BSE Barcelona School of Economics
WORKING PAPERS

Attitudes Towards Success and Failure

BSE Working Paper 1336

March 2022 (Revised March 2025)

Larbi Alaoui, Antonio Penta

Attitudes towards success and failure*

Larbi Alaoui[†]
UPF and BSE

Antonio Penta[‡]
ICREA, UPF and BSE

March 21, 2025

Abstract

Individuals often attach a special meaning to attaining a certain goal, that marks the difference between what they consider a success or a failure. Within a standard von Neumann-Morgenstern Expected Utility setting, with an exogenous reference point that separates success from failure, we define attitudes towards success and failure as features of preferences over lotteries. The distinctive feature of our definitions is that they all concern a local reversal of risk attitude, between risk-aversion and risk-lovingness, for lotteries that go across the reference point. Our findings provide a unified view of several models of reference-dependent preferences, as well as novel representations. We order the intensity of each attitude,and characterize these orderings in terms of properties of the utility representations, with indices analogous to the Arrow-Pratt index of risk aversion.

Keywords: expected utility; loss aversion; aspirations; risk aversion; reference-dependence.

*We are grateful to José Apesteguia, Aurelien Baillon, Pedro Barros, Han Bleichrodt, Antonio Cabrales, Tommaso Denti, Christian Fons-Rosen, Yoram Halevy, Shachar Kariv, Pietro Ortoleva, Luciano Pomatto, Kareen Rozen, Charlie Sprenger, Jakub Steiner and Peter Wakker for their helpful comments, and Malachy J. Gavan, Charlotte Hervy, Nathan J. Jones and Pia Ennuschat for excellent Research Assistanship. We also thank seminar audiences at the Collegio Carlo Alberto, TSE, Univ. of Toronto, UCL, Columbia, Princeton, Alicante, Erasmus Univ. of Rotterdam, Univ. of Nottingham, UPF, CUHK-HKU-HKUST, Univ. Carlos III, NOVA, UM6P, EUI, Univ. of Liverpool, Heidelberg, Warwick, NYUAD, the BSE Summer Forum, BRIC and RUD. The BSE benefited from the financial support from the Severo Ochoa Program for Centers of Excellence in R&D, grant CEX2019-000915-S. Alaoui gratefully acknowledges financial support from the Spanish Ministry of Science and Innovation under projects ECO2014-56154-P, PGC2018-098949- B-I00, grant RYC-2016-21127, and the BBVA Foundation (Leonardo Grant). The BBVA Foundation takes no responsibility for the opinions, statements and contents of this project, which are entirely the responsibility of its authors. Penta acknowledges the support of the ERC Consolidator grant #101089139.

[†]Universitat Pompeu Fabra and BSE. E-mail: larbi.alaoui@upf.edu

[‡]ICREA, Universitat Pompeu Fabra and BSE. E-mail: antonio.penta@upf.edu

1 Introduction

Individuals often attach a special meaning to attaining a certain goal, and whether or not they get past a threshold marks the difference between what they perceive as being successful and what they perceive as having failed. This general metaphor of success and failure emerges in different forms, both in economics and in psychology, in ways that are not fully understood. The study of personality, for instance, considers attitudes towards achieving a goal (in tenacity, perseverance and conscientiousness; cf. Deary et al. (2009)), as does the vast literature on reference dependence, with the notion of gains and losses (Kahneman and Tversky (1979, 1992)).[1] In development economics too, the idea of aspirations (Genicot and Ray (2017)) relates to achieving a goal that one aspires to. In business and finance, benchmarking and aspiration level models (Payne et al. (1980, 1981)) refer to the objective of reaching a specific target, and they are closely related to the idea of "gambling for resurrection" that is familiar in political science.

The formal models above are apparently disconnected, but they all share the feature that their utility representations induce reversals of the decision maker's risk-attitude, around the critical threshold: for instance, an otherwise risk-averse agent may be risk-loving over lotteries that can mark the difference between a failure and a success. Concerning the personality traits mentioned above, which are increasingly used in empirical economics (e.g., Heckman and Rubinstein (2001), Almlund et al. (2011), Gill and Prowse (2016), Proto and Rustichini (2015), Proto et al. (2019, 2022), Heckman et al. (2021), etc.), but for which there is currently no agreed upon formal representation, we perhaps have an intuitive understanding of their meaning, but we lack a definition of these traits in terms of preferences and choice, the way we have for instance with risk and time attitudes. More broadly, we do not know whether the shared metaphor of success and failure, and the corresponding attitudes which are studied in these different areas of research, are in some way linked to each other, and what are their fundamental economic underpinnings.

The aim of this paper is to analyze various attitudes towards success and failure underlying the concepts alluded to in several branches of research. We do this while remaining within the standard expected utility (EU) framework, which we supplement with an *exogenous* reference point, x_0, that serves as the threshold between success and failure: any outcome above x_0 is a success, any outcome below is a failure. Within this setting, we define the various attitudes in terms of primitive preferences over monetary lotteries, just as is done for risk-aversion, and we obtain representations of each attitude in terms of properties of the corresponding Bernoulli utility function. These representations highlight the connections between the various attitudes, and show that the seemingly distinct concepts discussed in different research fields intercon-

[1] Reference-dependence has a long history in economics (see, e.g., Markowitz (1952) and Kahneman and Tversky (1979)). It has been explored from several angles, typically departing from the von Neumann-Morgenstern axioms (e.g., Masatlioglu and Ok (2005, 2014), Masatlioglu and Raymond (2016), Wakker (2010), etc.). Models with endogenous reference points include Kőszegi and Rabin (2006, 2007), Kıbrıs et al. (2023), and have been used various contexts (e.g., Rosato and Tymula (2019) and Rosato (2023)). Recent work on the determination of reference points includes Cerulli-Harms et al. (2019) and Baillon et al. (2020), and for neuroscience-based models, see Steiner and Stewart (2016) and Netzer et al. (2022). Here, in contrast, we maintain the vNM axioms, and take the reference point to be exogenous. (cf. O'Donoghue and Sprenger (2018) and references therein.

nect in ways that may not have seemed obvious, *prima facie*. We also define orderings over the intensity of these attitudes, and characterize them in terms of transformations over their utility representations, thereby providing a means to perform comparative statics exercises, in a manner analogous to risk-aversion indices such as the Arrow-Pratt coefficient. These orderings and their corresponding indices thus provide tractable models of decision-making that can be used to capture economically relevant personality traits using standard economics notions and techniques.

Motivated by the common feature of the disparate models mentioned above, we aim to understand how different attitudes towards success and failure affect a decision maker's willingness to take risk, and induce reversals between risk-lovingness and risk-aversion around the threshold between perceived failure and success. To this end, we maintain all the von Neumann-Morgenstern axioms for an EU representation, as well as monotonicity (i.e., more money is preferred to less). Remaining within EU allows us to focus purely on attitudes towards success and failure and to pinpoint their determinants, while abstracting away from other factors, such as reweighting of the probabilities. As we will show, doing so leads to key insight into these attitudes, that would have been easy to miss otherwise.

To isolate the role played by the critical threshold in inducing reversals of the individual's risk-attitude, we further assume that, at least over some (arbitrarily small) left- and right-neighborhood of the threshold, the agent can be (weakly) risk-averse or risk-loving, but does not switch from risk-aversion to risk-lovingness for lotteries that are 'on the same side' of x_0. In contrast, all the attitudes we consider do entail a switch of risk attitude for lotteries that go *across* the threshold, in the sense that they attach positive probability to an outcome $x' > x_0$ and to an outcome $x < x_0$. The various attitudes differ in the way that these reversals manifest themselves. The idea in our approach is to take such reversals, which are present in several models of reference-dependence, as the defining feature of these behavioral phenomena. As our results show, this novel perspective provides a unified view on seemingly unrelated models of reference-dependent preferences, and sheds a new light on familiar notions and patterns of behavior. This includes the shape of the utility function used in *prospect theory* (leading to loss aversion and diminishing sensitivity), the *aspiration* representation used in Development Economics (e.g., Genicot and Ray (2017)), and the discontinuous model from the Finance literature (cf., Payne et al. (1980, 1981), Diecidue and Van De Ven (2008)). It further allows for a choice-based identification of the reference point, precisely for its inducement of the reversals in risk attitude.

The first two attitudes we consider are what we call *failure avoidance* and *success attachment*. Both attitudes posit a reversal of the risk-attitude for binary lotteries that go across x_0: over such lotteries, individuals would be risk-averse for some and risk-loving for others. The difference between the two is given by the source of the reversal, which could be primarily driven by the potential failures, or by the potential successes. *Failure avoidance* concerns the agent's desire to avoid the failure region, no matter by how small a margin. The idea is that, for any potential failure x in some left-neighborhood of x_0, the agent is willing to take a risk in order to attain a potential success $x' > x_0$, *no matter how small*, as long as the probability of failure is high enough. *Success attachment* instead captures the agent's desire to end up in the success region, no matter how small the potential failure might be. Symmetrically to failure avoidance, the idea is that for any potential success x' in some right-neighborhood of x_0, the agent is willing to take a risk in order to avoid a potential failure $x < x_0$, *no matter how close* x is to the critical threshold, as long as the probability of failure is above a certain threshold.

Our first results characterize the shape of the utility function for both attitudes, and jointly reveal a striking finding: failure avoidance and success attachment cannot co-exist, except when there is a discontinuity at x_0. They also reveal that a special case of failure avoidance, without success attachment, characterizes the hugely influential representation used in prospect theory, namely that of the kinked S-shaped utility function (Kahneman and Tversky (1979)). An important offshoot of our analysis therefore is to provide a characterization of this functional form in terms of the primitives that are standard in the theory of decisions under risk, i.e,. the agent's preferences over lotteries.[2] Hence, independent of one's position in the debate on the specification of the outcome space (e.g., Rabin (2000), Rubinstein (2002); see also O'Donoghue and Sprenger (2018)), our results formally describe how key behavioral phenomena that are commonly associated with prospect theory – namely, *loss aversion* and *diminishing sensitivity* – may be captured by risk preferences within a completely standard expected utility setting. As we discuss below, this brings in new insights about both the prospect theory utility representation and on loss aversion, especially with respect to how to perform interpersonal comparisons over these notions. Our approach therefore provides a new perspective that complements the active literature on this topic.[3]

We then introduce attitudes that mirror failure avoidance and success attachment, which we call *failure resignation* and *success seeking.* The difference between these attitudes and the previous ones is that, rather than having reversals in which risk aversion is 'at the top' (i.e., for high probability of success), the switch occurs in the opposite direction, with risk-aversion 'at the bottom'. In the case of failure resignation, for instance, the agent is willing to take a risk to pursue an arbitrarily small success, only when the probability of success is high enough. With success seeking, instead, the individual is unmotivated to take a risk to avoid an arbitrarily small failure, unless the likelihood of success is high enough. This last attitude is perhaps not as prevalent as the others, but we characterize it for completeness.

Equipped with the representation of the four attitudes, we then study how they interconnect with one another. In doing so, we obtain a complete map of how they determine the shape of

[2] A large empirical literature has explored loss aversion under risk, eliciting subjects' preferences over lotteries (see, e.g., Camerer et al. (1997), Abdellaoui et al. (2007), Choi et al. (2007), Crawford and Meng (2011), Imas (2016), Imas et al. (2017), Bernheim and Sprenger (2020), Ellis et al. (2022), etc.). Theoretical investigations have typically focused on settings with certainty or specified properties directly in the space of utility representations (cf. Wakker (2010) and references therein). Our results, in contrast, are in terms of preferences over lotteries.

[3] While it is a generally underappreciated fact that prospect-theory utility functions are valid Bernoullis within the EU-framework, we are not the first to make this observation. Neilson (2002), for instance, discusses the prospect theory shape within an EU setting, and studies comparative statics for various notions of loss aversion directly in the space of the utility representations. Our notion of loss aversion differs from those in Neilson (2002) for being defined in terms of the preferences over lotteries, and for yielding a representation consistent with the now commonly used definitions of this notion (cf. Köbberling and Wakker (2005)). Rozen (2010) explicitly allows for S-shaped utility functions in deterministic domains in an intertemporal context, within a time consistent and axiomatically derived representation. Shalev (1997) instead derives a special kind of loss aversion, distinct from the one discussed in this paper, in a time-inconsistent representation. For recent work on loss aversion, within a model of Cautious Utility, see Cerreia-Vioglio et al. (2022).

the utility function, when they are displayed both individually and jointly (when possible). This is best seen graphically (see Figure 3, p. 12), but we briefly discuss some of the findings. As previously mentioned, the kinked S-shaped utility function is a special case of failure avoidance that precludes success attachment, and also cannot co-exist with failure resignation or success seeking. Another noteworthy intersection is that of success attachment and failure resignation, which characterizes an *aspiration* representation that has been used in development economics (e.g., Genicot and Ray (2017, 2020)). This intersection instead cannot co-exist with failure avoidance or success seeking. Finally, unlike failure avoidance or success attachment, which can only co-exist with a utility function that is *discontinuous* at x_0, as in the so called "aspiration level" models in the finance literature (e.g., Payne et al. (1980, 1981), Diecidue and Van De Ven (2008)), failure resignation and success seeking are mutually exclusive attitudes.

Following the definition of attitudes towards success and failure, we introduce orderings over the degrees of each attitude. We show that what may seem to be natural orderings would in fact lack crucial features. For instance, consider the case of failure avoidance, and the special case of the kinked S-shaped utility function that is typically used in *prospect theory.* It may seem natural to rank an agent with a sharper kink to be more failure avoidant (as with the standard definition of loss aversion, cf. Köbberling and Wakker (2005)). But this would be incomplete: an agent with a sharper kink, all else being equal, is more loss averse, but he also exhibits *less* manifestations of failure avoidance. A sharper kink therefore does not suffice to rank individuals by this attitude. We rectify this issue by defining our rankings directly in terms of the (more transparent) primitive preferences, and obtain as a result indices in the utility space that characterize each ranking. These indices involve both a ranking of the sharpness of the kink at the threshold, and measures of concavity of the Bernoulli utility functions around it.

Put together, our approach serves to understand attitudes towards success and failure, to characterize them in terms of properties of the Bernouilli utility function, and to provide choice-based indices for their intensity. It further shows at a foundational level the ways in which seemingly disparate models that involve some form of reference-dependent preferences are linked. Our map of attitudes towards success and failure includes several of the most influential representations of reference-dependence, as well as novel ones. As we demonstrate, the ways in which these preferences are connected depends fundamentally on the direction and source of reversal of risk attitude around the reference point.

The rest of the paper is organized as follows: Section 2 introduces the general framework and the maintained axioms. Section 3 introduces the four attitudes towards success and failure, as well as the corresponding representation theorems. Section 4 discusses the joint implications of the main representation theorems, and discusses some special cases of interest, such as the kinked S-shape utility of prospect theory, aspirations, and the discontinuous case. Section 5 focuses on the interpersonal comparisons of the four attitudes (both their behavioral definitions and their utility characterizations). Section 6 concludes.

2 Model

We let $\mathbb{R}$ denote the space of monetary outcomes, and let L denote the set of simple lotteries over monetary outcomes, with typical elements $p, q, r \in L$. For any $x, x' \in \mathbb{R}$, we let $\Delta(x, x') \subseteq L$ denote the set of lotteries with support included in $\{x, x'\}$. With a slight abuse of notation, in that case we let p denote both the lottery itself, as well as the probability $p \in [0,1]$ attached to the high prize, $x' \geq x$ (x receives probability $(1-p)$). For any lotteries $p, q \in L$, we let Ep and Eq denote their respective expected values, and for any $x \in \mathbb{R}$, we let $\delta_x \in L$ denote the degenerate lottery which assigns probability one to x.

We assume that the decision maker (DM)'s preferences are represented by a weak order $\succsim$, with symmetric and asymmetric parts $\succ$ and $\sim$, respectively. For any $p \in L$, we let $CE(p) \in \mathbb{R}$ denote the certainty equivalent of p, if it exists, as the degenerate lottery $\delta_{CE(p)}$ which satisfies $\delta_{CE(p)} \sim p$. We maintain throughout all the von Neumann and Morgenstern (1944, vNM) axioms for an expected utility (EU) representation, as well as monotonicity:

- [**Weak Order:**] $\succsim$ is complete and transitive.
- [**Independence:**] For any $p, q, r \in L$ and $\alpha \in [0,1]$, $p \succsim q$ if and only if $\alpha p + (1-\alpha)r \succsim \alpha q + (1-\alpha)r$.
- [**Archimedean Property:**] For all $p, q, r \in L$ such that $p \succ q \succ r$, there exists an $a, b \in (0,1)$ such that $ap + (1-a)r \succ q \succ bp + (1-b)r$.
- [**Monotonicity:**] $\delta_{x'} \succ \delta_x$ if and only if $x' > x$.

We recall that the first three axioms (i.e., the standard vNM axioms) hold if and only if the preferences have an EU representation, i.e. there exists a Bernoulli utility function $u : X \to \mathbb{R}$ such that $p \succsim q$ if and only if $\sum_{x \in X} p(x)u(x) \geq \sum_{x \in X} q(x)u(x)$. We also remind the reader that these axioms alone do not impose any particular structure on u. In particular, they impose neither continuity of u nor concavity or convexity properties, but they do identify the utility function uniquely, up to positive affine transformations. If monotonicity is also assumed, then u is increasing.

As explained, we aim to understand how different attitudes towards success and failure affect a decision maker's willingness to take risk, and induce reversals between risk-lovingness and risk-aversion, for lotteries that can mark the difference between what he perceives as a failure or as a success. To this end, we let $x_0 \in \mathbb{R}$ denote the (exogenous) *threshold*, that separates successes ($x' > x_0$) from failures ($x < x_0$). All the attitudes we consider do entail a switch of risk attitude for lotteries that go *across* the threshold: that is, for lotteries that attach positive probability to outcomes $x' > x_0$ and $x < x_0$, the agent would be risk-averse for some probability of success, and risk-loving for others.[4]

In order to isolate the role of the threshold, we further assume that at least for some left- and right-neighborhoods of the threshold, the agent can be (weakly) risk-averse or risk-loving, but

[4]As we further discuss below, it will be easy to see from the definitions that the position of the threshold can be identified from choice. We nonetheless fix x_0 from the outset in order to simplify the notation.

does not switch between one and the other for lotteries that are supported over such intervals, and "on the same side" of x_0. Hence, all our notions will have in common the following features: (i) There is *no reversal* of the agent's risk-attitude over lotteries whose prizes are all *on the same side* of the reference point; (ii) There is a *reversal* of the agent's risk-attitude for lotteries *across* the reference point: over such lotteries, the agent would be risk-averse for some and risk-loving for others. The various attitudes will differ in the way that such reversals manifest themselves.

To formalize these ideas, for any interval (x_f, x_s) that contains x_0, and for any $x \in [x_f, x_s]$, let $S_{x_f}^{x_s}(x, x_0)$ denote the set of all $y \in [x_f, x_s]$ which are on the same side of x^0 as x.[5] Then, we can define the property of *Same-Side No Reversal* over an interval:

Definition 1 (Same-Side No Reversal)**.** *Let $x_f < x_0 < x_s$. Preferences $\succsim$ display* Same Side No-Reversal (SSNR) over the interval (x_f, x_s) *if, for any $x \in [x_f, x_s]$, $\nexists x' \in S_{x_f}^{x_s}(x, x_0)$ s.t. $\delta_{Eq} \succ q$ and $p \succ \delta_{Ep}$ for some $p, q \in \Delta(x, x')$.*

3 Attitudes Towards Success and Failure

In this section we introduce attitudes towards success and failures, and the corresponding representation theorems. All such attitudes posit a *reversal* of the risk-attitude for binary lotteries that go across x_0 – that is, lotteries that assign prize $x' > x_0$ with probability $p \in (0, 1)$, and $x < x_0$ otherwise: over such lotteries, individuals would be risk-averse for some and risk-loving for others. The difference between them is given by the source of the reversal, which could be primarily driven by the potential failures, or by the potential successes, and by the direction of such reversal. We start with *failure avoidance* and *success attachment*, before moving to their "duals", *success seeking* and *failure resignation*.

3.1 Failure Avoidance and Success Attachment: Model-Free Definitions

We introduce next the formal definition of *failure avoidance* and *success attachment*. As discussed, *failure avoidance* concerns the agent's desire to avoid the failure region, no matter by how small a margin. The idea is that, for any potential failure x in some left-neighborhood of x_0, the agent is willing to take a risk in order to attain a potential success $x' > x_0$, *no matter how small*, as long as the probability of failure is high enough. But once the probability of success is high enough, he reverts instead to being risk-averse. We formalize these ideas as follows:

Definition 2 (Failure Avoidance)**.** *Preferences $\succsim$ display* failure avoidance at $x_0 \in \mathbb{R}$ *if* $\exists x_f, x_s : x_f < x_0 < x_s$ *s.t.: (i) $\succsim$ display SSNR over (x_f, x_s); and (ii) $\forall x \in [x_f, x_0)$, $\exists \bar{x} \in (x_0, x_s] : \forall x' \in (x_0, \bar{x}]$, $\exists p, q \in \Delta(x, x')$ such that $p > q$, $\delta_{Ep} \succ p$ and $q \succ \delta_{Eq}$.*

The formal definition of *success attachment* is completely symmetrical to that of *failure avoidance*, with the roles of failures and successes swapped:

Definition 3 (Success Attachment)**.** *Preferences $\succsim$ display* success attachment at $x_0 \in \mathbb{R}$ *if* $\exists x_f, x_s : x_f < x_0 < x_s$ *s.t.: (i) $\succsim$ display SSNR over (x_f, x_s); and (ii) $\forall x' \in (x_0, x_s]$, $\exists \underline{x} \in [x_f, x_0) : \forall x \in [\underline{x}, x_0)$, $\exists p, q \in \Delta(x, x')$ s.t. $p > q$, $\delta_{Ep} \succ p$ and $q \succ \delta_{Eq}$.*

[5]Formally, $S_{x_f}^{x_s}(x, x_0) := \{y \in [x_f, x_s] : sign(x - x_0) = sign(y - x_0)\}$.

In both attitudes, the agent switches from risk-lovingness to risk-aversion, as the probability of success increases. But notice that the order of quantifiers over failures and successes is different. This captures that, with failure avoidance, the agent's objective is to pursue (via his willingness to take some risk) *any* success to get out of the failure region, while success attachment denotes the willingness to take some risk in order to avoid *any* failure. The difference between the two is thus given by the ultimate source of the reversal.

We note that these notions (including the SSNR requirement, as well as the attitudes we introduce in Section 3.3), are local notions, in the sense that they refer to properties of the agent's preferences for lotteries supported on some neighborhood of the threshold. In fact, the definitions refer to a particular attitude *at a threshold* x_0, with no implication that x_0 is the *only* point at which the agent displays a specific attitude. So, for instance, different thresholds may be relevant for the same agent, and trigger the same or different attitudes at different levels. For example, the same gambler may display failure avoidance for gambles that can mark the difference between 'winning something' and 'losing' (i.e., for $x_0 = \$0$), and display instead success attachment over gambles that may take him right above or right below some other salient threshold, e.g., for $\hat{x}_0 = \$1M$.[6] The definitions above also clarify that, while such thresholds are exogenous in our model, in the sense that they don't depend on the menu of choices that are presented to the agent, their position can be identified from choice: a particular outcome $\hat{x} \in \mathbb{R}$ is a 'threshold' if and only if the agent's preferences over lotteries around it satisfy the kind of reversals of risk-attitude that are entailed by the definitions (as well as the SSNR property on either side of it – cf. Definitions 1-5).

3.2 Failure Avoidance and Success Attachment: Representation Theorems

Before moving to the representation theorems, it is useful to first introduce some notation. Given the Bernoulli utility function that represents the agent's preferences (its existence is ensured by the vNM axioms), $u : X \to \mathbb{R}$, and the threshold $x_0 \in \mathbb{R}$, we let $u^-(x_0) := \lim_{x \to x_0^-} u(x)$, $u^+(x_0) := \lim_{x \to x_0^+} u(x)$, and for any $[x_f, x_s]$ and any $x \in [x_f, x_s] \setminus \{x_0\}$, we let

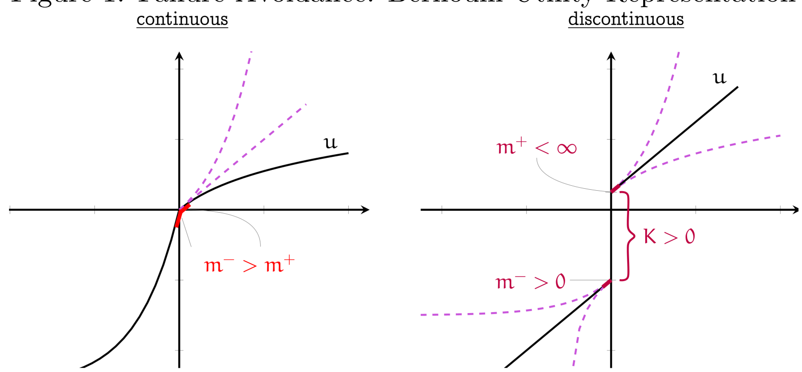
$$m(x) = \begin{cases} \frac{u(x)-u^+(x_0)}{x-x_0} & \text{if } x > x_0 \\ \frac{u(x)-u^-(x_0)}{x-x_0} & \text{if } x < x_0 \end{cases}$$

denote the average slope of the utility function in the interval (x_0, x) or (x, x_0), depending on whether $x > x_0$ or $x < x_0$. (By monotonicity, $m(x) \geq 0$ for any x). Also define $m^- := \lim_{x \to x_0^-} m(x)$ and $m^+ := \lim_{x \to x_0^+} m(x)$, and $K := u^+(x_0) - u^-(x_0)$.

Theorem 1 (Failure Avoidance: Representation). *Under vNM plus monotonicity,* $\succsim$ *displays* failure avoidance at x_0 *if and only if there exist* $x_s, x_f \in \mathbb{R} : x_f < x_0 < x_s$ *such that either: (i)* u *is continuous on* (x_f, x_s)*, strictly convex on* (x_f, x_0)*, either concave or convex on* (x_0, x_s)*, and such that* $m^- > m^+$*; or (ii)* u *is discontinuous at* x_0*, it is either convex or concave on each interval* (x_f, x_0) *and* (x_0, x_s)*, and such that* $m^+ < \infty$*.*

[6] Markowitz (1952) provides an early argument for the existence of multiple points of risk-attitude reversals.

Figure 1: Failure Avoidance: Bernoulli Utility Representation



Graphical illustration of the possible Bernoulli utility functions under failure avoidance.

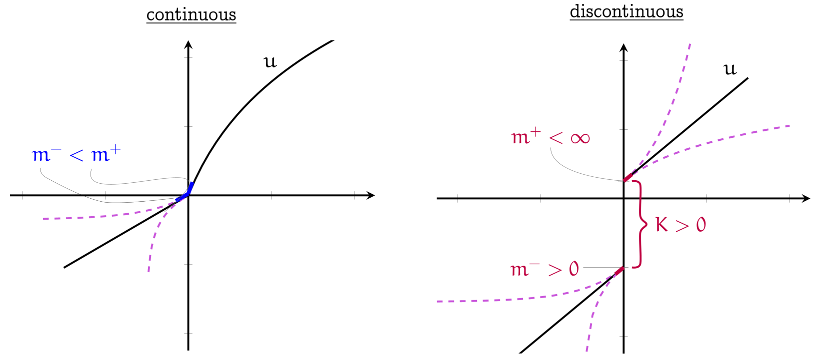
The logic is as follows (see Fig. 1). Consider first the continuous case. With failure avoidance, the agent is strictly risk-loving, over lotteries that go across the threshold, whenever the probability of the outcome in the failure *region* is sufficiently high. For this reason, his utility must be strictly convex on that region (locally, on an interval (x_l, x_0)). But since there must be a reversal from risk-lovingness to risk-aversion, this convexity must be counteracted by some form of concavity. This concavity does *not* come from concave utility on the success region, because the switch from risk-lovingness to risk-aversion occurs for binary lotteries that include any success, no matter how small (i.e., how close to x_0). Instead, it must come at the threshold itself, and it must come in form of a kink. Hence, $m^- > m^+$. The requirement on the success region is simply that due to SSNR, that there is concavity or convexity locally, without any imposition on which of the two it is.

This intuition captures two important features of the continuous case, whose logic will be adapted to all the attitudes we discuss: (i) since the attitude requires a reversal, it requires countervailing forces, one that provides risk-lovingness (convexity) and one that provides risk-aversion (concavity), and (ii) there is an asymmetry between the two regions. In the failure region, there is convexity over an *interval*, because the agent aims to avoid that entire region. An analogous curvature (concavity) is not required on the success region; instead, there is a *kink* at the threshold x_0 that captures the concavity.

Considering now the discontinuous case, the switch from risk-lovingness at the bottom to risk-aversion at the top comes from the discontinuity at x_0 directly. Intuitively, on the left of the discontinuity, the agent is willing to take a risk to 'jump up' to the success side. But if success is likely enough, so that it is likely enough to be on the success side, then the agent is risk-averse so as to avoid being on the left of the jump down. As for the requirement of convexity or concavity on either side, this is again due to SSNR. Note that, unlike the continuous case, there is no apparent asymmetry between the two regions in the logic above. This is because here, the discontinuity trumps the need for curvature on the left and the kink at the thresholds.

Theorem 2 (Success Attachment: Representation)**.** *Under vNM plus monotonicity,* $\succsim$ *displays* Success Attachment at x_0 *if and only if there exist* $x_s, x_f \in \mathbb{R} : x_f < x_0 < x_s$ *such that either: (i)* u *is discontinuous at* x_0*, it is either convex or concave on each interval* (x_f, x_0) *and* (x_0, x_s)*,*

Figure 2: Success Attachment: Bernoulli Utility Representation



Graphical illustration of the possible Bernoulli utility functions under success attachment.

and such that $m^- < \infty$; or (ii) u is continuous on (x_f, x_s), strictly concave on (x_0, x_s), either concave or convex on (x_0, x_f), and such that $m^- < m^+$.

A similar logic to that of the previous theorem holds (see Fig. 2). Here as well, there is an asymmetry for the continuous case between the two regions. But now, note that it is the success region over which the agent is risk-averse due to success attachment, and so here the curvature on an interval must be on the right. As it is risk-aversion, this corresponds to concavity of u. The counteraction to obtain a reversal can again not come from curvature on the failure region, as the agent is aiming not to be in any failure, no matter how small. Hence, it must again come from a kink at x_0. Since now the kink must counteract concavity, it must provide convexity. In other words, it must be that $m^- < m^+$. As for the discontinuous case, the logic here is identical, as is the result.

While the definitions of the two attitudes are related, note that there is a conflict in the continuous case of the representation theorems. The kink for failure avoidance must provide concavity ($m^- > m^+$) to counteract the convexity in the failure region, but for success attachment it must provide convexity ($m^- < m^+$) to counteract convexity in the success region. These two are incompatible, and so the two attitudes cannot coexist in the continuous case. They can only coexist in the discontinuous case, where the jump itself is responsible for the reversal, and provides the two counteracting forces.

Corollary 1. *Under vNM plus monotonicity, $\succsim$ displays* both *Success Attachment and Failure Avoidance at x_0 if and only if it is discontinuous at x_0.*

Notice also that a special case of the continuous representation of failure avoidance is one that takes the classical form, as in Fig. 1, of loss aversion: u is convex in the failure region, concave in the success region, and the slope on the left is steeper than on the right ($m^- > m^+$). It is only a special case because, while convexity and this kink shape are required for failure avoidance, concavity at the right is not. By the corollary above, it is immediate that loss aversion is incompatible with success attachment. We will return to this point more formally once all of the attitudes have been defined.

3.3 Failure Resignation and Success Seeking

We next introduce attitudes that mirror failure avoidance and success attachment, in the sense that rather than having reversals in which risk-aversion is 'at the top' (i.e., for high probability of success), the switch occurs in the opposite direction:

Definition 4 (Failure Resignation)**.** *Preferences* $\succsim$ *display* failure resignation at $x_0 \in \mathbb{R}$ *if* $\exists x_f, x_s : x_f < x_0 < x_s$ *s.t.: (i)* $\succsim$ *display SSNR over* (x_f, x_s)*; and (ii)* $\forall x \in [x_f, x_0)$, $\exists \bar{x} \in (x_0, x_s] : \forall x' \in (x_0, \bar{x}]$, $\exists p, q \in \Delta(x, x')$ *such that* $p < q$, $\delta_{Ep} \succ p$ *and* $q \succ \delta_{Eq}$.

Definition 5 (Success Seeking)**.** *Preferences* $\succsim$ *display* success seeking at $x_0 \in \mathbb{R}$ *if* $\exists x_f, x_s : x_f < x_0 < x_s$ *s.t.: (i)* $\succsim$ *display SSNR over* (x_f, x_s)*; and (ii)* $\forall x' \in (x_0, x_s]$, $\exists \underline{x} \in [x_f, x_0) : \forall x \in [\underline{x}, x_0)$, $\exists p, q \in \Delta(x, x')$ *such that* $p < q$, $\delta_{Ep} \succ p$ *and* $q \succ \delta_{Eq}$.

In words, with failure resignation the agent is willing to take a risk to pursue an arbitrarily small success, as long as the probability of success is high enough. With success seeking, instead, the individual is willing to take a risk to avoid an arbitrarily small failure, if success is sufficiently likely. The next results are analogous to the previous representation theorems:

Theorem 3 (Failure Resignation: Representation)**.** *Under vNM plus monotonicity,* $\succsim$ *displays* Failure Resignation *at* x_0 *if and only if there exist* $x_s, x_f \in \mathbb{R} : x_f < x_0 < x_s$ *such that: u is continuous on* (x_f, x_s)*, strictly concave on* (x_f, x_0)*, either concave or convex on* (x_0, x_s)*, and such that* $m^+(x_0) > m^-(x_0)$.

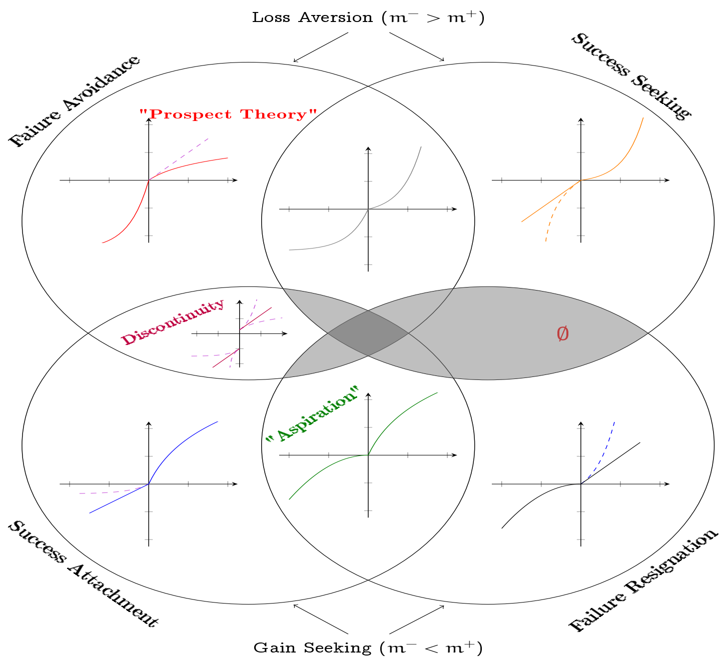
Theorem 4 (Success Seeking: Representation)**.** *Under vNM plus monotonicity,* $\succsim$ *displays* Failure Resignation *at* x_0 *if and only if there exist* $x_s, x_f \in \mathbb{R} : x_f < x_0 < x_s$ *such that: u is continuous on* (x_f, x_s)*, strictly convex on* (x_0, x_s)*, either concave or convex on* (x_0, x_f)*, and such that* $m^-(x_0) > m^+(x_0)$.

The logic of these results is completely analogous to those we discussed in the previous section, adequately adjusting the roles of convexity/concavity and the restrictions on the 'kink'. In particular, where failure avoidance requires convexity in the failure region and the kink to have $m^- > m^+$ to counteract it, failure resignation requires concavity and $m^- < m^+$. Likewise, where success attachment requires concavity in the success region and $m^- < m^+$, success seeking requires convexity in the success region and $m^- > m^+$. As a consequence, unlike the attitudes discussed in the previous sections, Success Seeking and Failure Resignation are mutually exclusive. This is because here too the continuous cases cannot coexist, since they require kinks in different directions, and furthermore they do not have discontinuous analogues where they can. I

4 Attitudes Towards Success and Failure: A Full Map

It is especially informative at this point is to reflect on the full picture that emerges from Theorems 1-4 considered jointly, and the corollaries that follow for all possible combinations of conjunctions and disjunctions of the four attitudes, which we summarize in Fig. 3.

Figure 3: Attitudes Towards Success and Failure: A Full Map



Graphical illustration of the logical relationship between the four representation theorems.

4.1 Special Cases of Interest

As previously mentioned, some special cases of our representation are especially significant, and have emerged in different contexts in different parts of the literature.

Prospect Theory Utility Function and Loss Aversion: A widely used representation within economics and psychology corresponds to the case, typically with $x_0 = 0$, where the utility function is convex on the losses (failure) and concave on the gains (success), and that it has a kink around the reference point such that $m^- > m^+$. The first feature is typically referred to as *diminishing sensitivity*, the second as *loss aversion*, to capture the idea that losses loom larger than commensurate gains. This representation is widely used in cumulative prospect theory and in the related literature (e.g., Kahneman and Tversky (1979); see also Abdellaoui (2000), Wakker (2010), O'Donoghue and Sprenger (2018), and references therein, both with a non-linear rank dependent weighting function and with a linear weighting function (the latter is especially common in applications). In the following we maintain a linear weighting function, as the vNM axioms are maintained throughout.[7]

[7]We maintain linear probability weighting here to focus on attitudes towards success and failure and to isolate them from other confounds, such as probability distortion. Allowing for non-linear weighting functions and exploring the possible composition effects would be an interesting extension for future research. We note, however, that doing so in this setting would require separate reweighting functions for failures and successes. Otherwise, given the appropriate adaptations of SSNR and our attitudes to an RDU setting, the weighting function would have to be both convex and concave (and, hence, linear), thereby reducing to EU. To avoid such a high number

A frequent specification of this functional form takes $v : \mathbb{R}_+ \to \mathbb{R}$ to be a concave increasing function defined on the gains domain, and lets $u : \mathbb{R} \to \mathbb{R}$ be such that:

$$u(x) := \begin{cases} v(x - x_0) \text{ if } x > x_0 \\ -\lambda v(-(x - x_0)) \text{ if } x < x_0 \end{cases}, \tag{1}$$

where the parameter $\lambda > 1$ is used to capture the notion of loss aversion (e.g., Wakker (2010), Imas (2016), etc.), and is equivalent to the ratio m^-/m^+ in our notation.

Despite the prominence of this representation, to the best of our knowledge no axiomatic characterization has been provided for it in terms of the underlying preferences over lotteries. Rather, there have been several important contributions in the axiomatic literature on CPT focusing on the distinct point of separating the utility function from the probability weighting function, which in our case is taken to be linear (see, e.g., Wakker (2010)). Exceptions to this broad tendency include Wakker and Zank (2002), who obtain a loss-aversion shape in a representation with power utility functions, and Schmidt and Zank (2012), who axiomatize an S-shaped utility function in a setting with non-linear probability weighting, but remain silent on the kink, i.e. on loss aversion (in the commonly used sense of Köbberling and Wakker (2005)). Likewise, Wakker and Tversky (1993) contains a global definition that is distinct from the kink (see Abdellaoui et al. (2007) for a discussion on these definitions). Neilson (2002) discusses various definitions of loss aversion, and studies comparative statics (including of the S-shape of the utility function) directly in terms of the utility representation. The exercise therefore is different from ours, as are the notions of loss aversion he considers.

Theorems 1-4 jointly provide a characterization of loss aversion (in the commonly used sense of the kink, cf. (Köbberling and Wakker, 2005)) in terms of preferences over lotteries, thereby complementing the literature along several dimensions. First, focusing only on loss aversion, *without* imposing diminishing sensitivity, it is immediately clear from Fig. 3 that the case of loss aversion (i.e., a continuous utility function with $m^-/m^+ > 1$) is completely characterized by the union of failure avoidance and success seeking, with none of the other attitudes. Formally:

Corollary 2. *Under the maintained vNM and monotonicity axioms, preferences are represented by a Bernoulli utility function that displays* Loss Aversion at x_0 *(i.e. continuous, with* $m^- < m^+$*) if and only if they display Failure Avoidance or Success Seeking, but none of the other attitudes.*

As previously mentioned, loss aversion is incompatible with success attachment, which joint with failure avoidance would entail a discontinuity. But it is also incompatible with failure resignation, whose kink goes in the opposite direction. If we further impose diminishing sensitivity, to obtain the kinked S-shape utility function of Prospect Theory, then we must further rule out Success Seeking, for which the utility function must be convex in the success region. This characterization therefore provides insight into the nature of preferences that are compatible with the Prospect Theory utility representation. Formally:

Corollary 3. *Under the maintained vNM and monotonicity axioms, preferences are represented by a Bernoulli utility function that displays the* Prospect Theory utility representation at x_0

of parameters, we do not include the RDU extension in this paper, so as to focus on the key innovations first.

(i.e., continuous, convex on a left-neighborhood of x_0, concave on a right-neighborhood, and with $m^- < m^+$) if and only if they display Failure Avoidance, but none of the other attitudes.

This result also shows that a behavioral characterization of phenomena that are commonly associated with prospect theory (such as *loss aversion* and *diminishing sensitivity*) can be given within a completely standard expected utility setting. Hence, setting aside the important and long-debated issue of whether outcomes should be regarded as total wealth, prospects, or the result of other forms of narrow bracketing (see, e.g., Rabin (2000), Rabin and Thaler (2001), Rubinstein (2002), etc.), our results formally show that loss aversion may be captured by standard risk preferences under the vNM axioms. That is, it need not involve other components of Prospect Theory, such as non-linear probability weighting or rank-dependence. Besides providing a formal result concerning this point of debate, Corollary 3 may thus also serve as a preliminary step to identify the behavioral foundation of other components of Prospect Theory, which have not been fully understood in terms of their distinct roles in accommodating deviations from the expected utility benchmark.

The Discontinuous Case: The discontinuous model has a long tradition within the finance literature, in which it is often referred to as the *aspiration level model* (see, e.g., Payne et al. (1980, 1981)), and has been studied both theoretically and experimentally.

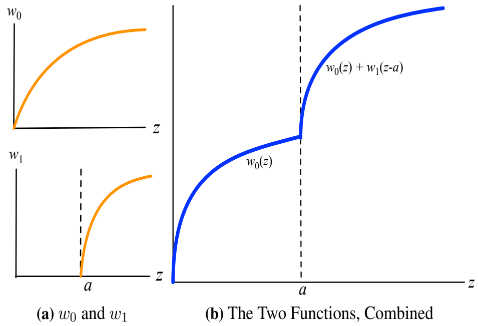
Within the decision theoretic literature, Diecidue and Van De Ven (2008) also present a model of decision under risk with a Bernoulli utility that is discontinuous at what they call 'aspiration level'. Like ours, their model also falls within expected utility. The key axioms in that paper (other than vNM and stochastic dominance), however, are not in terms of preferences over lotteries, but they are formulated directly as continuity properties of the Bernoulli utility function. Hence, another outgrowth of Theorems 1 and 2 – namely, Corollary 1 – is to provide a fully preference-based foundation to the discontinuous utility function, and hence to the broader finance literature on the 'aspiration level model'.[8]

From an empirical viewpoint, several findings in the literature are suggestive of the existence of discontinuities. Fishburn (1977, p.122), for instance, reports that similar preferences are often found in the literature, which can be represented by a 'pronounced change in the shape of their utility function.' Within finance, Mezias (1988) provides evidence in this sense in the pricing of securities in the stock market, when there is a fixed and predetermined benchmark return (similar evidence was provided by earlier work, e.g. Swalm (1966)). The influential paper by Chevalier and Ellison (1999) is also consistent with a discontinuity around the 'benchmark' return, although in that case the phenomenon may be at least partly due to a discontinuity in the reward scheme of the managers, in addition to the possible discontinuity in their primitive preferences. A few papers have further tested experimentally the existence of discontinuities at specific points (typically at $x_0 = 0$, as customary within the finance literature), with contrasting results.[9]

[8] A distinct model of discontinuity, due to the consumption of 'values', is Minardi et al. (2024).

[9] For instance, Payne (2005) find evidence in support of the discontinuity hypothesis, with findings replicated by Venkatraman et al. (2009, 2014). Markle et al. (2018) find evidence suggestive of discontinuities in a context of marathon running. Diecidue et al. (2015), instead, find no evidence of discontinuities at $x_0 = 0$.

Figure 4: A model of aspirations



The functional form in the aspiration model of Genicot and Ray (2017)

Experimental evidence aside, the discontinuous representation is often convenient to capture in a parsimonious way the key feature of the attitudes above, namely the risk-attitude reversals around the threshold. Alaoui and Fons-Rosen (2021), for instance, use a utility function with a discontinuity at the threshold to represent the effects of 'tenacity' on a gambling task, so as to capture the cost of failure. Their experimental analysis relate subjects' behavior in the task with *grit*, as measured by Duckworth and Quinn (2009).

Aspirations: A large literature has studied the origins and implications of *aspirations*, modeled as reference points that serve as a dividing line between achievement and failure (see Genicot and Ray (2020) for a survey of the literature). The focus of that literature is largely on the determinants of such reference points, and on the interplay between individual behavior and economic development, which affects the former through its effect on aspirations, and hence preferences (e.g., Ray (1998, 2006), Appadurai (2004), Genicot and Ray (2017), etc.). The literature has studied various mechanisms for the determination of aspirations levels. As discussed in Genicot and Ray (2020), the key ideas of this notion of aspirations can be modeled by a utility functions that is concave on both sides of the reference point, with a 'convex kink' at the aspiration threshold (i.e., with $m^- < m^+$), as in the representation that is characterized by the intersection of *Success Attachment* and *Failure Resignation*. In the model of Genicot and Ray (2017), for instance, crossing the threshold is "celebrated" by an additional separable payoff. (see Fig. 4).

Corollary 4. *Under the maintained vNM and monotonicity axioms, preferences are represented by a Bernoulli utility function that displays* an aspiration point at x_0 *(i.e., continuous, concave on a left- and right-neighborhood of x_0, with $m^- < m^+$) if and only if they display both failure resignation and success attachment.*

This result shows that the key feature that aspiration models typically capture in a risk-less setting – namely, the sudden increase in marginal utility past the aspiration threshold – can be given a behavioral characterization in a standard choice setting with risk.

Other Cases: The remaining cases, which are characterized by success seeking, or by failure resignation without success attachment, are perhaps not as frequently encountered, but they complete the map of possible attitudes. It is worth mentioning though that, motivated by the classic paper by Friedman and Savage (1948) – who observe the existence of decision makers who simultaneously buy insurance for moderate risks and tickets for actuarially unfair lotteries – Markowitz (1952) argues for a utility function over gains and losses (as opposed to wealth levels), with a pattern of risk lovingness followed by risk aversion as the stakes increase for gains, and the opposite for losses. This suggestion is consistent with the pattern characterized by success seeking without failure avoidance at $x_0 = 0$.

The empirical literature on loss aversion has also produced some evidence of behavior consistent with such representations, again for the $x_0 = 0$ threshold. In the experiment conducted by Schmidt and Traub (2002), for instance, 24 percent of subjects behave exactly opposite to loss aversion, i.e., as if they focus more on gains than on losses. In a decision context involving health outcomes and no risk, Bleichrodt and Pinto (2000) instead find that the proportion of such *gain-seeking* subjects is very low, between 0 and 2.5 percent.

4.2 Discussion and Variations

In practice, it is essentially impossible to test exactly whether an individual's utility function is continuous or differentiable at a particular point. So, just as it is impossible to literally test *global* risk-aversion, and as it is standard in the lab to elicit subjects' preferences over a 'grid' of outcomes, so the representations in Theorems 1-4 could only be tested up to some neighborhood around the threshold. This can be given a formal foundation by providing weaker versions of Def. 2-5 which are not referred to a specific threshold x_0, but to some threshold within a (small) interval.[10] The corresponding representation, and hence the predictions that are directly testable, are exactly those that can be obtained from those in the theorems above, for x and x' that do not converge to x_0 (as it would be in the continuum), but to $x_0 \pm \epsilon$, where ϵ denotes the smallest available discrete increment. For failure avoidance, for instance, it would still be the case that the agent would be risk-loving for lotteries concentrated on the left of $x_0 - \epsilon$, and for each $x \leq x_0 - \epsilon$ on the discrete grid, and for $x' = x + \epsilon$, there would exist a probability $p^* \in \Delta(x, x')$ such that the agent is risk-loving for all $p < p^*$ and risk-averse for all $p > p^*$.

Foundational considerations notwithstanding, a substantial body of literature has tested the discontinuous model, sometimes finding evidence in favor of a discontinuity (see, e.g., Payne (2005), Venkatraman et al. (2009, 2014) and Markle et al. (2018). This is typically done through maximum likelihood estimation of the parameters of a utility function which include discontinuity parameters at the relevant thresholds. The same applies to the point of non-differentiability in the entire literature on loss aversion and prospect theory.

[10]We are thankful to Antonio Cabrales for this suggestion.

5 Interpersonal Comparisons

In this section we provide model-free definitions to rank individuals by the intensity of the four attitudes we introduced. We first focus on *failure avoidance*, which thanks to its close connection with the prospect theory representation, is best suited to explain the key features that an adequate ordering of this attitude must satisfy. The corresponding notions for the other attitudes follow a similar logic, and are left to the online appendix.

Intuitively, an individual is *more failure avoidant* than another one if, compared to the preferences of the latter, his preferences satisfy the following two requirements: i) first, there is a smaller set of lotteries which he regards as 'net successes', and (ii) there is a smaller set of lotteries over which he is unwilling to take a risk in order to get out of the failure region. Formally, for any $x < x_0$ and $x' > x_0$, we define the following sets:

$$\mathcal{S}_i(x,x') := cl\left\{p \in \Delta(x,x') : CE_i(p) \text{ exists and } CE_i(p) > x_0\right\}, \tag{2}$$

$$RA_i(x,x') := cl\left\{p \in \Delta(x,x') : CE_i(p) \text{ exists and } CE_i(p) < Ep\right\}. \tag{3}$$

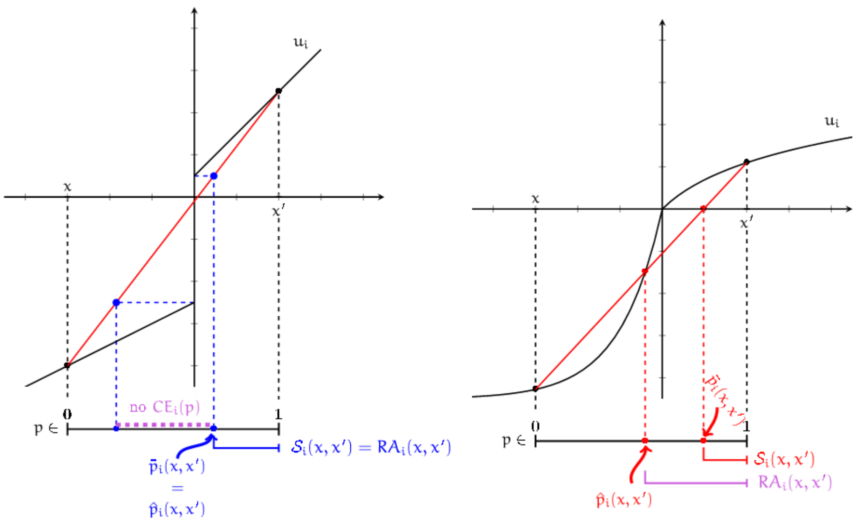
In words, the $\mathcal{S}_i(x,x')$ set represents (the closure of) the set of lotteries which he regards as *net successes*, in the sense that their certainty equivalent is larger than x_0. The $RA_i(x,x')$ set instead represents the set of lotteries over which failure avoidance is not manifested, in the sense that the agent is *not* willing to take risk in order to avoid the potential failure, provided that a certainty equivalent exists. In the continuous case, these sets could be equivalently defined, respectively, as $\mathcal{S}_i(x,x') = \{p \in \Delta(x,x') : p \succsim \delta_{x_0}\}$ and $RA_i(x,x') = \{p \in \Delta(x,x') : \delta_{Ep} \succsim p\}$, which have a straightforward interpretation.[11] For the discontinuous case, however, the formulations above add the further requirement that $CE_i(p)$ exists (which of course is not guaranteed for all p, when u is discontinuous).

To gain some intuition as to why it is desirable to specify this further requirement in the discontinuous case, note that implicit in Def. 2 there is the idea that the agent starts out from being risk-averse for high $p \in [0,1]$ – or, in certainty equivalents terms, they start out by having $Ep > CE_i(p)$ for sufficiently high p. Their desire to avoid failure is what may upset their risk-aversion, and in particular the ranking $Ep > CE_i(p)$, either by turning it into the opposite direction, or (for the case of a discontinuous Bernoulli utility function) by first preventing the existence of $CE(p)$. So, either the inversion of the inequality, or the non-existence region for the CE, are manifestations of a desire to avoid failure. Since, under Def. 2, $CE_i(p) < Ep$ implies that $CE_i(q) < Eq$ for all $q > p$, the set $RA_i(x,x')$ thus represents the set of lotteries over which this phenomenon is *not* (yet) manifested, and similarly $\mathcal{S}_i(x,x')$ represents the set of lotteries that are viewed as net successes, before the discontinuity (and, hence, the non-existence of the certainty equivalence) has kicked in. Fig. 5 illustrates the $\mathcal{S}_i$ and RA_i sets for preferences that display failure avoidance, both in the discontinuous and in the loss aversion case.

The next definition states that an agent is more failure avoidant than another one if he is both more reluctant to regard a lottery as a net success (i.e., a smaller $\mathcal{S}_i$ set), and if he manifests a desire to avoid failure for a larger set of lotteries (i.e., a smaller RA_i set), for all the x and x'

[11] In fact, the equivalence between the two formulations would hold for any u_i that is right-continuous at x_0.

Figure 5: Illustration of the $\mathcal{S}_i$ and RA_i sets



The $\mathcal{S}_i$ and RA_i sets for the *discontinuous* (left) and *loss aversion* case (right).

which identify the phenomenon of failure avoidance (as per Def. 2):

Definition 6. *Let preferences* $\succsim_1$ *and* $\succsim_2$ *both satisfy the conditions in Def. 2 with respect to the same* $x_0 \in \mathbb{R}$*. Then,* $\succsim_1$ displays (weakly) more failure avoidance than $\succsim_2$ *if,* $\exists x_f, x_s : x_f < x_0 < x_s$ *s.t.* $\forall x \in [x_f, x_0)$*,* $\exists \bar{x} \in (x_0, x_s]$ *s.t., for each* $x' \in (x_0, \bar{x}]$*, the following conditions are satisfied: (i)* $\mathcal{S}_1(x, x') \subseteq \mathcal{S}_2(x, x')$*, and (ii)* $RA_1(x, x') \subseteq RA_2(x, x')$*.*

The next result provides necessary and sufficient conditions on the relationship between two Bernoulli utility functions, for their corresponding preferences to be ranked by their failure avoidance, as we just defined.[12]

Theorem 5 (Failure Avoidance: Interpersonal Comparisons)**.** *Let* $\succsim_1$ *and* $\succsim_2$ *both satisfy the conditions in Def. 2 with respect to the same* $x_0 \in \mathbb{R}$ *and such that* $m_i^+ > 0$ *and* $m_i^- < \infty$ *for both* $i = 1, 2$*. Then,* $\succsim_1$ *displays* more failure avoidance *than* $\succsim_2$ ***only if*** *one of the following applies:*

1. $\frac{K_1}{m_1^+} > \frac{K_2}{m_2^+}$,

2. $\frac{K_1}{m_1^+} = \frac{K_2}{m_2^+} > 0$ *and* $\frac{m_1^-}{m_1^+} \geq \frac{m_2^-}{m_2^+}$,

3. $\frac{K_1}{m_1^+} = \frac{K_2}{m_2^+} = 0$, $\frac{m_1^-}{m_1^+} \geq \frac{m_2^-}{m_2^+}$, *and* $\left(lim_{x \to x_0^-} \frac{[m_1^- - m_1(x)]/m_1^-}{[m_2^- - m_2(x)]/m_2^-} \right) \geq \frac{1 - m_1^+/m_1^-}{1 - m_2^+/m_2^-}$.

These conditions are also ***sufficient*** *if all the inequalities hold strictly.*

[12]Theorem 10 in the online Appendix provides tight (but harder to read) if and only if conditions.

The conditions in this theorem have a straightforward interpretation. First, the condition $\frac{K_1}{m_1^+} \geq \frac{K_2}{m_2^+}$ says that the size of the discontinuity at x_0, normalized by m_i^+, is larger for 1 than for 2. Hence, this result implies that the first determinant of the relative failure avoidance is the size of the *normalized discontinuity*. In case of ties in this first component, if the utility functions are discontinuous, then the ranking is determined by the sharpness of the *kink* of the utility function around x_0, which is captured by the ratio m_i^-/m_i^+: the larger the ratio, the stronger the failure avoidance. If instead the functions are continuous, then agent 1 displays stronger failure avoidance than agent 2 if not only its utility function displays a sharper kink ($\frac{m_1^-}{m_1^+} \geq \frac{m_2^-}{m_2^+}$), but also if u_1 is *sufficiently more convex* than u_2 in some left-neighborhood of x_0.

To see that this is the content of the limit condition in point 3 of the result, note that $lim_{x \to x_0^-} \frac{[m_1^- - m_1(x)]/m_1^-}{[m_2^- - m_2(x)]/m_2^-} \geq 1$ is equivalent to requiring that u_1 is more convex than u_2 (in the Arrow-Pratt sense) in some left-neighborhood of x_0. Condition 3 strengthens this requirement by requiring the limit of this ratio to not be just larger than one, but also larger than $\frac{1-m_1^+/m_1^-}{1-m_2^+/m_2^-}$, which is a measure of the ratio of the kinks of the two utility functions (which in turn is also required to be larger than one, under the condition $\frac{m_1^-}{m_1^+} \geq \frac{m_2^-}{m_2^+}$). Intuitively, a sharper kink determines a stronger concavity on some right-neighborhood of x_0; the condition in point 3 requires that u_1 not only has a sharper kink, but it is also sufficiently more convex on the failures than u_2, so as to offset the stronger concavity on the successes associated with its sharper kink.

The intuition above is perhaps easiest to see in the case of differentiable utility functions, in which the conditions above take an easy-to-interpret form, analogous to the classical Arrow-Pratt indices of risk-aversion. Letting Du_i^- and Du_i^+ denote the left- and right-derivatives of u_i at x_0, and $D^2u_i^-$ the second left-derivative at x_0, we have:

Theorem 6 (F.A. Indices under Differentiability)**.** *Suppose that $(\succsim_i)_{i=1,2}$ are such that $m_i^+ > 0$ and $m_i^- < \infty$ and u_i is twice differentiable in some left- and right-neighborhoods of x_0. Then: $\succsim_1$ displays* more failure avoidance *than $\succsim_2$* ***only if*** *one of the following applies:*

1. $\frac{K_1}{Du_1^+} > \frac{K_2}{Du_2^+}$,

2. $\frac{K_1}{Du_1^+} = \frac{K_2}{Du_2^+} > 0$ *and* $\frac{Du_1^-}{Du_1^+} \geq \frac{Du_2^-}{Du_2^+}$,

3. $\frac{K_1}{Du_1^+} = \frac{K_2}{Du_2^+} = 0$, $\frac{Du_1^-}{Du_1^+} \geq \frac{Du_2^-}{Du_2^+}$ *and* $\frac{D^2u_1^-}{Du_1^- - Du_1^+} \geq \frac{D^2u_2^-}{Du_2^- - Du_2^+}$.

These conditions are also ***sufficient*** *if all the inequalities hold strictly.*

5.1 Ordering Failure Avoidance: Discussion

In this section we discuss the role of the two components that make up our definition of interpersonal comparison of failure avoidance, in terms of both the $\mathcal{S}_i$ and RA_i sets.

First, as can be seen from proof of Theorem 5, the following holds:

Lemma 1. *Let u_i be discontinuous at x_0 and represent preferences that exhibit* failure avoidance *at x_0. Then, $\exists x_f < x_0$ s.t.: $\forall x \in (x_f, x_0)$, $\exists \bar{x} > x_0 : \forall x' \in (x_0, \bar{x})$, $\mathcal{S}_i(x, x') = RA_i(x, x')$.*

That is, in the case of discontinuous representation, the $\mathcal{S}_i$ and RA_i sets coincide. Hence, the two conditions involved in Def. 6 are equivalent to each other if both u_1 and u_2 are discontinuous at x_0. For continuous utility functions, however, the two conditions are distinct. Hence, dropping either part of Def. 6 would have no bearing on the ranking of discontinuous utility functions, and it would yield a more complete order over the continuous utility functions. Either of these more complete orders, however, would not yield a satisfactory ranking of *failure avoidance.* To see this, first suppose that part (ii) is dropped from Def. 6, so that the ranking is solely based on the $\mathcal{S}_i$ sets. Then, the following holds:

Lemma 2 (Ordering Kinks)**.** *Let u_1 and u_2 be continuous utility functions. Then: $m_1^-/m_1^+ > m_2^-/m_2^+$ only if $\exists x_f < x_0$ s.t.: $\forall x \in (x_f, x_0)$, $\exists \bar{x} > x_0 : \forall x' \in (x_0, \bar{x})$, $\mathcal{S}_1(x, x') \subset \mathcal{S}_2(x, x')$. The converse holds with $m_1^-/m_1^+ \geq m_2^-/m_2^+$.*

This result follows from the proof of Theorem 5. Intuitively, given two agents with continuous utility functions, function u_1 has a *sharper kink* than u_2 at x_0 if and only if, for any x and x' (with the order of quantifiers as in Def. 2), the set of lotteries $p \in \Delta(x, x')$ that 1 regards as 'net successes' is a subset of those that 2 regards as 'net successes'.

Now, consider a sequence of utility functions $(u^{(n)})_{n\in\mathbb{N}}$ such that, for each $n \in \mathbb{N}$,

$$u^{(n)}(x) = \begin{cases} \hat{u}(x) & \text{if } x \geq x_0 \\ -(2m^+ - \alpha(n)) \cdot (x_0 - x)^{\frac{1}{(1-\alpha(n))}} & \text{if } x < x_0 \end{cases},$$

where $\hat{u}$ is an arbitrary concave function, with corresponding m^+, and $\alpha(n)$ is a decreasing sequence such that $\alpha(1) < min\{m^+, 1\}$ and such that $lim_{n\to\infty}\alpha(n) = 0$. As $n \to \infty$, the kink gets sharper along this sequence, and hence it is increasing in the ranking induced by part (i) of Def. 6, but $\hat{u}^{(n)}$ approaches risk neutrality over the loss domain, and hence at the limit, $u^* := lim_{n\to\infty}\hat{u}^{(n)}$, the u^* function is globally concave, and hence there is no failure avoidance. Thus, an order based on part (i) of Def. 6 alone would allow the possibility of sequences of increasingly failure avoidant preferences which converge to preferences that display no failure avoidance. This would not be a desirable property for an adequate ordering of failure avoidance.

Alternatively, suppose that part (i) is dropped, so that utility functions are ranked based on part (ii) of Def. 6 alone. The proof of Theorem 5 also shows the following:

Lemma 3. *Let u_1 and u_2 be continuous utility functions that exhibit* failure avoidance *at x_0. Then:*

$$\left(lim_{x\to x_0^-}\frac{[m_1^- - m_1(x)]/m_1^-}{[m_2^- - m_2(x)]/m_2^-}\right) > \frac{1 - m_1^+/m_1^-}{1 - m_2^+/m_2^-}.$$

only if $\exists x_f < x_0$ s.t.: $\forall x \in (x_f, x_0)$, $\exists \bar{x} > x_0 : \forall x' \in (x_0, \bar{x})$, $RA_1(x, x') \subset RA_2(x, x')$. The converse holds with the weak inequality.

Now, let $\hat{u}$ be a continuous utility function which satisfies the condition of the representation theorem, and which is linear in the success region. Next, consider a sequence of utility functions

$(u^{(n)})_{n\in\mathbb{N}}$ such that, for each $n \in \mathbb{N}$,

$$u^{(n)}(x) = \begin{cases} \hat{u}(x) & \text{if } x \le x_0 \\ \alpha(n) \cdot \hat{u}(x) & \text{if } x > x_0 \end{cases},$$

where $\alpha(n)$ is an increasing sequence of real numbers such that $\alpha(1) = 1$ and $lim_{n\to\infty}\alpha(n) = \frac{\hat{m}^-}{\hat{m}^+}$. Then, it can be verified that the sequence $u^{(n)}$ is increasing in the order defined by part (ii) of Def. 6, and yet $u^* := \lim_{n\to\infty} u^{(n)}$ is globally convex, and hence does not display any failure avoidance. Thus, just like the case discussed above, also an order only based on part (ii) of Def. 6 would allow for sequences of increasingly failure avoidant preferences which converge to preferences that display no failure avoidance at all. This, again, would not be a desirable feature for a conceptually sound notion of comparative failure avoidance.

5.2 Ordering Loss Aversion

The discussion above is also significant in relation to established notions of comparative loss aversion, which rank loss aversion by the sharpness of the kink, so that agent 1 is more loss averse than agent 2 if and only if $\frac{m_1^-}{m_1^+} > \frac{m_2^-}{m_2^+}$ (cf. Köbberling and Wakker (2005), Abdellaoui et al. (2007); see also Wakker (2010) and references therein) – or, in the parametric specification of eq. (1), if and only if $\lambda_1 > \lambda_2$. To the best of our knowledge, such interpersonal comparisons have been defined only in the space of the utility representation, but a characterization of such orderings in terms of primitive preferences is lacking. The next result provides such a characterization, and hence it also serves to open another perspective on loss aversion, in terms of preferences over lotteries:

Proposition 1 (Ordering Loss Aversion)**.** *Let u_1 and u_2 be continuous utility functions. Then: there exist $\underline{x} < x_0$ and $\bar{x} > x_0$ such that, for all $x \in (\underline{x}, x_0)$ and for all $x' \in (x_0, \bar{x})$, $\mathcal{S}_1(x, x') \subset \mathcal{S}_2(x, x')$ if $m_1^-/m_1^+ > m_2^-/m_2^+$ and only if $m_1^-/m_1^+ \ge m_2^-/m_2^+$.*

Namely, agent 1 has a sharper kink at x_0 than agent 2 if for all failures $x < x_0$ and successes $x' > x_0$ in some neighborhood of x_0, $\mathcal{S}_1(x, x') \subset \mathcal{S}_2(x, x')$ – meaning that the binary lotteries across the threshold that 1 views as *net successes* are a subset of those that agent 2 views as such. We note that the order of quantifiers in Proposition 1 is slightly different from that of Lemma 2, in that it is symmetric on both sides of the threshold. The reason is that Proposition 1 follows directly from the joint implications of Lemma 2 (in which the ordering stems from the possibility an arbitrarily small success, as in failure avoidance) and of an analogous result for when the ordering stems from the possibility of an arbitrarily small failure (as in Success Seeking) that also involves a condition on the nestedness of the $\mathcal{S}$-sets, but for an order of quantifiers that is symmetric with respect to that involved in Def. 2 (see Section C.1). In other words, within the spirit of our approach, Proposition 1 captures the union of ordering loss aversion as driven by the mere chance of success or by the mere chance of failure.

Suppose that, in addition to ordering loss aversion, one wishes such an ordering while remaining exactly within the context of the Prospect Theory utility representation, in the sense that the kink is sharpened while at the same time maintaining convexity on the left and concavity on the right. Then, the characterization of *"Ordering Loss Aversion within Prospect Theory"* is the one provided by Lemma 2, with the added restriction that the preferences display failure avoidance and *only* failure avoidance. If, in contrast, one wishes to rank loss aversion per se, independent of the S-shape of the utility representation, then Proposition 1 is the appropriate characterization. This also clarifies a subtlety about the intuitive view of what sharpening the kink does. Namely, it is clear that a sharper kink instills extra risk aversion in the preferences.[13] The $RA_i(x, x')$ and the $\mathcal{S}_i(x, x')$ sets both seem to capture the same qualitative idea, but in slightly different ways. The $\mathcal{S}_i(x, x')$ set identifies the set of lotteries supported on $\{x, x'\}$ that are at least as good as the threshold: as risk aversion increases, a larger set of lotteries ends up in the lower contour set of δ_{x_0}, and hence $\mathcal{S}_i(x, x')$ gets smaller. The $RA_i(x, x')$ set instead is exactly the set of lotteries over which the agent displays risk aversion, and hence it increases as risk aversion increses. But while they are obviously related, they do not coincide, and hence it is possible (as in Def. 6) that both sets get larger at the same time. Lemma 2 shows that, as long as loss aversion is identified with the presence of a kink at the threshold (as in the classical behavioral literature), increasing its intensity does amount to a local increase in risk aversion (as argued, for instance, by Cerreia-Vioglio et al. (2022)). Its exact ordering, however, is captured by shrinking of the $\mathcal{S}_i(x, x')$ sets, and not by the enlargment of the $RA_i(x, x')$ sets.

The discussion in the previous section also demonstrates that, as long as *loss aversion* is defined as something to be ranked solely by the sharpness of the kink, then it is distinct from our notion of *failure avoidance*. In particular, while a more loss averse agent 1 will have $\mathcal{S}_1$ to be a subset of that of a less loss averse agent 2, it need not be the case that RA_1 will also be a subset of RA_2. In fact, the increased sharpness of the kink on its own provides a force in the *opposite* direction. This is because a sharper kink leads to a more concave function, which on its own implies that there are fewer lotteries over which the agent is willing to take a risk to avoid failure. In the limit, failure avoidance disappears altogether. Hence, while the first requirement of Definition 6 is satisfied, the second requirement is violated. This is why point 3 of Theorem 5 requires a sufficient increase in the convexity to offset the concavity associated with the sharper kink. Hence, ranking loss aversion merely by the sharpness of the kink is appropriate for capturing the idea of *'losses looming larger than gains'* (cf. Köbberling and Wakker (2005) and Abdellaoui et al. (2007)). But it does not adequately capture a ranking of the *reversals* of the risk-attitude, which is the focus of this paper. Such a ranking requires an opposition of forces, in that any force that leads to an increase in risk-aversion must be countervailed by a force in the opposite direction.

These results also show the advantage of remaining within EU for our objectives, as the distinction between the ranking of failure avoidance and the standard one of loss aversion might have been easy to miss in a more complex setting, in which the centrality of the opposing forces may have been less transparent, when interacting with other factors.[14]

An analogous exercise to that of ordering failure avoidance can be conducted for all the

[13]Theorem 5 in Köbberling and Wakker (2005) formalizes precisely this argument.

[14]While in this paper we remain within the vNM model, it may be interesting for future research to analyze what our model would imply in settings which allow for a non-linear reweighting of probabilities. Such a reweighting provides another source of reversal of risk-attitude, independent of the shape of the utility function.

remaining attitudes. Much of the reasoning above carries through, *mutatis mutandis*, to the definitions and results of these attitudes. We leave these results to the online appendix.

6 Conclusions

This paper aims to understand, at a fundamental level, attitudes towards success and failure that are crucial to decision-making, as evidenced by their emergence in several influential fields. Within a standard expected-utility setting, we provide characterizations of these attitudes in terms of properties of the Bernoulli utility function. This exercise serves several purposes: First, it reveals the interconnection between different models of reference-dependent preferences. Second, it provides a decision theoretic foundation to important representations used in economics, finance and psychology (including influential models of *aspirations* and *loss aversion*), for which a standard preference-based characterization was lacking. This not only favors more direct comparisons of these models with standard expected utility notions, but it also uncovers subtleties which may be easily overlooked by only looking at the space of utility representations. A case in point is provided by the rankings that we introduce in order to perform interpersonal comparisons on the intensity of each attitude. The indices we develop, which are akin to the Arrow-Pratt indices used for studying risk aversion, shed a new light on seemingly intuitive notions of comparative statics that are directly based on the utility representation of reference-dependent models.

The distinctive feature of our approach is to identify the core of such behavioral phenomena in the *reversals* of the decision maker's risk-attitude over lotteries that go across a reference point. This novel perspective enables a unified view of several influential models of reference dependence, and it also provides a direct way of identiying reference points through choice, by the occurrences of such reversals around them.

Attitudes towards success and failures are the focus of central notions in the literature on personality traits, such as grit, tenacity, conscientiousness and neuroticism (e.g., Deary et al. (2009)). The key methodology in this literature involves 'indices' that are essentially scores on non-incentivized questionnaires, often based on self-reported scales, which are intended to capture various aspects of personality. The empirical economics literature has shown that these measurements are often predictive of systematic differences in behavior and measures of economic performance (e.g., Heckman and Rubinstein (2001), Almlund et al. (2011), Burks et al. (2015), Gill and Prowse (2016, 2021), Proto et al. (2019, 2022), Heckman et al. (2021), etc.). But several features of the psychology measurements make it difficult to perform a direct translation of those concepts into tractable economics concepts: First, the lack of precise and agreed-upon definitions of terms such as grit, tenacity, conscientiousness, etc., as something that is separate from their way of measurement. Second, the high dimensionality of the objects involved in each trait.

In contrast, albeit limited in its richness, the straightjacket of economic analysis has proven very successful in providing rigorous definitions of behavioral notions, which can be used both to make theoretical predictions and for empirical measurement. Our approach to attitudes towards success and failure mimics the development of the risk analysis program, building from the bottom up notions that are directly expressed in terms of primitive preferences and within the dictamen of the economics methodology.

Further empirical research is needed to assess to what extent the attitudes formalized in this paper are correlated with the psychology measures of personality traits.[15] However, while our notions inevitably miss some of the richness of the psychology definitions, they are amenable both to measurement based on choice data, and to theoretical and counterfactual analysis. In this sense, our framework may prove useful to incorporate, within a standard economic model, behavioral manifestations of personality traits that have proven relevant in empirical analysis but that have hitherto appeared to be elusive to formal modeling and structural analysis.

Appendix

A Representation Theorems: Proofs

Proof of Theorem 1:

Step 1. First note that, under the vNM axioms, SSNR (Def. 1) holds if and only if there exist intervals $[x_f, x_0)$ and $(x_0, x_s]$, with $x_f < x_0 < x_s$, such that u is either concave or convex on $[x_f, x_0)$, and either concave or convex on $(x_0, x_s]$. We also know that we cannot have global concavity nor global convexity, since the first would imply that $Ep \succsim p$ for all p and the second would imply that $p \succsim Ep$ for all p, contrary to Def. 2.

Step 2. [The discontinuous case] For the sufficiency part, suppose that $m^+(x_0) < \infty$. Then, because u is discontinuous at x_0, and letting $K := lim_{x \to x_0^+} u(x) - lim_{x \to x_0^-} u(x) > 0$, there exists $x_s > x_0$ and $x_f < x_0$ s.t. $\frac{u(x')-u(x)}{x'-x} > \max\{m^+(x_0), m(x_s)\}$ for all $x \in (x_f, x_0)$ and for all $x' \in (x_0, x_s)$. Hence, for any $x \in (x_f, x_0)$ and for all $x' \in (x_0, x_s)$, the straight line connecting $u(x)$ to $u(x')$ never intersects $u(\cdot)$ (other than at its extremes), and is such that it is below $u(\cdot)$ on the interval (x_0, x_s), and above it on (x_f, x_0), which implies that the agent is risk averse for all $p \in \Delta(x, x')$ such that $Ep > x_0$, and risk-loving otherwise. Hence, preferences satisfy Def. 2.

For the necessity part, given Step 1, the only thing which is left to prove for the discontinuous case is that $m^+(x_0) < \infty$. But if $m^+(x_0) = \infty$ (note that this is only possible if u is concave in the success region), then for any $x < x_0$, we can find $\bar{x} > x_0$ s.t. $\forall x' \in (x_0, \bar{x})$, $\frac{u(x')-u(x)}{x'-x} < m(x')$. Hence, the straight line connecting $u(x)$ to $u(x')$ is always above $u(\cdot)$, and never intersects it other than at its extremes, contrary to Def. 2.

Step 3. [The continuous case] Given Step 1 above, for the case in which u is continuous, the remaining possibilities are the following:

1. u is concave on both the failures (i.e., on $[x_f, x_0)$) and successes (i.e., on $(x_0, x_s]$), but with $m^-(x_0) < m^+(x_0)$ (otherwise it would be concave on $[x_f, x_s]$);
2. u is concave on the failure and convex on the success region;
3. u is convex on the failure and convex the success region, with $m^-(x_0) > m^+(x_0)$ (otherwise it would be convex on $[x_f, x_s]$);
4. u is convex on the failure and concave on the success region.

We show that Cases 1 and 2 can be discarded, and that failure avoidance holds if and only either a) Case 3 holds with u being strictly convex on some interval $[\tilde{x}_l, x_0]$ or Case 4 holds with $m^-(x_0) > m^+(x_0)$ and u being strictly convex on some interval $[\tilde{x}_l, x_0]$.

Case (1) can be discarded geometrically. First, since $m^-(x_0) < m^+(x_0)$, by continuity of u, there exists $\hat{x}_l \in [x_f, x_0)$ such that, $m(\hat{x}_l) < m^+(x_0)$. We now show that for any $x \in [\hat{x}_l, x_0)$, there is no $\bar{x} > x_0$ such that for all $x' \in (x_0, \bar{x}]$, $\exists p > p' \in \Delta(x, x')$ such that $Ep \succ p$ and $q \succ Eq$. To this end, note that since $m(\hat{x}_l) < m^+(x_0)$,

[15] Jagelka (2024), for instance, recently explored the correlation between various psychological traits and standard notions of risk preferences. Further extending that agenda so as to account for the novel notions put forward by this paper is part of our ongoing research (funded by ERC consolidator grant 101089139).

it follows from the continuity of u that $\exists \bar{x} > x_0$ such that, for all $x' \in [x_0, \bar{x}]$, $m(\hat{x}_l) < \frac{u(x')-u(\hat{x}_l)}{x'-\hat{x}_l}$, and such that the line segment connecting $u(\hat{x}_l)$ to $u(x')$ does not cross the utility function on $[x_0, x')$. Moreover, on $[\hat{x}_l, x_0]$, by concavity the slope $m(x)$ is decreasing which implies $m(x) \leq m(\hat{x}_l)$, and by the utility function being below the line segment between $u(\hat{x}_l)$ to $u(x')$, it implies $\frac{u(x')-u(\hat{x}_l)}{x'-\hat{x}_l} < \frac{u(x')-u(x)}{x'-x}$. Therefore, it must be that $m(x) < \frac{u(x')-u(x)}{x'-x}$ for any $x \in [\hat{x}_l, x_0)$ and $x' \in [x_0, \bar{x}]$.

For the reversal in Def. 2 to hold for $\hat{x}_l$, for any $x' \in (x_0, \bar{x}]$, the line segment connecting $u(\hat{x}_l)$ to $u(x')$ must cross the utility function somewhere in $(\hat{x}_l, x')$. By the previous argument, it cannot be in $[x_0, x')$, and so it must be in $(\hat{x}_l, x_0)$. Let x^* denote such a point. Note that since $x^* \in (\hat{x}_l, x_0)$, the slope $m(x^*) < \frac{u(x')-u(x^*)}{x'-x^*}$, so that the line segment crosses the utility function from below. Since this must hold for any such point, it must be that there is at most one crossing. Moreover, since it crosses from below, there exists a lottery p_t on $\Delta(\hat{x}_l, x'$ such that $p_t \sim Ep_t$, and such that for all $p > p_t$, $p \succsim Ep$, and for all $q < p_t$, $Eq \succsim q$. Since the same argument would hold replacing $\hat{x}_l$ with any $\hat{x}_l' \in [\hat{x}_l, x_0)$, this implies that Condition 2 of failure avoidance cannot hold for this case.

Case (2). For this case, we consider three subcases:

- if $m^-(x_0) < m^+(x_0)$, then it can be discarded based on the same argument as above, since that argument did not rely on the shape of the function on $(x_0, x_s]$.
- If $m^-(x_0) > m^+(x_0)$, then by continuity, there exists an $\hat{x}_l \in [x_f, x_0)$ such that $m(\hat{x}_l) > m^+(x_0)$, and there exists a $\bar{x} > x_0$ such that $\frac{u(x')-u(\hat{x}_l)}{x'-\hat{x}_l} > m^+(x_0)$ for all $x' \in (x_0, \bar{x}]$. Since, the slope m is decreasing on $[\hat{x}_l, x_0)$ by concavity in that interval, and since the above inequality holds, then for any such x', $Ep \succsim p$,
- If $m^-(x_0) = m^+(x_0)$, it is easy to verify geometrically that for any $x' \in [x_f, x_0)$ and $x'' \in (x_0, x_s]$, $\exists p^* \in \Delta(x', x'')$ such that $p^* \sim Ep^*$ and such that $p \succsim Ep$ for all $p > p^*$, and $Ep \succsim p$ for all $p < p^*$, thereby violating the reversal condition in Def. 2.

Cases (1) and (2) are thus both discarded. We are left with Cases (3) and (4), in which the utility is convex in the failures domain.

Case (3). The necessity of the condition $m^-(x_0) > m^+(x_0)$ follows directly from the fact that Condition 2 in Def. 2 can only hold if u is not convex on the entire interval $[x_f, x_s]$. Moreover, if the utility function is linear on any interval $[\tilde{x}_l', x_0]$, where $\tilde{x}_l' < x_0$, then for any $x' > x_0$, the line segment going from $\tilde{x}_l'$ to x' will be below the utility function, and hence $Ep \succeq p$ for any binary lottery $p^* \in \Delta(\tilde{x}_l'')$. As this is true for any such interval, there must be some $\tilde{x}_l$ for which u is strictly convex on $[\tilde{x}_l, x_0]$. We next show that u being as in case (3), with strict convexity on some interval $[\tilde{x}_l, x_0]$, is also sufficient to satisfy the conditions in Def. 2. Since $m^-(x_0) > m^+(x_0)$, by continuity of u $\exists \hat{x}_l \in [\tilde{x}_l, x_0)$ such that $m(\hat{x}_l) > m^+(x_0)$, and there exists a $\bar{x} > x_0$ such that $\frac{u(x')-u(\hat{x}_l)}{x'-\hat{x}_l} > m^+(x_0)$ for all $x' \in (x_0, \bar{x}]$. Since u is strictly convex on $[\hat{x}_l, x_0)$, the slope $m(x)$ is strictly increasing on $[\hat{x}_l, x_0)$, and since the above inequality holds, then $\forall x' \in (x_0, \bar{x}]$, $\exists p^* \in \Delta(\hat{x}_l, x')$ such that $p^* \sim Ep^*$, and such that $Ep \succ p$ whenever $p > p^*$, and $p \succ Ep$ whenever $p < p^*$. The same argument would hold replacing $\hat{x}_l$ with any $x_f \in [\hat{x}_l, x_0)$, which implies that Def. 2.

Case (4). For Case 4, we will show that it is consistent with Def. 2 if and only if (i) $m^-(x_0) > m^+(x_0)$, and (ii) there is an $\tilde{x}_l$ such that u is strictly convex on the interval $[\tilde{x}_l, x_0]$. To this end, we consider three subcases:

- Suppose that $m^-(x_0) < m^+(x_0)$. Taking any $\hat{x}_l \in [x_f, x_0)$, by convexity $m(x) < m^-(x)$ for any $x \in \hat{x}_l, x_0)$, and so $m(x) < m^+(x_0)$. Moreover, by continuity there exists an $\bar{x} > x_0$ such that $\frac{u(x')-u(\hat{x}_l)}{x'-\hat{x}_l} < m^+(x_0)$ for all $x' \in (x_0, \bar{x}]$. Hence, a line segment from $\hat{x}_l$ to $\bar{x}$ is above (weakly at the endpoints) the utility function on $\hat{x}_l, \bar{x}$, and so $p \succsim Ep$ for all p on $\Delta(\hat{x}_l, x')$. The reversal condition in Def. 2 therefore is violated.
- If $m^-(x_0) = m^+(x_0)$, then a similar logic to the one above applies. First, observe that if there is no $\hat{x}_l \in [x_f, x_0)$ such that $m(\hat{x}_l) < m^-_{\hat{x}_0}$ (i.e., if it is locally linear on the failure region), then the function $[x_f, x_s]$ is weakly concave, and will not satisfy failure avoidance. Hence, it must be that there exists a $\hat{x}_l \in [\tilde{x}_l, x_0)$ such that $m(\hat{x}_l) < m^-_{\hat{x}_0} = m^+(x_0) =$. The rest of the argument is then identical to the case above.
- If $m^-(x_0) > m^+(x_0)$ then the logic from Case 3 above applies, and $\succsim$ display failure avoidance. Moreover, as in the case of $m^-(x_0) = m^+(x_0)$, the utility function cannot be locally linear on any interval $[\hat{x}_l, x_0]$, since it would imply that the function on $[\hat{x}_l, x_s]$ is weakly concave, and hence he reversal condition in Def.

2 could not be satisfied. Noting that this holds for $\hat{x}_l$ arbitrarily close to x_0, it must then be that there is an interval $[\tilde{x}_l', x_0]$ on which u is strictly convex.

Lastly, since we have covered all possible cases for continuous u, it must be that for a continuous u, Def. 2 holds if and if only there exists an $x_f < x_0 < x_s$ such that u is strictly convex on $[x_f, x_0)$, concave or convex on (x_0, x_s) and such that $m_{x_0}^0 > m_{x_0}^+$. ■

The proofs of Theorems 2-4 are completely specular to that of Theorem 1, inverting the role of convexity and concavity, and the order of quantifiers in the success and failure regions, according to the corresponding definitions.

B Interpersonal Comparisons: Proofs

Proof of Theorem 5:

Consider the following objects:

$$\bar{p}_i(x,x') := \inf \left\{ p \in \Delta\left(\{x,x'\}\right) : CE_i\left(p\right) \text{ exists and } CE_i\left(p\right) > x_0 \right)\right\} \text{ and} \tag{4}$$

$$\hat{p}_i(x,x') := \inf \left\{ p : CE_i(p) \text{ exists and } CE_i(p) < Ep \right\}. \tag{5}$$

First note that, from the definition of the $S_i\left(x,x'\right)$ and $RA_i\left(x,x'\right)$ sets, it is clear that $S_1\left(x,x'\right) \subset S_2\left(x,x'\right)$ if and only if $\bar{p}_1(x,x') > \bar{p}_2(x,x')$, and $RA_1\left(x,x'\right) \subset RA_2\left(x,x'\right)$ if and only if $\hat{p}_1(x,x') > \hat{p}_2(x,x')$. The following observation follows immediately:

Lemma 4. *Let $\succsim_1$ and $\succsim_2$ satisfy the properties of Def. 2 with respect to x_0. Then:*

1. *Part (i) of Def. 6 holds if and only if $\exists x_f, x_s : x_f < x_0 < x_s$ s.t. $\forall x \in [x_f, x_0)$, $\exists \bar{x} \in (x_0, x_s]$ such that, for each $x' \in (x_0, \bar{x}]$, $\bar{p}_1(x,x') > \bar{p}_2(x,x')$.*
2. *Part (ii) of Def. 6 holds if and only if $\exists x_f, x_s : x_f < x_0 < x_s$ s.t. $\forall x \in [x_f, x_0)$, $\exists \bar{x} \in (x_0, x_s]$ such that, for each $x' \in (x_0, \bar{x}]$, $\hat{p}_1(x,x') > \hat{p}_2(x,x')$.*

Hence, the proof of the theorem will crucially rely on understanding the properties of the objects defined in equations (4) and (5).

First notice that, letting u_i denote the Bernoulli utility functions which represent preferences $\succsim_1$ and $\succsim_2$, as per Theorem 1, $\bar{p}_i(x,x')$ can be equivalently defined as follows:

$$\bar{p}_i(x,x') := \frac{u_i^+(x_0) - u_i(x)}{u_i(x') - u_i(x)}. \tag{6}$$

The equivalence of (4) and (6) follows directly from observing that (6) implies $\bar{p}_i(x,x') \cdot u_i(x') + (1 - \bar{p}_i(x,x')) \cdot u_i(x) = u_i^+(x_0)$, and the fact that, under the maintained assumptions, u_i is both continuous and strictly increasing on (x_0, x').

Lemma 5. *For any $x < x_0$ and $x' > x_0$, $\bar{p}_1(x,x') > \bar{p}_2(x,x')$ if and only if*

$$\frac{K_1}{m_1\left(x'\right)} + \frac{m_1(x)}{m_1\left(x'\right)}(x_0 - x) > \frac{K_2}{m_2\left(x'\right)} + \frac{m_2(x)}{m_2\left(x'\right)}(x_0 - x).$$

Proof. Exploiting the representation theorem, and the notation introduced above, $\bar{p}_i(x,x')$ can be rewritten as follows:

$$\bar{p}_i(x,x') := \frac{u_i^+(x_0) - u_i(x)}{u_i(x') - u(x)} = \frac{K_i + m_i\left(x\right)\left(x_0 - x\right)}{K_i + m_i\left(x\right)\left(x_0 - x\right) + m_i\left(x'\right)\left(x' - x_0\right)}$$

Re-arranging terms and simplifying, it can be shown that $\bar{p}_1(x,x') > \bar{p}_2(x,x')$ if and only if $\frac{K_1 + m_1(x)(x_0 - x)}{m_1(x')} > \frac{K_2 + m_2(x)(x_0 - x)}{m_2(x')}$.

***** Part (i): Characterization, Necessary and Sufficient conditions *****

Lemma 6 (Part (i): Characterization)**.** *Part (i) of Def. 6 holds if and only if there exists $\underline{x} < x_0$, such that $\forall x \in (\underline{x}, x_0)$, there exists $\bar{x} > x_0$, s.t, for all $x' \in (x_0, \bar{x})$:*

$$\frac{K_1}{m_1(x')} + \frac{m_1(x)}{m_1(x')}(x_0 - x) > \frac{K_2}{m_2(x')} + \frac{m_2(x)}{m_2(x')}(x_0 - x). \tag{7}$$

Proof. This result follows directly from Lemma 5 and part 1 of Lemma 4.

The next results provide necessary and sufficient conditions for part (i) of Def. 6:

Lemma 7 (Part (i): Necessity)**.** *If $m_i^+(x_0) > 0$ for both $i = 1, 2$, then:*

1. *Part (i) of Def. 6 implies $\frac{K_1}{m_1^+(x_0)} \geq \frac{K_2}{m_2^+(x_0)}$.*
2. *If $m_i^-(x_0) < \infty$ for both $i = 1, 2$, and $\frac{K_1}{m_1^+(x_0)} = \frac{K_2}{m_2^+(x_0)}$, then Part (i) of Def. 6 implies $\frac{m_1^-(x_0)}{m_1^+(x_0)} \geq \frac{m_2^-(x_0)}{m_2^+(x_0)}$*

Proof. For part 1, note that, by definition, $m_i(x) \cdot (x_0 - x) = u_i(x) - u_i^-(x_0)$, and hence Lemma 6 implies that there exists $\underline{x} < x_0$, such that $\forall x \in (\underline{x}, x_0)$, there exists $\bar{x} > x_0$, s.t, for all $x' \in (x_0, \bar{x})$: $\frac{K_1}{m_1(x')} + \frac{(u_1^-(x_0) - u_1(x_n))}{m_1(x')} > \frac{K_2}{m_2(x')} + \frac{(u_2^-(x_0) - u_2(x_n))}{m_2(x')}$, and hence we can construct a sequence $(x_n, \bar{x}_n, x'_n)_{n \in \mathbb{N}}$ converging to (x_0, x_0) such that, for every n, we have

$$\frac{K_1}{m_1(x')} + \frac{(u_1^-(x_0) - u_1(x_n))}{m_1(x')} > \frac{K_2}{m_2(x'_n)} + \frac{(u_2^-(x_0) - u_2(x_n))}{m_2(x'_n)}$$

Taking limits, and particularly for $x_n \to x_0$, since the $(u_i^-(x_0) - u_i(x_n))$ terms on both sides go to 0, and both the $u_i(\cdot)$ and $m_i(\cdot)$ functions are continuous on the relevant range, with $m_i(x') \to m_i^+$ (which, by assumption is non-zero), we obtain the condition in the statement: $\frac{K_1}{m_1^+(x_0)} \geq \frac{K_2}{m_2^+(x_0)}$.

For part 2, Lemma 6 again implies that there exists $\underline{x} < x_0$, such that $\forall x \in (\underline{x}, x_0)$, there exists $\bar{x} > x_0$, s.t, for all $x' \in (x_0, \bar{x})$:

$$\frac{K_1}{m_1(x')} + \frac{m_1(x)}{m_1(x')}(x_0 - x) > \frac{K_2}{m_2(x')} + \frac{m_2(x)}{m_2(x')}(x_0 - x). \tag{8}$$

Holding x fixed, and taking limits as $x' \to x_0^+$, this holds if:

$$\frac{K_1}{m_1^+(x_0)} + \frac{m_1(x)}{m_1^+(x_0)}(x_0 - x) > \frac{K_2}{m_2^+(x_0)} + \frac{m_2(x)}{m_2^+(x_0)}(x_0 - x)$$

and only if

$$\frac{K_1}{m_1^+(x_0)} + \frac{m_1(x)}{m_1^+(x_0)}(x_0 - x) \geq \frac{K_2}{m_2^+(x_0)} + \frac{m_2(x)}{m_2^+(x_0)}(x_0 - x).$$

Using the hypothesis $\frac{K_1}{m_1^+(x_0)} = \frac{K_2}{m_2^+(x_0)}$, and dividing everything by $(x_0 - x)$, this holds if $\frac{m_1(x)}{m_1^+(x_0)} > \frac{m_2(x)}{m_2^+(x_0)}$ and only if $\frac{m_1(x)}{m_1^+(x_0)} \geq \frac{m_2(x)}{m_2^+(x_0)}$. Hence, taking limits as $x \to x_0^-$, we have $\frac{m_1^-(x_0)}{m_1^+(x_0)} \geq \frac{m_2^-(x_0)}{m_2^+(x_0)}$.

Lemma 8 (Part (i): Sufficiency)**.** *If $m_i^+(x_0) > 0$ for both $i = 1, 2$, then:*

1. *$\frac{K_1}{m_1^+(x_0)} > \frac{K_2}{m_2^+(x_0)}$ implies that Part (i) of Def. 6 holds.*
2. *If $m_i^-(x_0) < \infty$ for both $i = 1, 2$, and $\frac{K_1}{m_1^+(x_0)} = \frac{K_2}{m_2^+(x_0)}$, then $\frac{m_1^-(x_0)}{m_1^+(x_0)} > \frac{m_1^-(x_0)}{m_2^+(x_0)}$ implies that Part (i) of Def. 6 holds.*

Proof. For Part 1, if $\frac{K_1}{m_1^+(x_0)} > \frac{K_2}{m_2^+(x_0)}$, then $\forall \varepsilon > 0$, $\exists \bar{x} > x_0$ s.t. $\frac{K_1}{m_1(x')} - \frac{K_2}{m_2(x')} > \varepsilon$ for all $x' \in (x_0, \bar{x})$. Hence, $\underline{x} < x_0$ can be chosen close enough to x_0 to ensure that $\left[\frac{m_1(x_n)}{m_1(x'_n)} - \frac{m_2(x_n)}{m_2(x'_n)}\right](x_0 - x_n) < \varepsilon$ for all $x \in (\underline{x}, x_0)$, satisfying the condition stated in Lemma 6.

For Part 2, if $\frac{m_1^-(x_0)}{m_1^+(x_0)} > \frac{m_2^-(x_0)}{m_2^+(x_0)}$, then $\exists \underline{x} < x_0$ s.t.

$$\frac{m_1(x)}{m_1^+(x_0)}(x_0 - x) > \frac{m_2(x)}{m_2^+(x_0)}(x_0 - x)$$

for all $x' \in (\underline{x}, x_0)$. Furthermore, since $\frac{K_1}{m_1^+(x_0)} = \frac{K_2}{m_2^+(x_0)}$, for any such $x \in (\underline{x}, x_0)$, $\exists \bar{x} > x_0$ s.t. $\forall x' \in (x_0, \bar{x})$

$$\frac{K_1}{m_1(x')} + \frac{m_1(x)}{m_1(x')}(x_0 - x) > \frac{K_2}{m_2(x')} + \frac{m_2(x)}{m_2(x')}(x_0 - x).$$

Then, the result follows from Lemma 6.

***** Part (ii): Characterization, Necessary and Sufficient conditions *****

We turn next to the characterization of part (ii) of Def. 6. As discussed above, this requires focusing on the ranking of the $\hat{p}_i(x, x')$. However, as shown by next lemma, $\hat{p}_i(x, x')$ is the same as $\bar{p}_i(x, x')$ when u_i is discontinuous at x_0:

Lemma 9. *If u_i is discontinuous, then $\hat{p}_i(x, x') = \bar{p}_i(x, x')$, and $E\bar{p}_i(x, x') > x_0$.*

Proof. As shown in Fig. 5,, p.17, the two numbers coincide in the discontinuous case merely because the existence of the certainty equivalent, common to both definitions, is the binding constraint for both notions when u is discontinuous.

Lemma 10. *If both u_1 and u_2 are discontinuous, the following are equivalent:*

1. *Part (i) of Def. 6 holds*
2. *Part (ii) of Def. 6 holds*
3. *There exists $\underline{x} < x_0$, such that $\forall x \in (\underline{x}, x_0)$, there exists $\bar{x} > x_0$, s.t, for all $x' \in (x_0, \bar{x})$:*

$$\frac{K_1}{m_1(x')} + \frac{m_1(x)}{m_1(x')}(x_0 - x) > \frac{K_2}{m_2(x')} + \frac{m_2(x)}{m_2(x')}(x_0 - x).$$

Proof. This follows from part 2 of Lemma 4, Lemma 9, and Lemma 6.

Lemma 11. *If u_1 is continuous and u_2 is discontinuous, then part (ii) of Def. part (ii) of Def. 6 does* not *hold.*

Proof. For any $x < x_0$ and $x' > x_0$, if u_1 is continuous then $E\hat{p}_1(x, x') < x_0$, and if u_2 is discontinuous implies $E\hat{p}_2(x, x') = E\bar{p}_2(x, x') > x_0$. It follows that $\hat{p}_1(x, x') < \hat{p}_2(x, x')$, which (by Lemma 4) implies that part (ii) of Def. 6 does not hold.

Lemma 12 (Part (ii): Summary of discontinuous case)**.** *The following holds:*

1. *If u_1 and u_2 are discontinuous, $m_i^+(x_0) > 0$ for both $i = 1, 2$, then part (ii) of Def. 6 holds if $\frac{K_1}{m_1^+} > \frac{K_2}{m_2^+}$, or if $\frac{K_1}{m_1^+} = \frac{K_2}{m_2^+} > 0$ and $\frac{m_1^-}{m_1^+} \geq \frac{m_2^-}{m_2^+}$; and only if either $\frac{K_1}{m_1^+} > \frac{K_2}{m_2^+}$, or $\frac{K_1}{m_1^+} = \frac{K_2}{m_2^+} > 0$ and $\frac{m_1^-}{m_1^+} \geq \frac{m_2^-}{m_2^+}$.*
2. *If u_1 is discontinuous and u_2 is continuous, then part (ii) of Def. 6 holds (and hence, by Lemma 10, point (i) holds as well).*
3. *If u_1 is continuous and u_2 is discontinuous, then part (ii) of Def. 6 does not hold.*

Proof. We consider each point separately:

1. For the ($\Rightarrow$) direction, note that if u_1 and u_2 are discontinuous, part (ii) of Def. 6 implies part (i) of Def. 6, and hence $\frac{K_1}{m_1^+} \geq \frac{K_2}{m_2^+}$ by Lemma 7. This of course can either be $\frac{K_1}{m_1^+} > \frac{K_2}{m_2^+}$ or $\frac{K_1}{m_1^+} = \frac{K_2}{m_2^+}$. If the latter, it also needs to satisfy $\forall x \in (\underline{x}, x_0)$, there exists $\bar{x} > x_0$, s.t, for all $x' \in (x_0, \bar{x})$: $\frac{K_1}{m_1(x')} + \frac{m_1(x)}{m_1(x')}(x_0 - x) > \frac{K_2}{m_2(x')} + \frac{m_2(x)}{m_2(x')}(x_0 - x)$, or it would contradict Lemma 10 (particularly, the fact that point 2 implies point 3). It then follows from continuity of u aside from at x_0 that in the limit, $\frac{K_1}{m_1^+} + \frac{m_1^-}{m_1^+}(x_0 - x) \geq \frac{K_2}{m_2^+} + \frac{m_2^-}{m_2^+}(x_0 - x)$, and hence that $\frac{m_1^-}{m_1^+} \geq \frac{m_2^-}{m_2^+}$.

   For the ($\Leftarrow$) direction, in the case in which $\frac{K_1}{m_1^+(x_0)} > \frac{K_2}{m_2^+(x_0)}$ holds, the result follows from Lemma 8. In the other case, it follows from $\frac{m_1^-}{m_1^+} > \frac{m_2^-}{m_2^+}$ and the continuity of u aside from at x_0 that there exists $\underline{x} < x_0$, such that $\forall x \in (\underline{x}, x_0)$, there exists $\bar{x} > x_0$ s.t. $\frac{K_1}{m_1(x')} + \frac{m_1(x)}{m_1(x')}(x_0 - x) > \frac{K_2}{m_2(x')} + \frac{m_2(x)}{m_2(x')}(x_0 - x)$. The statement then follows from lemma 10 (particularly, from the fact that point 3 implies point 2).

2. If u_1 is discontinuous and u_2 is continuous, then for any $x < x_0$ and $x' > x_0$, we have $E\hat{p}_1 > x_0 > E\hat{p}_2$, which implies $\hat{p}_1 > \hat{p}_2$, and hence the result follows from the second part of Lemma 4.
3. This is just the statement of Lemma 11.

Part 1 of this lemma concludes the proof of parts 1 and 2 of the Theorem. We next focus on the rest of the proof of part 3 of the theorem. Hence, to complete the characterization of Part (ii) of Def. 6, we need to characterize the condition in part 2 of Lemma 4.

Step 1: First note that, if both u_1 and u_2 are continuous, then $K_1 = K_2 = 0$. In this case, by continuity, $CE_i(p)$ exists for all p, and hence, letting $E\hat{p}_i(x,x') = \hat{p}_i(x,x') \cdot x' + (1 - \hat{p}_i(x,x')) \cdot x$, the cutoff probability $\hat{p}_i(x,x')$ can be written in implicit form as:

$$E\hat{p}_i(x,x') = u_i^{-1}\left(E\hat{p}_i(x,x')\right), \tag{9}$$

From (9), and from the definition of $m_i(x)$, the continuity of u_i implies the following:

$$lim_{x' \to x_0^+} E\hat{p}_i(x,x') = x_0^-, \tag{10}$$

$$lim_{x' \to x_0^+} \hat{p}_i(x,x') = 1, \tag{11}$$

$$lim_{x' \to x_0^+} m_i\left(E\hat{p}_i(x,x')\right) = m_i^-(x_0), \tag{12}$$

$$lim_{x' \to x_0^+} m_i(x') = m_i^+(x_0). \tag{13}$$

With these results in hand, we proceed to the next Lemma.

Lemma 13. *Under the maintained assumptions of the representation theorem, if u_i is continuous, then:*

$$\left(\frac{x_0 - E\hat{p}_i}{x' - E\hat{p}_i}\right)\left(\frac{(x'-x)}{(x_0-x)}\right) = \frac{m_i(x) - m_i(x')}{\left[m_i(E\hat{p}_i)\frac{(x_0-x)}{(x'-x)} - m_i(x')\right]}. \tag{14}$$

Proof. To simplify notation, in the following we will write $\hat{p}_i$ instead of $\hat{p}_i(x,x')$, with the understanding however that $\hat{p}_i$ should still be regarded as a function of x'. From (9), it is easy to see that for any $x < x_0$ and $x' > x_0$, $\hat{p}_i$ satisfies the following condition:

$$m_i(E\hat{p}_i) \cdot \frac{x_0 - E\hat{p}_i}{x' - E\hat{p}_i} + m_i(x') \cdot \frac{x' - x_0}{x' - E\hat{p}_i} = m_i(x) \cdot \frac{x_0 - x}{x' - x} + m_i(x') \cdot \frac{x' - x_0}{x' - x} \tag{15}$$

adding and subtracting $m_i(x') \cdot \frac{x_0 - E\hat{p}_i}{x' - E\hat{p}_i}$ from the LHS, and rearranging terms, we obtain:

$$\left(\frac{x_0 - E\hat{p}_i}{x' - E\hat{p}_i}\right)\left[m_i(E\hat{p}_i) - m_i(x')\right] + m_i(x')\left(\frac{x_0 - E\hat{p}_i + x' - x_0}{x' - E\hat{p}_i}\right) = m_i(x)\left(\frac{x_0 - x}{x' - x}\right) + m_i(x')\left(\frac{x' - x_0}{x' - x}\right)$$

$$\left(\frac{x_0 - E\hat{p}_i}{x' - E\hat{p}_i}\right)\left[m_i(E\hat{p}_i) - m_i(x')\right] + m_i(x')\left(\frac{x' - E\hat{p}_i}{x' - E\hat{p}_i}\right) = m_i(x)\left(\frac{x_0 - x}{x' - x}\right) + m_i(x') \cdot \frac{x' - x_0}{x' - x}$$

$$\left(\frac{x_0 - E\hat{p}_i}{x' - E\hat{p}_i}\right)\left[m_i(E\hat{p}_i) - m_i(x')\right] = m_i(x) \cdot \frac{x_0 - x}{x' - x} + m_i(x')\left[\frac{x' - x_0}{x' - x} - 1\right]$$

$$\left(\frac{x_0 - E\hat{p}_i}{x' - E\hat{p}_i}\right)\left[m_i(E\hat{p}_i) - m_i(x')\right] = m_i(x) \cdot \frac{x_0 - x}{x' - x} - m_i(x')\left[\frac{x - x_0}{x' - x}\right]$$

$$\left(\frac{x_0 - E\hat{p}_i}{x' - E\hat{p}_i}\right)\left(\frac{(x'-x)}{(x_0-x)}\right)\left[m_i(E\hat{p}_i)\frac{(x_0-x)}{(x'-x)} - m_i(x')\right] = m_i(x) - m_i(x')$$

$$\left(\frac{x_0 - E\hat{p}_i}{x' - E\hat{p}_i}\right)\left(\frac{(x'-x)}{(x_0-x)}\right) = \frac{m_i(x) - m_i(x')}{\left[m_i(E\hat{p}_i)\frac{(x_0-x)}{(x'-x)} - m_i(x')\right]}$$

Lemma 14. *If both u_1 and u_2 are continuous, $\hat{p}_1(x,x') > \hat{p}_2(x,x')$ if and only if*

$$\frac{[m_1(E\hat{p}_1) - m_1(x)]}{m_1(x')} - \frac{[m_2(E\hat{p}_1) - m_2(x)]}{m_2(x')} > \left(\frac{m_1(E\hat{p}_1)}{m_1(x')} - \frac{m_2(E\hat{p}_1)}{m_2(x')}\right)\left[1 - \left(\frac{x_0 - E\hat{p}_1}{x' - E\hat{p}_1}\right)\left(\frac{x'-x}{x_0-x}\right)\right].$$

Proof. Using the characterization of $\hat{p}_i(x,x')$ in (15), we have:

$$m_1(x)\left(\frac{x_0-x}{x'-x}\right)+m_1(x')\left(\frac{x'-x_0}{x'-x}\right)=m_1(E\hat{p}_1)\left(\frac{x_0-E\hat{p}_1}{x'-E\hat{p}_1}\right)+m_1(x')\left(\frac{x'-x_0}{x'-E\hat{p}_1}\right),\text{ or}$$

$$m_1(x')\left[\left(\frac{x'-x_0}{x'-x}\right)-\left(\frac{x'-x_0}{x'-E\hat{p}_1}\right)\right]=\left[m_1(E\hat{p}_1)-m_1(x)\right]\left(\frac{x_0-x}{x'-x}\right)-m_1(E\hat{p}_1)\left[\left(\frac{x_0-x}{x'-x}\right)-\left(\frac{x_0-E\hat{p}_1}{x'-E\hat{p}_1}\right)\right] \quad (16)$$

and, for $\alpha=\frac{m_1(x')}{m_2(x')}$, we also have that $\hat{p}_1(x,x')>\hat{p}_2(x,x')$ if and only if

$$\alpha m_2(x)\left(\frac{x_0-x}{x'-x}\right)+\alpha m_2(x')\left(\frac{x'-x_0}{x'-x}\right)>\alpha m_2(E\hat{p}_1)\left(\frac{x_0-E\hat{p}_1}{x'-E\hat{p}_1}\right)+\alpha m_2(x')\left(\frac{x'-x_0}{x'-E\hat{p}_1}\right),\text{ or}$$

$$\alpha m_2(x')\left[\left(\frac{x'-x_0}{x'-x}\right)-\left(\frac{x'-x_0}{x'-E\hat{p}_1}\right)\right]>\left[\alpha m_2(E\hat{p}_1)-\alpha m_2(x)\right]\left(\frac{x_0-x}{x'-x}\right)-\alpha m_2(E\hat{p}_1)\left[\left(\frac{x_0-x}{x'-x}\right)-\left(\frac{x_0-E\hat{p}_1}{x'-E\hat{p}_1}\right)\right] \quad (17)$$

Notice that, by the definition of α, the LHS of equation (16) is the same as the LHS of (17). Hence, equalizing both sides we obtain that $\hat{p}_1(x,x')>\hat{p}_2(x,x')$ if and only if

$$\begin{aligned}&\left[m_1(E\hat{p}_1)-m_1(x)\right]\left(\frac{x_0-x}{x'-x}\right)-m_1(E\hat{p}_1)\left[\left(\frac{x_0-x}{x'-x}\right)-\left(\frac{x_0-E\hat{p}_1}{x'-E\hat{p}_1}\right)\right]\\&>\alpha\left[m_2(E\hat{p}_1)-m_2(x)\right]\left(\frac{x_0-x}{x'-x}\right)-\alpha m_2(E\hat{p}_1)\left[\left(\frac{x_0-x}{x'-x}\right)-\left(\frac{x_0-E\hat{p}_1}{x'-E\hat{p}_1}\right)\right]\end{aligned}$$

re-arranging, substituting for $\alpha=\frac{m_1(x')}{m_2(x')}$, and dividing everything by $m_1(x')$, this is equivalent to:

$$\frac{\left[m_1(E\hat{p}_1)-m_1(x)\right]}{m_1(x')}-\frac{\left[m_2(E\hat{p}_1)-m_2(x)\right]}{m_2(x')}>\left(\frac{m_1(E\hat{p}_1)}{m_1(x')}-\frac{m_2(E\hat{p}_1)}{m_2(x')}\right)\left[1-\left(\frac{x_0-E\hat{p}_1}{x'-E\hat{p}_1}\right)\left(\frac{x'-x}{x_0-x}\right)\right].$$

Lemma 15. *If both u_1 and u_2 are continuous, $\hat{p}_1(x,x')>\hat{p}_2(x,x')$ if and only if*

$$\frac{m_1(y)-m_1(x)}{m_2(y)-\beta m_2(x)}>\frac{m_1(y)-m_1(x')}{m_2(y)-m_2(x')}+(1-\beta)\gamma(x,x',y), \quad (18)$$

where $\gamma(x,x',y)=\frac{m_1(y)m_1(x)m_2(x')-m_2(x)m_1(x'^2)}{m_1(x')(m_2(y)-m_2(x'))(m_2(y)-\beta m_2(x))}$.

Proof. Let $\beta:=\frac{x_0-x}{x'-x}$, and note that $\beta\in(0,1)$ and $\beta\to1$ as $x'\to x_0$. Also let $y=E\hat{p}_1$, and note that $y\to x_0$ as $x'\to x_0$ (these facts will be useful in the lemmas that follow). Then, from Lemma 13, we have that:

$$\left(\frac{x_0-y}{x'-y}\right)\left(\frac{(x'-x)}{(x_0-x)}\right)=\frac{m_1(x)-m_1(x')}{\beta m_1(y)-m_1(x')}. \quad (19)$$

Substituting this in the condition of Lemma 14, using eq. (19), and re-arranging, we obtain the following (details are in the online appendix):

$$m_1(y)\left(\beta m_2(x)-m_2(x')\right)-m_1(x')m_2(x)+\frac{(1-\beta)m_1(x)m_1(y)m_2(x')}{m_1(x')}>m_2(y)(m_1(x)-m_1(x'))-m_1(x)m_2(x'). \quad (20)$$

rearranging now Equation 28 (and writing γ rather than $\gamma(x,x',y)$, we have:

$$\begin{aligned}&m_1(y)m_2(y)-m_1(y)m_2(x')-m_1(x)m_2(y)+m_1(x)m_2(x')>\\&\quad m_1(y)m_2(y)-\beta m_1(y)m_2(x)-m_1(x')m_2(y)+\beta m_1(x')m_2(x)-\gamma(1-\beta)\left(m_2(y)-m_2(x')\right)\left(m_2(y)-\beta m_2(x)\right)\\&\qquad\Longleftrightarrow\\&\quad m_1(y)\left(\beta m_2(x)-m_2(x')\right)-\beta m_1(x')m_2(x)>\\&\qquad m_2(y)\left(m_1(x)-m_1(x')\right)-m_1(x)m_2(x')-\gamma(1-\beta)(m_2(y)-m_2(x'))(m_2(y)-\beta m_2(x)).\end{aligned}$$

Using that $-\beta m_1(x')m_2(x)=(1-\beta)m_1(x')m_2(x)-m_1(x')m_2(x)$, we obtain:

$$m_1(y)\left(\beta m_2(x)-m_2(x')\right)-m_1(x')m_2(x)+\left[\gamma(1-\beta)(m_2(y)-m_2(x'))(m_2(y)-\beta m_2(x))+(1-\beta)m_1(x')m_2(x)\right]>$$
$$m_2(y)\left(m_1(x)-m_1(x')\right)-m_1(x)m_2(x'). \tag{21}$$

For Inequality 20 to hold if and only Inequality 21 holds, it must be that (the details of the algebraic manipulations are in the online appendix):

$$\gamma=\frac{m_1(y)m_1(x)m_2(x')-m_2(x)m_1(x'^2)}{m_1(x')\left(m_2(y)-m_2(x')\right)\left(m_2(y)-\beta m_2(x)\right)}, \tag{22}$$

which concludes the proof of the lemma.

Lemma 16. *[Part (ii): Necessity]If both u_1 and u_2 are continuous, and if $m_i^-<\infty$ for both $i=1,2$, then Part (ii) of Def.6 implies*

$$lim_{x\to x_0^-}\frac{[m_1^- - m_1(x)]/m_1^-}{[m_2^- - m_2(x)]/m_2^-}\geq\frac{1-m_1^+/m_1^-}{1-m_2^+/m_2^-}.$$

Proof. Part 2 of Lemma 4 and Lemma 15 imply that, if part (ii) of Def.6 holds, then there exists $\underline{x}<x_0$, such that $\forall x\in(\underline{x},x_0)$, there exists $\bar{x}>x_0$, s.t., for all $x'\in(x_0,\bar{x})$,

$$\frac{m_1(E\hat{p}_1(x,x'))-m_1(x)}{m_2(E\hat{p}_1(x,x'))-\beta m_2(x)}>\frac{m_1(E\hat{p}_1(x,x'))-m_1(x')}{m_2(E\hat{p}_1(x,x'))-m_2(x')}+\left[1-\beta(x,x')\right]\gamma(x,x'), \tag{23}$$

where $\beta=\frac{x_0-x}{x'-x}$, and $\gamma(x,x')=\frac{m_1(E\hat{p}_1(x,x'))m_1(x)m_2(x')-m_2(x)m_1(x'^2)}{m_1(x')(m_2(E\hat{p}_1(x,x'))-m_2(x'))(m_2(E\hat{p}_1(x,x'))-\beta(x,x')m_2(x))}$. For any such $x\in(\underline{x},x_0)$, taking limits as $x'\to 0^+$, and using the limits in equations (10)-(13), the condition in (23) converges to the following:

$$\frac{m_1^-(x_0)-m_1(x)}{m_2^-(x_0)-m_2(x)}\geq\frac{m_1^-(x_0)-m_1^+(x_0)}{m_2^-(x_0)-m_2^+(x_0)},$$

dividing both sides by m_1^-/m_2^-, this yields

$$\frac{\left[m_1^-(x_0)-m_1(x)\right]/m_1^-(x_0)}{\left[m_2^-(x_0)-m_2(x)\right]/m_2^-(x_0)}\geq\frac{1-m_1^+(x_0)/m_1^-(x_0)}{1-m_2^+(x_0)/m_1^-(x_0)}.$$

Since this needs to hold for all $x\in(\underline{x},x_0)$, it also holds for $x\to x_0^-$, i.e.

$$\left(lim_{x\to x_0^-}\frac{[m_1^- - m_1(x)]/m_1^-}{[m_2^- - m_2(x)]/m_2^-}\right)\geq\frac{1-m_1^+/m_1^-}{1-m_2^+/m_2^-},$$

which completes the proof of the Lemma.

Lemma 17. *[Part (ii): Sufficiency]If both u_1 and u_2 are continuous, and if $m_i^-<\infty$ for both $i=1,2$, then*

$$\left(lim_{x\to x_0^-}\frac{[m_1^- - m_1(x)]/m_1^-}{[m_2^- - m_2(x)]/m_2^-}\right)>\frac{1-m_1^+/m_1^-}{1-m_2^+/m_2^-}$$

implies that Part (ii) of Def.6 holds

Proof. If $\left(lim_{x\to x_0^-}\frac{[m_1^- - m_1(x)]/m_1^-}{[m_2^- - m_2(x)]/m_2^-}\right)>\frac{1-m_1^+/m_1^-}{1-m_2^+/m_2^-}$, then there exists $\underline{x}<x_0$, such that $\forall x\in(\underline{x},x_0)$,

$$\frac{\left[m_1^-(x_0)-m_1(x)\right]/m_1^-(x_0)}{\left[m_2^-(x_0)-m_2(x)\right]/m_2^-(x_0)}\geq\frac{1-m_1^+(x_0)/m_1^-(x_0)}{1-m_2^+(x_0)/m_1^-(x_0)}, \text{ that is}$$
$$\frac{m_1^-(x_0)-m_1(x)}{m_2^-(x_0)-m_2(x)}\geq\frac{m_1^-(x_0)-m_1^+(x_0)}{m_2^-(x_0)-m_2^+(x_0)}$$

But since functions $\beta(\cdot)$, $\gamma(\cdot)$, $m_i(\cdot)$ and $E\hat{p}_1(\cdot)$ are all continuous in x', for any such $\forall x \in (\underline{x}, x_0)$, there exists $\bar{x} > x_0$, s.t., for all $x' \in (x_0, \bar{x})$,

$$\frac{m_1(E\hat{p}_1(x,x')) - m_1(x)}{m_2(E\hat{p}_1(x,x')) - \beta m_2(x)} > \frac{m_1(E\hat{p}_1(x,x')) - m_1(x')}{m_2(E\hat{p}_1(x,x')) - m_2(x')} + \left[1 - \beta(x,x')\right]\gamma(x,x'). \tag{24}$$

The result then follows from Lemma 15.

Lemmas 4, 16 and 17, together with Lemma 6 (noting that for $K_1 = K_2 = 0$, the expression in the lemma reduces to $\frac{m_1(x)}{m_1(x')} > \frac{m_2(x)}{m_2(x')}$, which by continuity holds if $\frac{m_1^-}{m_1^+} > \frac{m_2^-}{m_2^+}$, and only if $\frac{m_1^-}{m_1^+} \geq \frac{m_2^-}{m_2^+}$), prove the theorem.

Proof of Theorem 6.

Proof. Note that with differentiability, $Du_i^+ = m_i^+$, $Du_i^- = m_i^-$, and that $D^2u_i^- = \lim_{x \to x_0^-} \frac{m_i^- - m_i(x)}{x_0 - x}$ for $i = \{1, 2\}$. Parts 1 and 2 of the theorem therefore follow directly from parts 1 and 2, respectively, of Theorem 5. Part 3 follows from the following:

$$lim_{x \to x_0^-} \frac{[m_1^- - m_1(x)]/m_1^-}{[m_2^- - m_2(x)]/m_2^-} \geq \frac{1 - m_1^+/m_1^-}{1 - m_2^+/m_2^-} \Longleftrightarrow \qquad lim_{x \to x_0^-} \frac{[\frac{m_1^- - m_1(x)}{x_0 - x}]/m_1^-}{[\frac{m_2^- - m_2(x)}{x_0 - x}]/m_2^-} \geq \frac{(m_1^- - m_1^+)/m_1^-}{(m_2^- - m_2^+)/m_2^-}$$

$$\Longleftrightarrow$$

$$\frac{D^2u_1^-}{D^2u^+} \geq \frac{Du_1^- - Du_1^+}{Du_2^- - Du_2^+} \Longleftrightarrow \qquad \frac{D^2u_1^-}{Du_1^- - Du_1^+} \geq \frac{D^2u_2^-}{Du_2^- - Du_2^+}.$$

Proof of Lemma 2, Lemma 3 and Proposition 1.

Proof. Lemma 3 follows from Lemma 16 and 17. Then, using Lemma 6 from the proof of Theorem 5, and using that $K_1 = K_2 = 0$ since both u_1, u_2 are continuous, we have

$$S_1(x,x') \subset S_2(x,x') \Leftrightarrow \bar{p}_1(x,x') > \bar{p}_2(x,x') \Leftrightarrow \frac{m_1(x)}{m_1(x')}(x_0 - x) > \frac{m_2(x)}{m_2(x')}(x_0 - x).$$

dividing both sides by $(x_0 - x)$, this holds if and only if $\frac{m_1(x)}{m_1(x')} > \frac{m_2(x)}{m_2(x')}$.

Holding x fixed, and taking limits as $x' \to x_0^+$, this holds if $\frac{m_1(x)}{m_1^+(x_0)} > \frac{m_2(x)}{m_2^+(x_0)}$ and only if $\frac{m_1(x)}{m_1^+(x_0)} \geq \frac{m_2(x)}{m_2^+(x_0)}$. Taking now limits as $x \to x_0^-$, this holds if $\frac{m_1^-(x_0)}{m_1^+(x_0)} > \frac{m_2^-(x_0)}{m_2^+(x_0)}$ and only if $\frac{m_1^-(x_0)}{m_1^+(x_0)} > \frac{m_2^-(x_0)}{m_2^+(x_0)}$, which proves Lemma 2.

To prove Proposition 1, and given Lemma 2, it suffices to also prove that this results holds when the order of limits is reversed. Specifically, starting with $\frac{m_1(x)}{m_1(x')} > \frac{m_2(x)}{m_2(x')}$, holding x' fixed and taking the limits as $x \to x_0^-$, this holds if $\frac{m_1^-(x_0)}{m_1(x')} > \frac{m_2^-(x_0)}{m_2(x')}$ and only if $\frac{m_1^-(x_0)}{m_1(x')} \geq \frac{m_2^-(x_0)}{m_2(x')}$. Taking now limits as $x' \to x_0^+$, this holds if $\frac{m_1^-(x_0)}{m_1^+(x_0)} > \frac{m_2^-(x_0)}{m_2^+(x_0)}$ and only if $\frac{m_1^-(x_0)}{m_1^+(x_0)} > \frac{m_2^-(x_0)}{m_2^+(x_0)}$, which concludes the proof of Proposition 1.

References

Abdellaoui, M. (2000): "Parameter-free elicitation of utility and probability weighting functions," *Management science*, 46, 1497–1512.

Abdellaoui, M., H. Bleichrodt, and C. Paraschiv (2007): "Loss aversion under prospect theory: A parameter-free measurement," *Management science*, 53, 1659–1674.

Alaoui, L. and C. Fons-Rosen (2021): "Know when to fold'em: The flip side of grit," *European Economic Review*, 136, 103736.

Almlund, M., A. L. Duckworth, J. Heckman, and T. Kautz (2011): "Personality psychology and economics," in *Handbook of the Economics of Education*, Elsevier, vol. 4, 1–181.

Appadurai, A. (2004): "The Capacity to Aspire: Culture and the Terms of Recognition," *Culture and Public Action*.

Baillon, A., H. Bleichrodt, and V. Spinu (2020): "Searching for the reference point," *Management Science*, 66, 93–112.

Bernheim, B. D. and C. Sprenger (2020): "On the empirical validity of cumulative prospect theory: Experimental evidence of rank-independent probability weighting," *Econometrica*, 88, 1363–1409.

Bleichrodt, H. and J. L. Pinto (2000): "A parameter-free elicitation of the probability weighting function in medical decision analysis," *Management science*, 46, 1485–1496.

Burks, S. V., C. Lewis, P. A. Kivi, A. Wiener, J. E. Anderson, L. Götte, C. G. DeYoung, and A. Rustichini (2015): "Cognitive skills, personality, and economic preferences in collegiate success," *Journal of Economic Behavior & Organization*, 115, 30–44.

Camerer, C., L. Babcock, G. Loewenstein, and R. Thaler (1997): "Labor supply of New York City cabdrivers: One day at a time," *The Quarterly Journal of Economics*, 112, 407–441.

Cerreia-Vioglio, S., D. Dillenberger, and P. Ortoleva (2022): "Caution and Reference Effects," Tech. rep., Working Paper.

Cerulli-Harms, A., L. Goette, and C. Sprenger (2019): "Randomizing endowments: An experimental study of rational expectations and reference-dependent preferences," *American Economic Journal: Microeconomics*, 11, 185–207.

Chevalier, J. and G. Ellison (1999): "Career concerns of mutual fund managers," *The Quarterly Journal of Economics*, 114, 389–432.

Choi, S., R. Fisman, D. Gale, and S. Kariv (2007): "Consistency and heterogeneity of individual behavior under uncertainty," *American economic review*, 97, 1921–1938.

Crawford, V. P. and J. Meng (2011): "New York City cab drivers' labor supply revisited: Reference-dependent preferences with rational-expectations targets for hours and income," *American Economic Review*, 101, 1912–1932.

Deary, I. J., G. Matthews, and M. C. Whiteman. (2009): *Personality Traits*, New York: Cambridge University Press, 3rd ed.

Diecidue, E., M. Levy, and J. van de Ven (2015): "No aspiration to win? An experimental test of the aspiration level model," *Journal of Risk and Uncertainty*, 51, 245–266.

Diecidue, E. and J. Van De Ven (2008): "Aspiration level, probability of success and failure, and expected utility," *International Economic Review*, 49, 683–700.

Duckworth, A. L. and P. D. Quinn (2009): "Development and validation of the Short Grit Scale (GRIT–S)," *Journal of personality assessment*, 91, 166–174.

Ellis, K., S. Kariv, and E. Ozbay (2022): "What can the demand analyst learn from machine learning?" Tech. rep., Working Paper.

Fishburn, P. C. (1977): "Mean-risk analysis with risk associated with below-target returns," *The American Economic Review*, 67, 116–126.

Friedman, M. and L. J. Savage (1948): "The utility analysis of choices involving risk," *Journal of political Economy*, 56, 279–304.

Genicot, G. and D. Ray (2017): "Aspirations and inequality," *Econometrica*, 85, 489–519.

——— (2020): "Aspirations and economic behavior," *Annual Review of Economics*, 12, 715–746.

Gill, D. and V. Prowse (2016): "Cognitive ability, character skills, and learning to play equilibrium: A level-k analysis," *Journal of Political Economy*, 124, 1619–1676.

Gill, D. and V. L. Prowse (2021): "The Creativity Premium," *Available at SSRN 3855808*.

Heckman, J. J., T. Jagelka, and T. Kautz (2021): "Some contributions of economics to the study of personality," in *Handbook of Personality: Theory and Research*, ed. by O. P. John and R. W. Robins, The Guilford Press, 853–892, 4th ed.

Heckman, J. J. and Y. Rubinstein (2001): "The importance of noncognitive skills: Lessons from the GED testing program," *American economic review*, 91, 145–149.

Imas, A. (2016): "The realization effect: Risk-taking after realized versus paper losses," *American Economic Review*, 106, 2086–2109.

Imas, A., S. Sadoff, and A. Samek (2017): "Do people anticipate loss aversion?" *Management Science*, 63, 1271–1284.

Jagelka, T. (2024): "Are economists' preferences psychologists' personality traits? A structural approach," *Journal of Political Economy*, 132, 910–970.

Kahneman, D. and A. Tversky (1979): "Prospect theory: An analysis of decision under risk," *Econometrica*, 47, 363–391.

——— (1992): "Advances in prospect theory: Cumulative representation of uncertainty," *Journal of Risk and uncertainty*, 5, 297–323.

Köbberling, V. and P. P. Wakker (2005): "An index of loss aversion," *Journal of Economic Theory*, 122, 119–131.

Kőszegi, B. and M. Rabin (2006): "A model of reference-dependent preferences," *The Quarterly Journal of Economics*, 121, 1133–1165.

——— (2007): "Reference-dependent risk attitudes," *American Economic Review*, 97, 1047–1073.

Kibris, O., Y. Masatlioglu, and E. Suleymanov (2023): "A theory of reference point formation," *Economic Theory*, 75, 137–166.

Markle, A., G. Wu, R. White, and A. Sackett (2018): "Goals as reference points in marathon running: A novel test of reference dependence," *Journal of Risk and Uncertainty*, 56, 19–50.

Markowitz, H. (1952): "The utility of wealth," *Journal of political Economy*, 60, 151–158.

Masatlioglu, Y. and E. A. Ok (2005): "Rational choice with status quo bias," *Journal of economic theory*, 121, 1–29.

——— (2014): "A canonical model of choice with initial endowments," *Review of Economic Studies*, 81, 851–883.

Masatlioglu, Y. and C. Raymond (2016): "A behavioral analysis of stochastic reference dependence," *American Economic Review*, 106, 2760–2782.

Mezias, S. J. (1988): "Aspiration level effects: An empirical investigation," *Journal of Economic Behavior & Organization*, 10, 389–400.

Minardi, S., F. Wang, and I. Gilboa (2024): "Consumption of Values," *Management Science*.

Neilson, W. S. (2002): "Comparative risk sensitivity with reference-dependent preferences," *Journal of Risk and Uncertainty*, 24, 131–142.

Netzer, N., A. J. Robson, J. Steiner, and P. Kocourek (2022): "Endogenous risk attitudes," .

O'Donoghue, T. and C. Sprenger (2018): "Reference-dependent preferences," in *Handbook of behavioral economics: Applications and foundations 1*, Elsevier, vol. 1, 1–77.

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

Payne, J. W., D. J. Laughhunn, and R. Crum (1980): "Translation of gambles and aspiration level effects in risky choice behavior," *Management Science*, 26, 1039–1060.

——— (1981): "Further Tests of Aspiration Level Effects in Risky Choice Behavior, 27 MGMT," .

Proto, E. and A. Rustichini (2015): "Life satisfaction, income and personality," *Journal of Economic Psychology*, 48, 17–32.

Proto, E., A. Rustichini, and A. Sofianos (2019): "Intelligence, personality, and gains from cooperation in repeated interactions," *Journal of Political Economy*, 127, 1351–1390.

——— (2022): "Intelligence, errors, and cooperation in repeated interactions," *The Review of Economic Studies*, 89, 2723–2767.

Rabin, M. (2000): "Risk Aversion and Expected-Utility Theory: A Calibration Theorem," *Econometrica*, 68, 1281–1292.

Rabin, M. and R. H. Thaler (2001): "Anomalies: risk aversion," *Journal of Economic perspectives*, 15, 219–232.

Ray, D. (1998): *Development economics*, Princeton University Press.

——— (2006): "Aspirations, poverty, and economic change," *Understanding poverty*, 1, 409–421.

Rosato, A. (2023): "Loss aversion in sequential auctions," *Theoretical Economics*, 18, 561–596.

Rosato, A. and A. A. Tymula (2019): "Loss aversion and competition in Vickrey auctions: Money ain't no good," *Games and Economic Behavior*, 115, 188–208.

Rozen, K. (2010): "Foundations of intrinsic habit formation," *Econometrica*, 78, 1341–1373.

Rubinstein, A. (2002): "Comments on the risk and time preferences in economics," *Working Paper*.

Schmidt, U. and S. Traub (2002): "An experimental test of loss aversion," *Journal of risk and Uncertainty*, 25, 233–249.

Schmidt, U. and H. Zank (2012): "A genuine foundation for prospect theory," *Journal of Risk and Uncertainty*, 45, 97–113.

Shalev, J. (1997): "Loss aversion in a multi-period model," *Mathematical Social Sciences*, 33, 203–226.

Steiner, J. and C. Stewart (2016): "Perceiving prospects properly," *American Economic Review*, 106, 1601–1631.

Swalm, R. O. (1966): "Utility theory-insights into risk taking," *Harvard business review*, 44, 123–136.

Venkatraman, V., J. W. Payne, J. R. Bettman, M. F. Luce, and S. A. Huettel (2009): "Separate neural mechanisms underlie choices and strategic preferences in risky decision making," *Neuron*, 62, 593–602.

Venkatraman, V., J. W. Payne, and S. A. Huettel (2014): "An overall probability of winning heuristic for complex risky decisions: Choice and eye fixation evidence," *Organizational Behavior and Human Decision Processes*, 125, 73–87.

von Neumann, J. and O. Morgenstern (1944): "Theory of Games and Economic Behavior (Princeton University, Princeton, NJ)," .

Wakker, P. and A. Tversky (1993): "An axiomatization of cumulative prospect theory," *Journal of risk and uncertainty*, 7, 147–175.

Wakker, P. P. (2010): *Prospect theory: For risk and ambiguity*, Cambridge university press.

Wakker, P. P. and H. Zank (2002): "A simple preference foundation of cumulative prospect theory with power utility," *European Economic Review*, 46, 1253–1271.

C Online Appendix

C.1 Ordering the Remaining Attitudes

An analogous exercise to that of ordering failure avoidance can be conducted for all the remaining attitude. Much of the reasoning above carries through, *mutatis mutandis*, to the definitions and results of these attitudes, as we now discuss.

Concerning success attachment, the first requirement will be that the more success-attaching agent will have a smaller set of lotteries that he regards as net *failures*, in the sense of being worse than the certain x_0. In the continuous case, this would be identical to saying that there is a larger set of lotteries that he regards as net successes, and so the first requirement is simply the reverse of that for failure avoidance. But, as discussed above, for the discontinuous case it is important to account for the existence of certainty equivalents, and hence the notion of *net failure* is adequately captured by a set, $\mathcal{F}_i(x,x')$, whose definition is specular to $\mathcal{S}_i(x,x')$ above:

$$\mathcal{F}_i(x,x') := cl\left\{p \in \Delta(x,x') : CE_i(p) \text{ exists and } CE_i(p) < x_0\right\} \tag{25}$$

Similarly, we define the set $RL_i(x,x')$ of lotteries over which success attachment is *not* manifested, symmetrically to the $RA_i(x,x')$ sets above:

$$RL_i(x,x') := cl\left\{p \in \Delta(x,x') : CE_i(p) \text{ exists and } CE_i(p) > Ep\right\}. \tag{26}$$

The ranking over success attachment is thus defined as follows:

Definition 7. *Let preferences $\succsim_1$ and $\succsim_2$ both satisfy the conditions in Def.3 with respect to the same $x_0 \in \mathbb{R}$. Then, $\succsim_1$ displays (weakly)* more success attachment than *$\succsim_2$ if there exist $x_f, x_s : x_f < x_0 < x_s : \forall x' \in (x_0, x_s]$, $\exists \underline{x} \in [x_f, x_0)$ such that, for each $x \in [\underline{x}, x_0)$, both the following conditions are satisfied: (i) $\mathcal{F}_1(x,x') \subseteq \mathcal{F}_2(x,x')$, and (ii) $RL_1(x,x') \subseteq RL_2(x,x')$.*

Analogous of Theorems 5 and 6 hold for this definition too. Here we only reproduce the statement of the differentiable case, which is easier to read and most useful in applications:

Theorem 7 (Success Attachment: Interpersonal Comparisons)**.** *Suppose that $(\succsim_i)_{i=1,2}$ are such that $Du_i^- > 0$ and $Du_i^+ < \infty$ and u_i is twice differentiable in some left- and right-neighborhoods of x_0. Then: $\succsim_1$ displays* more success avoidance than *$\succsim_2$* ***only if*** *one of the following applies:*

1. $\frac{K_1}{Du_1^-} > \frac{K_2}{Du_2^-}$,

2. $\frac{K_1}{Du_1^-} = \frac{K_2}{Du_2^-} > 0$ *and* $\frac{Du_1^+}{Du_1^-} \geq \frac{Du_2^+}{Du_2^-}$,

3. $\frac{K_1}{Du_1^-} = \frac{K_2}{Du_2^-} = 0$, $\frac{Du_1^+}{Du_1^-} \geq \frac{Du_2^+}{Du_2^-}$ *and* $\frac{D^2u_1^+}{Du_1^- - Du_1^+} \geq \frac{D^2u_2^+}{Du_2^- - Du_2^+}$.[16]

These conditions are also ***sufficient*** *if all the inequalities hold strictly.*

[16] Note that, given the restrictions imposed by Theorem 2, both the numerators and the denominators on both sides of the latter inequality are negative.

For the remaining two attitudes, Success Seeking and Failure Resignation, things are simpler, due to the fact they only admit a continuous representation, and hence the certainty equivalent existence requirement in the definitions of the $\mathcal{S}$, $\mathcal{F}$, RA and RL sets are moot. As a consequence, the $\mathcal{F}$ and RL sets are, respectively, the complements of the $\mathcal{S}$ and RA sets, and hence $\mathcal{F}_1 \subseteq \mathcal{F}_2$ if and only if $\mathcal{S}_2 \subseteq \mathcal{S}_1$, and $RA_1 \subseteq RA_2$ if and only if $RL_2 \subseteq RL_1$. The definitions of the orderings for these two attidues therefore may be equivalently expressed in several ways.

Definition 8. *Let preferences $\succsim_1$ and $\succsim_2$ both satisfy the conditions in Def.4 with respect to the same $x_0 \in \mathbb{R}$. Then, $\succsim_1$ displays (weakly)* more failure resignation than *$\succsim_2$ if thre exist $x_f, x_s : x_f < x_0 < x_s$ s.t. $\forall x \in [x_f, x_0)$, $\exists \bar{x} \in (x_0, x_s]$ such that, for each $x' \in (x_0, \bar{x}]$, both the following conditions are satisfied: (i) $\mathcal{F}_1(x, x') \subseteq \mathcal{F}_2(x, x')$, and (ii) $RA_1(x, x') \subseteq RA_2(x, x')$.*

Definition 9. *Let preferences $\succsim_1$ and $\succsim_2$ both satisfy the conditions in Def.3 with respect to the same $x_0 \in \mathbb{R}$. Then, $\succsim_1$ displays (weakly)* more success seeking than *$\succsim_2$ if there exist $x_f, x_s : x_f < x_0 < x_s$ s.t.: $\forall x' \in (x_0, x_s]$, $\exists \underline{x} \in [x_f, x_0)$ such that, for each $x \in [\underline{x}, x_0)$, both the following conditions are satisfied: (i) $\mathcal{S}_1(x, x') \subseteq \mathcal{S}_2(x, x')$, and (ii) $RL_1(x, x') \subseteq RL_2(x, x')$.*

The next results provide characterize these orderings in the space of utility representations. They are completely analogous to the previous two theorems, with the only difference that they only account for the continuous case, and hence $K_i = 0$ for both agents:

Theorem 8 (Failure Resignation: Interpersonal Comparisons)**.** *Suppose that $(\succsim_i)_{i=1,2}$ are such that $Du_i^- > 0$ and $Du_i^+ < \infty$ and u_i is twice differentiable in some left- and right-neighborhoods of x_0. Then: $\succsim_1$ displays* more success avoidance than *$\succsim_2$* ***only if*** *both (i) $\frac{Du_1^+}{Du_1^-} \geq \frac{Du_2^+}{Du_2^-}$ and (ii) $\frac{D^2u_1^-}{Du_1^- - Du_1^+} \geq \frac{D^2u_2^-}{Du_2^- - Du_2^+}$. These conditions are also* ***sufficient*** *if all the inequalities hold strictly.*

Theorem 9 (Success Seeking: Interpersonal Comparisons)**.** *Suppose that $(\succsim_i)_{i=1,2}$ are such that $Du_i^- < \infty$ and $Du_i^+ > 0$ and u_i is twice differentiable in some left- and right-neighborhoods of x_0. Then: $\succsim_1$ displays* more success seeking than *$\succsim_2$* ***only if*** *both (i) $\frac{Du_1^-}{Du_1^+} \geq \frac{Du_2^-}{Du_2^+}$ and (ii) $\frac{D^2u_1^+}{Du_1^- - Du_1^+} \geq \frac{D^2u_2^+}{Du_2^- - Du_2^+}$. These conditions are also* ***sufficient*** *if all the inequalities hold strictly.*

C.2 Interpersonal Comparisons: A Tight Characterization

The next result provides a tight characterization of the ranking of agents' failure avoidance (as per Def. 6), in terms of the key elements in the main representation theorem:

Theorem 10. *Let preferences $\succsim_1$ and $\succsim_2$ both satisfy the conditions in Def. 2 with respect to the same $x_0 \in \mathbb{R}$. Then, $\succsim_1$ displays* more failure avoidance than *$\succsim_2$ if and only if there exists $\underline{x} < x_0$, such that $\forall x \in (\underline{x}, x_0)$, there exists $\bar{x} > x_0$, s.t., for all $x' \in (x_0, \bar{x})$, one of the following applies:*

1. *$K_1 > 0$ and $\frac{K_1}{m_1(x')} - \frac{K_2}{m_2(x')} > \left[\frac{m_2(x)}{m_2(x')} - \frac{m_1(x)}{m_1(x')}\right](x_0 - x)$.*
2. *$K_1 = K_2 = 0$, $\frac{m_1(x)}{m_1(x')} > \frac{m_2(x)}{m_2(x')}$ and*

$$\frac{m_1(E\hat{p}_1(x,x')) - m_1(x)}{m_2(E\hat{p}_1(x,x')) - \beta m_2(x)} > \frac{m_1(E\hat{p}_1(x,x')) - m_1(x')}{m_2(E\hat{p}_1(x,x')) - m_2(x')} + \left[1 - \beta(x,x')\right]\gamma(x,x'), \tag{27}$$

where $\beta = \frac{x_0 - x}{x' - x}$, *and* $\gamma(x,x') = \frac{m_1(E\hat{p}_1(x,x'))m_1(x)m_2(x') - m_2(x)m_1(x'^2)}{m_1(x')(m_2(E\hat{p}_1(x,x')) - m_2(x'))(m_2(E\hat{p}_1(x,x')) - \beta(x,x')m_2(x))}$.

Proof. Lemma 10 above proves part 1 of the theorem, while Lemma 4, together with Lemma 6 (concerning the $\bar{p}_i$ ranking, noting that for $K_1 = K_2 = 0$, the expression in the lemma reduces to $\frac{m_1(x)}{m_1(x')} > \frac{m_2(x)}{m_2(x')}$) and Lemma 15 (concerning the $\hat{p}_i$ ranking) prove part 2 of the theorem.

C.3 Detailed Proof of Lemma 15

Lemma 15: If both u_1 and u_2 are continuous, $\hat{p}_1(x,x') > \hat{p}_2(x,x')$ if and only if

$$\frac{m_1(y) - m_1(x)}{m_2(y) - \beta m_2(x)} > \frac{m_1(y) - m_1(x')}{m_2(y) - m_2(x')} + (1-\beta)\gamma(x,x',y), \tag{28}$$

where $\gamma(x,x',y) = \frac{m_1(y)m_1(x)m_2(x') - m_2(x)m_1(x'^2)}{m_1(x')(m_2(y) - m_2(x'))(m_2(y) - \beta m_2(x))}$.

Proof: Let $\beta := \frac{x_0 - x}{x' - x}$, and note that $\beta \in (0,1)$ and $\beta \to 1$ as $x' \to x_0$. Also let $y = E\hat{p}_1$, and note that $y \to x_0$ as $x' \to x_0$ (these facts will be useful in the lemmas that follow). Then, from Lemma 13, we have that:

$$\left(\frac{x_0 - y}{x' - y}\right)\left(\frac{(x' - x)}{(x_0 - x)}\right) = \frac{m_1(x) - m_1(x')}{\beta m_1(y) - m_1(x')}. \tag{29}$$

Substituting this notation in the condition of Lemma 14, and particularly using eq. (29), we obtain

$$\frac{m_1(y) - m_1(x)}{m_1(x')} - \frac{m_2(y) - m_2(x)}{m_2(x')} > \left(\frac{m_1(y)}{m_1(x')} - \frac{m_2(y)}{m_2(x')}\right)\left(1 - \frac{m_1(x) - m_1(x')}{\beta m_1(y) - m_1(x')}\right). \tag{30}$$

Next, re-arrange (30) to:

$$\left(\frac{m_1(y)}{m_1(x')} - \frac{m_2(y)}{m_2(x')}\right) - \left(\frac{m_1(x)}{m_1(x')} - \frac{m_2(x)}{m_2(x')}\right) > \left(\frac{m_1(y)}{m_1(x')} - \frac{m_2(y)}{m_2(x')}\right)\left(1 - \frac{m_1(x) - m_1(x')}{\beta m_1(y) - m_1(x')}\right)$$

$$\iff$$

$$\left(\frac{m_1(y)}{m_1(x')} - \frac{m_2(y)}{m_2(x')}\right)\left(\frac{m_1(x) - m_1(x')}{\beta m_1(y) - m_1(x')}\right) > \left(\frac{m_1(x)}{m_1(x')} - \frac{m_2(x)}{m_2(x')}\right)$$

$$\iff$$

$$\frac{1}{m_1(x')}\left(\frac{m_1(y)(m_1(x) - m_1(x'))}{\beta m_1(y) - m_1(x')} - m_1(x)\right) > \frac{1}{m_2(x')}\left(\frac{m_2(y)(m_1(x) - m_1(x'))}{\beta m_1(y) - m_1(x')} - m_2(x)\right)$$

$$\iff$$

$$\frac{m_1(y)(m_1(x) - m_1(x')) - m_1(x)(\beta m_1(y) - m_1(x'))}{m_1(x')(\beta m_1(y) - m_1(x'))} > \frac{m_2(y)(m_1(x) - m_1(x')) - m_2(x)(\beta m_1(y) - m_1(x'))}{m_2(x')(\beta m_1(y) - m_1(x'))}$$

$$\iff$$

$$\frac{m_1(y)\left((1-\beta)m_1(x) - m_1(x')\right) + m_1(x)m_1(x')}{m_1(x')} > \frac{m_2(y)(m_1(x) - m_1(x')) - m_2(x)(\beta m_1(y) - m_1(x'))}{m_2(x')}$$

$$\iff$$

$$\frac{m_1(x')(m_1(x)-m_1(y))+(1-\beta)m_1(y)m_1(x)}{m_1(x')} > \frac{m_2(y)(m_1(x)-m_1(x'))-m_2(x)(\beta m_1(y)-m_1(x'))}{m_2(x')}$$

$$\Longleftrightarrow$$

$$m_1(x)-m_1(y)+\frac{(1-\beta)m_1(y)m_1(x)}{m_1(x')} > \frac{m_2(y)(m_1(x)-m_1(x'))+m_2(x)m_1(x')}{m_2(x')}-\frac{\beta m_2(x)m_1(y)}{m_2(x')}$$

$$\Longleftrightarrow$$

$$m_1(y)\left(\frac{\beta m_2(x)}{m_2(x')}-1\right)+m_1(x) > \frac{m_2(y)(m_1(x)-m_1(x'))+m_2(x)m_1(x')}{m_2(x')}-\frac{(1-\beta)m_1(y)m_1(x)}{m_1(x')}$$

$$\Longleftrightarrow$$

$$\frac{m_1(y)\left(\beta m_2(x)-m_2(x')\right)+m_1(x)m_2(x')}{m_2(x')} > \frac{m_2(y)(m_1(x)-m_1(x'))+m_2(x)m_1(x')-\frac{(1-\beta)m_1(x)m_1(y)m_2(x')}{m_1(x')}}{m_2(x')}$$

and re-arranging further, which we obtain the following (details are in the online appendix):

$$m_1(y)\left(\beta m_2(x)-m_2(x')\right)-m_1(x')m_2(x)+\frac{(1-\beta)m_1(x)m_1(y)m_2(x')}{m_1(x')} > m_2(y)(m_1(x)-m_1(x'))-m_1(x)m_2(x'). \quad (31)$$

rearranging now Equation 28 (and writing γ rather than $\gamma(x,x',y)$, we have:

$$\begin{aligned} & m_1(y)m_2(y)-m_1(y)m_2(x')-m_1(x)m_2(y)+m_1(x)m_2(x') > \\ & \quad m_1(y)m_2(y)-\beta m_1(y)m_2(x)-m_1(x')m_2(y)+\beta m_1(x')m_2(x)-\gamma(1-\beta)\left(m_2(y)-m_2(x')\right)\left(m_2(y)-\beta m_2(x)\right) \\ & \Longleftrightarrow \\ & m_1(y)\left(\beta m_2(x)-m_2(x')\right)-\beta m_1(x')m_2(x) > \\ & \quad m_2(y)\left(m_1(x)-m_1(x')\right)-m_1(x)m_2(x')-\gamma(1-\beta)(m_2(y)-m_2(x'))(m_2(y)-\beta m_2(x)). \end{aligned}$$

Using that $-\beta m_1(x')m_2(x)=(1-\beta)m_1(x')m_2(x)-m_1(x')m_2(x)$, we obtain:

$$\begin{aligned} & m_1(y)\left(\beta m_2(x)-m_2(x')\right)-m_1(x')m_2(x)+\left[\gamma(1-\beta)(m_2(y)-m_2(x'))(m_2(y)-\beta m_2(x))+(1-\beta)m_1(x')m_2(x)\right] > \\ & \quad m_2(y)\left(m_1(x)-m_1(x')\right)-m_1(x)m_2(x'). \end{aligned} \quad (32)$$

For Inequality 31 to hold if and only Inequality 32 holds, it must be that:

$$\gamma(1-\beta)(m_2(y)-m_2(x'))(m_2(y)-\beta m_2(x))+(1-\beta)m_1(x')m_2(x)=\frac{(1-\beta)m_1(x)m_1(y)m_2(x')}{m_1(x')}$$

$$\Longleftrightarrow \quad (33)$$

$$\gamma=\frac{\frac{m_1(y)m_1(x)m_2(x)}{m_1(x')}-m_1(x')m_2(x)}{(m_2(y)-m_2(x'))(m_2(y)-\beta m_2(x))} \quad (34)$$

$$\Longleftrightarrow \quad (35)$$

$$\gamma=\frac{m_1(y)m_1(x)m_2(x')-m_2(x)m_1(x'^2)}{m_1(x')\left(m_2(y)-m_2(x')\right)\left(m_2(y)-\beta m_2(x)\right)}, \quad (36)$$

which concludes the proof of the lemma. QED.