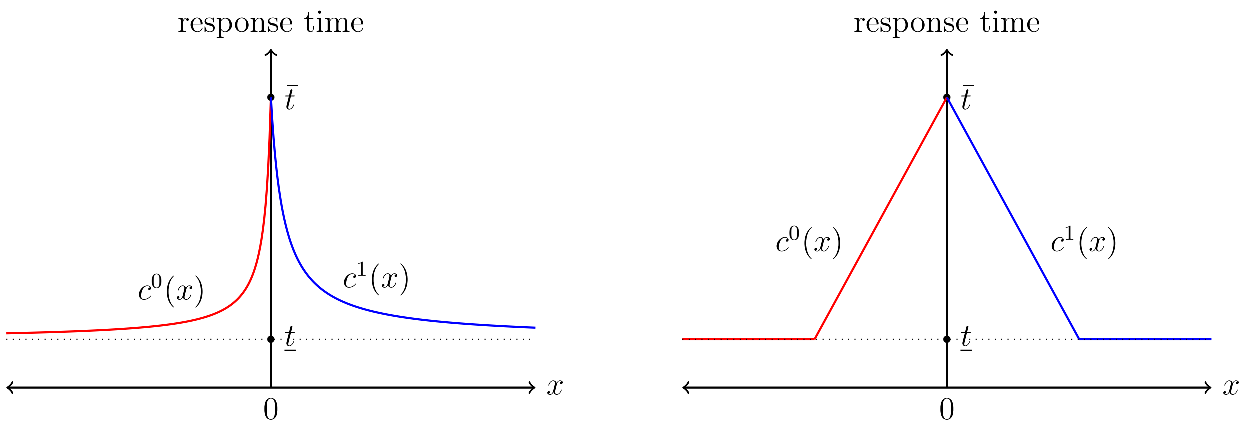
2 General Theory

In this section, we develop our general theoretical framework and present our main result, which shows how and under which conditions distributional properties can be detected using response time data.

2.1 Binary Response Model

We first introduce the binary response model (e.g., Manski, 1988). There is a random variable $\tilde{x}$ with values $x \in \mathbb{R}$ that induce binary responses by comparison with a decision threshold. We normalize the threshold to zero without loss of generality. Thus, the response is $i = 0$ if $\tilde{x}$ takes a value $x \leq 0$ and $i = 1$ if $\tilde{x}$ takes a value $x > 0$. We describe the distribution of the latent variable $\tilde{x}$ by a cumulative distribution function (cdf) G, which we assume to be continuous. It follows that the probabilities of the two responses are $p^0 = G(0)$ and $p^1 = 1 - G(0)$.

The model has different applications and interpretations. For example, the latent variable could be a random utility difference $\tilde{x} = u(1) - u(0) + \tilde{\epsilon}(1) - \tilde{\epsilon}(0)$ between two options, inducing stochastic choices of a single agent (as in Alós-Ferrer, Fehr and Netzer, 2021). In another application, $\tilde{x}$ could describe the distribution of happiness in a population of agents, inducing frequencies of responses to a binary survey question about happiness (as in Liu and Netzer, 2023). The same logic applies to other survey questions, where the responses could be driven



(a) $c^0(x) = 0.1 + \frac{1}{1-2x}, c^1(x) = 0.1 + \frac{1}{1+x}$.

(b) $c^0(x) = c^1(-x) = \max\{0.1x + 1.1, 0.1\}$.

Figure 1: Examples of chronometric functions.

Notes: The left panel in the figure depicts an asymmetric chronometric function c that asympotically approaches the fastest response time $\underline{t}$, while the right panel shows a symmetric one that attains $\underline{t}$ at finite absolute values of the latent variable. In both panels, the red and blue curves correspond to the restrictions of c to $\mathbb{R}_-$ and $\mathbb{R}_+$, respectively.

by a distribution of political attitudes or preference parameters in the population. In yet another application, $\tilde{x}$ could capture the random quality of papers that are submitted to a journal, inducing the editor's decision to accept or reject (as in Card et al., 2024). The same logic applies to other settings where quality determines a binary decision, such as whether to invest in an innovation project. The latent variable could also describe the difference between the willingness to pay $\tilde{v}$ for a product among consumers and the product price p, $\tilde{x} = \tilde{v} - p$, inducing the demand for the product at that price (in the spirit of Cotet, Zhao and Krajbich, 2025).

We now follow Alós-Ferrer, Fehr and Netzer (2021) and Liu and Netzer (2023) and assume that the realized value x of $\tilde{x}$ not only determines the response but also the response time, with larger absolute values implying faster responses, in line with the well-established chronometric effect. Formally, we denote by $c : \mathbb{R} \to [\underline{t}, \bar{t}]$ the chronometric function, where $0 \leq \underline{t} < \bar{t} < \infty$. The function c maps each realized value x into a response time $c(x)$. It is assumed to be continuous, strictly increasing on $\mathbb{R}_-$ and strictly decreasing on $\mathbb{R}_+$ whenever $c(x) > \underline{t}$, and to satisfy $c(0) = \bar{t}$ and $\lim_{x\to-\infty} c(x) = \lim_{x\to+\infty} c(x) = \underline{t}$. Figure 1 illustrates two examples of chronometric functions that adhere to all these conditions.

The restriction of c to $x \in \mathbb{R}_-$ is denoted c^0. This function c^0 has a well-defined inverse $(c^0)^{-1} : (\underline{t}, \bar{t}] \to \mathbb{R}_-$ that is continuous and strictly increasing. We extend it to $\underline{t}$ by setting $(c^0)^{-1}(\underline{t}) = -\infty$ if $c(x) > \underline{t}$ for all $x \in \mathbb{R}_-$ and $(c^0)^{-1}(\underline{t}) = \max\{x \in \mathbb{R}_- \mid c(x) = \underline{t}\}$ otherwise.

Analogously, the restriction of c to $x \in \mathbb{R}_+$ is denoted c^1, with a continuous and strictly decreasing inverse $(c^1)^{-1} : (\underline{t}, \overline{t}] \to \mathbb{R}_+$, which we extend to $\underline{t}$ by setting $(c^1)^{-1}(\underline{t}) = +\infty$ or $(c^1)^{-1}(\underline{t}) = \min\{x \in \mathbb{R}_+ \mid c(x) = \underline{t}\}$, as appropriate.

In addition to response probabilities, the model (G, c) induces distributions of response times. We denote by F^i the cdf of response times conditional on a response of $i = 0, 1$. Since a response $i = 0$ at time t or earlier arises if $x \leq (c^0)^{-1}(t)$, we obtain that

$$p^0 F^0(t) = G((c^0)^{-1}(t)) \tag{1}$$

for all $t \in [\underline{t}, \overline{t}]$, where we use the convention $G(-\infty) = 0$. Analogously, a response $i = 1$ at time t or earlier arises if $x \geq (c^1)^{-1}(t)$, so that

$$p^1 F^1(t) = 1 - G((c^1)^{-1}(t)) \tag{2}$$

for all $t \in [\underline{t}, \overline{t}]$, where we use $G(+\infty) = 1$. The induced response-time cdfs F^i are continuous on $[\underline{t}, \overline{t}]$ and satisfy $F^i(\overline{t}) = 1$. In summary, the binary response model (G, c) induces the data $(p, F) = (p^0, p^1, F^0, F^1)$ according to (1) and (2).

2.2 Detecting Properties

We now ask what an analyst can learn from observed data about the underlying binary response model and, in particular, about the distribution G of the latent variable. Taking observed data as given, different binary response models could have generated those data, so inference about the model is not straightforward. This is true especially if the analyst is not willing to make potentially strong assumptions about the form of the chronometric function or the latent distribution. We ask whether there are some properties that all models which are consistent with the data satisfy. These respective properties are then detected from the data rather than assumed by the analyst.

Consider a profile of data $(p_j, F_j)_j = (p^0_j, p^1_j, F^0_j, F^1_j)_j$ indexed by $j \in J$. The set J could be a singleton, e.g., when studying the choices of a single agent between two options or the responses of a single group of agents to one binary survey question. In this case, we typically omit the index j. If we observe the choices of a single agent for multiple pairs of options, then the set J would describe the different binary choice problems. Alternatively, the index could capture different agents or combinations of agents and choice problems. In a survey application, the index could describe different questions or different demographic groups. Since J is not necessarily finite, it could also model income levels $j \in J \subseteq \mathbb{R}_+$ of survey participants. In the other applications discussed earlier, the index could describe different journal editors or different prices at which market demand is observed. We assume that each F_j^i is continuous and satisfies $F_j^i(\underline{t}) = 0$ and $F_j^i(\bar{t}) = 1$.[1]

Denote by $\mathscr{G}$ the set of all possible profiles $(G_j)_j$ of cdfs, where each individual G_j satisfies our assumptions from the previous subsection. Similarly, denote by $\mathscr{C}$ the set of all possible profiles $(c_j)_j$ of chronometric functions that individually satisfy our previous assumptions. For our main result, we leave $\mathscr{G}$ unrestricted but allow for a possibly restricted set $\mathscr{C}^* \subseteq \mathscr{C}$ of admissible chronometric functions. For example, each profile $(c_j)_j \in \mathscr{C}^*$ may have to satisfy that all functions are symmetric across responses (i.e., $c_j^0(-x) = c_j^1(x)$ for all $x \in \mathbb{R}_+$). This restriction embodies the assumption that the chronometric effect is identical for the two choice options. Another possible restriction would be that the functions are identical across indices (i.e., $c_j = c$ for all $j \in J$), reflecting the assumption that the chronometric effect is the same for all groups j. Other restrictions could be functional forms such as piece-wise linearity, or combinations of multiple of these assumptions.

Definition 1. Given data $(p_j, F_j)_j$ and a set $\mathscr{C}^*$ of admissible chronometric functions, a property $\mathbf{P}$ of the distributions $(G_j)_j$ is *detected* if all $((G_j)_j, (c_j)_j) \in \mathscr{G} \times \mathscr{C}^*$ that induce $(p_j, F_j)_j$—in the sense of (1) and (2) for all $j \in J$ —have in common that $(G_j)_j$ satisfies $\mathbf{P}$.

This idea of detection formalizes that property $\mathbf{P}$ is inferred from the data instead of being imposed on the distributions by assumption. It corresponds to a standard notion of non-parametric identification in econometrics (e.g., Manski, 1988; Bond and Lang, 2019). The revealed preference approach in choice theory (Samuelson, 1938; Arrow, 1958) embodies the same logic (see, e.g., Benkert and Netzer, 2018; Alós-Ferrer, Fehr and Netzer, 2021).[2]

The ability to detect a property $\mathbf{P}$ will depend on the extent to which $\mathbf{P}$ is invariant to transformations. For example, the property that a single cdf G is strictly increasing (full support) is invariant to all strictly increasing transformations, because $G(\psi(x))$ is still strictly increasing in x, for any strictly increasing function ψ. This large class of admissible transformations will make it possible to detect the property under relatively mild assumptions on the chronometric function. Let Ψ be a set of profiles $(\psi_j)_j$ of functions $\psi_j : \mathbb{R} \to \mathbb{R}$ that are bijective and strictly increasing, hence continuous. The set Ψ can embody various constraints, such as the restriction that each profile $(\psi_j)_j \in \Psi$ is composed of identical functions (i.e., $\psi_j = \psi$ for all $j \in J$). For any $(G_j)_j \in \mathscr{G}$ and any $(\psi_j)_j \in \Psi$, the composition $(G_j \circ \psi_j)_j$ is another profile of cdfs in $\mathscr{G}$.

[1] Even though our approach is very general, it does not allow describing correlations between choices or response times across different indices j. We leave an extension that allows for such correlations across decision problems for future work.

[2] We assume here implicitly that there exists at least one $((G_j)_j, (c_j)_j) \in \mathscr{G} \times \mathscr{C}^*$ that induces the data. As we will see later, this is always the case when $\mathscr{G}$ is unrestricted. In Subsection 2.5.2 we will study the case with restricted $\mathscr{G}^* \subseteq \mathscr{G}$ and provide conditions for rationalizability of the data.

Definition 2. A property $\mathbf{P}$ of $(G_j)_j$ is *invariant to transformations* Ψ if $(G_j \circ \psi_j)_j$ also has property $\mathbf{P}$, for all $(\psi_j)_j \in \Psi$.

As another example, the property of (G_1, G_2) that G_1 first-order stochastically dominates G_2 is invariant to profiles (ψ_1, ψ_2) of strictly increasing transformations that are identical for the two distributions (i.e., $\psi_1 = \psi_2$), but not to distribution-specific transformations. Classifying properties by their invariance to transformations parallels an approach used in social choice theory to describe the measurability and interpersonal comparability of utilities that different welfare functions require (d'Aspremont and Gevers, 2002).

2.3 Generating Chronometric Functions

We consider sets $\mathscr{C}^*$ of chronometric functions that are generated by a representative profile $(c_j^*)_j \in \mathscr{C}$ and a set of transformations Ψ as introduced above.

Definition 3. The pair $((c_j^*)_j, \Psi)$ *generates* $\mathscr{C}^*$ if for each $(c_j)_j \in \mathscr{C}^*$ there exists $(\psi_j)_j \in \Psi$ such that $(c_j)_j = (c_j^* \circ \psi_j)_j$.

Only transformations ψ_j that satisfy $\psi_j(0) = 0$ yield well-defined chronometric functions, while Ψ could contain functions without that property. More generally, the definition does not require that all elements of Ψ must be used for the construction of $\mathscr{C}^*$.

To illustrate the concept, we discuss several examples of sets $\mathscr{C}^*$ and how they can be generated. Consider the set of all profiles of chronometric functions which approach $\underline{t}$ asymptotically in the limit but never reach $\underline{t}$. This set is generated by a simple representative member $(c_j^*)_j$, for example the symmetric hyperbolic form

$$c_j^*(x) = \underline{t} + \frac{1}{|x| + 1/(\bar{t} - \underline{t})}, \tag{3}$$

identical for each $j \in J$, together with the unrestricted set of all profiles of transformations. To see why, just note that any desired $(c_j)_j$ can be obtained from $(c_j^*)_j$ by using the transformations $(\psi_j)_j$ given by

$$\psi_j(x) = \begin{cases} (c_j^{*,1})^{-1}(c_j(x)) & \text{if } x > 0, \\ (c_j^{*,0})^{-1}(c_j(x)) & \text{if } x \leq 0, \end{cases}$$

for each $j \in J$. Similarly, the set of all profiles of chronometric functions which have $c_j(x) = \underline{t}$ for large but finite absolute values of x can be generated by a simple representative member, for example the symmetric linear form

$$c_j^*(x) = \begin{cases} \underline{t} & \text{if } (\overline{t} - \underline{t}) < x, \\ \overline{t} - x & \text{if } 0 < x \leq (\overline{t} - \underline{t}), \\ \overline{t} + x & \text{if } -(\overline{t} - \underline{t}) \leq x \leq 0, \\ \underline{t} & \text{if } x < -(\overline{t} - \underline{t}), \end{cases} \tag{4}$$

identical for each $j \in J$, together with the unrestricted set of all profiles of transformations.[3]

If we combine either (3) or (4) with the smaller set of transformations that are symmetric around zero ($\psi_j(-x) = -\psi_j(x)$ for all $x \in \mathbb{R}_+$), we can generate the respective sets of profiles of chronometric functions that are symmetric across responses. If we combine (3) or (4) only with transformations that are identical across indices, we generate only profiles of chronometric functions that are identical across indices. We will discuss all these and additional cases in our applications in Section 3.

2.4 Main Result

Given observed data $(p_j, F_j)_j$ and a representative profile $(c_j^*)_j$ of chronometric functions, we can derive the empirical distribution functions

$$H_j(x) = \begin{cases} 1 - p_j^1 F_j^1(c_j^*(x)) & \text{if } x > 0, \\ p_j^0 F_j^0(c_j^*(x)) & \text{if } x \leq 0, \end{cases} \tag{5}$$

for all $j \in J$. Each H_j is a well-defined and continuous cdf. It would be the true cdf of the binary response model inducing (p_j, F_j) if c_j^* was the true chronometric function.[4]

Of course, c_j^* may be different from the true chronometric function. However, the proof of our main result shows that whenever any admissible chronometric function c_j can be written as $c_j = c_j^* \circ \psi_j$ for some ψ_j, the corresponding distribution G_j can be written as $G_j = H_j \circ \psi_j$. Therefore, if $(H_j)_j$ satisfies a property that is invariant to Ψ, then any $(G_j)_j$ that is compatible with the data also satisfies that property, under the assumption that $((c_j^*)_j, \Psi)$ generates $\mathscr{C}^*$. This gives rise to the following result.

[3] An example of a set $\mathscr{C}^*$ which cannot be generated by any $((c_j^*)_j, \Psi)$ is one which contains some profiles in which c_j reaches $c_j(x) = \underline{t}$ and other profiles in which $c_j(x) > \underline{t}$ throughout, for some $j \in J$. However, this can be addressed by our generalization in Subsection 2.5.1 which allows for the generation of sets by more than one representative profile $(c_j^*)_j$.

[4] See Alós-Ferrer, Fehr and Netzer (2021, Online Appendix p. 5) for an analogous construction.

Theorem 1. *Suppose $\mathscr{C}^*$ is generated by $((c_j^*)_j, \Psi)$. If $(H_j)_j$ satisfies a property* $\mathbf{P}$ *that is invariant to transformations* Ψ*, then* $\mathbf{P}$ *is detected.*

Proof. Let $((G_j)_j, (c_j)_j) \in \mathscr{G} \times \mathscr{C}^*$ be any model that induces $(p_j, F_j)_j$. Using the definition of $(H_j)_j$ and the conditions (1) and (2) for inducing the data, we obtain that

$$H_j(x) = \begin{cases} G_j((c_j^1)^{-1}(c_j^{*1}(x))) & \text{if } x > 0, \\ G_j((c_j^0)^{-1}(c_j^{*0}(x))) & \text{if } x \leq 0, \end{cases}$$

for each $j \in J$. Since $((c_j^*)_j, \Psi)$ generates $\mathscr{C}^*$, there exists $(\psi_j)_j \in \Psi$ such that $c_j = c_j^* \circ \psi_j$, for each $j \in J$. We therefore have $c_j^0(x) = c_j^{*0}(\psi_j(x))$ for any $x \leq 0$ and $c_j^1(x) = c_j^{*1}(\psi_j(x))$ for any $x > 0$. It follows that $(c_j^0)^{-1}(t) = \psi_j^{-1}((c_j^{*0})^{-1}(t))$ and $(c_j^1)^{-1}(t) = \psi_j^{-1}((c_j^{*1})^{-1}(t))$ for all $t \in [\underline{t}, \overline{t}]$. It then follows that

$$H_j(x) = \begin{cases} G_j(\psi_j^{-1}(x)) & \text{if } x > 0, \\ G_j(\psi_j^{-1}(x)) & \text{if } x \leq 0. \end{cases}$$

Note that this is also true when $\bar{x}_j^* := (c_j^{*0})^{-1}(\underline{t})$ is finite and we consider any $x \leq \bar{x}_j^*$. In that case, $\bar{x}_j := (c_j^0)^{-1}(\underline{t}) = \psi_j^{-1}(\bar{x}_j^*)$ is also finite, and we have $G_j(x) = 0$ for all $x \leq \bar{x}_j$ because (G_j, c_j) induces data with $F_j^0(\underline{t}) = 0$. For all $x \leq \bar{x}_j^*$, we therefore have $G_j((c_j^0)^{-1}(c_j^{*0}(x))) = G_j(\bar{x}_j) = 0 = G_j(\psi_j^{-1}(\bar{x}^*)) = G_j(\psi_j^{-1}(x))$. The analogous argument applies to $x \geq (c_j^{*1})^{-1}(\underline{t})$. To summarize, we have $H_j = G_j \circ \psi_j^{-1}$, for each $j \in J$. If $(H_j)_j$ satisfies a property $\mathbf{P}$ that is invariant to transformations Ψ, then $(H_j \circ \psi_j)_j = (G_j \circ \psi_j^{-1} \circ \psi_j)_j = (G_j)_j$ also satisfies $\mathbf{P}$, hence $\mathbf{P}$ is detected. □

As we will show in Section 3, Theorem 1 is easy to apply. If we are interested in some distributional property, we only need to check specific empirical distributions $(H_j)_j$ that are based on a representative profile $(c_j^*)_j$ of chronometric functions. These distributions can be constructed from the observed data. If they satisfy a property that is invariant to transformations Ψ, then this property is detected for the entire class of chronometric functions that $((c_j^*)_j, \Psi)$ generates. Often, the resulting detection conditions can be expressed directly and transparently in terms of the observable data.

The theorem can also be used to detect whether a property is violated. Denote by $\neg\mathbf{P}$ the property that property $\mathbf{P}$ is not true. The theorem immediately implies that if the empirical distributions $(H_j)_j$ satisfy $\neg\mathbf{P}$ and this property is invariant to transformations Ψ, then $\neg\mathbf{P}$ is detected. In words, the violation of a property is detected when $(H_j)_j$ violates the property and the violation is invariant to transformations Ψ. We remark that $\mathbf{P}$ and $\neg\mathbf{P}$ can be invariant to the same set of transformations or to different sets of transformations.

Section 3 contains examples of both cases. Note further that detecting $\neg\mathbf{P}$ is stronger than not detecting $\mathbf{P}$. A property is not detected if *at least one* model that is compatible with the data violates the property, while detecting $\neg\mathbf{P}$ requires that *all* models that are compatible with the data violate the property.[5] Again, Section 3 contains examples of both cases.

2.5 Extensions

2.5.1 Multiple Representative Chronometric Functions

Our first extension covers the case where $\mathscr{C}^*$ is not generated from one representative profile of chronometric functions $(c_j^*)_j$ but from multiple profiles $(c_j^{k*})_j$ where $k \in K$. We say that $((c_j^{k*})_j^k, \Psi)$ generates $\mathscr{C}^*$ if for each $(c_j)_j \in \mathscr{C}^*$ there exist $(\psi_j)_j \in \Psi$ and $k \in K$ such that $(c_j)_j = (c_j^{k*} \circ \psi_j)_j$. For example, if we want to generate a function c_j that never reaches $\underline{t}$, we need to start from a representative function c_j^{1*} with that same property, like (3). If we simultaneously want to generate another function that has $c_j(x) = \underline{t}$ for large absolute values of x, we need a different representative function c_j^{2*} with that property, like (4).

Instead of a single empirical profile $(H_j)_j$, we obtain one profile $(H_j^k)_j$ for each $k \in K$, defined as in (5) using the respective function c_j^{k*}. Theorem 1 then becomes that property $\mathbf{P}$ is detected if $(H_j^k)_j$ satisfies $\mathbf{P}$ for each $k \in K$ (under the otherwise identical assumptions). In words, we simply need to repeat the previous procedure for each $k \in K$ separately, and we achieve detection if we achieve it for each $k \in K$.

Analogously, a violation of $\mathbf{P}$ is detected if $(H_j^k)_j$ violates $\mathbf{P}$ for each $k \in K$. When $(H_j^k)_j$ satisfies $\mathbf{P}$ for some k but not for others, then the data is compatible with some distributions that satisfy $\mathbf{P}$ and with others that violate it, so that we can detect neither $\mathbf{P}$ nor its violation $\neg\mathbf{P}$.[6] We will illustrate this case in Section 3.

2.5.2 Distributional Assumptions

Our second extension covers the case where the set of admissible distributions is restricted to some $\mathscr{G}^* \subseteq \mathscr{G}$. Any such restriction reflects prior knowledge of the analyst about properties of the distributions (e.g., Manski, 1988). For example, in an application where each G_j describes the willingness to pay of consumers minus the corresponding product price, the different distributions in the profile $(G_j)_j$ should all be horizontal shifts (by the difference in product prices) of one identical distribution.

[5] See Liu and Netzer (2023, p. 3303) for a discussion of this difference in the context of first-order stochastic dominance of happiness distributions.

[6] To be precise, this statement is correct when each (c_j^{k*}) for $k \in K$ can actually be used for generating an element of $\mathscr{C}^*$. We could always add profiles (c_j^{k*}) that are not required for generating $\mathscr{C}^*$. The corresponding empirical functions $(H_j^k)_j$ tell us nothing about properties of the true distributions.

Knowledge about the distributions can make detection easier. To illustrate, assume that we want to detect whether the mean of G_1 is larger than the mean of G_2. Without making distributional assumptions, this property is invariant to identical positive affine transformations of the form $\psi_j(x) = a + bx$ with $b > 0$ for $j = 1, 2$, but not to identical monotone transformations in general. Hence, detecting it requires strong assumptions on the set of admissible chronometric functions. Consider instead the property that the median of G_1 is larger than the median of G_2. This property is invariant to all identical monotone transformations and can therefore be detected under milder conditions. We can compensate these milder conditions by assuming that all distributions in $\mathscr{G}^*$ are symmetric around their means, an assumption frequently made in conventional binary response models. Since mean and median coincide in that case, a detected ranking of the medians implies a detected ranking of the means.

A question that arises when working with restrictions on both $\mathscr{G}^*$ and $\mathscr{C}^*$ is whether the given data are *rationalizable*, i.e., whether there exist distributions and chronometric functions in the restricted sets that induce the data.[7] In Appendix A we study this question and provide necessary and sufficient conditions for rationalizability.

2.5.3 Heterogeneity and Noise

Before applying our method to data, it is natural to ask about heterogeneity in response speed and about noise or measurement error. Alós-Ferrer, Fehr and Netzer (2021) and Liu and Netzer (2023) have shown that several of their results have appropriate extensions that facilitate empirical testing. In Appendix B, we follow the same approach and model the response time of an individual with realized latent value x by $t = c_j(x) \cdot \eta \cdot \epsilon$, where $c_j(x)$ is the chronometric function, $\eta > 0$ captures the individual's general response speed, and $\epsilon > 0$ is a noise term comprising uncontrolled factors and measurement error. We assume that the terms (x, η) are drawn from a joint distribution, possibly j-specific, with marginal cdf $G_j(x)$ whose properties we are interested in. Moreover, we assume that the analyst has access to a baseline response, such as a demographic survey question, where the response time is given by $t_b = \phi \cdot \eta$ with intensity parameter $\phi > 0$. The terms (ϵ, ϕ) are also drawn from a joint distribution, possibly j-specific, but independent from the other variables.

The baseline response allows us to normalize response times in the decision problem of

[7] Echenique and Saito (2017) study rationalizability for the case of deterministic responses and response times. Alós-Ferrer, Fehr and Netzer (2021) discuss the difficulty of the problem when responses and response times are stochastic. For a single binary choice problem, they show that all stochastic choice functions with response times are rationalizable when the distributions are unrestricted. For multiple binary choice problems, they discuss necessary conditions for rationalizability as well as the possibility of rationalization by generalized random utility models.

interest to $\hat{t} = t/t_b = c_j(x) \cdot \epsilon/\phi$, which takes care of the individual heterogeneity embodied in η. However, the analyst who works with such normalized data $(p_j, \hat{F}_j)_j$ to derive candidate distributions $(\hat{H}_j)_j$ according to (5) is still misspecified, as she ignores the variation embodied in ϵ and ϕ.

In Appendix B, we derive conditions under which the distributions $(\hat{H}_j)_j$ still inherit a property $\mathbf{P}$ from $(G_j)_j$ and can therefore be used for testing the null hypothesis that $\mathbf{P}$ is satisfied by the true distributions. In addition to invariance properties similar to our main analysis, the key requirement is invariance to multiplicative convolutions of the form

$$\hat{G}_j(x) = \int G_j(xz)\, d\mu_j(z).$$

This invariance is satisfied by many of the properties that we will study in Section 3, in particular those which reduce to "pointwise arguments" in the integral. For single distributions, examples include full support and an asymmetry condition that implies a unique sign of the mean. For multiple distributions, examples include an asymmetry condition that implies a unique ranking of the means and which is invariant to j-specific noise distributions μ_j, as well as several properties that are invariant under common noise $\mu_j = \mu$, such as first-order stochastic dominance and weaker conditions that imply a ranking of the means. In particular, this covers all the properties that we will use to empirically test the hypothesis of decreasing marginal happiness of income in Section 4.

There are also properties for which these convolutions are not innocuous. It is known, for instance, that single-crossing of cdfs does not aggregate in general (Quah and Strulovici, 2012). When attempting to test such a condition in noisy data, one may have to look for restricted classes of latent distributions and noise terms under which testing remains possible.

3 Theoretical Applications

In this section, we apply Theorem 1 to study distributional properties that have been of interest in diverse applications. We replicate and generalize existing results from the literature and present many new results. We first consider properties of single distributions in Subsection 3.1, followed by properties of collections of distributions in Subsection 3.2.

3.1 Single Distribution

3.1.1 Full Support

We start with an application that is primarily didactic. Suppose we are interested in the property of full support of the latent distribution, i.e., the cdf G being strictly increasing on the entire $\mathbb{R}$.

This property (and also its violation) is invariant to strictly increasing transformations. We denote the set of all these transformations (still assumed to be also bijective and hence continuous) by Ψ_{all}. Suppose that we want to allow all chronometric functions that approach $\underline{t}$ asymptotically in the limit but never reach $\underline{t}$. We impose no other assumptions on their shape, such as symmetry across the two different responses. Denote the set of all these functions by $\mathscr{C}^*_{a.all}$ (where a stands for asymptotic). It is generated by the representative member (3) together with the transformations Ψ_{all}. We now construct an empirical function H according to (5) using c^* from (3), which yields

$$H(x) = \begin{cases} 1 - p^1 F^1\left(\underline{t} + \frac{1}{x+1/(\overline{t}-\underline{t})}\right) & \text{if } x > 0, \\ p^0 F^0\left(\underline{t} + \frac{1}{-x+1/(\overline{t}-\underline{t})}\right) & \text{if } x \leq 0. \end{cases} \tag{6}$$

Theorem 1 tells us that full support is detected if H is strictly increasing in x, and a violation of full support is detected if H is not strictly increasing in x. Importantly, here and in subsequent applications, we can express this condition directly in terms of the observed data. Full support is detected if $p^i > 0$ and F^i has full support on $[\underline{t}, \overline{t}]$, for both $i = 0, 1$, and a violation of full support is detected otherwise.

Suppose instead that we had reasons to believe that all chronometric functions reach $c(x) = \underline{t}$ for finite absolute values of x, again without making any other assumptions on their shape. Denote this set by $\mathscr{C}^*_{f.all}$ (where f stands for finite) and observe that it is generated by (4) together with Ψ_{all}. Theorem 1 now tells us that we need to check whether

$$H(x) = \begin{cases} 1 & \text{if } (\overline{t} - \underline{t}) < x, \\ 1 - p^1 F^1(\overline{t} - x) & \text{if } 0 < x \leq (\overline{t} - \underline{t}), \\ p^0 F^0(\overline{t} + x) & \text{if } -(\overline{t} - \underline{t}) \leq x \leq 0, \\ 0 & \text{if } x < -(\overline{t} - \underline{t}), \end{cases} \tag{7}$$

is strictly increasing in x. This is not the case, and we therefore detect that the distribution violates full support. Intuitively, since the chronometric functions reach $\underline{t}$ but the response time distributions have no atoms at $\underline{t}$, the latent distributions cannot have full support.

If we want to allow the union $\mathscr{C}^*_{a.all} \cup \mathscr{C}^*_{f.all}$ of chronometric functions, we can apply the extension discussed in Subsection 2.5.1. If (6) is not strictly increasing, we detect that all distributions that are compatible with the data do not have full support. If (6) is strictly increasing, then we detect neither full support nor a violation of it, because the data are compatible with some distributions that have full support and others that do not.

3.1.2 Sign of Mean

Our first economically relevant application concerns the sign of the mean. In a random utility application where $\tilde{x} = u(1) - u(0) + \tilde{\epsilon}(1) - \tilde{\epsilon}(0)$ and the errors have mean zero, the mean equals $u(1) - u(0)$. Detecting the sign of the mean is therefore the same as deducing the agent's true (non-distorted) ordinal preference between the two options.

The sign of the mean is invariant to linear transformations of the form $\psi(x) = bx$ for $b > 0$. Denote the set of these transformations by Ψ_{lin}. Unfortunately, the sets of chronometric functions which can be generated using Ψ_{lin} are rather restrictive. For example, when starting from the symmetric linear function (4), we can generate the set of all symmetric linear functions. To detect the sign of the mean $u(1) - u(0)$ assuming this class, by Theorem 1 we just need to calculate the sign of the mean of H defined in (7), which can easily be done empirically.[8]

We can achieve more robust detection by working with a property that is sufficient for a positive mean (the argument for a negative mean is analogous). Consider the asymmetry property that $G(-x) \leq 1 - G(x)$ for all $x \in \mathbb{R}_+$. This property implies that the mean of G is positive (see Alós-Ferrer, Fehr and Netzer, 2021). It is invariant to all transformations that are symmetric around zero but not necessarily linear. Denote this set by Ψ_{sym}.

Suppose that we once more allow chronometric functions that approach $\underline{t}$ asymptotically, but further restrict attention to those which are symmetric across responses. The set of all these functions, denoted $\mathscr{C}^*_{a.sym}$, is generated by (3) together with Ψ_{sym}. Intuitively, this set captures the assumption that the chronometric effect is identical for the two choice options. We obtain that the desired asymmetry of G is detected if H defined in (6) exhibits the desired asymmetry. Taken together and expressed directly in terms of the observed data, it follows that

$$p^0 F^0(t) \leq p^1 F^1(t) \text{ for all } t \in [\underline{t}, \overline{t}] \tag{8}$$

[8]Some distributions do not have a mean. This can be dealt with either by using the procedure described in Subsection 2.5.2 to restrict the set of distributions to those which have a mean, or by refining the desired property, for example to "the mean exists and is positive." Analogous arguments apply whenever a property is not well-defined for all possible distributions.

is a sufficient condition for a revealed preference $u(0) \leq u(1)$. We remark that the same condition is obtained when considering the set $\mathscr{C}^*_{f.sym}$ of all symmetric chronometric functions that reach $\underline{t}$ and that analogous statements hold for revealed strict preferences.

Condition (8) is the same as in Theorem 1 of Alós-Ferrer, Fehr and Netzer (2021). They discuss in detail that observing choice frequencies $p^0 \leq p^1$ is not sufficient for a preference $u(0) \leq u(1)$ to be revealed without the additional assumptions on error distributions that are made in conventional logit or probit models. However, if the inequality holds for all response times, as stated in (8), then a preference is robustly revealed without distributional assumptions.[9] They then report that slightly more than 60% of all stochastic choices in the data of Clithero (2018) do robustly reveal a preference.

Several authors have by now used the condition to answer open questions in different settings. Alós-Ferrer, Garagnani and Fehr (2024) show that a sizable fraction of choices that violate stochastic transitivity in different experiments reveal non-transitive preferences and thus cannot be explained by transitive preferences together with noise. Similarly, Alós-Ferrer et al. (2024) show that common ratio and common consequence effects reflect preferences rather than noise, and Castillo and Vitaku (2024) confirm that choice reversals in line with salience theory reflect true preference reversals.

Our approach here suggests generalizations of condition (8). For example, if we have reasons to believe that the chronometric effect is different for the two choice options, we can construct H based on a representative chronometric function c^* that is asymmetric across the options, and we obtain a modified version of (8) which reflects our prior knowledge of the asymmetry. Assume, for example, that the representative function satisfies $c^{*1}(x) = m(c^{*0}(-x))$ for some $m : [\underline{t}, \overline{t}] \to [\underline{t}, \overline{t}]$ and all $x \in \mathbb{R}_+$. Then,

$$p^0 F^0(t) \leq p^1 F^1(m(t)) \text{ for all } t \in [\underline{t}, \overline{t}]$$

is sufficient for detecting $u(0) \leq u(1)$. If responses $i = 1$ are known a priori to be faster than responses $i = 0$, formalized by $m(t) \leq t$ for all t, then the right-hand side is smaller than in (8) and the inequality is harder to satisfy. The converse is true if responses $i = 0$ are faster. If we want to allow for some degree of asymmetry without knowing the details, we can construct multiple functions H^k using different representative functions c^{k*} with varying degrees of asymmetry. For a revealed preference, all modified versions of condition (8) have to hold simultaneously, ultimately resulting in a requirement that the gap between the left- and right-hand sides of (8) must be large enough.

As a side remark—and to illustrate the logic of additional distributional assumptions

[9] Alós-Ferrer, Fehr and Netzer (2021) allow all chronometric functions that are symmetric. One minor difference is that they assume the chronometric function to be unbounded while we assume that $\overline{t}$ is finite.

discussed in Subsection 2.5.2—let us finally assume that any admissible G is symmetric around its mean. In that case, we can try to detect the sign of the median instead of the mean, because the two are identical. Whatever representative chronometric function c^* we use in our construction of H—and therefore irrespective of which $\mathscr{C}^*$ we want to generate—the sign of the median of H equals the sign of $p^1 - p^0$. This mirrors a well-known result (stated, for example, as Proposition 2 in Alós-Ferrer, Fehr and Netzer, 2021): under the assumption of symmetric noise distributions, choice frequencies reveal preferences without the need to rely on response time data.

3.1.3 Inequality

Our method can be used to detect inequality or dispersion of a distribution, which has applications across multiple fields. For instance, within the literature on subjective well-being, there is a substantial interest in understanding the inequality of happiness (e.g., Stevenson and Wolfers, 2008). Similarly, researchers have studied societal polarization by measuring the dispersion of individual attitudes towards social and political issues (DiMaggio, Evans and Bryson, 1996; Evans, 2003). In the context of market competition, the spread of consumer preferences has direct implications for the optimal pricing and advertising strategies of firms (Johnson and Myatt, 2006; Hefti, Liu and Schmutzler, 2022). Unfortunately, standard measures of inequality like the Gini index require cardinal information, which makes their application to ordered response data questionable (see the discussion in Dutta and Foster, 2013). Response times may serve as a source of cardinal information, even when decisions are binary such as whether to buy or not buy a product (Cotet, Zhao and Krajbich, 2025).

The Lorenz curve is a convenient graphical representation of a distribution's inequality, and interesting measures of inequality like the Gini index are based on the Lorenz curve (Atkinson, 1970; Cowell, 2011). For any distribution G, the Lorenz curve is defined by

$$L(q, G) = \frac{\int_0^q G^{-1}(x)dx}{\int_0^1 G^{-1}(x)dx} \quad \text{for all } q \in [0, 1],$$

where $G^{-1}(x) := \inf\{y \mid G(y) \geq x\}$ denotes the left inverse of G. In the context of subjective well-being, $L(q, G)$ could be understood as the proportion of total happiness allocated to the least happy $100q$ percent of the population. How far the Lorenz curve falls below the 45-degree line is an indication of how unequal the distribution is. The curve is invariant to all linear transformations Ψ_{lin} of the distribution G. Therefore, if we plot the Lorenz curve for an empirical function H like in (6) or (7), or based on any other representative function c^*, exactly this Lorenz curve (and any measure based on it like the Gini index) is detected for the class of chronometric functions that are linearly generated from c^*. As before, we can repeat this procedure for various different representative functions c^{k*} to obtain bounds on the true Lorenz curve for larger sets of chronometric functions.

We just remark that analogous arguments apply to other distributional properties that are invariant to linear transformations, like skewness or kurtosis.

3.1.4 Unimodality

We close the section on single distributions with the property that the distribution is unimodal with mode at zero, i.e., G is convex below zero and concave above zero, and strictly so except when $G(x) \in \{0, 1\}$. Unimodality is of interest once more in political applications where G describes the distribution of political attitudes in a population. Unimodality reflects a centered population where more extreme positions receive less support. An ongoing debate in political science concerns whether political attitudes follow unimodal distributions or are polarized and better described by bimodal distributions (see Lelkes, 2016; Vaeth, 2023). Analysts often want to learn about these properties from survey responses, but as Vaeth (2023) points out, ordinal responses are inadequate to test for properties like uni- or bimodality of the underlying distribution.

The property of unimodality is invariant to (sigmoid) transformations that satisfy $\psi(0) = 0$ and are weakly convex below zero and weakly concave above zero, the set of which is denoted Ψ_{sig}. Starting from a representative chronometric function c^*, the transformations Ψ_{sig} allow us to generate all chronometric functions which are weakly "more convex" than c^* on $\mathbb{R}_-$ and on $\mathbb{R}_+$ (because c^* is increasing on $\mathbb{R}_-$ but decreasing on $\mathbb{R}_+$). We will use this insight in a slightly different way than before. Assume that the observed cdfs F^i are strictly increasing on $[\underline{t}, \overline{t}]$. Then, construct from the data the representative chronometric function

$$c^*(x) = \begin{cases} \underline{t} & \text{if } 1 < x, \\ (F^1)^{-1}(1-x) & \text{if } 0 < x \leq 1, \\ (F^0)^{-1}(1+x) & \text{if } -1 \leq x \leq 0, \\ \underline{t} & \text{if } x < -1, \end{cases} \tag{9}$$

which is a member of $\mathscr{C}^*_{f.all}$. With this function, the empirical distribution H becomes

$$H(x) = \begin{cases} 1 & \text{if } 1 < x, \\ 1 - p^1 + p^1 x & \text{if } 0 < x \leq 1, \\ p^0 + p^0 x & \text{if } -1 \leq x \leq 0, \\ 0 & \text{if } x < -1, \end{cases}$$

which is piecewise linear. Applying any strictly sigmoid transformation to (9) generates a chronometric function in $\mathscr{C}^*_{f.all}$ that is piecewise more convex and results in a unimodal H. By the same logic, applying any inverse-sigmoid transformation generates a piecewise more concave chronometric function and a distribution H that is not unimodal. The constructed function (9) therefore delimits sets of chronometric functions for which we can detect or reject unimodality. It follows from Theorem 1 that unimodality is detected for all those functions that are more convex than (9) and rejected for those that are more concave. This approach does not cover all possible chronometric functions but may yield expressive results. For example, if (9) plotted from the data is already strongly convex, then it appears unlikely that the true chronometric function is even more convex, and we may be able to reject the assumption of unimodality of the distribution.

3.2 Multiple Distributions

3.2.1 First-Order Stochastic Dominance

As a first application involving more than one distribution, consider the property of (G_1, G_2) that G_1 first-order stochastically dominates G_2, i.e., $G_1(x) \leq G_2(x)$ for all $x \in \mathbb{R}$. In the context of happiness surveys, Bond and Lang (2019) have pointed out that conventional probit or logit models make that assumption when comparing two (or more) groups of survey participants. The assumption is crucial for the results of these models, as it yields a ranking of the groups' average happiness for any choice of the happiness scale. Without the assumption, the sign of estimated parameters can often be flipped by using a different scale. For example, rich survey participants may be less happy on average than poor participants despite responding to be happy more frequently. It is difficult to test FOSD using only response data.

For our approach here, observe that FOSD (and also its violation) is invariant to all profiles (ψ_1, ψ_2) of increasing transformations that satisfy $\psi_1 = \psi_2$. Denote the set of all these profiles by $\Psi_{all.i}$ (where i stands for identical across the index j). Let $\mathscr{C}^*_{a.all.i}$ be the set of all profiles of chronometric functions that approach $\underline{t}$ asymptotically and are identical across indices. This set is generated by the representative chronometric function (3) for all $j \in J$ together with $\Psi_{all.i}$. Intuitively, it embodies the assumption that the chronometric effect is identical for all groups but otherwise unrestricted. By Theorem 1, we now need to check whether H_1 first-order stochastically dominates H_2, where H_j is defined as in (6) using the observed data $(p_j^0, p_j^1, F_j^0, F_j^1)$ of group $j = 1, 2$. Expressed directly in terms of the data, we detect FOSD if

$$p_1^0 F_1^0(t) - p_2^0 F_2^0(t) \leq 0 \leq p_1^1 F_1^1(t) - p_2^1 F_2^1(t) \tag{10}$$

for all $t \in [\underline{t}, \bar{t}]$, and a violation of FOSD otherwise. We remark that the same condition is obtained when considering the set $\mathscr{C}^*_{f.all.i}$ of identical chronometric functions that reach $\underline{t}$ and that an analogous statement holds when a strict inequality for some x is required in the definition of FOSD.

Condition (10) equals conditions (i) and (ii) of Proposition 2 in Liu and Netzer (2023).[10] As they point out, for $t = \bar{t}$ condition (10) reduces to $p_1^0 \leq p_2^0$, which is the condition under which conventional probit or logit models conclude that group $j = 1$ is happier than group $j = 2$. To arrive at this conclusion without making distributional assumptions, the inequalities in (10) must hold for all t. Liu and Netzer (2023) test these inequalities using data from an online survey. They show that the null hypothesis of FOSD often cannot be rejected, in particular in cases where the probit model yields significant parameter estimates, indicating that the results of conventional models often seem to be robust at least qualitatively.

Castillo and Vitaku (2024) have subsequently used the condition to confirm the robustness of their test of salience theory.

Our approach here suggests possible generalizations. We already discussed that subjects may differ in their decision speed and how this can be addressed by normalizing response times using a baseline response. There may be cases where this approach does not work, for example when speed heterogeneity in a survey is question-specific and cannot be neutralized by normalization. A different approach that accounts for group-specific decision speed is to construct the functions H_j based on group-specific representative chronometric functions c_j^*. The result would be asymmetric versions of (10). Imprecise knowledge of group differences could once more be captured by working with multiple functions H_j^k and checking multiple corresponding conditions, giving rise to a more demanding but more robust test.

[10] Proposition 2 in Liu and Netzer (2023) characterizes detection of first-order stochastic dominance using general ordered response models for surveys with more than two response categories. Their condition (iii) applies to the intermediate response categories and coincides with the demanding condition by Bond and Lang (2019), as response times are not monotone and hence not informative in the intermediate categories.

3.2.2 Ranking of Means

Suppose that we want to detect whether the mean of G_1 is larger than the mean of G_2. A first application where the comparison of means matters is once more the case of surveys, where we want to learn whether one group is happier than another on average. We will discuss a second application at the end of this subsection.

The ranking of the means is invariant to identical positive affine transformations of the form $\psi_j(x) = a + bx$ for $b > 0$. Since we can only use the subset of those transformations which satisfy $a = 0$ when generating chronometric functions, we restrict attention to the set $\Psi_{lin.i}$ of identical linear transformations right away. The set of chronometric functions that can be generated by linear transformations is rather restrictive, as we discussed before. However, if the restriction holds, the analysis can be remarkably simple. As an example, if we have reasons to believe that the chronometric functions are symmetric, linear, and identical for both $j = 1, 2$, then we can simply compare the means of H_1 and H_2, where H_j is defined as in (7) using the observed data $(p_j^0, p_j^1, F_j^0, F_j^1)$ of group $j = 1, 2$.

It is again possible to achieve a more robust detection using conditions that are sufficient for a ranking of the means. One sufficient condition is of course first-order stochastic dominance of G_1 over G_2, which we discussed before. Hence, (10) is a sufficient condition for detecting that the mean of G_1 is larger than the mean of G_2, using only the assumption that the chronometric function is the same in both groups $j = 1, 2$. Another sufficient condition would be the detection that the mean of G_1 is positive while the mean of G_2 is negative, which we discussed in Subsection 3.1.2. It thus follows that

$$p_2^1 F_2^1(t) - p_2^0 F_2^0(t) \leq 0 \leq p_1^1 F_1^1(t) - p_1^0 F_1^0(t) \tag{11}$$

for all $t \in [\underline{t}, \overline{t}]$ is also sufficient for detecting a ranking of the means. Comparing conditions (10) and (11) is instructive. In (10), we compare response times between the groups j (by calculating a difference of the distributions) but not between the response categories i. The result therefore generates detection under the assumption that the chronometric function is identical for the two groups but not necessarily symmetric across the two responses. In (11), we compare response times between response categories i but not between groups j. It therefore generates detection under the assumption that the chronometric function is symmetric across responses but not necessarily identical between the two groups.[11]

[11]Formally, the condition that we are aiming to detect with (11) is invariant to transformations that are symmetric around zero but possibly different between the groups, the set of which could be denoted $\Psi_{sym.d}$ (where d stands for different across the index j). These transformations can, for example, be used to generate the sets $\mathscr{C}^*_{a.sym.d}$ or $\mathscr{C}^*_{f.sym.d}$ of profiles of chronometric functions with the described properties. Condition (11) and the following condition (12) were first derived in our earlier working paper Liu and Netzer (2020).

We now discuss a third sufficient condition for a ranking of the means. Consider the property that $G_1(x)+G_1(-x) \leq G_2(x)+G_2(-x)$ for all $x \in \mathbb{R}_+$. This condition implies that the mean of G_1 is larger than the mean of G_2 (see Appendix C). Furthermore, the condition is invariant to transformations that are identical for both $j = 1, 2$ and symmetric around zero. Denote this set by $\Psi_{sym.i}$.

Starting from the representative functions (3) for all $j \in J$, we can use $\Psi_{sym.i}$ to generate the set $\mathscr{C}^*_{a.sym.i}$ of all chronometric functions that approach $\underline{t}$ asymptotically and are symmetric across responses and identical across indices. Using this set, Theorem 1 implies that our desired inequality condition is detected if it holds for the functions H_1 and H_2 that are constructed based on (6). Taken together and expressed directly in terms of the observed data, it follows that

$$p_1^0 F_1^0(t) - p_2^0 F_2^0(t) \leq p_1^1 F_1^1(t) - p_2^1 F_2^1(t) \tag{12}$$

for all $t \in [\underline{t}, \bar{t}]$ is another sufficient condition for detecting that the mean of G_1 is larger than the mean of G_2. We remark that the same condition is obtained when allowing the respective set $\mathscr{C}^*_{f.sym.i}$ of chronometric functions that reach $\underline{t}$ and that an analogous statement holds for detecting a strict inequality of means.

Inequality (12) is a weaker requirement than the directly comparable conditions (10) or (11). However, as it implements a comparison of response times across the groups and across the responses, it generates detection only under the stronger combination of the assumptions required in (10) and (11), namely that the chronometric functions are symmetric across responses and identical between groups.

We now discuss another application that involves the comparison of two means. Suppose that we observe the choices of a single agent between the two options x and z and between the two options y and z. Can we infer the agent's preference between x and y from these choices? Sometimes this is possible based on transitivity of preferences, for example if x is chosen over z and z is chosen over y. If, by contrast, both x and y are chosen over z, then we cannot rank x and y directly. Krajbich, Oud and Fehr (2014) noted, however, that the preference can be deduced from response times under the assumption of a monotone chronometric effect. When the choice of x over z is faster than the choice of y over z, then $u(x) - u(z)$ must be larger than $u(y) - u(z)$ and we can conclude that $u(y) \leq u(x)$ (see also Echenique and Saito, 2017). Alós-Ferrer, Fehr and Netzer (2021) provide a generalization of this argument for stochastic choice under the assumption of symmetric utility distributions.

Following their setting, suppose that the random utility difference between x and z is described by a cdf G_{xz} with mean $u(x) - u(z)$, and the random utility difference between y

and z is described by G_{yz} with mean $u(y) - u(z)$. Deducing a revealed preference for x over y can be rephrased as detecting that the mean of G_{xz} is larger than the mean of G_{yz}. We can put several of our above results to work. First, one sufficient condition is that G_{xz} first-order stochastically dominates G_{yz}, which we detect (under the above-described assumptions on the chronometric functions) when

$$p^z_{xz}F^z_{xz}(t) - p^z_{yz}F^z_{yz}(t) \leq 0 \leq p^x_{xz}F^x_{xz}(t) - p^y_{yz}F^y_{yz}(t)$$

for all $t \in [\underline{t}, \bar{t}]$, where lower indices describe the binary choice problem and upper indices describe the chosen option. Another sufficient condition is that the mean of G_{xz} is positive while the mean of G_{yz} is negative, which we detect (under different assumptions on the chronometric functions) when

$$p^y_{yz}F^y_{yz}(t) - p^z_{yz}F^z_{yz}(t) \leq 0 \leq p^x_{xz}F^x_{xz}(t) - p^z_{xz}F^z_{xz}(t)$$

for all $t \in [\underline{t}, \bar{t}]$. Finally, a weaker sufficient condition for detecting the out-of-sample preference (but under stricter assumptions on the chronometric functions) is

$$p^z_{xz}F^z_{xz}(t) - p^z_{yz}F^z_{yz}(t) \leq p^x_{xz}F^x_{xz}(t) - p^y_{yz}F^y_{yz}(t)$$

for all $t \in [\underline{t}, \bar{t}]$. To our knowledge, none of these conditions has been studied in the context of individual choice. We emphasize that we obtain the revealed preference between x and y without making any assumptions on the shape of the utility distributions, but remark that a revealed preference translates into an out-of-sample prediction of choice probabilities only with additional distributional assumptions such as symmetry.

We can also follow Alós-Ferrer, Fehr and Netzer (2021) and assume right away that G_{xz} and G_{yz} are symmetric around their means. Since mean and median coincide in this case, we can instead try to detect whether the median of G_{xz} is larger than that of G_{yz}. This property is invariant to the set $\Psi_{all.i}$ of all transformations that are identical for the two distributions. We can therefore detect the property assuming either $\mathscr{C}^*_{a.all.i}$ or $\mathscr{C}^*_{f.all.i}$, which means that we only need to assume that the chronometric effect is the same in the two binary decision problems. Consider then the case where $p^x_{xz} > 1/2$ and $p^y_{yz} > 1/2$, which under symmetry implies that both $u(x) - u(z)$ are $u(y) - u(z)$ are strictly positive, so that a preference between x and y does not follow from transitivity. Define θ_{xz} and θ_{yz} as percentiles of the response time distributions when x or y were chosen over z, respectively, as follows:

$$F^x_{xz}(\theta_{xz}) = \frac{1}{2p^x_{xz}} \quad \text{and} \quad F^y_{yz}(\theta_{yz}) = \frac{1}{2p^y_{yz}}.$$

It is now an easy exercise to show that the median of H_{xz} is larger than the median of H_{yz}, where both functions are constructed either as in (6) or as in (7), if and only if

$$\theta_{xz} \leq \theta_{yz}. \tag{13}$$

Analogous statements hold for strict inequalities and for the case $p^x_{xz} < 1/2$ and $p^y_{yz} < 1/2$.

Inequality (13) is the condition stated in Theorem 2 of Alós-Ferrer, Fehr and Netzer (2021) for a revealed preference $u(y) \leq u(x)$ under the assumption of symmetric distributions. As these authors discuss in detail, (13) formalizes that the choice of x over z is faster than the choice of y over z, in a setting with stochastic choices and response times. They then use the result to make out-of-sample predictions in the data of Clithero (2018) and show that about 80% of the predictions are correct, as compared to only about 74% for a conventional logit model.

Alós-Ferrer and Garagnani (2024) have subsequently used the result to predict subjects' choices under risk in different experiments. They show that it outperforms the predictions of various structural models of behavior.

Our approach here adds to the existing results in different ways. First, it clarifies the intuition underlying (13) as a condition which detects the ranking of two medians. Second, it therefore shows that the same out-of-sample elicitation of preferences obtains under any other distributional assumption which implies that the ranking of medians is the same as the ranking of means, not just symmetry. Third, it shows that the assumption of a symmetric chronometric function, which Alós-Ferrer, Fehr and Netzer (2021) made throughout their paper, is in fact not necessary for this result to hold. Finally, it once more paves the way for generalizations which allow for different chronometric functions across binary choice problems. For example, if the representative chronometric functions in the choice problems are $c^*_{xz} = c^*$ and $c^*_{yz} = m \circ c^*$ for some functions c^* and m, then the condition analogous to (13) becomes

$$\theta_{xz} \leq m^{-1}(\theta_{yz}). \tag{14}$$

If we have reasons to believe that the choices of y are faster than those of x for any fixed utility difference to z (for example because y and z are easier to compare than x and z, see Gonçalves, 2024), then the transformation m^{-1} makes the right hand side of (14) larger and the inequality easier to satisfy. The opposite is true when the choices of x are faster. General robustness considerations could be taken into account once more by requiring a sufficiently large gap between θ_{xz} and θ_{yz}.

3.2.3 Likelihood-Ratio Dominance

Consider the property that G_1 likelihood-ratio dominates G_2, defined by the inequality

$$(G_1(x) - G_1(x''))(G_2(x) - G_2(x')) \leq (G_1(x) - G_1(x'))(G_2(x) - G_2(x''))$$

for all $x'' < x' < x$ (see Wang and Lehrer, 2024). If G_1 and G_2 are absolutely continuous, the property can be equivalently expressed in terms of their density functions g_1 and g_2 as $g_1(x')g_2(x) \leq g_1(x)g_2(x')$ for any $x' < x$ (see Shaked and Shanthikumar, 2007). Likelihood-ratio dominance, which is stronger than FOSD, has proven useful in a range of economic applications that involve monotone comparative statics under uncertainty (Milgrom, 1981; Athey, 2002). Suppose that the goal is to maximize the expected value of a function $\pi(y, x)$ by choice of y when x is distributed according to G_j. If $\pi(y, x)$ satisfies a single-crossing property and G_1 likelihood-ratio dominates G_2, then the optimal choice of y is larger for G_1 than for G_2 (under appropriate technical conditions, see Athey, 2002). For example, a firm may introduce a new product in two different countries, each of which is characterized by a distribution of consumer types. Higher types imply a higher demand for the product in a way that the firm's profit $\pi(p, x)$ as a function of price p and type x is single-crossing. If the firm's market research (e.g., a survey that asks potential consumers whether they like the product) allows the detection of likelihood-ratio dominance of the type distributions of the two countries, the firm knows in which country to charge a higher price. Other applications of the concept include optimal investment decisions and bidding in auctions.

Like FOSD, likelihood-ratio dominance (and its violation) is invariant to all profiles of transformations in $\Psi_{all,i}$. We thus proceed as for FOSD by constructing (H_1, H_2) defined in either (6) or (7) and then verifying whether H_1 likelihood-ratio dominates H_2. For example, checking the respective inequalities for all $x'' < x' < x \leq 0$ can be expressed in terms of the data as

$$(p_1^0 F_1^0(t) - p_1^0 F_1^0(t''))(p_2^0 F_2^0(t) - p_2^0 F_2^0(t')) \leq (p_1^0 F_1^0(t) - p_1^0 F_1^0(t'))(p_2^0 F_2^0(t) - p_2^0 F_2^0(t''))$$

for all $t'' < t' < t$. The full condition is particularly easy to express when $p_j^i > 0$ and the response time distributions F_j^i admit strictly positive densities f_j^i. In that case, we detect likelihood-ratio dominance if the empirical likelihood-ratio

$$\frac{p_1^i f_1^i(t)}{p_2^i f_2^i(t)}$$

is weakly increasing in t for $i = 0$ and weakly decreasing in t for $i = 1$, and

$$\frac{p_1^0 f_1^0(\bar{t})}{p_2^0 f_2^0(\bar{t})} \leq \frac{p_1^1 f_1^1(\bar{t})}{p_2^1 f_2^1(\bar{t})}$$

holds. Otherwise, a violation of likelihood-ratio dominance is detected.

Our approach extends analogously to the detection of hazard-rate dominance and reversed hazard-rate dominance, both of which lie between likelihood-ratio dominance and FOSD, and which are also commonly used in applied settings (e.g., Kiefer, 1988; Maskin and Riley, 2000; Wang and Lehrer, 2024).

3.2.4 Ranking of Inequality

We can return to our discussion of inequality from Subsection 3.1.3 and study the comparison of two distributions in terms of their dispersion. The subjective well-being literature has been interested in how inequality of happiness changed over time and across nations (Kalmijn and Veenhoven, 2005; Stevenson and Wolfers, 2008; Dutta and Foster, 2013). There is also a great interest in political polarization trends (DiMaggio, Evans and Bryson, 1996; Evans, 2003). Furthermore, firms have an interest in learning about changes in the spread of consumer preferences to optimally adapt their pricing and advertising strategies (Johnson and Myatt, 2006; Hefti, Liu and Schmutzler, 2022).

Consider first the problem of detecting whether G_1 has a smaller variance than G_2, which is a convenient way of ranking the dispersion of distributions due to its ability to provide a complete order. Just as for the ranking of means, this property is invariant only to transformations (ψ_1, ψ_2) that are linear and additionally satisfy $\psi_1 = \psi_2$. Comparing instead the inequality based on Lorenz curves has the benefit of avoiding direct scale comparisons, as the two distributions' Lorenz curves are invariant to transformations (ψ_1, ψ_2) that are linear but not necessarily identical. For example, we can say that G_1 Lorenz-dominates G_2 if $L(q, G_1) \geq L(q, G_2)$ holds for all $q \in [0, 1]$ (see Shaked and Shanthikumar, 2007). Then, under the assumption of knowing the chronometric functions up to group-specific linear transformations, detecting a Lorenz-dominance relationship between G_1 and G_2 (or the absence thereof) is equivalent to checking whether the property holds for the respective empirical functions (H_1, H_2).

Another measure that can be used for assessing the relative dispersion of distributions is the concept of single-crossing dominance (Diamond and Stiglitz, 1974; Hammond, 1974). Formally, G_1 single-crossing dominates G_2 if there exists x^* such that $G_1(x) \leq G_2(x)$ if $x \leq x^*$ and $G_1(x) \geq G_2(x)$ if $x \geq x^*$. Intuitively, the condition reflects that G_1 assigns less probability weight to values at the tails of the distribution compared to G_2. Although the property of single-crossing dominance is sensitive to group-specific transformations, it has the advantage of being robust to non-linear transformations. Indeed, just like FOSD, the property (and its violation) is invariant to all profiles of transformations in $\Psi_{all,i}$. Therefore, according to Theorem 1, single-crossing dominance is detected for a relatively large class of chronometric functions if H_1 single-crossing dominates H_2, using either (6) or (7), and a violation thereof is detected otherwise. The condition can again be expressed directly in terms of the data. It requires that $p_1^i F_1^i(t)$ and $p_2^i F_2^i(t)$ cross at most once for one of the two responses $i \in \{0, 1\}$ and not at all for the other response.

Finally, it is known that either Lorenz-dominance or single-crossing dominance imply second-order stochastic dominance (SOSD) when G_1 has a higher mean than G_2 (Thistle, 1989). Therefore, combining several of the results derived so far allows us to detect SOSD.

3.2.5 Correlation

Our last theoretical application expands upon the baseline model by investigating the correlation between a latent variable $x \in \mathbb{R}$ (e.g., happiness, political attitude, or willingness to pay) and an observable variable represented by the index $j \in \mathbb{R}$ (e.g., income, social media usage, or experimental treatments). Such an application can be of value in the context of opinion surveys on topics that are sensitive and where subjects hesitate to provide certain answers that conflict with social norms (e.g., Coffman, Coffman and Ericson, 2017), which can result in limited variation in the response data. Our approach may compensate for the lack of power of traditional analysis by replacing response variation with variation in response times. This can also be useful for experimental economists who are interested in the effect of different treatments on subjects' behavior. A null result in behavior does not necessarily establish the absence of an effect. If choices are discrete, such a null result can arise when the effect exists but is not large enough to shift behavior, given the parameters chosen by the experimenter. The effect may still be detectable in response times.

We define a cumulative distribution function Γ over the indices $j \in J \subseteq \mathbb{R}$, which is observable, and interpret $(G_j)_j$ as the conditional distributions of x given each j. The joint distribution of (x, j) is fully determined by $(G_j)_j$ and Γ. Further, just like in the baseline model, given observed data $(p_j, F_j)_j$ and a representative profile $(c_j^*)_j$ of chronometric functions, we can derive empirical distribution functions $(H_j)_j$ as defined in (5). These functions together with Γ also give rise to a well-defined joint distribution of (x, j).

A first attempt to quantify the association between x and j is to employ the standard

Pearson correlation coefficient

$$\rho = \frac{Cov(x, j)}{\sqrt{Var(x)\; Var(j)}}, \tag{15}$$

where Cov indicates the covariance function and Var indicates the variance function. The value of this coefficient is invariant to all profiles $(\psi_j)_j$ of positive affine transformations that are identical for all $j \in J$. Consequently, when the chronometric functions are known up to identical linear transformations, we can ascertain the exact correlational pattern by computing (15) for the joint distribution of (x, j) constructed via the functions $(H_j)_j$ and the marginal distribution Γ.

To circumvent the linearity restriction, one might opt for measuring the rank correlation between x and j, which is particularly natural when x and j are ordinal variables (Kendall, 1955). Intuitively, the rank of a variable is preserved under any monotone transformation, thereby allowing for more robust detection. For example, consider Spearman's (1904) rho, defined as

$$\rho_s = \frac{Cov(G(x), \Gamma(j))}{\sqrt{Var(G(x))\; Var(\Gamma(j))}},$$

where the function G represents the marginal distribution of x and is given by

$$G(x) = \int_J G_j(x)\, d\Gamma(j) \text{ for all } x \in \mathbb{R}.$$

Since the rank of x within the population remains unchanged under any profile $(\psi_j)_j$ of strictly increasing transformations that are identical across j, so does the value of Spearman's rho. A similar observation holds for another rank-based correlation measure, Kendall's (1955) tau, which is defined as

$$\rho_\tau = \mathbb{E}\left[\mathbb{1}_{\{(x-x')(j-j')>0\}}\right] - \mathbb{E}\left[\mathbb{1}_{\{(x-x')(j-j')<0\}}\right],$$

where (x', j') is distributed independently of (x, j) but with the same joint distribution. Consequently, the rank correlation patterns for our variables of interest are detected under a fairly general class of chronometric functions, whenever they hold for the joint distribution induced by the appropriate functions $(H_j)_j$ and Γ.[12]

[12]The invariance property shared by Spearman's rho and Kendall's tau is related to the fact that they both depend only on the bivariate Copula of the two random variables, which is invariant to monotone transformations; see, e.g., Fan and Patton (2014) and Haugh (2016). The population versions of Spearman's rho and Kendall's tau that we adopt here were taken from Haugh (2016).

4 Empirical Application

Easterlin (2005) postulates that "[f]ew generalizations in the social sciences enjoy such wide-ranging support as that of diminishing marginal utility of income", where he interprets utility explicitly as "subjective well-being" (p. 243). Easterlin then criticizes this notion of decreasing marginal happiness by showing that it does not generalize from cross-sectional data to time series data. Oswald (2008) and Kaiser and Oswald (2022) point out that we do not even know whether marginal happiness is decreasing in income for cross-sectional data. The empirically observed relationship $e : W \rightarrow R$ between income and reported happiness is the composition of a first function $h : W \rightarrow H$ that maps income into happiness and a second function $r : H \rightarrow R$ that maps happiness into reported happiness. Observing that $e(w) = r(h(w))$ is concave in w does not imply that $h(w)$ is concave in w and therefore does not establish decreasing marginal happiness. The observed concave relationship may just as well be due to a concave reporting function r. Much of the existing evidence for decreasing marginal happiness is thus based on the assumption of a linear reporting function. Other studies have estimated more advanced ordered response models. However, as Bond and Lang (2019) have shown, the estimated relationship between income and happiness in this case depends on the distributional assumptions of the model, which again impose an arbitrary cardinal scale. Kaiser and Oswald (2022) conclude that the problem is "fundamental, little recognized, and so far unsolved" (p. 3).

To formalize the income-happiness relationship in our framework, let $(G_w)_w$ denote the family of happiness distributions for all possible income levels $w \in W \subseteq \mathbb{R}_+$ and define $\mu(w) = \int_{\mathbb{R}} x dG_w(x)$ to be the average happiness of agents with income w. The question of interest is whether μ is a concave function of w. It appears natural to answer this question by plotting the average response in a happiness survey against income, as many studies have done and which typically results in a concave relationship. We repeat this exercise here using data from an online survey conducted on MTurk by Liu and Netzer (2023). Their data have responses from 3'743 subjects to the binary question about whether they are "rather happy" or "rather unhappy." Our analysis is complicated by the fact that household income is reported only in three broad bins: below \$40'000, between \$40'000 and \$69'999, and above \$70'000. We need to associate each of these bins with a unique income value, denoted w_L, w_M and w_H. We do this by eight different methods, ranging from the simple use of bin midpoints to an advanced procedure that predicts individual incomes based on observables using estimation results from auxiliary PSID data (Survey Research Center, 2021). Details of these methods are described in Table 2 in Appendix D. Figure 2 plots the average response to the happiness question (labeling the responses 0 and 1, so that this average coincides

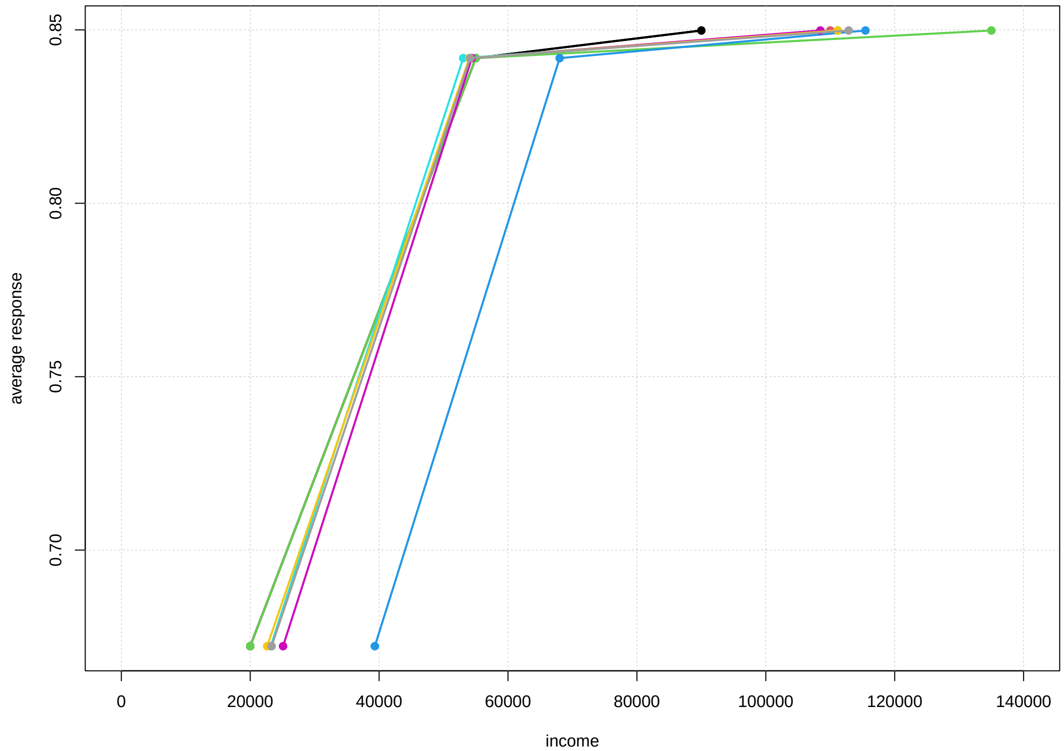


Figure 2: The relationship between income and reported happiness.

Notes: The eight colored curves correspond to the different methods used to determine the average income within each bin.

with the relative frequency of the "rather happy" response) on the y-axis against the three imputed income levels on the x-axis. As expected, the relationship between income and reported happiness is concave, for all eight methods that we use to determine the three income levels. Like in Easterlin (2005) and the literature cited therein, this is a bivariate relation without any additional controls, but the results are analogous when conditioning on various socio-demographic variables.

Our point here is that the concavity in Figure 2 does not establish concavity of μ, which in our context reduces to the single inequality

$$\alpha\mu(w_L) + (1-\alpha)\mu(w_H) \leq \mu(w_M), \tag{16}$$

where α is such that $\alpha w_L + (1-\alpha)w_H = w_M$. Following the logic of Bond and Lang (2019), we can explain the data in Figure 2 with happiness distributions (G_L, G_M, G_H) for which (16) is violated, for example distributions that become more right-skewed as income grows and which therefore go along with high average happiness among the rich. It is impossible to reject or verify such distributional assumptions based on response data alone.

According to (16), decreasing marginal happiness is equivalent to an appropriate ranking of the means of distributions, for which we have developed criteria in Subsection 3.2.2. The left-hand side of (16) is the average happiness in a mixed population composed of a fraction α of subjects with low income and a fraction $1-\alpha$ of subjects with high income. Under the assumption that all income groups have the same chronometric function, which most of our criteria from Subsection 3.2.2 require anyway, we can pool the two extreme income groups with appropriate weights and investigate (16) using response time data.

Two issues arise when applying our results empirically. First, and as discussed earlier, real data may exhibit individual heterogeneity in response speed. We therefore follow Liu and Netzer (2023) and obtain normalized response times by subtracting in logs the response time to the marital-status question. This normalization corroborates further the assumption of identical chronometric functions in the different income groups. Our results in Subsection 2.5.3 then imply that all the distributional properties which we will investigate below can be tested using the respective response time-based detection conditions: under the null hypothesis that the property holds, its detection condition must be satisfied in the normalized data even when there is noise or measurement error (independent and common to all groups). Second, real data is finite. If a detection condition is violated in the data, this may be because of finite sampling rather than a violation of the underlying property. For example, the fastest normalized response in our data comes from a "rather unhappy" subject in the middle income group. With a sufficiently extreme chronometric function, we can explain such data with arbitrarily small average happiness in the middle income group, thus violating (16), even though the data point may be a finite sample artifact. We address this issue by employing bootstrap-based statistical hypothesis tests.

Consider first the distributional property that $G_M(x)+G_M(-x) \leq G_P(x)+G_P(-x)$ for all $x \in \mathbb{R}_+$, where $G_P(x) = \alpha G_L(x) + (1-\alpha)G_H(x)$ is the cdf of happiness in the pooled group. This property implies the ranking of means in (16) and is therefore sufficient for concavity of μ. Its detection condition (12) requires that the chronometric functions are symmetric across responses and identical between groups. In our present context, it becomes the testing condition

$$p^0_M \hat{F}^0_M(t) - p^0_P \hat{F}^0_P(t) \leq p^1_M \hat{F}^1_M(t) - p^1_P \hat{F}^1_P(t) \tag{17}$$

for all $t \geq 0$, where $p^i_P \hat{F}^i_P(t) = \alpha p^i_L \hat{F}^i_L(t) + (1-\alpha) p^i_H \hat{F}^i_H(t)$. Intuitively, this inequality rules out examples like the increasingly right-skewed distributions discussed above, as these would

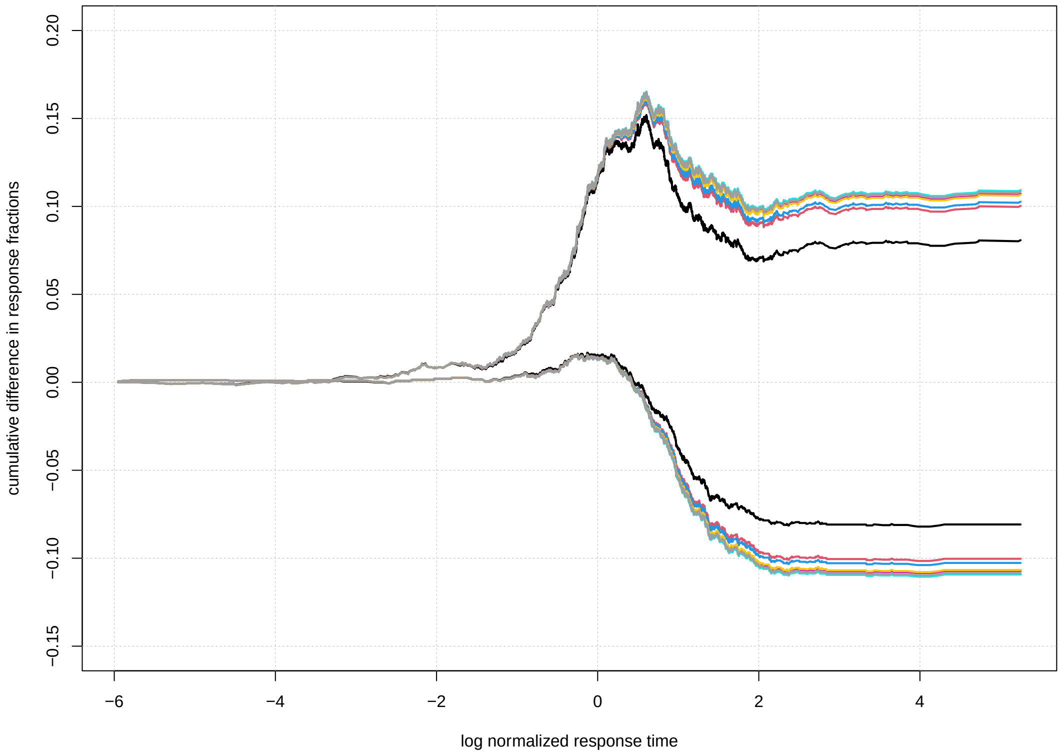


Figure 3: Empirical conditions for testing the income-happiness relation.

Notes: The curves at the top represent the empirical functions $p_M^1 \hat{F}_M^1(t) - p_P^1 \hat{F}_P^1(t)$, while those at the bottom represent $p_M^0 \hat{F}_M^0(t) - p_P^0 \hat{F}_P^0(t)$. Different colors indicate the varying methods used to determine the average income within each bin.

generate relatively fast happy and slow unhappy responses in the pooled group. Figure 3 plots the left-hand side and the right-hand side of (17), for all eight methods used to impute incomes. Inequality (17) is indeed not satisfied exactly in the data, because the left-hand side exceeds the right-hand side for some small t. However, the crossing of the functions appears to be minor, so that the question of statistical significance arises. To obtain p-values for the null hypothesis that (17) holds, we employ a bootstrap-based method as in Liu and Netzer (2023) that rests on a test for conventional first-order stochastic dominance by Barrett and Donald (2003). Table 1 shows that the p-values are very large. We clearly cannot reject our sufficient condition for a ranking of the means in line with decreasing marginal happiness.

We can also test the stronger property that G_M first-order stochastically dominates G_P, which is sufficient for decreasing marginal happiness but whose detection condition (10) does

Method	α		p-value	
		(17)	(18)	(19)
1	0.500	0.9548	0.3380	0.0000
2	0.610	0.9501	0.4105	0.0000
3	0.660	0.9457	0.4397	0.0000
4	0.623	0.9481	0.4215	0.0000
5	0.660	0.9417	0.4405	0.0000
6	0.649	0.9427	0.4364	0.0000
7	0.646	0.9445	0.4350	0.0000
8	0.656	0.9482	0.4391	0.0000

Table 1: Summary of tests for concavity.

not require symmetry of the chronometric functions. The testing condition here becomes

$$p_M^0 \hat{F}_M^0(t) - p_P^0 \hat{F}_P^0(t) \leq 0 \leq p_M^1 \hat{F}_M^1(t) - p_P^1 \hat{F}_P^1(t) \tag{18}$$

for all $t \geq 0$, which in Figure 3 requires that the functions are separated by zero. While not satisfied exactly in the data, Table 1 shows that the p-values of the null hypothesis that (18) holds are smaller than for (17) but still large.[13] Thus, we cannot reject even this strong sufficient condition for decreasing marginal happiness that applies under weak assumptions on the chronometric function.

For completeness, Table 1 also reports p-values for the null hypothesis that

$$p_P^1 \hat{F}_P^1(t) - p_P^0 \hat{F}_P^0(t) \leq 0 \leq p_M^1 \hat{F}_M^1(t) - p_M^0 \hat{F}_M^0(t) \tag{19}$$

holds for all $t \geq 0$, which implements (11) and can be used to test whether the left-hand side of (16) is negative while the right-hand side is positive. This condition is of less interest. It is very strong, and its advantage of allowing group-specific chronometric functions has no bite because our approach of pooling groups requires identical chronometric functions (after normalization) anyway. The null hypothesis that (19) holds is clearly rejected.

We could also investigate the opposite hypothesis that marginal happiness increases with income, which corresponds to reversing all the inequalities studied so far. Following the same procedures, we find that the reversed versions of (17), (18), and (19) are rejected at all plausible levels of statistical significance (all $p = 0.0000$).

[13]The test implements the procedure of Barrett and Donald (2003) together with a joint hypothesis correction by Romano and Wolf (2016) to account for the fact that condition (18) contains two inequalities. We refer the reader to Liu and Netzer (2023) for details.

To summarize, our results support the idea that marginal happiness is decreasing in income, in a cross-sectional data set. While our tests avoid several of the pitfalls noted at the beginning of this section, some limitations remain. Most notably, we were only able to test the concavity of expected happiness for three income levels implied by the survey bins, rather than across the entire income distribution. Additionally, the income-happiness relation is bivariate without controlling for potential confounding factors. Future research that addresses these limitations holds great promise.

5 Conclusion

The goal of this paper is to provide a systematic account of the information that response time data contain. We approach the problem by phrasing it as one of identification in the context of binary response models. Our main result relates the set of identifiable distributional properties to the set of admissible chronometric functions. The fundamental idea is that the joint distribution of responses and response times identifies a composition of the latent distribution and the chronometric function. Properties of the distribution that are preserved under a given set of transformations can therefore be identified if the chronometric function is known up to these transformations. Several existing results in the literature follow as corollaries and can be generalized. Many new results emerge. To illustrate the applicability of our approach, we empirically test the hypothesis of decreasing marginal happiness of income and find that it cannot be rejected.

Our theoretical applications in Section 3 are merely examples of the scope of the method and not an exhaustive list. Additional properties that one could study include general linear relationships between observable variables and the latent variable like in regression models, as well as the extent to which responses are potentially distorted by the framing of a decision problem. Similarly, our empirical study in Section 4 is only one straightforward application showing how the method can be used to contribute to long-standing debates. Other applications that we have in mind include the study of polarization using surveys on political attitudes, as well as optimal product pricing using data on purchase decisions from online platforms. There are also several possible extensions of our framework that merit investigation. These include correlations between multiple latent variables, the case with more than two choice options, and the use of response times in settings where these times are affected by additional factors like player types or decision modes as in Rubinstein (2007, 2013, 2016).

References

Alós-Ferrer, Carlos, and Michele Garagnani. 2022. "Strength of preference and decisions under risk." *Journal of Risk and Uncertainty*, 64: 309–329.

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

Alós-Ferrer, Carlos, Ernst Fehr, and Nick Netzer. 2021. "Time Will Tell: Recovering Preferences when Choices Are Noisy." *Journal of Political Economy*, 129(6): 1828–1877.

Alós-Ferrer, Carlos, Ernst Fehr, Helga Fehr-Duda, and Michele Garagnani. 2024. "Common Ratio and Common Consequence Effects Arise from True Preferences." Mimeo.

Alós-Ferrer, Carlos, Michele Garagnani, and Ernst Fehr. 2024. "Identifying Nontransitive Preferences." Mimeo.

Arrow, Kenneth J. 1958. "Rational Choice Functions and Orderings." *Economica*, 26(102): 121–127.

Athey, Susan. 2002. "Monotone Comparative Statics Under Uncertainty." *The Quarterly Journal of Economics*, 117(1): 187–223.

Atkinson, Anthony B. 1970. "On the Measurement of Inequality." *Journal of Economic Theory*, 2(3): 244–263.

Barrett, Garry F., and Stephen G. Donald. 2003. "Consistent Tests for Stochastic Dominance." *Econometrica*, 71(1): 71–104.

Benkert, Jean-Michel, and Nick Netzer. 2018. "Informational Requirements of Nudging." *Journal of Political Economy*, 126(6): 2323–2355.

Bond, Timothy N., and Kevin Lang. 2019. "The Sad Truth About Happiness Scales." *Journal of Political Economy*, 127(4): 1629–1640.

Card, David, Stefano DellaVigna, Chenxi Jiang, and Dmitry Taubinsky. 2024. "Understanding Expert Choices Using Decision Time." Mimeo.

Castillo, Marco, and Valon Vitaku. 2024. "Salience in Choice Under Risk: An Experimental Investigation." Mimeo.

Chabris, Christopher F., Carrie L. Morris, Dmitry Taubinsky, David Laibson, and Jonathon P. Schuldt. 2009. "The Allocation of Time in Decision-Making." *Journal of the European Economic Association*, 7(2-3): 628–637.

Clithero, John 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.

Coffman, Katherine B, Lucas C Coffman, and Keith M Marzilli Ericson. 2017. "The Size of the LGBT Population and the Magnitude of Antigay Sentiment Are Substantially Underestimated." *Management Science*, 63(10): 3168–3186.

Cotet, Miruna, Wenjia Joyce Zhao, and Ian Krajbich. 2025. "Deliberation During Online Bargaining Reveals Strategic Information." *Proceedings of the National Academy of Sciences*, 122(7): e2410956122.

Cowell, Frank Alan. 2011. *Measuring Inequality.* Oxford University Press.

d'Aspremont, Claude, and Louis Gevers. 2002. "Social welfare functionals and interpersonal comparability." In *Handbook of Social Choice and Welfare*. Vol. 1, , ed. K.J. Arrow, A. K. Se and K. Suzumura, Chapter 10, 459–541. Elsevier.

Diamond, Peter, and Joseph Stiglitz. 1974. "Increases in Risk and in Risk Aversion." *Journal of Economic Theory*, 8(3): 337–360.

DiMaggio, Paul, John Evans, and Bethany Bryson. 1996. "Have American's Social Attitudes Become More Polarized?" *American Journal of Sociology*, 102(3): 690–755.

Dutta, Indranil, and James Foster. 2013. "Inequality of Happiness in the US: 1972–2010." *Review of Income and Wealth*, 59(3): 393–415.

Easterlin, Richard A. 2005. "Diminishing Marginal Utility of Income? Caveat Emptor." *Social Indicators Research*, 70(3): 243–255.

Echenique, Federico, and Kota Saito. 2017. "Response Time and Utility." *Journal of Economic Behavior & Organization*, 139: 49–59.

Evans, John H. 2003. "Have Americans' Attitudes Become More Polarized? – An Update." *Social Science Quarterly*, 84(1): 71–90.

Fan, Yanqin, and Andrew J Patton. 2014. "Copulas in Econometrics." *Annual Review of Economics*, 6(1): 179–200.

Fudenberg, Drew, Philipp Strack, and Tomasz Strzalecki. 2018. "Speed, accuracy, and the optimal timing of choices." *American Economic Review*, 108(12): 3651–3684.

Gonçalves, Duarte. 2024. "Speed, Accuracy, and Complexity." Mimeo.

Haile, Philip A., Ali Hortaçsu, and Grigory Kosenok. 2008. "On the Empirical Content of Quantal Response Equilibrium." *American Economic Review*, 98(1): 180–200.

Hammond, John S. 1974. "Simplifying the Choice Between Uncertain Prospects Where Preference Is Nonlinear." *Management Science*, 20(7): 1047–1072.

Haugh, Martin. 2016. "An Introduction to Copulas." Lecture Notes, `www.columbia.edu/~mh2078/QRM/Copulas.pdf`.

Hefti, Andreas, Shuo Liu, and Armin Schmutzler. 2022. "Preferences, Confusion and Competition." *The Economic Journal*, 132(645): 1852–1881.

Johnson, Justin P, and David P Myatt. 2006. "On the Simple Economics of Advertising, Marketing, and Poduct Design." *American Economic Review*, 96(3): 756–784.

Kaiser, Caspar, and Andrew J. Oswald. 2022. "Inequality, Well-Being, and the Problem of the Unknown Reporting Function." *Proceedings of the National Academy of Sciences*, 119(50): e2217750119.

Kalmijn, Wim, and Ruut Veenhoven. 2005. "Measuring Inequality of Happiness in Nations: In Search for Proper Statistics." *Journal of Happiness Studies*, 6(4): 357–396.

Kellogg, W. N. 1931. "The Time of Judgment in Psychometric Measures." *American Journal of Psychology*, 43(1): 65–86.

Kendall, Maurice George. 1955. *Rank Correlation Methods.* New York: Hafner Publishing Co.

Kiefer, Nicholas M. 1988. "Economic Duration Data and Hazard Functions." *Journal of Economic Literature*, 26(2): 646–679.

Konovalov, Arkady, and Ian Krajbich. 2019. "Revealed Strength of Preference: Inference From Response Times." *Judgment & Decision Making*, 14(4): 381–394.

Krajbich, Ian, Bastiaan Oud, and Ernst Fehr. 2014. "Benefits of Neuroeconomic Modeling: New Policy Interventions and Predictors of Preference." *American Economic Review: Papers & Proceedings*, 105(5): 501–506.

Krajbich, Ian, Dingchao Lu, Colin Camerer, and Antonio Rangel. 2012. "The Attentional Drift-Diffusion Model Extends to Simple Purchasing Decisions." *Frontiers in Psychology*, 3. Article 193.

Lelkes, Yptach. 2016. "Mass Polarization: Manifestations and Measurements." *Public Opinion Quarterly*, 80(S1): 392–410.

Liu, Shuo, and Nick Netzer. 2020. "Happy Times: Identification From Ordered Response Data." University of Zurich, Department of Economics, Working Paper No. 371.

Liu, Shuo, and Nick Netzer. 2023. "Happy Times: Measuring Happiness Using Response Times." *American Economic Review*, 113: 3289–3322.

Manski, Charles F. 1988. "Identification of Binary Response Models." *Journal of the American Statistical Association*, 83(403): 729–738.

Maskin, Eric, and John Riley. 2000. "Asymmetric Auctions." *The Review of Economic Studies*, 67(3): 413–438.

Matzkin, Rosa L. 1992. "Nonparametric and Distribution-Free Estimation of the Binary Threshold Crossing and The Binary Choice Models." *Econometrica*, 60(2): 239–270.

Milgrom, Paul R. 1981. "Good News and Bad News: Representation Theorems and Applications." *The Bell Journal of Economics*, 12(2): 380–391.

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

Moss, Aaron J, Cheskie Rosenzweig, Jonathan Robinson, and Leib Litman. 2020. "Demographic Stability on Mechanical Turk Despite COVID-19." *Trends in Cognitive Sciences*, 24(9): 678–680.

Moss, Aaron J, Cheskie Rosenzweig, Jonathan Robinson, Shalom N Jaffe, and Leib Litman. 2023. "Is It Ethical to Use Mechanical Turk for Behavioral Research? Relevant Data From a Representative Survey of Mturk Participants and Wages." *Behavior Research Methods*, 55(8): 4048–4067.

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

Oswald, Andrew J. 2008. "On the Curvature of the Reporting Function From Objective Reality to Subjective Feelings." *Economics Letters*, 100(3): 369–372.

Palmer, John, Alexander C. Huk, and Michael N. Shadlen. 2005. "The Effect of Stimulus Strength on the Speed and Accuracy of a Perceptual Decision." *Journal of Vision*, 5: 376–404.

Quah, John K.-H., and Bruno Strulovici. 2012. "Aggregating the Single Crossing Property." *Econometrica*, 80(5): 2333–2348.

Robinson, Jonathan, Cheskie Rosenzweig, Aaron J Moss, and Leib Litman. 2019. "Tapped Out or Barely Tapped? Recommendations for How to Harness the Vast and Largely Unused Potential of the Mechanical Turk Participant Pool." *PLoS One*, 14(12): e0226394.

Romano, Joseph P., and Michael Wolf. 2016. "Efficient Computation of Adjusted p-Values for Resampling-Based Stepdown Multiple Testing." *Statistics and Probability Letters*, 113: 38–40.

Rubinstein, Ariel. 2007. "Instinctive and Cognitive Reasoning: A Study of Response Times." *The Economic Journal*, 117: 1243–1259.

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

Rubinstein, Ariel. 2016. "A Typology of Players: Between Instinctive and Contemplative." *The Quarterly Journal of Economics*, 131(2): 859–890.

Samuelson, Paul A. 1938. "A Note on the Pure Theory of Consumer's Behaviour." *Economica*, 5(17): 61–71.

Shaked, Moshe, and J George Shanthikumar. 2007. *Stochastic Orders.* Springer.

Spearman, Charles. 1904. "The Proof and Measurement of Association Between Two Things." *The American Journal of Psychology*, 15(1): 72–101.

Stevenson, Betsey, and Justin Wolfers. 2008. "Happiness Inequality in the United States." *The Journal of Legal Studies*, 37(S2): S33–S79.

Survey Research Center, Institute for Social Research, University of Michigan. 2021. "Panel Study of Income Dynamics, public use dataset." Produced and distributed by the Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI.

Thistle, Paul D. 1989. "Ranking Distributions with Generalized Lorenz Curves." *Southern Economic Journal*, 5(1): 1–12.

Vaeth, Martin. 2023. "Voter Learning, Unidimensional Ideology, and Polarization." Mimeo.

Wang, Tao, and Ehud Lehrer. 2024. "Weighted Utility and Optimism/Pessimism: A Decision-Theoretic Foundation of Various Stochastic Dominance Orders." *American Economic Journal: Microeconomics*, 16(1): 210–223.

Appendices

A Rationalizability

In this appendix, we study the question of *rationalizability* of the data: when $\mathscr{G}^*$ and $\mathscr{C}^*$ are both restricted sets of distributions and chronometric functions, respectively, does there exist $((G_j)_j, (c_j)_j) \in \mathscr{G}^* \times \mathscr{C}^*$ that induces the given data $(p_j, F_j)_j$?

We give an answer to this question for the case in which both $\mathscr{G}^*$ and $\mathscr{C}^*$ are generated using sets of transformations with properties like before. Formally, let $\mathscr{C}^*$ be the *maximal set* generated by $((c_j^*)_j, \Psi)$, that is,

$$\mathscr{C}^* = \{(c_j)_j \in \mathscr{C} \mid (c_j)_j = (c_j^* \circ \psi_j)_j \text{ for some } (\psi_j)_j \in \Psi\}.$$

In words, not only can every $(c_j)_j \in \mathscr{C}^*$ be derived from $(c_j^*)_j$ using a composition with some $(\psi_j)_j \in \Psi$, but the composition of $(c_j^*)_j$ with any $(\psi_j)_j \in \Psi$ yields an element of $\mathscr{C}^*$. This requires that $\psi_j(0) = 0$ holds for all transformations that are part of Ψ. Analogously, let

$$\mathscr{G}^* = \{(G_j)_j \in \mathscr{G} \mid (G_j)_j = (G_j^* \circ \phi_j)_j \text{ for some } (\phi_j)_j \in \Phi\}$$

be the maximal set generated by $((G_j^*)_j, \Phi)$. For example, if $(G_j^*)_j$ are strictly increasing cdfs and Φ is the set of all profiles of strictly increasing transformations, then $\mathscr{G}^*$ becomes the set of all profiles of strictly increasing cdfs. If, in addition, $(G_j^*)_j$ are symmetric around their means, then combined with a suitably defined set of symmetric transformations we obtain a set of profiles of distributions that are also symmetric.

For the following result, denote by $\Phi \circ \Psi^{-1} = \{(\phi_j \circ \psi_j^{-1})_j \mid (\phi_j)_j \in \Phi \text{ and } (\psi)_j \in \Psi\}$ the set of all profiles of functions which are compositions of the elements of Φ and the inverse elements of Ψ. Any such function is bijective and strictly increasing, hence continuous.

Theorem 2. *Let $\mathscr{C}^*$ and $\mathscr{G}^*$ be the maximal sets generated by $((c_j^*)_j, \Psi)$ and $((G_j^*)_j, \Phi)$, respectively. Then, data $(p_j, F_j)_j$ is rationalizable if and only if there exists $(L_j)_j \in \Phi \circ \Psi^{-1}$ such that $(G_j^* \circ L_j)_j = (H_j)_j$.*

Proof. Only-if-statement. Suppose there exists $((G_j)_j, (c_j)_j) \in \mathscr{G}^* \times \mathscr{C}^*$ that induces $(p_j, F_j)_j$. As shown in the proof of Theorem 1, there exists $(\psi_j)_j \in \Psi$ such that $(H_j)_j = (G_j \circ \psi_j^{-1})_j$, because $((c_j^*)_j, \Psi)$ generates $\mathscr{C}^*$. It also holds that there exists $(\phi_j)_j \in \Phi$ such that $(G_j)_j = (G_j^* \circ \phi_j)_j$, because $((G_j^*)_j, \Phi)$ generates $\mathscr{G}^*$. Now define $L_j = (\phi_j \circ \psi_j^{-1})$ for all $j \in J$, so that $(L_j)_j \in \Phi \circ \Psi^{-1}$ holds. Furthermore,

$$(G_j^* \circ L_j)_j = (G_j^* \circ \phi_j \circ \psi_j^{-1})_j = (G_j \circ \psi_j^{-1})_j = (H_j)_j.$$

If-statement. Suppose there exists $(L_j)_j \in \Phi \circ \Psi^{-1}$ that satisfies $(G_j^* \circ L_j)_j = (H_j)_j$. Let $(\phi_j)_j \in \Phi$ and $(\psi_j)_j \in \Psi$ be such that $(L_j)_j = (\phi_j \circ \psi_j^{-1})_j$. Then $(G_j)_j := (G_j^* \circ \phi_j)_j \in \mathscr{G}^*$ because $\mathscr{G}^*$ is the maximal set generated by $((G_j^*)_j, \Phi)$, and $(c_j)_j := (c_j^* \circ \psi_j)_j \in \mathscr{C}^*$ because $\mathscr{C}^*$ is the maximal set generated by $((c_j^*)_j, \Psi)$. The data induced by $((G_j)_j, (c_j)_j)$ are, for all $t \in [\underline{t}, \bar{t}]$ and $j \in J$,

$$\begin{aligned}
\hat{p}_j^0 \hat{F}_j^0(t) &= G_j((c_j^0)^{-1}(t)) \\
&= G_j^*(\phi_j((c_j^0)^{-1}(t))) \\
&= G_j^*(\phi_j(\psi_j^{-1}((c_j^{*,0})^{-1}(t)))) \\
&= G_j^*(L_j((c_j^{*,0})^{-1}(t))) \\
&= H_j((c_j^{*,0})^{-1}(t)) \\
&= p_j^0 F_j^0(t),
\end{aligned}$$

where the third equality has been established in the proof of Theorem 1. The analogous argument shows that $((G_j)_j, (c_j)_j)$ induces also $p_j^1 F_j^1(t)$, hence the data is rationalizable. □

For rationalizability, we need to check whether there exist functions $(L_j)_j \in \Phi \circ \Psi^{-1}$ which satisfy the implicit condition

$$G_j^*(L_j(x)) = H_j(x) \text{ for all } x \in \mathbb{R}.$$

This is easier than it may appear. For example, if both H_j and G_j^* are strictly increasing, then the candidate L_j is unique and given by $L_j = (G_j^*)^{-1} \circ H_j$, which is a bijective and strictly increasing function based on observables. It remains to be checked whether this function can be written as an admissible composition $\phi_j \circ \psi_j^{-1}$. This is always the case, for example, if Φ is the unrestricted set of all profiles of transformations, because $\Phi \circ \Psi^{-1}$ is unrestricted in that case as well. In other cases, the function L_j will be unique in some intervals and can be extended outside these intervals in a way that guarantees bijectivity and monotonicity. One then needs to check whether L_j coincides with an admissible composition $\phi_j \circ \psi_j^{-1}$ wherever it is uniquely defined.

Other cases are even easier. For example, if H_j takes the value 0 for finite values $x < 0$ but G_j^* does not, then a function L_j satisfying $G_j^* \circ L_j = H_j$ cannot be strictly increasing and hence cannot be a composition $\phi_j \circ \psi_j^{-1}$. This implies that the data is not rationalizable. Intuitively, since G_j^* extends to infinity and the data has no atoms at response time $\underline{t}$, it can only be rationalized by a chronometric function that satisfies $c_j(x) > \underline{t}$ for all $x < 0$, and hence H_j cannot reach 0.

B Robustness to Heterogeneity and Noise

Suppose response time is given by $t = c_j(x) \cdot \eta \cdot \epsilon$ in the decision problem of interest and by $t_b = \phi \cdot \eta$ in the baseline problem, where $\eta > 0$ captures the individual's general response speed, $\epsilon > 0$ comprises different sources of noise, and $\phi > 0$ parameterizes the baseline problem. The terms x and η may be correlated and follow a j-specific distribution. The cdf of the marginal distribution of x is denoted by $G_j(x)$ and is assumed to be continuous. The terms ϵ and ϕ may also be correlated and follow a j-specific distribution described by a joint probability measure μ_j, but they are assumed to be independent of the other variables.

In a slight deviation from the main model, we let $\underline{t} = 0$ and $\bar{t} = \infty$, as heterogeneity and noise may spread out response times to arbitrary values. The chronometric function $c_j : \mathbb{R} \to [0, \infty)$ is described by $c_j^0 : (-\infty, 0) \to [0, \infty)$ and $c_j^1 : (0, +\infty) \to [0, \infty)$, assumed to be continuous, strictly increasing/decreasing throughout, and to approach the respective limits asymptotically. Continuity of $G_j(x)$ implies that $x = 0$ is a zero probability event and allows us to leave $c_j(0)$ unspecified.

We consider normalized response times $\hat{t} = t/t_b = c_j(x) \cdot \epsilon/\phi$ and denote the cdfs of their (conditional on choice) distributions by $\hat{F}_j^i$. The model $(G_j, c_j, \mu_j)_j$ then generates the data

$$p_j^0 \hat{F}_j^0(t) = \int_{\text{supp } \mu_j} G_j((c_j^0)^{-1}(t\phi/\epsilon)) \, d\mu_j(\phi, \epsilon)$$

and

$$p_j^1 \hat{F}_j^1(t) = 1 - \int_{\text{supp } \mu_j} G_j((c_j^1)^{-1}(t\phi/\epsilon)) \, d\mu_j(\phi, \epsilon),$$

for all $t \geq 0$ and $j \in J$.

Suppose the analyst relies on the representative chronometric function $c_j^*(x) = 1/|x|$ and therefore constructs the empirical distributions

$$\hat{H}_j(x) = \begin{cases} 1 - p_j^1 \hat{F}_j^1(1/x) & \text{if } x > 0, \\ p_j^0 & \text{if } x = 0, \\ p_j^0 \hat{F}_j^0(-1/x) & \text{if } x < 0, \end{cases} \tag{20}$$

from those data, for all $j \in J$. The following result shows how these distributions, which are misspecified as they ignore heterogeneity and noise, are related to the model fundamentals.

Proposition 1. *Given $(G_j, c_j, \mu_j)_j$, the distributions $(\hat{H}_j)_j$ defined in (20) can be written as*

$$\hat{H}_j(x) = \int_{supp\ \mu_j} K_j(x\epsilon/\phi)\, d\mu_j(\phi, \epsilon), \tag{21}$$

where $(K_j)_j = (G_j \circ \psi_j^*)_j$ with

$$\psi_j^*(x) = \begin{cases} (c_j^1)^{-1}(1/x) & \text{if } x > 0, \\ 0 & \text{if } x = 0, \\ (c_j^0)^{-1}(-1/x) & \text{if } x < 0, \end{cases}$$

for all $j \in J$.

Proof. Consider any $j \in J$. For $x = 0$ we have

$$\int_{\text{supp } \mu_j} K_j(x\epsilon/\phi)\, d\mu_j(\phi, \epsilon) = K_j(0) = G_j(0) = p_j^0 = \hat{H}_j(0).$$

For any $x < 0$ we have

$$\begin{aligned} \int_{\text{supp } \mu_j} K_j(x\epsilon/\phi)\, d\mu_j(\phi, \epsilon) &= \int_{\text{supp } \mu_j} G_j((c_j^0)^{-1}(-(1/x)(\phi/\epsilon)))\, d\mu_j(\phi, \epsilon) \\ &= p_j^0 \hat{F}_j^0(-1/x) = \hat{H}_j(x). \end{aligned}$$

For any $x > 0$ we analogously have

$$\begin{aligned} \int_{\text{supp } \mu_j} K_j(x\epsilon/\phi)\, d\mu_j(\phi, \epsilon) &= \int_{\text{supp } \mu_j} G_j((c_j^1)^{-1}((1/x)(\phi/\epsilon)))\, d\mu_j(\phi, \epsilon) \\ &= 1 - p_j^1 \hat{F}_j^1(1/x) = \hat{H}_j(x), \end{aligned}$$

which establishes (21). □

Proposition 1 extends the main insight from the proof of Theorem 1 to the case with noise. Indeed, if there is no noise (i.e., ϵ/ϕ is fixed at 1), then it implies that $\hat{H}_j = K_j = G_j \circ \psi_j^*$, where ψ_j^* corresponds to ψ_j^{-1} in the proof of Theorem 1.

In the general case, the result can be applied to check whether $(\hat{H}_j)_j$ is suitable for testing the hypothesis that $(G_j)_j$ satisfies a property **P**. The first requirement is that **P** of $(G_j)_j$ is invariant to ψ_j^*, so that $(K_j)_j$ also satisfies **P**. Without knowledge of the chronometric function beyond monotonicity, this requires invariance to all increasing transformations. With the knowledge of symmetry, it requires invariance to symmetric transformations, and so on. The second (and crucial) requirement is invariance of $\mathbf{P}$ of $(K_j)_j$ to the multiplicative convolutions (21), so that $(\hat{H}_j)_j$ inherits $\mathbf{P}$.

To illustrate the argument, consider the property that a single distribution G is strictly increasing, as discussed in Subsection 3.1.1. This property is invariant to all strictly increasing transformations, and furthermore it is invariant to (21) for arbitrary noise measures μ, based on the observation that increasing x increases $K_j(x\epsilon/\phi)$ pointwise for each (ϕ, ϵ) in the integral. Hence strict monotonicity of G carries over to strict monotonicity of $\hat{H}$ without any additional restrictions on the model. Consequently, if $\hat{H}$ is not strictly increasing, the hypothesis that G is strictly increasing can be rejected.

As a second illustration, consider the property of (G_1, G_2) that $G_1(-x) + G_1(x) \leq 1 \leq G_2(-x) + G_2(x)$ for all $x \geq 0$, which implies that the mean of G_1 is larger than the mean of G_2, as discussed in Subsection 3.2.2. This property is invariant to symmetric transformations that can differ for $j = 1, 2$, hence requiring the assumption that the chronometric effect is identical for the two options but not necessarily for the two groups. The property is then also invariant to (21) for arbitrary group-specific noise measures μ_j, again based on a pointwise argument in the integrals. Consequently, $(\hat{H}_1, \hat{H}_2)$ can be used to test the desired property under the assumption of symmetry even for group-specific chronometric functions and noise distributions. By contrast, the weaker property $G_1(-x) + G_1(x) \leq G_2(-x) + G_2(x)$ for all $x \geq 0$, which also implies that the mean of G_1 is larger than the mean of G_2 (also discussed in Subsection 3.2.2), additionally requires the assumption of identical chronometric functions and identical noise distributions in the two groups.

An example of a property that is generally not invariant to (21) is single-crossing of (K_1, K_2), as discussed in Subsection 3.2.4. This is true even with identical noise distributions $\mu_1 = \mu_2 = \mu$, as single-crossing does not aggregate well (see Quah and Strulovici, 2012). An interesting question in this case is whether invariance can be established under plausible constraints on μ. We leave this to future research.

C Ranking of Means

In this appendix, we show that the inequality $G_1(x) + G_1(-x) \leq G_2(x) + G_2(-x)$ for all $x \in \mathbb{R}_+$ implies that the mean of G_1, denoted μ_1, is larger than the mean of G_2, denoted μ_2 (assuming that both means exist). Using the fact that

$$\mu_j = -\int_{-\infty}^{0} G_j(x)dx + \int_{0}^{+\infty} [1 - G_j(x)]dx,$$

we have

$$\begin{aligned}\mu_1 - \mu_2 &= \int_{-\infty}^{0} [G_2(x) - G_1(x)]dx + \int_{0}^{+\infty} [1 - G_1(x) - 1 + G_2(x)]dx \\ &= \int_{0}^{+\infty} [G_2(-x) - G_1(-x)]dx + \int_{0}^{+\infty} [G_2(x) - G_1(x)]dx \\ &= \int_{0}^{+\infty} [G_2(x) + G_2(-x) - G_1(x) - G_1(-x)]dx \geq 0.\end{aligned}$$

D Decreasing Marginal Happiness

The survey of Liu and Netzer (2023) has income data in three broad bins: below \$40'000, between \$40'000 and \$69'999, and above \$70'000. Here, we describe the eight different methods that we use to put a unique value on each of these bins. The first three methods use the midpoints of the bins, with different upper bounds on the open-ended top bin. Method 4 predicts each subject's income based on observable covariates and uses these predictions for averaging within each bin. The last four methods use information about the income distribution of MTurk subjects from different sources to derive the average income in each bin. The weights α and $1 - \alpha$ are then always chosen such that the weighted average of the low and high bin is equal to the middle bin. Table 2 summarizes the implied incomes and weights associated with each method.

Method	Implied Incomes (w_L, w_M, w_H)	Weights ($\alpha, 1 - \alpha$)
1	(20′000, 55′000, 90′000)	(0.500, 0.500)
2	(20′000, 55′000, 110′000)	(0.610, 0.390)
3	(20′000, 55′000, 135′000)	(0.660, 0.340)
4	(39′334, 68′013, 115′459)	(0.623, 0.377)
5	(23′172, 53′048, 111′167)	(0.660, 0.340)
6	(25′094, 54′381, 108′469)	(0.649, 0.351)
7	(22′627, 53′979, 111′161)	(0.646, 0.354)
8	(23′318, 54′125, 112′854)	(0.656, 0.344)

Table 2: Summary of methods for constructing incomes (rounded to thousands).

Methods 1-3. We use the midpoints of the bins to assign a unique value to each bin. The difference between the three methods is the upper bound on the open-ended top bin. Method 1 uses an upper bound of \$110'000, method 2 uses an upper bound of \$150'000, and method 3 uses an upper bound of \$200'000. The resulting incomes associated with each bin are then

$20'000 and $55'000 for the low and middle bins, respectively, and $90'000, $110'000, and $135'000 for the high bin in methods 1, 2, and 3, respectively.

Method 4. We augment the survey data from Liu and Netzer (2023) with 2021 PSID data (Survey Research Center, 2021), where we use the following variables:

- ER81775 "TOTAL FAMILY INCOME-2020"
- ER78018 "SEX OF REFERENCE PERSON"
- ER78017 "AGE OF REFERENCE PERSON"
- ER78021 "# CHILDREN IN FU"
- ER78025 "REFERENCE PERSON MARITAL STATUS"
- ER81926 "COMPLETED ED-RP"

Based on the these variables, we estimate a simple linear wage model, regressing income on gender, age, kids, marital status, and education, where we recode the variables so coding and definition match the corresponding variables obtained in the survey. The estimation results then allow us to predict the income of each subject in the survey. We correct these estimates whenever they fall outside the bin that the subject reported to be in, by setting the value equal to the lower or upper bound of the bin, respectively. We use these predictions to determine the average income for each bin.

Methods 5-8. We use more granular income data about MTurk subjects drawn from the studies conducted in Robinson et al. (2019), Moss et al. (2023), and Moss et al. (2020) to determine the average income in each bin of the survey of Liu and Netzer (2023). These studies contain finer bins of income data. For each of these bins, we determine the average using the midpoint approach as in method 1, using an upper bound on the open-ended top bin of $200'000 for the first three studies and $300'000 for the last study.