# BINARY SINGLE-CROSSING RANDOM UTILITY MODELS

HENRIK PETRI*

ABSTRACT. We characterize the single-crossing random utility model (SCRUM) (Apesteguia, Ballester, and Lu (2017)) restricted to the (experimentally and empirically) important domain of binary choice menus. As corollaries we characterize some important subclasses of SCRUM (such as single-peaked and single-dipped random utility models) on this domain. In doing so, we address instances of the (still unsolved) binary random utility problem, which asks under what conditions there is a random utility model consistent with data from binary menus.

**Keywords:** stochastic choice, binary choice menus, single-crossing property, single-peaked preferences, single-dipped preferences, stochastic transitivity.

**JEL codes:** D01, D90.

## 1. INTRODUCTION

The random utility model (RUM) posits a probability distribution on a set of preferences. The probability of choosing an alternative is the sum of probabilities of preferences having the alternative as their top choice. The single-crossing random utility model (SCRUM), introduced by Apesteguia, Ballester, and Lu (2017), is a special case of RUM where the collection of preferences in its support satisfies the single-crossing condition. Roughly speaking, a collection of preferences is single-crossing w.r.t. an order $\succ$ if the preferences can be ordered such that if $x \succ y$ then lower ranked preferences prefer $y$ to $x$ and higher ranked preferences prefer $x$ to $y$. This condition arises naturally in a number of applications and is an effective way to control heterogeneity in underlying preferences. The tractability offered by the single-crossing condition implies a number of desirable properties of SCRUM such as unique representation/identification and a clean characterization.

*Department of Economics, University of Bath, 3 East, Bath BA2 7AY, UK, henrik@petri.se. I thank the advisory editor, two anonymous referees, Jose Apesteguia, Miguel Ballester, Jay Lu, Paola Manzini and Marco Mariotti for comments and valuable suggestions that helped improve the paper.

The main contribution of this paper is a characterization of the single-crossing random utility model restricted to the collection of binary menus. Our main property is a type of stochastic transitivity. Our motivation is threefold.

First, binary menus are ubiquitous in experimental and empirical work on discrete choice (see e.g. Tversky (1969), Hey and Orme (1994), Regenwetter, Dana, and Davis-Stober (2011), Halevy, Persitz, and Zrill (2018), Alós-Ferrer, Fehr, and Netzer (2021), Apesteguia and Ballester (2021a) and Apesteguia and Ballester (2021b)). There is an abundance of experimental designs in economics and psychology that make use of so called "Two Alternative Forced Choice" response tasks.[1] Further, binary data sets are readily available in practice due to the binary nature of many choice decisions. Whether to join or not the labour force, to invest in a safe or risky asset, to buy a health insurance or not, or whether to go to work by public or private transportation are just a few examples of binary choices.[2]

Second, although Apesteguia et al. (2017) show that binary menus are sufficient to identify the underlying parameters of SCRUM, the characterization provided by them does not directly apply to the domain of binary menus.[3] Indeed, both of their main properties, monotonicity and centrality vacuously hold on this domain. This paper complements their analysis by showing that there are also simple conditions that characterize SCRUM on binary menus.

Third, the corresponding problem for the random utility model is, despite quite much effort, still largely unsolved.[4] The literature has established various necessary conditions, but to date no set of necessary and sufficient conditions is known (see e.g. Gilboa (1990), Fishburn (1990) and Cohen and Falmagne (1990)). We address an instance of this problem, where the collection of preferences in the support of a random utility model satisfies the single-crossing condition.

[1] "Forced" is to emphasise that participants are forced to choose one of two options in each binary menu they are presented with. We refer to Tversky (1969), Regenwetter et al. (2011) and references therein for further details.

[2] See Bhattacharya (2021), Matzkin (1992) for a further discussion.

[3] Domain issues in the random utility framework are also explored in Apesteguia and Ballester (2021a).

[4] A similar, but distinct, problem is studied in Manzini and Mariotti (2018) (see also Manzini et al. (2019)). They ask: when is there a random utility with support consisting of two distinct utility functions consistent with data from all menus. The binary random utility problem asks: when is there a random utility (with no restriction on the support of the RUM) that is consistent with data from binary menus.

## 2. Binary single-crossing random utility models

2.1. **Notation and preliminary definitions.** We denote by $X$ a finite set of alternatives. Let $\succ$ be a strict linear order on $X$. A *binary stochastic choice function* (binary SCF) is a function $\rho : X \times X \to [0,1]$ such that $\rho(x,y) + \rho(y,x) = 1$ for all alternatives $x, y \in X$. Let $\mathcal{P}$ denote the set of strict linear orders on $X$. For all $x, y \in X$ denote by $\mathcal{P}(x,y) = \{P \in \mathcal{P} : xPy\}$. A binary stochastic choice function $\rho$ is a *random utility model* (RUM) if there is a measure $\mu$ on $\mathcal{P}$ such that $\rho(x,y) = \mu(\mathcal{P}(x,y))$ for all $x, y \in X$. The *support of* $\mu$ is the set of $P \in \mathcal{P}$ with $\mu(P) > 0$. A binary stochastic choice function $\rho$ is a *single-crossing random utility model* (SCRUM) if $\rho$ is a RUM with measure $\mu$ on $\mathcal{P}$ and there is an ordering $\{P_1, ..., P_T\}$ of the support of $\mu$ such that if $x \succ y$ and $s > t$ then $xP_ty$ implies $xP_sy$. Intuitively, $P_s$ is more aligned with $\succ$ than $P_t$. That is, the collection of preferences in the support of $\mu$ satisfies the single-crossing condition w.r.t. $\succ$.

2.2. **Axioms.** Our characterization of SCRUM on the collection of binary menus makes use of a simple stochastic transitivity property.

**Aligned stochastic transitivity (AST).** For all $x, y, z \in X$: If $x \succ y \succ z$ or $z \succ y \succ x$ then

$$\rho(x,z) \geq \min\{\rho(x,y), \rho(y,z)\}.$$

The intuition behind AST is quite simple.[5] Suppose that $x \succ y \succ z$. Then $\rho(x,y)$ is the measure of the set of preferences that are aligned with $x \succ y$. Similarly, $\rho(y,z)$ ($\rho(x,z)$) is the measure of the set of preferences that are aligned with $y \succ z$ ($x \succ z$). AST is a simple consequence of the ordered structure of the support of a SCRUM. To explain, let $\mathcal{S} = \{P_1, ..., P_n\}$ denote the support of a (binary) SCRUM. The structure of $\mathcal{S}$ then implies that either $\mathcal{S}(x,y) \subseteq \mathcal{S}(x,z)$ or $\mathcal{S}(y,z) \subseteq \mathcal{S}(x,z)$ and the conclusion of AST is immediate.[6] To see that $\mathcal{S}(x,y) \subseteq \mathcal{S}(x,z)$ or $\mathcal{S}(y,z) \subseteq \mathcal{S}(x,z)$, we reason by contradiction. If not, then there is a preference $P$ such that $xPy$ and $zPx$, so $zPxPy$. Further there is a preference $P'$ such that $yP'z$ and $zP'x$, so $yP'zP'x$. But, this is a contradiction to $\mathcal{S}$ being a single-crossing collection of preferences.[7]

[5]To provide some further intuition, appendix A discusses other (equivalent) formulations of this property.
[6]Where $\mathcal{S}(x,y) = \{P \in \mathcal{S} : xPy\}$ for all $x, y \in X$.
[7]If $\mathcal{S}$ is single-crossing w.r.t. $\succ$ and $x \succ y \succ z$, then $\mathcal{S}$ cannot contain preferences $P, P'$ such that the "middle" alternative $y$ is best (among $x, y, z$) according to one of these preferences and worst according to the other.

The following, equivalent, formulation of AST highlights an important aspect of AST that is perhaps not immediate from its original formulation above (we show that AST' and AST are equivalent properties in appendix A).

**Aligned Stochastic Transitivity' (AST').** For all $x, y, z \in X$: If $x \succ y \succ z$ then

$$\max\{\rho(x, y), \rho(y, z)\} \geq \rho(x, z) \geq \min\{\rho(x, y), \rho(y, z)\}.$$

In particular, it shows that AST requires $\rho(x, z)$ to be smaller than the maximum of $\rho(x, y)$ and $\rho(y, z)$ whenever $x \succ y \succ z$. As we will discuss later in section 2.3 this implies that AST is logically distinct from many other types of stochastic transitivity considered in the literature.

There is a related (but stronger) notion of stochastic transitivity requiring the stronger conclusion that $\rho(x, z)$ is larger than the maximum of $\rho(x, y)$ and $\rho(y, z)$. As shown in appendix B, this stronger notion characterizes a very restricted class of SCRUMs, where each preference in the support of the RUM is either completely aligned or completely unaligned with the order $\succ$.

2.3. **Discussion of related stochastic transitivity properties.** AST is reminiscent of various stochastic transitivity properties considered in the literature (see He and Natenzon (2020) and Caliari (2021) for an overview of previous literature). The statement of AST resembles the classic moderate stochastic transitivity property.

**Moderate stochastic transitivity (MST).** For all $x, y, z \in X$: If $\rho(x, y) \geq \frac{1}{2}$ and $\rho(y, z) \geq \frac{1}{2}$ then

$$\rho(x, z) \geq \min\{\rho(x, y), \rho(y, z)\}.$$

Note that MST requires the conclusion of AST to hold whenever there is a weak (majority) preference in favor of $x$ compared to $y$ and a weak (majority) preference in favor of $y$ compared to $z$. Despite their resemblance, AST is neither implied by nor implies MST. Example C.1 in appendix C gives an example of a binary SCRUM that fails to satisfy MST. Conversely, example 2.2 below (and the discussion in appendix C.4) shows that a large class of Luce/logit SCFs fails to satisfy AST, but these rules satisfy MST.

The dissimilarity between MST and AST is mainly due to the weak premises of MST. AST is indeed related to the following weaker (stronger premises) version of MST that only requires the conclusion of MST to hold in cases where there is a strict majority in favor of $x$ when compared to $y$ and a strict majority in favor of $y$ when compared to $z$.[8]

**Weak moderate stochastic transitivity (WMST).** For all $x, y, z \in X$: If $\rho(x, y) > \frac{1}{2}$ and $\rho(y, z) > \frac{1}{2}$ then

$$\rho(x, z) \geq \min\{\rho(x, y), \rho(y, z)\}.$$

Our AST property is easily read from this weaker property. Note again that the conclusions of both properties are the same. WMST says that whenever $x$ is "better" than $y$ and $y$ is "better" than $z$, the conclusion must hold. AST suggests that what matters is not the quality/majority order but rather the underlying/exogenous order $\succ$ of alternatives. It is therefore not very surprising that AST implies WMST.

**Proposition 2.1.** *If a binary stochastic choice function $\rho$ satisfies AST then it satisfies WMST.*

The proof of proposition 2.1 is straightforward and delegated to appendix C. As the next example shows, WMST is strictly weaker than our stochastic transitivity property and hence not equivalent to it. It shows that there are Luce/logit rules that fail to satisfy AST. Since every Luce rule satisfies WMST it follows that WMST is strictly weaker than AST.[9]

**Example 2.2.** Let $X = \{x, y, z\}$ be a set with three alternatives and an order $\succ$ such that $x \succ y \succ z$. Consider a utility function $u$ on $X$ defined by $u(x) = 3, u(y) = 2, u(z) = 1$. Let $\rho$ be a (binary) Luce rule with utility function $u$, i.e.

$$\rho(a, b) = \frac{u(a)}{u(a) + u(b)}$$

for all $a, b \in X$. The probability of choosing an alternative in this model is thus proportional to its utility value. We show that $\rho$ violates AST. Note that:

$$\rho(x, z) = \frac{u(x)}{u(x) + u(z)} = \frac{3}{4} > \frac{2}{3} = \frac{u(y)}{u(y) + u(z)} = \rho(y, z),$$
$$\rho(x, z) = \frac{u(x)}{u(x) + u(z)} = \frac{3}{4} > \frac{3}{5} = \frac{u(x)}{u(x) + u(y)} = \rho(x, y).$$

[8]This property is also discussed in Fishburn (1973) under the name partial stochastic transitivity.

[9]This example also shows that MST does not imply AST, since every Luce rule satisfies the stronger MST property.

Thus, $\rho(x,z) > \max\{\rho(x,y),\rho(y,z)\}$. This is a violation of AST. To see this, note that $x \succ y \succ z$ which by formulation AST' of AST implies that $\rho(x,z) \le \max\{\rho(x,y),\rho(y,z)\}$, a contradiction. ◁

Example 2.2 shows that there are binary Luce SCFs that fail to satisfy AST (for some exogenous order $\succ$). Although, the binary Luce SCF $\rho$ in example 2.2 violates AST w.r.t. the exogenous order $x \succ y \succ z$, there are other orders $\succ'$ s.t. $\rho$ satisfies AST w.r.t. $\succ'$. Indeed, one can show that $\rho$ satisfies AST w.r.t. the order $x \succ' z \succ' y$.[10] However, if the grand set of alternatives $X$ contains at least five alternatives, then one can show that the class of binary Luce SCFs (with strict utility) fails to satisfy AST for *any* order $\succ$ on $X$ (see appendix C.4 for details). So the behavior predicted by the binary Luce model and AST is, in this sense, very different.

The basic reason why the binary Luce SCF in example 2.2 fails to satisfy AST is that it satisfies *strong stochastic transitivity*. This property requires that if $\rho(x,y) > \frac{1}{2}$ and $\rho(y,z) > \frac{1}{2}$ then $\rho(x,z) > \max\{\rho(x,y),\rho(y,z)\}$ and is a characterizing property of a larger class of Fechnerian/simple scalability SCFs. As noted above, AST implies that $\rho(x,z)$ is *smaller* than $\max\{\rho(x,y),\rho(y,z)\}$. This inconsistency, modulo some technicalities, implies that the class of SCFs satisfying AST is disjoint from those satisfying strong stochastic transitivity. A corollary is hence that AST is inconsistent with any type of behavior generated by the (strict) Fechnerian model. We elaborate further on this in appendix C.4.

2.4. **Characterization.** We are now ready to state our main characterization result. The proof is constructive and follows similar lines of reasoning as in Apesteguia et al. (2017).

**Theorem 2.3.** *A binary stochastic choice function $\rho$ satisfies AST if and only if it is a SCRUM.*

*Proof.* For each $\theta \in (0,1]$ define a binary relation $P_\theta$ by $xP_\theta y$ if and only if $[x \succ y$ and $\rho(y,x) < \theta]$ or $[y \succ x$ and $\rho(x,y) \ge \theta]$. We claim that each $P_\theta$ is a strict linear order. The proof of asymmetry and completeness of $P_\theta$ is immediate. We claim that $P_\theta$ is transitive. Assume that $xP_\theta y$ and $yP_\theta z$. There are six mutually exhaustive cases:

(1) $x \succ y \succ z$. Then $\rho(y,x) < \theta$ and $\rho(z,y) < \theta$. Hence $\rho(x,y) \ge \theta$ and $\rho(y,z) \ge \theta$, so by AST it then follows that $\rho(x,z) \ge \min\{\rho(x,y),\rho(y,x)\} \ge \theta$. Hence $\rho(z,x) < \theta$ so $xP_\theta z$.

[10] To see this, note that $\rho(x,z) = \frac{3}{4} > \rho(x,y) = \frac{3}{5} > \rho(z,y) = \frac{1}{3}$, so $\max\{\rho(x,z),\rho(z,y)\} \ge \rho(x,y) \ge \min\{\rho(x,z),\rho(z,y)\}$ and hence AST (AST') holds.

(2) $x \succ z \succ y$. Then $\rho(y,x) < \theta \leq \rho(y,z)$. It follows that $\rho(z,x) < \theta$, since otherwise $\rho(z,x) \geq \theta$ and $\rho(y,z) \geq \theta$ together with AST implies that $\rho(y,x) \geq \theta$. A contradiction. Thus $xP_\theta z$.

(3) $y \succ x \succ z$. Then $\rho(x,y) \geq \theta > \rho(z,y)$. By AST it follows that $\rho(z,y) \geq \min\{\rho(x,y), \rho(z,x)\}$ hence $\theta > \rho(z,y) \geq \rho(z,x)$ implying that $xP_\theta z$.

(4) $y \succ z \succ x$. Then $\rho(z,y) < \theta \leq \rho(x,y)$. Thus $\rho(y,x) \leq 1-\theta < \rho(y,z)$. By AST it follows that $\rho(y,x) \geq \min\{\rho(y,z), \rho(z,x)\} = \rho(z,x)$. Hence $\rho(z,x) \leq 1-\theta$, or equivalently $\rho(x,z) \geq \theta$. Thus $xP_\theta z$.

(5) $z \succ y \succ x$. Then $\rho(x,y) \geq \theta$ and $\rho(y,z) \geq \theta$. Then AST immediately gives $\rho(x,z) \geq \min\{\rho(x,y), \rho(y,z)\} \geq \theta$. Thus $xP_\theta z$.

(6) $z \succ x \succ y$. Then $\rho(y,x) < \theta \leq \rho(y,z)$. Hence $\rho(x,y) > 1-\theta \geq \rho(z,y)$. It follows by AST that $1-\theta \geq \rho(z,y) \geq \min\{\rho(x,y), \rho(z,x))\} = \rho(z,x)$. Thus $\rho(x,z) \geq \theta$. So $xP_\theta z$.

Define a RUM $\mu$ by assigning to any $P \in \mathcal{P}$ the value $\mu(P) = \ell\{\theta : P_\theta = P\}$, where $\ell$ is the Lebesgue measure. We show that $\rho(x,y) = \mu(\mathcal{P}(x,y))$ for all $x,y \in X$ (the case $y \succ x$ is symmetric). Let $x,y \in X$ and assume w.l.o.g. that $x \succ y$. By construction of $P_\theta$ it then follows that for all $\theta$: $xP_\theta y$ if and only if $\rho(y,x) < \theta$. Hence $\rho(x,y) = 1 - \rho(y,x) = \ell\{\theta : \rho(y,x) < \theta\} = \ell\{\theta : xP_\theta y\} = \mu(\mathcal{P}(x,y))$. Similarly, it follows that $\rho(y,x) = \ell\{\theta : \rho(y,x) \geq \theta\} = \ell\{\theta : yP_\theta x\} = \mu(\mathcal{P}(y,x))$. It remains to prove that the RUM $\mu$ constructed above is single-crossing, but this follows directly from the proof of Claim 6 (p.671) in Apesteguia et al. (2017). This completes the proof of sufficiency.

We next show that AST is necessary for a binary SCRUM representation. Let $\rho$ be a binary SCRUM and let $x,y,z \in X$. Assume w.l.o.g. that $\rho(x,y) > 0$ and $\rho(y,z) > 0$. Let $\{P_1, ..., P_n\}$ be an ordering of the support of $\mu$ that satisfies the single-crossing condition w.r.t. $\succ$.

Case 1 ($x \succ y \succ z$): Since $\rho(x,y) > 0, \rho(y,z) > 0$ there are indices $k(x)$ and $k(y)$ s.t. $\rho(x,y) = \sum_{i=k(x)}^n \mu(P_i)$ and $\rho(y,z) = \sum_{i=k(y)}^n \mu(P_i)$ and such that $xP_iy$ if and only if $i \geq k(x)$ and $yP_iz$ if and only if $i \geq k(y)$. It follows by transitivity of the $P_i$ that $xP_iz$ for all $i \geq \max\{k(x), k(y)\}$. The conclusion of AST then follows since

$$\rho(x,z) \geq \sum_{i=\max\{k(x),k(y)\}}^n \mu(P_i) = \min\{\sum_{i=k(x)}^n \mu(P_i), \sum_{i=k(y)}^n \mu(P_i)\} =$$

$$= \min\{\rho(x,y), \rho(y,z)\}.$$

Case 2 ($z \succ y \succ x$): Since $\rho(x,y) > 0, \rho(y,z) > 0$ there are indices $k(x)$ and $k(y)$ s.t. $\rho(x,y) = \sum_{i=1}^{k(x)} \mu(P_i)$ and $\rho(y,z) = \sum_{i=1}^{k(y)} \mu(P_i)$ and such that $xP_iy$ if and only if $i \leq k(x)$ and $yP_iz$ if and only if $i \leq k(y)$. By transitivity of the $P_i$ it follows that $xP_iz$ for all $i \leq \min\{k(x), k(y)\}$. Hence,

$$\rho(x,z) \geq \sum_{i=1}^{\min\{k(x),k(y)\}} \mu(P_i) = \min\{\sum_{i=1}^{k(x)} \mu(P_i), \sum_{i=1}^{k(y)} \mu(P_i)\} =$$
$$= \min\{\rho(x,y), \rho(y,z)\}.$$

This completes the proof of the claim. □

## 3. Binary single-peaked and single-dipped random utility models

We next study and provide characterizations of two interesting subclasses of SCRUM. Let $\max(X,P)$ denote the best alternative in $X$ according to preference $P \in \mathcal{P}$.[11] A preference relation $P$ is *single-peaked* if $y \succ x \succ \max(X,P)$ or $\max(X,P) \succ x \succ y$ implies $xPy$. A random utility model $\mu$ is a *single-peaked random utility model* (SPRUM) if every preference $P$ in the support of $\mu$ is single-peaked. The following property is an observable implication of SPRUM on binary menus.

**Monotone increasing demand (MID).** For all $x, y, z \in X$: If $x \succ y \succ z$ or $z \succ y \succ x$ then $\rho(x,z) \geq \rho(x,y)$.

Formally, MID is a strengthening of AST. Intuitively, MID says that if the distance from $x$ to another alternative increases then the demand for $x$ relative to the other alternative must (weakly) increase as well. MID is related to concavity of the preferences in the support of the RUM. If $X$ is a subset of the real line, then MID is satisfied if each preference in the support of the RUM is represented by a quasiconcave utility function. Indeed, MID is reminiscent of a condition called convex substitutability in Lu and Saito (2021).[12] Convex substitutability requires that $\rho(x, \alpha x + (1-\alpha)y) \leq \rho(x,y)$ for all $x, y \in X$ and $\alpha \in (0,1)$ (where $X$ is convex) and is a necessary condition for pure characteristic models (which are random utility models with continuous utilities) with quasiconcave utility functions. Loosely

[11] I.e. $\max(X,P)Py$ for all $y \in X \setminus \{\max(X,P)\}$.

[12] I thank Jay Lu for suggesting the MID property and its interpretation.

speaking, convex substitutability requires the demand for $x$ to increase when the "convex distance" to another alternative increases.[13] To further illustrate this we next consider an example adapted from Lu and Saito (2021).

**Example 3.1.** Let $X = \{x, y, z\}$ where $x > y > z$ are real numbers situated in the interval $[0, 1]$ and $y > \frac{x+z}{2}$ (for concreteness think of the case $x = 1$, $y = \frac{3}{4}$ and $z = 0$). Assume that $\succ$ agrees with the order $>$, i.e. $x \succ y \succ z$. Define two utility functions $u_1, u_2$ by:

$$u_1(\alpha) = -(\alpha - y)^2,$$
$$u_2(\alpha) = -(\alpha - 2)^2.$$

Assume that $\rho$ puts equal probability on $u_1$ and $u_2$. To verify that the MID property holds we calculate the choice probabilities $\rho(x, y)$ and $\rho(x, z)$. Note that $u_2(x) > u_2(y) > u_2(z)$ since $u_2(\alpha)$ is increasing for all $\alpha \leq 2$. Further, we have $u_1(z) = -(z - y)^2 < -(x - y)^2 = u_1(x)$, where $-(x - y)^2 > -(z - y)^2$ follows since $y > \frac{x+z}{2}$. It is clear that $u_1(x) < 0 = u_1(y)$, and thus $\rho(x, y) = \frac{1}{2} < 1 = \rho(x, z)$. $\triangleleft$

The fundamental reason why the RUM in example 3.1 satisfies MID is (quasi)concavity of $u_1$ and $u_2$. Indeed, example 3.1 can be generalized to show that MID holds if each preference in the support of a RUM satisfies an abstract quasiconcavity condition. Recall that a utility function $u$ is quasiconcave if $u(y) \geq \min\{u(x), u(z)\}$, where $y = \lambda x + (1 - \lambda)z$. In analogy with this definition call a preference relation $P$ on X *quasiconcave* if for all $x, y, z \in X$ with $x \succ y \succ z$ or $z \succ y \succ x$ we have that $yPx$ or $yPz$.[14] Quasiconcavity of a preference implies that if $x$ is preferred to a "close" alternative $y$ then $x$ must also be preferred to a more "distant" alternative $z$. I.e., if $x \succ y \succ z$ and $xPy$ then $xPz$. If not, then we have $zPxPy$, a contradiction to quasiconcavity. By this it is immediate that any RUM with quasiconcave preferences satisfies MID. I.e. the demand for $x$ relative to a more distant alternative $z$ must be larger than the demand for $x$ relative to a close alternative $y$. This is because quasiconcavity implies that every agent that prefers $x$ to $y$ must also prefer $x$ to $z$.[15]

In appendix C.3 we show that quasiconcavity and single-peakedness are equivalent properties for a preference relation $P$. Given the preceding discussion and the resemblance between

[13]When $X$ is a convex subset of the real line with its usual order, then our notion of distance agrees with that of Lu and Saito (2021) and the MID property agrees with convex substitutability. A proof of this result is available from the author upon request.

[14]Equivalently, a preference relation $P$ is quasiconcave if it can be represented by a utility function $u : X \to \mathbb{R}$ such that for all $x, y, z \in X$: If $x \succ y \succ z$ or $z \succ y \succ x$ then $u(y) \geq \min\{u(x), u(z)\}$.

[15]More formally, note that MID holds since $\rho(x, y) = \sum_{P \in \mathcal{P}: xPy} \mu(P) \leq \sum_{P \in \mathcal{P}: xPz} \mu(P) = \rho(x, z)$.

quasiconcavity and single-peakedness it is perhaps not surprising that MID is a necessary property in a characterization of SPRUM. The following proposition shows that MID is also sufficient.

**Proposition 3.2.** *A binary stochastic choice function $\rho$ satisfies MID if and only if it is a single-peaked random utility model.*

*Proof.* We first show necessity. Let $x, y, z \in X$. Assume that $x \succ y \succ z$ (the case $z \succ y \succ x$ follows by similar reasoning). We show that for all $P \in \mathcal{P}$ in the support of $\mu$: $xPy$ implies $xPz$. There are three cases:

(1) $\max(X, P) \succ y$. Then $yPz$. So if $xPy$ then $xPz$.
(2) $\max(X, P) = y$. Then $\max(X, P) = y$ so $xPy$ cannot hold.
(3) $y \succ \max(X, P)$. Then $yPx$, so $xPy$ cannot hold.

Hence, it follows that $\rho(x, y) = \sum_{P:xPy} \mu(P) \leq \sum_{P:xPz} \mu(P) = \rho(x, z)$. We next show sufficiency. It is clear that MID implies AST. Hence $\rho$ has a RUM representation with preferences $P_\theta$ as constructed in the proof of theorem 2.3. We show that these preferences satisfy the single-peakedness condition. If $\max(P_\theta, X) \succ x \succ y$ then since $\max(P_\theta, X)P_\theta y$ it follows by construction of $P_\theta$ that $\rho(y, \max(P_\theta, X)) < \theta$. Hence, an application of MID gives $\rho(y, x) \leq \rho(y, \max(P_\theta, X)) < \theta$ implying that $xP_\theta y$. If $y \succ x \succ \max(X, P_\theta)$ then since $\max(P_\theta, X)P_\theta x$ it follows by construction of $P_\theta$ that $\rho(\max(P_\theta, X), x) \geq \theta$. Several applications of MID gives $\rho(x, y) \geq \rho(\max(P_\theta, X), y) \geq \rho(\max(P_\theta, X), x) \geq \theta$, so $xP_\theta y$. $\square$

We next consider a class of stochastic choice functions inversely related to SPRUM, the single-dipped random utility model. Let $\min(X, P)$ denote the worst alternative in $X$ according to preference $P \in \mathcal{P}$.[16] A preference relation $P$ is *single-dipped* if $x \succ y \succ \min(X, P)$ or $\min(X, P) \succ y \succ x$ implies $xPy$. A random utility model $\mu$ is a *single-dipped random utility model* (SDRUM) if every preference $P$ in the support of $\mu$ is single-dipped. Perhaps not surprisingly, this model is characterized by a property inversely related to MID.

**Monotone decreasing demand (MDD).** For all $x, y, z \in X$: If $x \succ y \succ z$ or $z \succ y \succ x$ then $\rho(x, z) \leq \rho(x, y)$.

To the best of our knowledge the MDD property is novel. Analogously to MID, this property says that the demand for $x$ (weakly) decreases when the distance to the other

[16] I.e. $yP\min(X, P)$ for all $y \in X \setminus \{\min(X, P)\}$.

alternative increases. A somewhat subtle point is that MDD implies AST (in contrast to MID that trivially implies AST). To see this, assume that $x \succ y \succ z$ or $z \succ y \succ x$. The MDD property then implies that $\rho(z, x) \leq \rho(z, y)$ (this follows by "relabeling" of alternatives). Since $\rho(z, x) \leq \rho(z, y)$ we have $\rho(x, z) \geq \rho(y, z) \geq \min\{\rho(x, y), \rho(y, z)\}$. Hence AST follows from MDD.

Similarly, but opposite to the case of MID, the MDD property can be interpreted as a convexity condition imposed on the preferences in the support of the RUM. Example 3.1 shows that MID and MDD are distinct properties since the MID property is satisfied with strict inequality in its conclusion (implying that MDD is violated). Example 3.1 can also be "reversed" to show that there are stochastic choices that satisfy MDD with strict inequality (and hence violate MID).[17]

Similar to SPRUMs, the single-dipped condition is closely related to quasiconvexity of the preferences in the support of the RUM. Define a preference $P$ to be *quasiconvex* if $x \succ y \succ z$ or $z \succ y \succ x$ implies that $xPy$ or $zPy$. Appendix C.3 shows that a preference $P$ is quasiconvex if and only if it is single-dipped. From the previous discussion it is quite clear that MDD is a necessary property in a characterization of SDRUM. The next proposition shows that this property alone characterizes SDRUM.

**Proposition 3.3.** *A binary stochastic choice function $\rho$ satisfies MDD if and only if it is a single-dipped random utility model.*

*Proof.* We first show necessity. Let $x, y, z \in X$. Assume that $x \succ y \succ z$ (the case $z \succ y \succ x$ follows by similar reasoning). We show that for all $P \in \mathcal{P}$ in the support of $\mu$: $xPz$ implies $xPy$. There are three cases:

(1) $\min(X, P) \succ y$. Then $zPy$. So if $xPz$ then $xPy$.
(2) $\min(X, P) = y$. Then $xPy$ always holds.
(3) $y \succ \min(X, P)$. Then $xPy$ always holds.

Hence, it follows that $\rho(x, y) = \sum_{P:xPy} \mu(P) \geq \sum_{P:xPz} \mu(P) = \rho(x, z)$. We next show sufficiency. It is shown in the main text that MDD implies AST (see paragraph following the statement of MDD). It thus follows that $\rho$ has a RUM representation with preferences $P_\theta$ as constructed in the proof of the theorem 2.3. We show that these preferences satisfy the single-dippedness condition. If $\min(P_\theta, X) \succ y \succ x$ then since $xP_\theta \min(P_\theta, X)$ it follows by

[17] To see this, let everything be the same as in example 3.1 except that $u_1(\alpha) = (\alpha - y)^2$ and $u_2(\alpha) = (\alpha - 2)^2$, so that $u_1$ and $u_2$ are convex on the interval $[0, 1]$. Similar arguments as in example 3.1 show that $\rho(x, z) = 0 < \frac{1}{2} = \rho(x, y)$.

construction of $P_\theta$ that $\rho(x, \min(P_\theta, X)) \geq \theta$. Hence, an application of MDD gives $\rho(x,y) \geq \rho(x, \min(P_\theta, X)) \geq \theta$ implying that $xP_\theta y$. If $x \succ y \succ \min(X, P_\theta)$ then since $yP_\theta \min(P_\theta, X)$ it follows by construction of $P_\theta$ that $\rho(\min(P_\theta, X), y) < \theta$. Several applications of MDD gives $\rho(y,x) \leq \rho(\min(P_\theta, X), x) \leq \rho(\min(P_\theta, X), y) < \theta$, so $xP_\theta y$. □

MID and MDD require the demand for $x$ to weakly decrease/increase when the distance to another alternative increases. Relatedly, one may also ask what happens if both MID and MDD hold, i.e. if we impose a condition requiring the demand for $x$ to remain unchanged when the distance to another alternative increases.

**Monotone indifferent demand (MIND).** For all $x, y, z \in X$: If $x \succ y \succ z$ or $z \succ y \succ x$ then $\rho(x,z) = \rho(x,y)$.

Given that MID corresponds to a case of quasiconcave preferences, and MDD corresponds to a case of quasiconvex preferences, it is perhaps not surprising that MIND corresponds to a case of linear preferences in the support of the RUM. As the following proposition shows, MIND implies that all preferences in the support of $\mu$ either linearly increases or decreases with $\succ$. In our setting this means that they are either completely aligned or completely unaligned with $\succ$.

**Proposition 3.4.** *A binary stochastic choice function $\rho$ satisfies MIND if and only if there is an $\alpha \in [0,1]$ such that for all $x, y \in X$: $\rho(x,y) = \alpha$ if $x \succ y$ (i.e. $\rho$ is a RUM with measure $\mu$ that puts probability $\alpha$ on $\succ$ and remaining probability $1-\alpha$ on its reverse order).*

Proposition 3.4 follows as a corollary to proposition 3.2 and proposition 3.3. To see this, note that MIND implies MID and MDD. By MID it follows that $\rho$ is a single-peaked RUM. Suppose that there is a preference $P$ in the support of the RUM that is not equal $\succ$ or its reverse order. Then there are $x \succ y \succ z$ such that $yPxPz$ or $yPzPx$. If $yPxPz$ then $\rho(x,y) < \rho(x,y) + \mu(P) \leq \rho(x,z)$. This violates MDD. If $yPzPx$ then $\rho(z,y) < \rho(z,y) + \mu(P) \leq \rho(z,x)$. A violation of MDD.

## 4. Concluding remarks

We characterized the single-crossing random utility model (and subclasses of it) using versions of stochastic transitivity. As noted, the representation of a SCRUM on binary menus is unique,[18] which justifies the collection of binary menus as an ideal environment

[18]This follows by inspection of the proof of proposition 1 (p.672) in Apesteguia et al. (2017).

to study SCRUMs. An avenue for future research is to experimentally test the properties introduced above by using similar two alternative forced choice designs as in Tversky (1969) and Regenwetter et al. (2011).

## References


C. Alós-Ferrer, E. Fehr, and N. Netzer. Time will tell: Recovering preferences when choices are noisy. *Journal of Political Economy*, 129(6):1828–1877, 2021.

J. Apesteguia and M. A. Ballester. Random utility models with ordered types and domains. 2021a.

J. Apesteguia and M. A. Ballester. Separating predicted randomness from residual behavior. *Journal of the European Economic Association*, 19(2):1041–1076, 2021b.

J. Apesteguia, M. A. Ballester, and J. Lu. Single-crossing random utility models. *Econometrica*, 85(2):661–674, 2017.

D. Bhattacharya. The empirical content of binary choice models. *Econometrica*, 89(1): 457–474, 2021.

D. Caliari. Imperfect discrimination and stochastic transitivity. 2021.

M. Cohen and J.-C. Falmagne. Random utility representation of binary choice probabilities: a new class of necessary conditions. *Journal of Mathematical Psychology*, 34(1):88–94, 1990.

P. C. Fishburn. *The theory of social choice*. Princeton University Press, 1973.

P. C. Fishburn. Binary probabilities induced by rankings. *SIAM Journal on Discrete Mathematics*, 3(4):478–488, 1990.

I. Gilboa. A necessary but insufficient condition for the stochastic binary choice problem. *Journal of Mathematical Psychology*, 34(4):371–392, 1990.

Y. Halevy, D. Persitz, and L. Zrill. Parametric recoverability of preferences. *Journal of Political Economy*, 126(4):1558–1593, 2018.

J. He and P. Natenzon. Moderate expected utility. *Working Paper*, 2020.

J. D. Hey and C. Orme. Investigating generalizations of expected utility theory using experimental data. *Econometrica*, pages 1291–1326, 1994.

J. Lu and K. Saito. Mixed logit and pure characteristics models. 2021.

P. Manzini and M. Mariotti. Dual random utility maximisation. *Journal of Economic Theory*, 177:162–182, 2018.

P. Manzini, M. Mariotti, and H. Petri. Corrigendum to "Dual random utility maximisation"[J. Econ. Theory 177 (2018) 162–182]. *Journal of Economic Theory*, 184:104944, 2019.

R. L. Matzkin. Nonparametric and distribution-free estimation of the binary threshold crossing and the binary choice models. *Econometrica*, pages 239–270, 1992.

M. Regenwetter, J. Dana, and C. P. Davis-Stober. Transitivity of preferences. *Psychological review*, 118(1):42, 2011.

A. Tversky. Intransitivity of preferences. *Psychological review*, 76(1):31, 1969.

A. Tversky and J. E. Russo. Substitutability and similarity in binary choices. *Journal of Mathematical psychology*, 6(1):1–12, 1969.

## Appendix A. Equivalent formulations of AST

The aim of this section is to show that our different formulations of the AST property are equivalent. Before doing so we state yet another formulation of AST.

**Aligned Stochastic Transitivity" (AST").** For all $x, y, z \in X$ and $\alpha \in [0, 1]$: If $x \succ y \succ z$ or $z \succ y \succ x$ then $\rho(x, y) \geq \alpha$ and $\rho(y, z) \geq \alpha$ implies $\rho(x, z) \geq \alpha$.

The following lemma shows that AST, AST' and AST" are equivalent properties and that AST can be replaced with either AST' or AST" in our characterization of binary SCRUMs.

**Lemma A.1.** *Let $\rho$ be a binary stochastic choice function. Then AST, AST' and AST" are equivalent properties.*

*Proof.* We show the implications AST' $\Rightarrow$ AST $\Rightarrow$ AST" $\Rightarrow$ AST'.

(AST' $\Rightarrow$ AST): Let $x, y, z \in X$ with $x \succ y \succ z$ or $z \succ y \succ x$. If $x \succ y \succ z$ then AST' implies $\rho(x, z) \geq \min\{\rho(x, y), \rho(y, z)\}$. If $z \succ y \succ x$ then AST' implies $\max\{\rho(y, x), \rho(z, y)\} \geq \rho(z, x)$, which is equivalent to $\rho(x, z) \geq \min\{\rho(x, y), \rho(y, z)\}$.

(AST $\Rightarrow$ AST"): Let $x, y, z \in X$ and $\alpha \in [0, 1]$ with $\rho(x, y) \geq \alpha$ and $\rho(y, z) \geq \alpha$. If $x \succ y \succ z$ or $z \succ y \succ x$ then AST implies that $\rho(x, z) \geq \min\{\rho(x, y), \rho(y, z)\} \geq \alpha$. Thus AST" holds.

(AST" $\Rightarrow$ AST'): Let $x, y, z \in X$ with $x \succ y \succ z$. Set $\alpha = \min\{\rho(x, y), \rho(y, z)\}$. Then $\rho(x, y) \geq \alpha$ and $\rho(y, z) \geq \alpha$, so AST" implies that $\rho(x, z) \geq \alpha = \min\{\rho(x, y), \rho(y, z)\}$.

Set $\alpha' = \min\{\rho(y,x), \rho(z,y)\}$ then $\rho(y,x) \geq \alpha'$ and $\rho(z,y) \geq \alpha'$, so another application of AST" gives $\rho(z,x) \geq \alpha' = \min\{\rho(y,x), \rho(z,y)\}$ and this implies that $\rho(x,z) \leq \max\{\rho(x,y), \rho(y,z)\}$. □

## Appendix B. Discussion of strong aligned stochastic transitivity

A property related to, but stronger than, our main stochastic transitivity property is the following.

**Strong aligned stochastic transitivity (SAST).** For all $x, y, z \in X$: If $x \succ y \succ z$ or $z \succ y \succ x$ then

$$\rho(x,z) \geq \max\{\rho(x,y), \rho(y,z)\}.$$

Note that SAST strengthens the conclusion of AST by requiring $\rho(x,z)$ to be larger than both $\rho(x,y)$ and $\rho(y,z)$. The following proposition shows that this condition is, in a way, too strong. It is only satisfied by a very restricted class of stochastic choice functions.

**Proposition B.1.** *A binary stochastic choice function $\rho$ satisfies SAST if and only if there is an $\alpha \in [0,1]$ such that for all $x, y \in X$: $\rho(x,y) = \alpha$ if $x \succ y$ (i.e. $\rho$ is a RUM with measure $\mu$ that puts probability $\alpha$ on $\succ$ and remaining probability $1-\alpha$ on its reverse order).*

*Proof.* To prove sufficiency it suffices (by proposition 3.4) to prove that SAST implies MIND. Let $x \succ y \succ z$ or $z \succ y \succ x$ then SAST implies that $\rho(x,z) \geq \max\{\rho(x,y), \rho(y,z)\} \geq \rho(x,y)$ and by "relabeling" that $\rho(z,x) \geq \max\{\rho(z,y), \rho(y,x)\} \geq \rho(y,x)$ so $\rho(x,z) \leq \rho(x,y)$. Hence $\rho(x,y) = \rho(x,z)$.

Conversely, let $\rho$ be such that there is an $\alpha \in [0,1]$ such that for all $x, y \in X$: $\rho(x,y) = \alpha$ if $x \succ y$. Let $x, y, z \in X$. If $x \succ y \succ z$ then $\rho(x,z) = \alpha = \rho(x,y) = \rho(y,z)$ so $\rho(x,z) \geq \max\{\rho(x,y), \rho(y,z)\}$. If $z \succ y \succ x$ then $\rho(x,z) = 1-\alpha = \rho(x,y) = \rho(y,z)$ so $\rho(x,z) \geq \max\{\rho(x,y), \rho(y,z)\}$. □

## Appendix C. Proofs and examples omitted from main text

### C.1. Example showing that MST is not implied by AST.

**Example C.1.** Let $X$ be a set with three alternatives $x \succ y \succ z$. Let $\rho$ have a SCRUM representation w.r.t. $\succ$. More precisely, $\rho$ is represented by a SCRUM $\mu$ with equal probabilities put on orders $P_1$ and $P_2$ that are such that $yP_1zP_1x$ and $xP_2yP_2z$. Since $\rho$ is a

SCRUM w.r.t. $\succ$ it follows by theorem 2.3 that $\rho$ satisfies AST. However, $\rho$ fails to satisfy MST, since $\rho(z,x) = \frac{1}{2}$ and $\rho(x,y) = \frac{1}{2}$, but $\rho(z,y) = 0 < \frac{1}{2} = \min\{\rho(z,x), \rho(x,y)\}$. $\triangleleft$

C.2. **Proof of proposition 2.1.** Let $x, y, z \in X$ with $\rho(x,y) > \frac{1}{2}$ and $\rho(y,z) > \frac{1}{2}$. There are six mutually exhaustive cases:

(1) $x \succ y \succ z$. Then AST gives $\rho(x,z) \geq \min\{\rho(x,y), \rho(y,z)\}$.
(2) $z \succ y \succ x$. Then AST gives $\rho(x,z) \geq \min\{\rho(x,y), \rho(y,z)\}$.
(3) $x \succ z \succ y$. Then AST implies that $\rho(y,x) \geq \min\{\rho(y,z), \rho(z,x)\}$. Hence, since $\rho(y,z) > \frac{1}{2} > \rho(y,x)$, it follows that $\rho(y,x) \geq \rho(z,x)$, which in turn implies that $\rho(x,z) \geq \rho(x,y) \geq \min\{\rho(x,y), \rho(y,z)\}$.
(4) $y \succ z \succ x$. This case follows by similar reasoning as the previous one.
(5) $z \succ x \succ y$. Then AST implies that $\rho(z,y) \geq \min\{\rho(z,x), \rho(x,y)\}$. Hence, since $\rho(x,y) > \frac{1}{2} > \rho(z,y)$, it follows that $\rho(z,y) \geq \rho(z,x)$. Hence $\rho(x,z) \geq \rho(y,z) \geq \min\{\rho(x,y), \rho(y,z)\}$.
(6) $y \succ x \succ z$. This case follows by similar reasoning as the previous one. □

C.3. **Proofs on quasiconcavity and quasiconvexity.** We first prove that a preference $P$ is single-peaked if and only if it is quasiconcave.

**Lemma C.2.** *A preference relation $P$ is single-peaked if and only if it is quasiconcave.*

*Proof.* Suppose that $P$ is single-peaked. Assume that $x \succ y \succ z$ (the case $z \succ y \succ x$ follows by similar reasoning). We show that $xPy$ implies $yPz$. There are three cases:

(1) $\max(X,P) \succ y$. Then $\max(X,P) \succ y \succ z$ so $yPz$.
(2) $\max(X,P) = y$. Then, clearly $yPz$.
(3) $y \succ \max(X,P)$. Then $x \succ y \succ \max(X,P)$ so $yPx$ and hence $xPy$ cannot hold.

Conversely, assume that $P$ is quasiconcave. Assume that $\max(X,P) \succ x \succ y$ (the case $y \succ x \succ \max(X,P)$ is similar). Assume by contradiction that $yPx$ then quasiconcavity of $P$ implies that $xP\max(X,P)$. A contradiction. Hence $xPy$. □

We next prove that single-dippedness and quasiconvexity are equivalent properties for a preference relation $P$.

**Lemma C.3.** *A preference relation $P$ is single-dipped if and only if it is quasiconvex.*

*Proof.* We first show necessity. Let $x, y, z \in X$. Assume that $x \succ y \succ z$ (the case $z \succ y \succ x$ follows by similar reasoning). We show that $yPx$ implies $zPy$. There are three cases:

(1) $\min(X,P) \succ y$. Then $\min(X,P) \succ y \succ z$ so $zPy$.
(2) $\min(X,P) = y$. Then $xPy$, a contradiction to $yPx$.
(3) $y \succ \min(X,P)$. Then $x \succ y \succ \min(X,P)$, so $xPy$. A contradiction to $yPx$.

Conversely, assume that $P$ is quasiconvex. Assume that $\min(X,P) \succ y \succ x$ (the case $x \succ y \succ \min(X,P)$ is similar). Assume by contradiction that $yPx$ then quasiconvexity of $P$ implies that $\min(X,P)Py$. A contradiction. Hence $xPy$. □

C.4. **Proof that the binary SCRUM and strict Fechnerian model are disjoint.** We will in this section show that the strict Fechnerian model is disjoint from SCRUM provided that $X$ contains five or more alternatives. A binary SCF $\rho$ will be called a strict Fechnerian SCF if it satisfies the following two properties.

**Stochastic completeness (SC).** For all $x, y \in X$: $\rho(x,y) \neq \rho(y,x)$.

**Strong stochastic transitivity (SST).** For all $x, y, z \in X$: If $\rho(x,y) > \frac{1}{2}$ and $\rho(y,z) > \frac{1}{2}$ then

$$\rho(x,z) > \max\{\rho(x,y), \rho(y,z)\}.$$

It follows by the main result in Tversky and Russo (1969) that a binary SCF $\rho$ satisfies the above properties if and only if there is a strict utility function $u$ such that $\rho(x,y) = F(u(x), u(y))$ for all $x, y \in X$ where $F$ is a real valued function which is strictly increasing in the first argument and strictly decreasing in the second. The following lemma will be needed to show that the strict Fechnerian model is disjoint from SCRUM.

**Lemma C.4.** *Let $a, b, c, d, e \in X$ with $a \succ b \succ c \succ d \succ e$ and let $u$ be a strict utility function on $X$. Then there are three elements $x, y, z \in X$ such that $x \succ y \succ z$ and either $u(x) > u(y) > u(z)$ or $u(z) > u(y) > u(x)$.*

*Proof.* Assume by way of contradiction that there are no triplets of alternatives satisfying the conclusion of the lemma. W.l.o.g. we assume that $u(a) < u(b)$ (the other case is symmetric). Then $u(b) > u(c)$, which in turn implies that $u(c) < u(d)$, implying that $u(d) > u(e)$. Further, $u(a) < u(b)$ implies that $u(b) > u(d)$. But, the latter inequality implies that $u(d) < u(e)$. We have arrived at the contradiction that $u(d) > u(e)$ and $u(d) < u(e)$. □

It follows as a corollary to the proposition below that the set of strict Fechnerian SCFs is disjoint from SCRUM (provided $|X| \geq 5$).

**Proposition C.5.** *If $|X| \geq 5$ then there is no strict Fechnerian SCF that satisfies AST.*

*Proof.* Assume that $\rho$ satisfies SC and SST. It is then straightforward to verify that there is a strict utility function $u$ such that $u(x) > u(y)$ if and only if $\rho(x, y) > \frac{1}{2}$. Let $\succ$ be an order on $X$. Since $|X| \geq 5$ there are five alternatives $a, b, c, d, e \in X$ such that $a \succ b \succ c \succ d \succ e$. Lemma C.4 implies that there are three alternatives $x \succ y \succ z$ such that $u(x) > u(y) > u(z)$ or $u(z) > u(y) > u(x)$. If $u(x) > u(y) > u(z)$ then $\rho(x, z) > \max\{\rho(x, y), \rho(y, z)\}$. A contradiction to formulation AST' of AST in appendix A. If $u(z) > u(y) > u(x)$ then $\rho(z, x) > \max\{\rho(z, y), \rho(y, x)\}$ so $\rho(x, z) < \min\{\rho(x, y), \rho(y, z)\}$, a contradiction to AST. $\square$