SINGLE-CROSSING RANDOM UTILITY MODELS*

JOSE APESTEGUIA† AND MIGUEL A. BALLESTER‡

Abstract. We propose a novel model of stochastic choice: the single-crossing random utility model (SCRUM). This is a random utility model in which the collection of utility functions satisfies the single-crossing property. We offer a characterization of SCRUMs based on three easy-to-check properties: Positivity, Monotonicity and Centrality. The identified collection of utility functions and associated probabilities is basically unique. We establish a stochastic monotone comparative result for the case of SCRUMs and study several generalizations of SCRUMs.

Keywords: Stochastic Choice; Single-Crossing Property; Random Utility Models; Monotone Comparative Statics.
JEL classification numbers: D00.

1. Introduction

In a random utility model, there is a collection of utilities and a probability mass function over them. The probability of choosing an option from the set of available alternatives is described by the sum of the probability masses associated with the utility functions that consider that option to be maximal. This is a flexible model that can be interpreted from the perspective of both an individual and a group of individuals. In the first case, the different utility functions may stand for different criteria, selves, or moods of the individual, with their corresponding probabilities describing their prevalence. Accordingly, individual choice here is understood as stochastic in nature.[1] In the second case, the utilities represent individuals that differ in their tastes and, hence, in their evaluation of the alternatives, and the probability masses describe how

Date: March, 2016.
* Financial support from the 2015 BBVA Foundation Grant to Researchers, the Spanish Ministry of Science and Innovation (ECO2014-5614-P and ECO2014-53051-P), the Catalan Agency for Research (2014-SGR-515 and 2014-SGR-694) and MOVE is gratefully acknowledged.
† ICREA, Universitat Pompeu Fabra and Barcelona GSE. E-mail: jose.apesteguia@upf.edu.
‡ Universitat Autonoma de Barcelona and Barcelona GSE. E-mail: miguelangel.ballester@uab.es.
[1]Agranov and Ortoleva (2016) provide recent experimental evidence supporting the stochastic nature of individual behavior.

prevalent these utilities are in the population. Here, the probability distribution over choices describes the frequency with which the different options are selected in the population.

Under both interpretations, the multi-self individual and the group of individuals, the random utility model crucially allows for heterogeneity in preferences. A property that has proven of great practical relevance in introducing structure into the modeling of preference heterogeneity is the classical single-crossing condition (see Mirrlees, 1971; Spence, 1974; Milgrom and Shannon, 1994). The single-crossing condition essentially assumes that there is an order over the alternatives $\prec$ and a collection of utility functions that is ordered such that whenever $x^L \prec x^H$, lower utilities in the collection prefer x^L and higher utilities prefer x^H. This condition has been critical in a number of diverse and relevant settings. For example, it permits to consider optimal taxation and market signaling problems (Mirrlees, 1971; Spence, 1974); to provide sharp comparative statics results (Milgrom and Shannon, 1994); to solve the problem of preference aggregation (Gans and Smart, 1996); to characterize equilibria in incomplete information games (Athey, 2001); to study the value of information (Persico, 2000); to address a number of issues in political economy (Persson and Tabellini, 2000); and to gain a deep understanding of the fundamental preference parameters such as risk, time and altruism (see Jewitt (1987), Benoît and Ok (2007) and Cox, Friedman and Sadiraj (2008), respectively).

In this paper, we propose and study a random utility model in which preference heterogeneity is modeled by way of the single-crossing condition. That is, we impose that the collection of utility functions involved in the random utility model satisfies the single-crossing condition. The model, therefore, endows heterogeneity with an intuitive structure and is sufficiently flexible to apply to a wide variety of settings, as illustrated in the previous paragraph. We call such a model the *single-crossing random utility model* (SCRUM).

In our first result, we characterize SCRUMs by way of three simple properties, i.e., Positivity, Monotonicity and Centrality, providing testable foundations to the model. Positivity imposes that the choice probability of every available alternative is away from zero. This is a standard property in the stochastic choice literature. Monotonicity is another classical property; it simply states that the probability of choosing an option from a set should not increase as more alternatives are considered. Finally, Centrality is a new property that exploits the structure that the single-crossing condition brings to the model. Consider three ordered alternatives. Centrality imposes that the probability of choosing either one of the alternatives in the extremes does not depend on the presence of the other extreme alternative. The intuition is that all of the reasons for choosing an extreme alternative in the triplet are inherited by the central alternative when the former is absent. Note that Centrality uses only the triplets and the binary sets, and hence it is a computationally easy property to check. Theorem 1 shows that these properties characterize SCRUMs. Furthermore, the proof of Theorem 1 is constructive; the collection of utilities and their associated weights are obtained from the stochastic revealed choices. In addition, it shows that the identification is basically unique; the utilities are unique up to ordinal transformations.

Milgrom and Shannon (1994) first formally introduce the single-crossing condition, with the purpose of studying how optimal choices vary when using different utility functions. They establish their famous monotone comparative statics result, whereby the optimal alternative of a lower utility in the collection of single-crossing utilities precedes the optimal alternative with higher utility. In Section 4, we revisit the question of monotone comparative statics, in the stochastic approach of SCRUMs. We construct a partial order on SCRUMs that describes how close they are to the order of alternatives $\prec$, and in Theorem 2, we establish how this partial order is linked to the corresponding stochastic choices. Concretely, we show that SCRUM^L is closer to $\prec$ than SCRUM^H if and only if the stochastic choices generated by SCRUM^L are first-order stochastically dominated by those of SCRUM^H. This is the stochastic analogue of the classical monotone comparative statics result. In fact, we show that the classical result is simply a special case of Theorem 2.

Section 5 is devoted to the study of two extensions of SCRUMs. First, in Section 5.1, we investigate the case in which there is not an exogenous, observable order over the alternatives. It may be the case that the analyst does not have enough information to assume a particular order over the alternatives, but the decision-maker actually contemplates one. In this case, the question arises as to whether there exists an order over the alternatives such that the revealed stochastic choice is a SCRUM with respect to it. In Theorem 3, we identify the properties that allow us to address this question. We show that Positivity, Monotonicity, and a slight variation of Centrality characterize this endogenous version of SCRUMs. Again, the proof is constructive; the unique endogenous order over the alternatives, up to symmetry, is obtained from the stochastic choices involving the triplets and the binary sets. Then, Theorem 1 can be used to construct the unique collection of utilities and associated probability weights that rationalize stochastic choice.

In Section 5.2, we extend SCRUMs to contemplate the possibility of dominated alternatives, that is, situations in which, no matter what, the decision-maker always finds one option superior to another. Accordingly, one would like to allow for the possibility of choosing the dominated alternative with zero probability, whenever the dominating alternative is available. Because this is in contrast with the property of Positivity, we need to introduce a variation of the property to allow for zero probabilities in these cases. Theorem 4 shows that Monotonicity, Centrality and a variation of Positivity characterize SCRUMs that allow for the presence of dominated alternatives.

Finally, Section 6 concludes by commenting on the relationship between SCRUMs and certain behavioral phenomena that has attracted a good deal of attention.

We close this section by relating our work to the relevant literature. Random utility models have a long tradition in economics. Block and Marschak (1960) provide an early and deep theoretical treatment of them but leave the characterization of the model as an open question. In subsequent contributions, Falmagne (1978), Barberà and Pattanaik (1986), and McFadden and Richter (1990) are able to solve the challenge posed by Block and Marschak and offer a full characterization of the model. However, the nature of the characterizations is algorithmic, and hence the properties are difficult to interpret and to operationalize. Notably, Gul and Pesendorfer (2006) revisit this question. They characterize the case in which alternatives are lotteries and the collection of utilities is formed by expected utility functions, and use properties that exploit the structure of expected utility. By doing so, they are able to provide intuitive properties that are analogues of the standard properties in the deterministic study of decision under risk and are easy to interpret. Recently, Lu and Saito (2016) provide foundations for the case where alternatives are consumption streams and utilities are discounted utility functions. In this paper, we contribute to the study of random utility models by endowing them with a flexible structure that makes them applicable to a number of diverse settings, and as in Gul and Pesendorfer and Lu and Saito, the special structure makes the model tractable and testable.

There are a number of recent papers studying variations of the model of Luce (1959), which is probably the most popular probabilistic choice model. Recently, Fudenberg and Strzalecki (2013) characterize a dynamic version of the Luce model. Gul, Natenzon, and Pesendorfer (2014) extend the Luce model to the consideration of stochastic-attribute-based choice. Fudenberg, Iijima and Strzalecki (2015) relax Luce's IIA axiom, the key axiom characterizing the model of Luce, to consider nonlinear perturbations of utility. In Section 3.2, we discuss how Luce-type models are substantially different from SCRUMs. Manzini and Mariotti (2014) study a decision-maker that first pays attention to a subset of the available alternatives following a random procedure and then maximizes a preference relation. Again, in Section 3.2, we argue that this model and ours are fundamentally different. Finally, Caplin and Dean (2016) characterize every revealed stochastic choice consistent with the optimal acquisition of costly information. Their model is very different in nature from all of these stochastic models, including ours, as theirs uses state-dependent choice data.

2. Single-Crossing Random Utility Models

Let X be a finite set of alternatives. A *stochastic choice function* is a mapping $p : X \times 2^X \setminus \emptyset \to [0,1]$ such that, for every menu $A \in 2^X \setminus \emptyset$, the following properties hold: (i) $p(x, A) > 0$ only if $x \in A$ and (ii) $\sum_{x \in A} p(x, A) = 1$. We interpret $p(x, A)$ as the probability of choosing alternative x from menu $A \subseteq X$.

The key notion in the paper is that of a *single-crossing random utility model* (SCRUM). To introduce it, we borrow two ingredients from the literature on the single-crossing condition. First, we assume that there is a given observable linear order $\prec$ on X.[2] Second, there is a collection of utility functions $\{U_t\}_{t=1}^T$ on X satisfying the single-crossing condition with respect to $(X, \prec)$. Namely, for every pair of alternatives $x^L \prec x^H$ and pair of utilities U_{t^L} and U_{t^H}, with $t^L < t^H$, if $U_{t^L}(x^L) < U_{t^L}(x^H)$, then $U_{t^H}(x^L) < U_{t^H}(x^H)$. That is, the single-crossing condition states that the ranking of any pair of alternatives reverses at most once in the ordered collection $\{U_t\}_{t=1}^T$. Then, low utilities prefer the low alternative, while high utilities prefer the high alternative, and there are no indifferences in utility functions. We begin by studying the case of undominated alternatives, by assuming that for every $x \in X$, there is t such that $U_t(x) > U_t(y)$ for every $y \in X \setminus \{x\}$.[3] Settings fitting these assumptions abound and

[2]Below, in Section 5.1, we show how we can dispense with this assumption and identify the structure of alternatives perceived by the decision-maker by employing the revealed choices.

[3]We relax this assumption in Section 5.2, when we incorporate into the analysis the possibility of the existence of dominated alternatives.

cover all the main preference parameters of interest. The following example illustrates this point.

Example 1. Consider a collection of vectors $(x_1^1, x_1^2), (x_2^1, x_2^2), \dots, (x_n^1, x_n^2)$ that lie in the same downward-slopping line. That is, there exist δ^1 and δ^2, with $\delta^1 \geq \delta^2 > 0$, such that $\delta^1 x_i^1 + \delta^2 x_i^2 = 1$. To facilitate the exposition, let $0 < x_1^1 < x_2^1 < \cdots < x_n^1$.

- *Risk.* $X = \{g_1, \dots, g_n\}$, where gamble g_i assigns equiprobable prizes x_i^1 and x_i^2, with $x_n^1 \leq \frac{1}{\delta^1+\delta^2}$.[4] Take $\{U_t\}_{t=1}^T$ to be a collection of CARA or CRRA expected utilities, ordered by the risk-aversion coefficient. This model is single-crossing, with gambles ordered as $g_1 \prec g_2 \prec \cdots \prec g_n$.
- *Time.* $X = \{s_1, \dots, s_n\}$, where stream s_i has present payout x_i^1 and future payout x_i^2. Let u be a strictly increasing monetary utility, and let $\{U_t\}_{t=1}^T$ be a collection of exponential discounted utilities using u and ordered by the delay-aversion coefficient. This model is single-crossing, with streams ordered as $s_1 \prec s_2 \prec \cdots \prec s_n$.
- *Social preferences.* $X = \{a_1, \dots, a_n\}$, where allocation a_i consists of a payout of x_i^1 dollars to oneself and x_i^2 dollars to another person. Consider a family of Andreoni-Miller altruism-CES utility functions $\{U_t\}_{t=1}^T$ with a fixed substitutability coefficient and ordered by the altruism coefficient. This model is single-crossing, with allocations ordered as $a_n \prec a_{n-1} \prec \cdots \prec a_1$.
- *Complementarities.* $X = \{b_1, \dots, b_n\}$, where bundle b_i is described by the consumption of good 1, x_i^1 and good 2, x_i^2, with $x_n^1 \leq \frac{1}{\delta^1+\delta^2}$. Consider a family of CES utility functions $\{U_t\}_{t=1}^T$ with a fixed share coefficient of $\frac{1}{2}$ and ordered by the substitutability coefficient. This model is single-crossing, with bundles ordered as $b_n \prec b_{n-1} \prec \cdots \prec b_1$. ♦

The last ingredient of a SCRUM is a strictly positive probability mass function μ over the set of utility functions $\{U_t\}_{t=1}^T$. The value $\mu(t)$ describes the probability with which the utility function U_t is realized, or in other words, it represents the weight, or the prevalence, of U_t in $\{U_t\}_{t=1}^T$.

We can now define the stochastic choice function associated with a SCRUM. Consider a menu $A \subseteq X$ and a utility function U_t, and denote by $m_t(A)$ the maximal alternative in A according to U_t, that is, $U_t(m_t(A)) > U_t(x)$ for every $x \in A \setminus \{m_t(A)\}$. The

[4]This guarantees that no gamble stochastically dominates another. An analogous argument applies to the case of complementarities below.

probability that the SCRUM stochastic choice function associates to the choice of x from A is simply the sum of the weights of all the utility functions for which x is maximal in A. Formally, for every menu A and alternative x, the SCRUM stochastic choice function is $p(x, A) = \sum_{t:x=m_t(A)} \mu(t)$.

3. Characterization of Single-Crossing Random Utility Models

3.1. **Characterization Result.** We now introduce a set of properties on stochastic choice functions that identify SCRUMs. The first property, Positivity, simply states that the probability of choosing any option in a menu is strictly larger than zero. This is a mild property, often imposed in the definition of stochastic choice models. Here we impose it as a property and show that, in conjunction with the two properties below, it is not only necessary but sufficient. In Section 5.2, we modify this property to address the case of dominated alternatives.

Positivity (POS). If $x \in A$, $p(x, A) > 0$.

The second property, Monotonicity, is a classical condition in the study of stochastic choice already employed in Block and Marschak (1960). It states that the probability of selecting an option does not increase when more alternatives are added to the menu.

Monotonicity (MON). If $B \subseteq A$, then $p(x, A) \leq p(x, B)$.

The third property, Centrality, uses the structure that the single-crossing condition brings to random utility models. It states that, in a triplet, the central alternative, according to the given order $\prec$, makes the two other extreme alternatives irrelevant to one another. Intuitively, given the ordered structure of the alternatives, the arguments for choosing an extreme alternative in the triplet are inherited by the central alternative when the former is absent.

Centrality (CEN). Let $x^L \prec x^C \prec x^H$. Then, $p(x^L, \{x^L, x^C, x^H\}) = p(x^L, \{x^L, x^C\})$ and $p(x^H, \{x^L, x^C, x^H\}) = p(x^H, \{x^C, x^H\})$.

It is clear that these three properties are necessarily satisfied by a SCRUM. Theorem 1 shows that these properties are not only necessary but sufficient. Given the simplicity of the properties, Theorem 1 shows, therefore, that SCRUMs are easily testable.

Theorem 1. *A stochastic choice function p satisfies POS, MON and CEN if and only if p is a SCRUM stochastic choice function.*

Proof of Theorem 1: The necessity of the axioms is straightforward. We prove the sufficiency of the axioms through a series of steps. To ease the exposition, we denote the alternatives in X by $1 \prec 2 \prec \cdots \prec |X|$.

Step 1. We claim that for every triplet of alternatives $\{i, j, k\}$ such that $i \prec j \prec k$, it is $p(i, \{i, j\}) \leq p(i, \{i, k\}) \leq p(j, \{j, k\})$. To see the first inequality, notice that by CEN, it must be $p(i, \{i, j\}) = p(i, \{i, j, k\})$. By MON, it must be $p(i, \{i, j, k\}) \leq p(i, \{i, k\})$, which implies $p(i, \{i, j\}) \leq p(i, \{i, k\})$. For the second inequality, notice that MON guarantees $p(i, \{i, k\}) = 1 - p(k, \{i, k\}) \leq 1 - p(k, \{i, j, k\})$. By CEN, it must also be $p(k, \{i, j, k\}) = p(k, \{j, k\})$, which implies $p(i, \{i, k\}) \leq 1 - p(k, \{j, k\}) = p(j, \{j, k\})$.

Step 2. We construct a collection of utility functions $\{U_t\}_{t=1}^{T}$ over X and assign a probability mass to each of them. Let $\Lambda = \{\lambda :$ there exist i, j such that $i \prec j$ and $p(i, \{i, j\}) = \lambda\} \cup \{1\}$. Denote the elements of Λ by $\lambda_1 < \lambda_2 < \cdots < \lambda_{|\Lambda|} = 1$. We construct $T = |\Lambda|$ utility functions. The utility of alternative i under utility function U_t is $U_t(i) = -i + |\{k : k \prec i, p(k, \{k, i\}) < \lambda_t\}| - |\{k : i \prec k, p(i, \{i, k\}) < \lambda_t\}|$. The mass of utility U_t is $\mu(t) = \lambda_t - \lambda_{t-1}$, where $\lambda_0 = 0$. Notice that POS guarantees that $\lambda_1 > 0$, and hence all masses are strictly positive.

Step 3. We claim that the collection of utility functions $\{U_t\}_{t=1}^{T}$ satisfies the single-crossing condition with respect to $\prec$. For every pair of distinct alternatives $i \prec j$, we define $t(i, j)$ as the integer such that $\lambda_{t(i,j)} = p(i, \{i, j\})$ and show that $U_t(i) > U_t(j)$ whenever $t \leq t(i, j)$ and $U_t(i) < U_t(j)$ whenever $t > t(i, j)$.

Step 3a. Let $t \leq t(i, j)$. Notice that this is equivalent to $p(i, \{i, j\}) \geq \lambda_t$. For any alternative k such that $k \prec i$, Step 1 guarantees that $p(k, \{k, i\}) \leq p(k, \{k, j\})$, and hence $|\{k : k \prec i, p(k, \{k, i\}) < \lambda_t\}| \geq |\{k : k \prec i, p(k, \{k, j\}) < \lambda_t\}|$. For any alternative k such that $j \prec k$, Step 1 guarantees that $p(j, \{j, k\}) \geq p(i, \{i, k\}) \geq p(i, \{i, j\}) \geq \lambda_t$, and hence $-|\{k : j \prec k, p(i, \{i, k\}) < \lambda_t\}| = 0 = -|\{k : j \prec k, p(j, \{j, k\}) < \lambda_t\}|$. Finally, for any alternative k such that $i \prec k \prec j$, Step 1 guarantees that $p(k, \{k, j\}) \geq p(i, \{i, j\}) \geq \lambda_t$. That is, $-i - |\{k : i \prec k \prec j, p(i, \{i, k\}) < \lambda_t\}| > -j = -j + |\{k : i \prec k \prec j, p(k, \{k, j\}) < \lambda_t\}|$. These three facts, together with $p(i, \{i, j\}) \geq \lambda_t$, lead to $U_t(i) > U_t(j)$, as desired.

Step 3b. Let $t > t(i, j)$. Notice that this is equivalent to $p(i, \{i, j\}) < \lambda_t$. For any alternative k such that $k \prec i$, Step 1 guarantees that $p(k, \{k, i\}) \leq p(k, \{k, j\}) \leq p(i, \{i, j\}) < \lambda_t$, and hence $|\{k : k \prec i, p(k, \{k, j\}) < \lambda_t\}| = i - 1 = |\{k : k \prec i, p(k, \{k, i\}) < \lambda_t\}|$. For any alternative k such that $j \prec k$, Step 1 guarantees that $p(i, \{i, k\}) \leq p(j, \{j, k\})$, which implies $-|\{k : j \prec k, p(j, \{j, k\}) < \lambda_t\}| \geq -|\{k :$

$j \prec k, p(i,\{i,k\}) < \lambda_t\}|$. Finally, for any alternative k such that $i \prec k \prec j$, Step 1 guarantees that $p(i,\{i,k\}) \leq p(i,\{i,j\}) < \lambda_t$, which implies $-j + |\{k : i \prec k \prec j, p(k,\{k,j\}) < \lambda_t\}| + 1 \geq -j+1 = -i - |\{k : i \prec k \prec j, p(i,\{i,k\}) < \lambda_t\}| > -i - |\{k : i \prec k \prec j, p(i,\{i,k\}) < \lambda_t\}| - 1$. These three facts, together with $p(i,\{i,j\}) < \lambda_t$, lead to $U_t(j) > U_t(i)$, as desired.

Step 4. We claim that for any set A such that $|A| \geq 3$, denoting its alternatives by $a_1 \prec a_2 \prec \cdots \prec a_{|A|}$, it is $p(a_1, A) = p(a_1, \{a_1, a_2\})$, $p(a_i, A) = p(a_i, \{a_{i-1}, a_i, a_{i+1}\})$ whenever $2 \leq i \leq |A|-1$ and $p(a_{|A|}, A) = p(a_{|A|}, \{a_{|A|-1}, a_{|A|}\})$. To show this, we use CEN repeatedly to obtain $1 = p(a_1,\{a_1,a_2,a_3\}) + p(a_2,\{a_1,a_2,a_3\}) + p(a_3,\{a_1,a_2,a_3\}) = p(a_1,\{a_1,a_2\}) + p(a_2,\{a_1,a_2,a_3\}) + p(a_3,\{a_2,a_3\}) = p(a_1,\{a_1,a_2\}) + p(a_2,\{a_1,a_2,a_3\}) + (1 - p(a_2,\{a_2,a_3\})) = p(a_1,\{a_1,a_2\}) + p(a_2,\{a_1,a_2,a_3\}) + (1 - p(a_2,\{a_2,a_3,a_4\})) = p(a_1,\{a_1,a_2\}) + p(a_2,\{a_1,a_2,a_3\}) + p(a_3,\{a_2,a_3,a_4\}) + p(a_4,\{a_2,a_3,a_4\}) = \ldots$

$\cdots = p(a_1,\{a_1,a_2\}) + \sum_{i=2}^{|A|-1} p(a_i,\{a_{i-1},a_i,a_{i+1}\}) + p(a_{|A|},\{a_{|A|-1},a_{|A|}\})$. By MON, it is $p(a_1,\{a_1,a_2\}) + \sum_{i=2}^{|A|-1} p(a_i,\{a_{i-1},a_i,a_{i+1}\}) + p(a_{|A|},\{a_{|A|-1},a_{|A|}\}) \geq \sum_{i=1}^{|A|} p(a_i, A) = 1$. Then, every summand involved in the left-hand side of the inequality must be equal to its corresponding counterpart in the right-hand side of the inequality, which proves the claim.

Step 5. Finally, we show that $p(x, A) = \sum_{t:x=m_t(A)} \mu(t)$ for every menu A and alternative x. Note that, by POS, this shows, in turn, that every alternative is maximal for at least one of the utility functions, concluding the proof. If $x \notin A$, the result is trivial by the definition of stochastic choice functions and maximal elements. If $|A| < 3$, the result is trivial by construction. We then assume $x \in A$, $|A| \geq 3$ and use the same notation as in Step 4.

Step 5a. Let $x = a_1$. By Steps 2, 3 and 4, it is $p(a_1, A) = p(a_1, \{a_1, a_2\}) = \lambda_{t(a_1,a_2)} = \sum_{t=1}^{t(a_1,a_2)} (\lambda_t - \lambda_{t-1}) = \sum_{t=1}^{t(a_1,a_2)} \mu(t)$. We know that for every $t \leq t(a_1,a_2)$, it is $U_t(a_1) > U_t(a_2)$, and for any a_j with $j > 2$, Step 1 guarantees that $t(a_1,a_j) \geq t(a_1,a_2)$, and hence it is also $U_t(a_1) > U_t(a_j)$. Thus, $a_1 = m_t(A)$. For every $t > t(a_1,a_2)$, it is $U_t(a_1) < U_t(a_2)$, and hence $a_1 \neq m_t(A)$. Thus, $a_1 = m_t(A)$ if and only if $t \leq t(a_1,a_2)$, which leads to $p(a_1, A) = \sum_{t=1}^{t(a_1,a_2)} \mu(t) = \sum_{t:a_1=m_t(A)} \mu(t)$, as desired.

Step 5b. Let $x = a_{|A|}$. By Steps 2, 3 and 4, it is $p(a_{|A|}, A) = p(a_{|A|}, \{a_{|A|-1}, a_{|A|}\}) = 1 - p(a_{|A|-1}, \{a_{|A|-1}, a_{|A|}\}) = 1 - \lambda_{t(a_{|A|-1},a_{|A|})} = 1 - \sum_{t=1}^{t(a_{|A|-1},a_{|A|})} (\lambda_t - \lambda_{t-1}) = 1 - \sum_{t=1}^{t(a_{|A|-1},a_{|A|})} \mu(t) = \sum_{t=t(a_{|A|-1},a_{|A|})+1}^{T} \mu(t)$. We know that for every $t \leq t(a_{|A|-1}, a_{|A|})$, it is $U_t(a_{|A|-1}) > U_t(a_{|A|})$, and hence $a_{|A|} \neq m_t(A)$. For every $t > t(a_{|A|-1}, a_{|A|})$, it is $U_t(a_{|A|-1}) < U_t(a_{|A|})$, and also, for any a_j with $j < |A| - 1$, Step 1 guarantees that $t(a_j, a_{|A|}) \leq t(a_{|A|-1}, a_{|A|})$, and hence it is also $U_t(a_j) < U_t(a_{|A|})$ and $a_{|A|} = m_t(A)$. Thus, $a_{|A|} = m_t(A)$ if and only if $t > t(a_{|A|-1}, a_{|A|})$, which leads to $p(a_{|A|}, A) = \sum_{t=t(a_{|A|-1},a_{|A|})+1}^{T} \mu(t) = \sum_{t:a_{|A|}=m_t(A)} \mu(t)$, as desired.

Step 5c. Let $x = a_i$, with $1 < i < |A|$. Step 4 and CEN guarantee that $p(a_i, A) = p(a_i, \{a_{i-1}, a_i, a_{i+1}\}) = 1-p(a_{i-1}, \{a_{i-1}, a_i, a_{i+1}\})-p(a_{i+1}, \{a_{i-1}, a_i, a_{i+1}\}) = 1-p(a_{i-1}, \{a_{i-1}, a_i\})-p(a_{i+1}, \{a_i, a_{i+1}\}) = p(a_i, \{a_i, a_{i+1}\})-p(a_{i-1}, \{a_{i-1}, a_i\})$, which is simply $\sum_{t=1}^{t(a_i,a_{i+1})} \mu(t)-\sum_{t=1}^{t(a_{i-1},a_i)} \mu(t) = \sum_{t=t(a_{i-1},a_i)+1}^{t(a_i,a_{i+1})} \mu(t)$. By Step 1, it is $t(a_j, a_i) \leq t(a_{i-1}, a_i)$ whenever $j \leq i-1$ and $t(a_i, a_j) \geq t(a_i, a_{i+1})$ whenever $j \geq i+1$, and then $a_i = m_t(A)$ if and only if $t(a_{i-1}, a_i) + 1 \leq t \leq t(a_i, a_{i+1})$. This implies $p(a_i, A) = \sum_{t=t(a_{i-1},a_i)+1}^{t(a_i,a_{i+1})} \mu(t) = \sum_{t:a_i=m_t(A)} \mu(t)$, as desired. ■

The proof of Theorem 1 explicitly constructs the pair $(\{U_t\}_{t=1}^T, \mu)$ that makes p a SCRUM stochastic choice function. The construction is intuitive and easy to implement in practice, as it is exclusively based on the order $\prec$ and the revealed stochastic choices in the binary sets. It proceeds as follows. First, construct U_1 as a utility representation of $\prec$. That is, $x \prec y$ if and only if $U_1(x) > U_1(y)$. Then, identify the pair of alternatives $\{x, y\}$ such that $x \prec y$ and has the lowest binary choice probability $p(x, \{x, y\})$. Construct U_2 by reproducing U_1 but reversing the utility values of alternatives x and y: $U_2(x) = U_1(y)$, $U_2(y) = U_1(x)$, and $U_2(a) = U_1(a)$ for any other alternative a. Once U_t is constructed, identify the pair of alternatives $\{z, w\}$, not considered before, such that $z \prec w$ and has the lowest binary choice probability $p(z, \{z, w\})$. Construct U_{t+1} from U_t by reversing the utility values of alternatives z and w. Proceed in this ordered way until exhausting all binary comparisons.[5] Regarding the associated probability masses, define $\mu(1)$ as the lowest binary choice probability $p(x, \{x, y\})$. Given $\mu(1), \ldots, \mu(t)$, define $\mu(t+1) = p(z, \{z, w\}) - \sum_{s=1}^{t} \mu(s)$. Proceed in this way, assigning to the last utility function the mass $\mu(T) = 1 - \sum_{s=1}^{T-1} \mu(s)$. The proof then shows that this construction satisfies the single-crossing condition, and that rationalizes the stochastic choice function p in the SCRUM sense.

Another important and notable aspect of the characterization result is that the identification result is tight. The pair $(\{U_t\}_{t=1}^T, \mu)$ is unique up to a set of basic transformations. First, note that replacing a utility function U_t with any ordinal transformation

[5]The proof reverses only consecutive alternatives in the utility functions and analyzes how, in the more delicate case in which more than one binary comparison must be reversed simultaneously, this procedure works correctly.

U'_t, respecting its mass $\mu(t)$, is inconsequential. Second, one can create copies of U_t, place them in between U_{t-1} and U_{t+1}, and distribute the mass $\mu(t)$ among the copies without modifying choice probabilities. Any other transformation would lead to a different stochastic choice function. Hence, the SCRUM $(\{U_t\}_{t=1}^{T}, \mu)$ rationalizing the stochastic choice function is generically unique. This is in contrast to unrestricted random utility models, which are well-known to have multiple representations (see Fishburn, 1998).

3.2. **Relation to Luce's IIA axiom and Stochastic Transitivity.** We now illustrate the relationship between SCRUMs and two classical properties in the study of stochastic choice. Let us start with Luce's Independence of Irrelevant Alternatives axiom (Luce-IIA). This is the key property characterizing the well-known model of Luce (1959). It essentially requires that the choice ratio of two alternatives is independent of the other alternatives that may be available, i.e., $\frac{p(x,A)}{p(y,A)} = \frac{p(x,B)}{p(y,B)}$. Clearly, Luce-IIA is not necessarily satisfied by SCRUMs, as the property is in direct conflict with CEN. Notice that when eliminating x^H from $\{x^L, x^C, x^H\}$, $p(x^H, \{x^L, x^C, x^H\})$ must be inherited by x^C according to CEN but distributed proportionally between x^L and x^C according to Luce-IIA. Contrary to CEN, Luce-IIA pays no attention whatsoever to the structure that alternatives have. This, in turn, shows that models extending the Luce model have a different structure from SCRUMs.

Stochastic transitivity is a cornerstone concept in the understanding of the models of stochastic choice. The literature distinguishes among three versions of the property, namely, weak, moderate and strong stochastic transitivity. To formally define them, consider three distinct alternatives x, y and z such that $p(x, \{x, y\}) \geq \frac{1}{2}$ and $p(y, \{y, z\}) \geq \frac{1}{2}$. Stochastic transitivity notions require $p(x, \{x, z\}) \geq \Psi$, where $\Psi = \frac{1}{2}$ for the weak notion, $\Psi = \min\{p(x, \{x, y\}), p(y, \{y, z\})\}$ for the moderate notion and $\Psi = \max\{p(x, \{x, y\}), p(y, \{y, z\})\}$ for the strong notion.[6]

It is well known that the standard random utility model does not satisfy even the weak version of stochastic transitivity (Block and Marschak, 1960). To illustrate, consider a Condorcet cycle. There are three utility functions U, V and W defined over three alternatives x, y and z, with: $U(x) > U(y) > U(z)$, $V(y) > V(z) > V(x)$ and $W(z) > W(x) > W(y)$. Assume $\mu(U) = \mu(V) = \mu(W) = \frac{1}{3}$. Then, it is clear that $p(x, \{x, y\}) = p(y, \{y, z\}) = \frac{2}{3} \geq \frac{1}{2}$ but $p(x, \{x, z\}) = \frac{1}{3}$, violating the three versions of

[6]Notice that the strong notion implies the moderate notion and the moderate notion implies the weak one.

stochastic transitivity. Interestingly, we now prove that the single-crossing condition endows random utility models with the moderate version of stochastic transitivity.[7]

Proposition 1. *Every SCRUM stochastic choice function satisfies Moderate Stochastic Transitivity.*

Proof of Proposition 1: Consider three distinct alternatives x, y and z and suppose that $p(x,\{x,y\}) \geq \frac{1}{2}$ and $p(y,\{y,z\}) \geq \frac{1}{2}$. We first prove that z cannot be central in the triplet $\{x,y,z\}$ according to $\prec$. Otherwise, the use of CEN and POS would imply that $p(x,\{x,z\}) = p(x,\{x,y,z\}) = 1 - p(y,\{x,y,z\}) - p(z,\{x,y,z\}) = 1 - p(y,\{y,z\}) - p(z,\{x,y,z\}) \leq \frac{1}{2} - p(z,\{x,y,z\}) < \frac{1}{2}$, which is absurd. Suppose now that x is the central alternative in $\{x,y,z\}$ according to $\prec$. In this case, CEN implies $p(x,\{x,z\}) = 1 - p(z,\{x,z\}) = 1 - p(z,\{x,y,z\})$. MON implies that $p(z,\{x,y,z\}) \leq p(z,\{y,z\})$, and hence $p(x,\{x,z\}) \geq 1 - p(z,\{y,z\}) = p(y,\{y,z\})$, as desired. Finally, suppose that y is the central alternative. In this case, it can be immediately seen that MON and CEN imply $p(x,\{x,z\}) \geq p(x,\{x,y,z\}) = p(x,\{x,y\})$, as desired. ■

It is easy to see that the strong version of stochastic transitivity is not satisfied by SCRUMs. To illustrate, consider three utility functions $\{U_t\}_{t=1}^{3}$ over three alternatives x, y and z, with: $U_1(y) > U_1(x) > U_1(z)$, $U_2(x) > U_2(y) > U_2(z)$ and $U_3(z) > U_3(x) > U_3(y)$. Let $\mu(1) = \frac{1}{5}$ and $\mu(2) = \mu(3) = \frac{2}{5}$. This is clearly a SCRUM, but $p(x,\{x,y\}) = \frac{4}{5}$, $p(y,\{y,z\}) = \frac{3}{5}$ and $p(x,\{x,z\}) = \frac{3}{5}$.

4. Stochastic Monotone Comparative Statics for SCRUMs

The single-crossing condition has been instrumental in establishing the so-called monotone comparative statics results. These are results characterizing the relationship between preference parameters and optimal choices. In the classical result, $\{U_t\}_{t=1}^{T}$ is a collection of utilities satisfying the single-crossing condition with respect to $(X,\prec)$, and it establishes that for every pair of utilities U_{t^L} and U_{t^H}, with $t^L < t^H$, the maximal alternative of U_{t^L} either precedes that of U_{t^H} in $\prec$, or they are the same alternative. This result has proven to be of great practical importance in a variety of applications,

[7] It can be shown that, under POS, neither MON nor CEN by themselves imply moderate stochastic transitivity. Moreover, note that this result implies that the model of Manzini and Mariotti (2014) is not a special case of ours, as theirs does not satisfy even weak stochastic transitivity. To see that SCRUMs are not a special case of their model, simply notice that their property i-Asymmetry is in direct conflict with Centrality.

such as political economy, labor economics, and the economics of education. In this section, we revisit the question of monotone comparative statics results from a stochastic perspective. The aim is to consider two SCRUMs that are ordered in terms of their collection of utilities and to establish monotone comparative results in their distributions of stochastic choices.

We then begin by constructing a partial order on SCRUMs that is based on how close their collections of utilities are to the order of alternatives $\prec$. To do so, we first introduce a partial order over utility functions. Given two utility functions U and V on X, we say that U is closer to $\prec$ than V, and write $U \trianglelefteq V$, if $i \prec j$ and $V(i) > V(j)$ imply $U(i) > U(j)$. In other words, U agrees with $\prec$ in the same pairwise comparisons as V, and maybe more. Note that the relation $\trianglelefteq$ is ingrained in the structure of SCRUMs; the construction of SCRUMs in the proof of Theorem 1 shows that for every U_{t^L} and U_{t^H}, with $t^L < t^H$, it is the case that $U_{t^L} \trianglelefteq U_{t^H}$. We now use the notion of the closeness relation $\trianglelefteq$ to partially order SCRUMs. To do so, we need to introduce some new notation. Note that we can alternatively describe the SCRUM $\alpha = (\{U_t\}_{t=1}^T, \mu)$ by way of a mapping F_α, from $[0,1]$ to utilities, given by $F_\alpha(\omega) = U_1$ whenever $\omega \in [0, \mu(1)]$ and $F_\alpha(\omega) = U_t$ whenever $\omega \in (\sum_{s=1}^{s=t-1} \mu(s), \sum_{s=1}^{s=t} \mu(s)]$. Now, given two SCRUMs $\alpha = (\{U_t\}_{t=1}^T, \mu)$ and $\beta = (\{V_t\}_{t=1}^S, \rho)$, we say that F_α is closer to $\prec$ than F_β, and write $F_\alpha \trianglelefteq F_\beta$, whenever $F_\alpha(\omega) \trianglelefteq F_\beta(\omega)$ for every $\omega \in [0,1]$. That is, the probability with which utilities closer to $\prec$ are realized in α is larger than in β.

Having established a partial order on SCRUMs, we need to do so on the stochastic choices of the models. Denote by p_α and p_β the stochastic choice functions associated to the SCRUMs α and β. Without loss of generality, let $1 \prec 2 \prec \cdots \prec |X|$. We say that p_α is first-order stochastically dominated by p_β if for every menu A and for every $i \in \{1, 2, \ldots, |X|\}$, it is $\sum_{j=1}^{i} p_\alpha(j, A) \geq \sum_{j=1}^{i} p_\beta(j, A)$. That is, the probability with which lower alternatives in $\prec$ are selected is larger in p_α than in p_β.

We are now in a position to establish our stochastic monotone comparative statics result. Theorem 2 shows that two SCRUMs are ordered in terms of the closeness relation if and only if their corresponding stochastic choices are ordered in terms of first-order stochastic dominance.[8]

Theorem 2. *$F_\alpha \trianglelefteq F_\beta$ if and only if p_α is first-order stochastically dominated by p_β.*

[8]In addition, Theorem 2 shows that SCRUMs are free from the non-monotonicity problems discuss in Apesteguia and Ballester (2016).

Proof of Theorem 2: We prove first that $F_\alpha \trianglelefteq F_\beta$ implies that p_α is first-order stochastically dominated by p_β, through a series of steps.

Step 1. We claim that for every menu A, if $U \trianglelefteq V$, then either $m_U(A) \prec m_V(A)$ or $m_U(A) = m_V(A)$. To see this, suppose by way of contradiction that $m_V(A) \prec m_U(A)$. The definition of a maximal element implies that $V(m_V(A)) > V(m_U(A))$. Given that $U \trianglelefteq V$, it must also be the case that $U(m_V(A)) > U(m_U(A))$, which is a contradiction.

Step 2. Denote by $t_\alpha(i,A)$ and $t_\beta(i,A)$ the last integers for which i is maximal in A according to α and β, respectively. We claim that $\sum_{j=1}^{i} p_\alpha(j,A) = \sum_{t=1}^{t_\alpha(i,A)} \mu(t)$ and $\sum_{j=1}^{i} p_\beta(j,A) = \sum_{t=1}^{t_\beta(i,A)} \rho(t)$. The construction of utilities in a SCRUM shows that $t^L < t^H$ implies $U_{t^L} \trianglelefteq U_{t^H}$. We can then use Step 1 to conclude the proof of this claim.

Step 3. By definition, $F_\beta(\sum_{t=1}^{t_\beta(i,A)} \rho(t)) = V_{t_\beta(i,A)}$. By assumption, $F_\alpha(\sum_{t=1}^{t_\beta(i,A)} \rho(t)) \trianglelefteq F_\beta(\sum_{t=1}^{t_\beta(i,A)} \rho(t))$, and hence, by Step 1, the maximal alternative in A according to $F_\alpha(\sum_{t=1}^{t_\beta(i,A)} \rho(t))$ precedes or equals alternative i, which is the maximal alternative in A according to $F_\beta(\sum_{t=1}^{t_\beta(i,A)} \rho(t))$. Then, $\sum_{t=1}^{t_\alpha(i,A)} \mu(t) \geq \sum_{t=1}^{t_\beta(i,A)} \rho(t)$, and therefore, by Step 2, $\sum_{j=1}^{i} p_\alpha(j,A) \geq \sum_{j=1}^{i} p_\beta(j,A)$, as desired.

We now prove that whenever p_α is first-order stochastically dominated by p_β, it must be $F_\alpha \trianglelefteq F_\beta$. Suppose by contradiction that this is not the case. Then, there exists ω such that it is not true that $F_\alpha(\omega) \trianglelefteq F_\beta(\omega)$. In other words, there exist $i \prec j$ such that $F_\beta(\omega)(i) > F_\beta(\omega)(j)$ but $F_\alpha(\omega)(i) < F_\alpha(\omega)(j)$. Given that p_α and p_β are SCRUM stochastic choice functions, it is immediately clear, from the construction of F, that $p_\alpha(i,\{i,j\}) < \omega \leq p_\beta(i,\{i,j\})$, proving that p_α is not first-order stochastically dominated by p_β, a contradiction. This concludes the proof. ■

The classical deterministic monotone comparative statics result is implied by Steps 1 and 2 in the proof of Theorem 2. Namely, we show there that in a single-crossing collection of utilities $\{U_t\}_{t=1}^T$, every pair of utilities U_{t^L} and U_{t^H}, with $t^L < t^H$, it is $U_{t^L} \trianglelefteq U_{t^H}$, and this in turn implies that $m_{U_{t^L}}(A) \prec m_{U_{t^H}}(A)$ or $m_{U_{t^L}} = m_{U_{t^H}}(A)$, for every A. In other words, our stochastic setting has the classical result as a special case.

Another particularly interesting special case of Theorem 2 is when the two SCRUMs share the same collection of utilities $\{U_t\}_{t=1}^T$ and their probability mass functions are related by first-order stochastic dominance. That is, let $\alpha^L = (\{U_t\}_{t=1}^T, \mu^L)$ and $\alpha^H = (\{U_t\}_{t=1}^T, \mu^H)$ be two SCRUMs, such that μ^L is first-order stochastically dominated by

μ^H.[9] Hence, μ^L places more probability mass on lower utility functions over the same single-crossing collection of utilities $\{U_t\}_{t=1}^T$, and then, clearly, $F_{\alpha^L} \trianglelefteq F_{\alpha^H}$. Theorem 2 implies that the stochastic choices p_{α^L} are first-order stochastically dominated, with respect to $\prec$, by the stochastic choices p_{α^H}.

Notably. Theorem 2 does not require that the SCRUMs hold the same collection of utility functions to establish the result. To illustrate, consider the following simple example.

Example 2. Consider two SCRUMs with four utility functions over four alternatives, where each utility function has a probability mass of 1/4. The following figure describes the two SCRUMs, where the columns represent the ordinal rankings, with upper alternatives denoting alternatives with higher utility values.

α	U_1	U_2	U_3	U_4
	1	2	3	4
	2	3	2	3
	3	1	1	2
	4	4	4	1

β	V_1	V_2	V_3	V_4
	1	2	3	4
	2	3	2	3
	3	4	4	2
	4	1	1	1

Note that $1 \prec 2 \prec 3 \prec 4$. We need to show that $F_\alpha \trianglelefteq F_\beta$. Because the two SCRUMs share the same probability masses, we only need to verify that $U_t \trianglelefteq V_t$, $t \in \{1, 2, 3, 4\}$. Because $U_1 = V_1$ and $U_4 = V_4$, we show it for $t \in \{2, 3\}$. Because U_t and V_t, $t \in \{2, 3\}$, coincide in the binary rankings of all the pairs $\{i, j\}$ such that $i \prec j$ and $V_t(i) > V_t(j)$, then it must be that $U_t \trianglelefteq V_t$. Theorem 2 then implies that p_α is first-order stochastically dominated by p_β. This is obvious for every set $A \neq \{1, 4\}$, as $p_\alpha(i, A) = p_\beta(i, A)$ for all i. For $\{1, 4\}$, $p_\alpha(1, \{1, 4\}) = \frac{3}{4} > \frac{1}{4} = p_\beta(1, \{1, 4\})$. ♦

5. Extensions

5.1. **Endogenous Order of Alternatives.**

Thus far, we have assumed that the order $\prec$ over the alternatives is known. This is the standard approach in the single-crossing literature because, typically, the order emerges naturally from the characteristics of the alternatives. On occasion, however, the order over the alternatives may not be so apparent, and one may wonder whether a stochastic choice function corresponds to a SCRUM for some underlying, unobserved order over the set of alternatives. That is, it may be the case that although there is not a clear order over the alternatives for

[9] μ^L is first-order stochastically dominated by μ^H whenever for every $\tilde{t} \in \{1, 2, \ldots, T\}$, it is $\sum_{t=1}^{\tilde{t}} \mu^L(t) \geq \sum_{t=1}^{\tilde{t}} \mu^H(t)$.

the analyst, the decision-maker actually contemplates one and behaves à la SCRUM. We address this question here. Formally, we say that p is a SCRUM* stochastic choice function whenever there is an order $\prec$ over X such that p is a SCRUM stochastic choice function for $(X, \prec)$.

The challenge is to discover the underlying order of the alternatives from the revealed data p. Let us define alternative x as central* in the triplet $\{x, y, z\}$ whenever $p(y, \{x, y, z\}) = p(y, \{x, y\})$ and $p(z, \{x, y, z\}) = p(z, \{x, z\})$. We know from Section 3 that, if p is a SCRUM stochastic choice function with respect to a given order over the alternatives, the central alternative in a triplet according to that order is central*. Suppose now that neither y nor z are central* in the triplet $\{x, y, z\}$. Then, for p to be a SCRUM* stochastic choice function, the following must hold. Obviously, x must be the central alternative in the triplet for the underlying order and thus, it must be central*. More interestingly, any $w \neq x$ must lie either to the left or to the right of x for the underlying order, which immediately implies that x is central* for either $\{x, y, w\}$ or $\{x, z, w\}$. This is the content of the following axiom.

Centrality* (CEN*). Let x, y, z and w be four distinct alternatives. If neither y nor z is central* in $\{x, y, z\}$, then x is central* in either $\{x, y, w\}$ or $\{x, z, w\}$.

This basic modification is sufficient, together with POS and MON, to characterize SCRUM* stochastic choice functions.

Theorem 3. *A stochastic choice function p satisfies POS, MON and CEN* if and only if it is a SCRUM* stochastic choice function.*

Proof of Theorem 3: The necessity of the axioms is straightforward. We prove the sufficiency of the axioms through a series of steps.

Step 1. We prove that, for every triplet, there is one and only one central* alternative. We first prove uniqueness. Assume by way of contradiction that there are at least two alternatives that are central* in the triplet $\{x, y, z\}$, say x and y. Then, it is $p(y, \{x, y, z\}) = p(y, \{x, y\})$ and $p(x, \{x, y, z\}) = p(x, \{x, y\})$, respectively. Hence, $p(y, \{x, y, z\}) + p(x, \{x, y, z\}) = p(y, \{x, y\}) + p(x, \{x, y\}) = 1$. This implies $p(z, \{x, y, z\}) = 0$, which contradicts POS. Hence, every triplet has at most one central* alternative. We now prove existence. Assume by way of contradiction that none of the alternatives is central* in the triplet $\{x, y, z\}$. By CEN*, x must be central in either $\{x, y, w\}$ or $\{x, z, w\}$. Indeed, we prove that it must be central* in both triplets. To see this, assume w.l.o.g. that x is central* in $\{x, y, w\}$, which, by uniqueness, implies that neither y nor w is central* in $\{x, y, w\}$. Then, we know that x must be central* in either $\{x, y, z\}$ or $\{x, z, w\}$. Because, by assumption, the former cannot be true, the latter must be. Then, however, the same symmetric argument can be applied to alternative y, which must be central* in both $\{x, y, w\}$ and $\{y, z, w\}$. This contradicts uniqueness in $\{x, y, w\}$, proving the claim.

Step 2. We prove that every alternative in a quadruple $\{x, y, z, w\}$ is central* in either zero or two triplets formed by alternatives in the quadruple. We first prove that an alternative cannot be central* in only one of these triplets. Suppose that x is central* in one of the triplets, say $\{x, y, z\}$. By Step 1, alternatives y and z are not central* in $\{x, y, z\}$, and hence CEN* implies that x must be central* in a triplet containing w. We now prove that an alternative cannot be central* in all the three triplets to which it belongs. Suppose by contradiction that x is central* in all these triplets. Then, Step 1 guarantees that y, z and w are never central* in the presence of x. However, Step 1 also guarantees that one of them must be central* in $\{y, z, w\}$. This alternative would be central* in only one triplet, which we just proved to be absurd.

Step 3. We make the set of alternatives X be a linearly ordered set $(X, \prec)$. We do so inductively by defining linearly ordered sets $\{(A^l, \prec^l)\}_{l=3}^{|X|}$ satisfying, for every $l \geq 3$: (i) $A^l \subseteq X$, with $|A^l| = l$, (ii) $A^l \subseteq A^{l+1}$ and (iii) the restriction of $\prec^{l+1}$ to A^l is $\prec^l$. Take any three distinct alternatives a, b and c and define $A^3 = \{a, b, c\}$. By Step 1, there is a unique central* alternative in the triplet, say b. Then, define $\prec^3$ by $a \prec^3 b \prec^3 c$. Setting a or c first in the order is without loss of generality. As will be seen, once this symmetry is broken, the rest of the algorithm respects it. We now describe how to add any alternative $x \in X \setminus A^l$ to the linearly ordered set $(A^l, \prec^l)$ forming a new linearly ordered set $(A^{l+1}, \prec^{l+1})$ with $A^{l+1} = A^l \cup \{x\}$ and the restriction of $\prec^{l+1}$ to A^l equating $\prec^l$. For the ease of exposition, denote the alternatives of A^l by $1 \prec^l 2 \prec^l \cdots \prec^l l$. If x is not central* in any triplet involving two consecutive alternatives in A^l, we place x either in the first or in the last position of $\prec^{l+1}$, respecting the order over the alternatives in A^l. We do so as follows. If 1 is central* in $\{x, 1, 2\}$, we define $x \prec^{l+1} 1 \prec^{l+1} 2 \cdots \prec^{l+1} l$. Otherwise, we define $1 \prec^{l+1} 2 \prec^{l+1} \cdots \prec^{l+1} l \prec^{l+1} x$.[10] If, on the contrary, x is central* for some triplet involving a pair of consecutive alternatives, let i_x be the first alternative in $(A^l, \prec^l)$ such that x is central* in $\{i_x, x, i_x + 1\}$. In this case, define $1 \prec^{l+1} \cdots \prec^{l+1} i_x \prec^{l+1} x \prec^{l+1} i_x + 1 \prec^{l+1} \cdots \prec^{l+1} l$, the restriction of which to

[10] In this case, notice that because neither x nor 1 is central* in $\{x, 1, 2\}$, Step 1 guarantees that 2 is.

A^l coincides with $\prec^l$. Clearly, this algorithm constructs a linear order $\prec=\prec^{|X|}$ over $X = A^{|X|}$.

Step 4. We claim that the constructed linear order $\prec$ has the property that for every triplet of alternatives, the central one according to $\prec$ is central*. This would prove that p satisfies CEN for $(X, \prec)$, which by Theorem 1 would conclude the proof. We prove this claim inductively over the collection of linearly ordered sets $\{(A^l, \prec^l)\}_{l=3}^{|X|}$. Notice that the result is trivial by construction for $(A^3, \prec^3)$. We then prove that the result is true for $l + 1$ whenever it is true for l. To see this, notice that the restriction of $\prec^{l+1}$ to A^l coincides with $\prec^l$. Hence, we only need to prove that the property holds true for every triplet of alternatives in A^{l+1} containing the alternative $x \in A^{l+1} \setminus A^l$.

Step 4a. Let $x \prec^{l+1} 1$. The claim is equivalent to proving that for every triplet $\{x, i, j\}$ with $1 \leq i < j \leq l$, i is central*. The construction guarantees that this is true for the triplet $\{x, 1, 2\}$. For every $2 < j$, as alternative 1 is central* in $\{x, 1, 2\}$, Step 2 guarantees that alternative 1 is central* in either $\{x, 1, j\}$ or $\{1, 2, j\}$. By the induction hypothesis, the latter cannot be true, proving the result. Whenever $1 < i < j$, the induction hypothesis guarantees that alternative i is central* in $\{1, i, j\}$. Step 2 guarantees that i must be central* in either $\{x, 1, i\}$ or $\{x, i, j\}$. We have already proven that 1 is the unique central* alternative in $\{x, 1, i\}$, leading to the result.

Step 4b. Let $l \prec^{l+1} x$. The claim requires us to prove that for every triplet $\{i, j, x\}$ with $1 \leq i < j \leq l$, j is central*. We prove it by induction on the value of $t = j - i$. Consider $t = 1$. We continue the proof by using another inductive argument on the value of i. We already proved the case in which $i = 1$ (see footnote 10). Suppose that the claim is true up to a certain alternative i, which means that alternative $i + 1$ is central* in $\{i, i+1, x\}$. Because we already know that $i+1$ is central* in $\{i, i+1, i+2\}$, Step 2 guarantees that $i + 1$ cannot be central* in $\{i + 1, i + 2, x\}$. By construction of the algorithm, x is also not central* in the latter triplet, and hence Step 1 guarantees that $i + 2$ is. This concludes the inductive argument on i and proves the claim for $t = 1$. Suppose now that the claim is true up to a certain t and consider $t + 1$. We prove the claim for every i such that $i + t + 1 \leq n$. We know that $i + t$ is central* in $\{i, i + t, x\}$ and also in $\{i, i + t, i + t + 1\}$. Hence, Step 2 guarantees that $i + t$ is not central* in $\{i + t, i + t + 1, x\}$, and given that x is also not, Step 1 guarantees that $i + t + 1$ must be. This concludes the inductive argument and the case.

Step 4c. Let $i_x \prec^{l+1} x \prec^{l+1} i_x + 1$, where i_x is the first alternative for which x is central* in $\{i_x, x, i_x + 1\}$. We first prove the claim for triplets of the form $\{i, j, x\}$ with

$1 \leq i < j \leq i_x$. Let $j = i_x$. Because, by construction, x is central* in $\{j, x, i_x + 1\}$, Step 1 guarantees that j cannot be. Because j is central* in $\{i, j, i_x + 1\}$, Step 2 guarantees that j is also central* in $\{i, j, x\}$, as desired. Whenever $j < i_x$, notice that j is central* in $\{i, j, i_x\}$, and we just proved that i_x is the unique central* alternative in $\{j, i_x, x\}$. Hence, Step 2 guarantees that j is central* in $\{i, j, x\}$. Second, the claim can be analogously proven for all the triplets of the form $\{x, i, j\}$, with $i_x + 1 \leq i < j \leq l$, and hence we omit the proof. Third, consider triplets of the form $\{i, x, j\}$ with $i \leq i_x$ and $i_x + 1 \leq j$. If $i = i_x$, notice that, by construction, x is central* in $\{i_x, x, i_x + 1\}$ but as proven above, x is not central* in $\{x, i_x + 1, j\}$. Hence, Step 2 guarantees that x is central* in $\{i, x, j\}$, as desired. A symmetric argument shows that x is central* in $\{i, x, j\}$ whenever $j = i_x + 1$. Finally, whenever $i < i_x$ and $i_{x+1} < j$, we have just proven that x is central* in $\{i_x, x, j\}$ but not in $\{i, i_x, x\}$, and hence Step 2 guarantees that x must be central* in $\{i, x, j\}$, as desired. ■

The proof of Theorem 3 explicitly constructs the underlying order $\prec$ from the revealed choices involving two and three alternatives. It uses an inductive argument by which, at every step, a new alternative x is incorporated into the order constructed in the previous step $1 \prec^l 2 \prec^l \cdots \prec^l l$. If x is central* for two consecutive alternatives in $\prec^l$, then it is incorporated between these alternatives to form $\prec^{l+1}$. If this is not the case, whenever alternative 1 is central in $\{x, 1, 2\}$, x is incorporated into the first position of the order; otherwise, x is placed at the end of the order. The rest of the proof shows that the constructed order has the property that for every ordered set of three alternatives, the central alternative is central*. Moreover, we argue in the proof that the constructed order is unique up to symmetry. Then, we appeal to Theorem 1, which guarantees the existence of a pair $(\{U_t\}_{t=1}^T, \mu, \prec)$ that is basically unique and that rationalizes p in the sense of SCRUMs.

It is immediate that the endogenous version of the model retains the properties of a SCRUM, which we have discussed throughout the paper. In particular, SCRUM* stochastic choice functions satisfy moderate stochastic transitivity, and moreover, once the underlying order over the alternatives is identified, our result on stochastic monotone comparative statics also readily applies here.

5.2. **Dominated Alternatives.** On occasion, there are alternatives in the analysis such that, under any circumstances, one is always considered superior to the other. Returning to our Example 1 above, this is the case, e.g., when x and y represent a pair of gambles, or streams of payoffs, or allocations, or bundles, and x first-order stochastically dominates y. In all these cases, every utility in $\{U_t\}_{t=1}^{T}$ regards x as superior to y. In a random utility model, this implies that, as y is never maximal in the presence of x, the probability of choosing y in this case would be zero. Note that this is in conflict with POS, which requires that the probability of choosing any alternative from any menu is always above zero. In this section, we extend the main model in this paper, the SCRUM, to allow for the incorporation of dominated alternatives into the analysis. We do so by extending the POS property to allow for zero choice probabilities in the presence of dominations.

We model domination by way of a binary relation $\rightarrow$ over the set of alternatives, where $x \rightarrow y$ represents the case when x dominates y, and assume $\rightarrow$ to be a partial order. That is, not all alternatives must be related by domination, and we require that if $x \rightarrow y$ and $y \rightarrow z$, it must also be the case that $x \rightarrow z$. Consequently, the given order over the alternatives $\prec$ is now also assumed to be partial. Further, we assume that the union of all the binary comparisons of $\rightarrow$ and $\prec$ exhausts all the possible ones and denote by $\prec'$ the linear order resulting from the union of $\rightarrow$ and $\prec$. That is, for every two distinct alternatives $x, y \in X$, it is either $x \prec' y$ or $y \prec' x$, and $x \prec' y$ means either that x precedes y, $x \prec y$, or x dominates y, $x \rightarrow y$. We now modify POS to allow for the treatment of dominated alternatives.

Positivity' (POS'). If $x \in A$ and there is no $y \in A$ such that $y \rightarrow x$, $p(x, A) > 0$. Otherwise, $p(x, A) = 0$.

POS' coincides with POS in the absence of dominated alternatives and requires that when there is an alternative x that is dominated by some other alternative y in A, the choice probability of x in A is zero. This is the only change we need to introduce in the identification of this variation of SCRUM.[11] Denote by SCRUM' the SCRUM-type model where alternatives are maximal for some utility function if and only if they are not dominated by any other alternative.

Theorem 4. *A stochastic choice function p satisfies POS', MON and CEN if and only if it is a SCRUM' stochastic choice function.*

[11]Notice that CEN is now weaker, as it only applies to triplets of alternatives where domination is not present.

Proof of Theorem 4: The necessity of the axioms is straightforward. We prove the sufficiency of the axioms through a series of steps. To ease the exposition, we denote the alternatives in X by $1 \prec' 2 \prec' \cdots \prec' |X|$.

Step 1. We prove that for every triplet such that $i \prec j \to k$ and $i \prec k$, it is $p(i,\{i,j\}) \leq p(i,\{i,k\})$, and for every triplet such that $i \to j \prec k$ and $i \prec k$, it is $p(j,\{j,k\}) \leq p(i,\{i,k\})$. To see the first case, notice that MON implies $p(i,\{i,j\}) = 1 - p(j,\{i,j\}) \leq 1 - p(j,\{i,j,k\}) = p(i,\{i,j,k\}) + p(k,\{i,j,k\})$. POS' guarantees that $p(k,\{i,j,k\}) = 0$, and hence by MON, it is $p(i,\{i,j\}) \leq p(i,\{i,j,k\}) \leq p(i,\{i,k\})$. To see the second case, notice that MON implies $p(j,\{j,k\}) = 1 - p(k,\{j,k\}) \leq 1 - p(k,\{i,j,k\}) = p(i,\{i,j,k\}) + p(j,\{i,j,k\})$. POS' guarantees that $p(j,\{i,j,k\}) = 0$, and hence by MON, it is $p(j,\{j,k\}) \leq p(i,\{i,j,k\}) \leq p(i,\{i,k\})$.

Step 2. We claim that the collection of utility functions $\{U_t\}_{t=1}^{T}$ constructed in the proof of Theorem 1 satisfies the single-crossing condition with respect to $\prec'$. We first consider pairs of alternatives $\{i,j\}$ such that $i \to j$ and show that for every t, it is $U_t(i) > U_t(j)$. Then, we consider the case in which $i \prec j$ and divide the analysis in $t \leq t(i,j)$ and $t > t(i,j)$, where $t(i,j)$ is defined as in the proof of Theorem 1.

Step 2a. Let $i \to j$. Consider any alternative k such that $k \prec' i$. If $p(k,\{k,j\}) < \lambda_t$, then it must be $k \prec j$ and it must also be $k \prec i$, as otherwise the transitivity of $\to$ would imply $k \to j$, a contradiction. Hence, the first relationship in Step 1 guarantees that $p(k,\{k,i\}) \leq p(k,\{k,j\}) < \lambda_t$ and, consequently, $|\{k : k \prec' i, p(k,\{k,i\}) < \lambda_t\}| \geq |\{k : k \prec' i, p(k,\{k,j\}) < \lambda_t\}|$. Consider now any alternative k such that $j \prec' k$. If $p(i,\{i,k\}) < \lambda_t$, it must be $i \prec k$. It must also be $j \prec k$, as otherwise the transitivity of $\to$ would imply $i \to k$, a contradiction. The second relationship in Step 1 implies that $p(j,\{j,k\}) \leq p(i,\{i,k\}) < \lambda_t$ and, therefore, $-|\{k : j \prec' k, p(i,\{i,k\}) < \lambda_t\}| \geq -|\{k : j \prec' k, p(j,\{j,k\}) < \lambda_t\}|$. Finally, consider any alternative k such that $i \prec' k \prec' j$. If $p(i,\{i,k\}) < \lambda_t$, it must be $i \prec k$. It cannot be $k \prec j$, as otherwise the transitivity of $\prec$ would imply $i \prec j$, a contradiction. Then, it must be $k \to j$ or equivalently $p(k,\{k,j\}) = 1 \geq \lambda_t$. Hence, it must be $-i - |\{k : i \prec' k \prec' j, p(i,\{i,k\}) < \lambda_t\}| > -j + |\{k : i \prec' k \prec' j, p(j,\{j,k\}) < \lambda_t\}|$. We can use these three inequalities together with the fact that $p(i,\{i,j\}) = 1 \geq \lambda_t$, as in Step 3a of the proof of Theorem 1, to show that $U_t(i) > U_t(j)$.

Step 2b. Let $i \prec j$ and $t \leq t(i,j)$. Consider any alternative k such that $k \prec' i$. If $p(k,\{k,j\}) < \lambda_t$, then it must be $k \prec j$. Notice that $k \to i$ would imply, by the second relationship in Step 1, that $p(i,\{i,j\}) \leq p(k,\{k,j\}) < \lambda_t$, which is a contradiction with $t \leq t(i,j)$. Hence, it must be $k \prec i \prec j$ and, as in Step 1 of the proof Theorem 1, $p(k,\{k,i\}) \leq p(k,\{k,j\}) < \lambda_t$. This implies again $|\{k : k \prec' i, p(k,\{k,i\}) < \lambda_t\}| \geq |\{k : k \prec' i, p(k,\{k,j\}) < \lambda_t\}|$. Consider now any alternative k such that $j \prec' k$. Notice that it must be $p(i,\{i,k\}) \geq \lambda_t$. Otherwise, it would be $i \prec k$ and both $j \prec k$ and $j \rightarrow k$ would imply, by using respectively the argument in Step 1 of the proof of Theorem 1 and the first relationship in Step 1 in this proof, $p(i,\{i,j\}) \leq p(i,\{i,k\}) < \lambda_t$, a contradiction. Hence, $-|\{k : j \prec' k, p(i,\{i,k\}) < \lambda_t\}| = 0 \geq -|\{k : j \prec' k, p(j,\{j,k\}) < \lambda_t\}|$. Finally, consider any alternative k such that $i \prec' k \prec' j$. If $p(i,\{i,k\}) < \lambda_t$, it must be $i \prec k$. Whenever $k \prec j$, we can use the argument in Step 1 of the proof of Theorem 1 to show that $p(k,\{k,j\}) \geq p(i,\{i,j\}) \geq \lambda_t$. If $k \rightarrow j$, POS' implies that $p(k,\{k,j\}) = 1 \geq \lambda_t$. Therefore, it must be $-i - |\{k : i \prec' k \prec' j, p(i,\{i,k\}) < \lambda_t\}| \geq -j + |\{i \prec' k \prec' j, p(k,\{k,j\}) < \lambda_t\}|$. The result follows in the same manner as above.

Step 2c. Let $i \prec j$ and $t > t(i,j)$. Consider any alternative k such that $k \prec' i$. If $p(k,\{k,i\}) < \lambda_t$, we have $k \prec i$, and the transitivity of $\prec$ guarantees that $k \prec j$. Step 1 in the proof of Theorem 1 guarantees that $p(k,\{k,j\}) \leq p(i,\{i,j\}) < \lambda_t$. Hence, $|\{k : k \prec' i, p(k,\{k,j\}) < \lambda_t\}| \geq |\{k : k \prec' i, p(k,\{k,i\}) < \lambda_t\}|$. Consider any alternative k such that $j \prec' k$. If $p(j,\{j,k\}) < \lambda_t$, then it is $i \prec j \prec k$, and hence, again, $p(i,\{i,k\}) \leq p(j,\{j,k\}) < \lambda_t$. That is, $-|\{k : j \prec' k, p(j,\{j,k\}) < \lambda_t\}| \geq -|\{k : j \prec' k, p(i,\{i,k\}) < \lambda_t\}|$. Finally, consider any alternative k such that $i \prec' k \prec' j$. Suppose that $p(i,\{i,k\}) \geq \lambda_t$. We claim that $i \prec k$ is not possible because then, by using Step 1 in Theorem 1 and the first relationship in Step 1 in this proof, $p(i,\{i,k\}) \leq p(i,\{i,j\}) < \lambda_t$, a contradiction with $t > t(i,j)$. Hence, it is $i \rightarrow k$, and the transitivity of $\rightarrow$ guarantees that we are under the second relationship in Step 1. Thus, $p(k,\{k,j\}) \leq p(i,\{i,j\}) < \lambda_t$, which implies that $-j + |\{k : i \prec' k \prec' j, p(k,\{k,j\}) < \lambda_t\}| + 1 > -i - |\{k : i \prec k \prec j, p(i,\{i,k\}) < \lambda_t\}| - 1$, and the result follows as in the previous steps.

Step 3. Given a set A, let B_A be the set of undominated alternatives in A, i.e., $B_A = \{x \in A :$ there does not exist $y \in A$ such that $y \rightarrow x\}$. We claim that for every $x \in B_A$, it is $p(x,A) = p(x,B_A)$. For any alternative $x \notin B_A$, POS' guarantees that $p(x,A) = 0$, and hence $1 = \sum_{x \in B_A} p(x,A)$. By MON, $1 = \sum_{x \in B_A} p(x,A) \leq \sum_{x \in B_A} p(x,B_A) = 1$, and the claim follows.

Step 4. We conclude the proof by showing that for every A and every x, it is $p(x,A) = \sum_{t:x=m_t(A)} \mu(t)$. This is trivial whenever $x \notin A$ and for any alternative

$x \in A \setminus B_A$, as in both cases $p(x, A) = 0$ and x is not maximal for any utility, as desired. Consider then any alternative $x \in B_A$. By applying Steps 4 and 5 of the proof of Theorem 1 over set B_A, we know that $p(x, B_A) = \sum_{t:x=m_t(B_A)} \mu(t)$. Clearly, $m_t(B_A) = m_t(A)$, and the result follows. ■

The main part of the proof consists in showing that the constructed utilities in the proof of Theorem 1 also work here for the case of $\prec'$, which allows for dominated alternatives. We do so by studying all the possible binary relations in a triplet. Finally, SCRUM', again, also has the desirable properties of SCRUM, regarding stochastic transitivity and stochastic monotone comparative statics.

6. Final Remarks

In this paper we have proposed and studied a new stochastic choice model that can be used in a wide variety of settings, namely, those in which the single-crossing property applies, and we have shown that is easily testable in practice. We close this paper by commenting on some limitations of the SCRUM and offering suggestions on how to overcome them.

Our model belongs to the tradition of the classical rational stochastic choice models of Luce (1959), Block and Marshak (1960), McFadden and Richter (1990), and more recently, of Gul and Pesendorfer (2006) or Fudenberg, IIjima and Strzalecki (2015). It is important to stress that there is a growing literature that rigorously introduces behavioral aspects into stochastic choice (see, e.g., Manzini and Mariotti, 2014; Caplin and Dean, 2015). Although some of these behavioral phenomena are not compatible with our rational methodological approach, our model can be used to quantify their prevalence in the data and, furthermore, could be easily extended to accommodate them. We illustrate these points with the cases of violations of stochastic dominance and the attraction and compromise effects.

Consider a violation of stochastic dominance. That is, let x and y be two lotteries such that x stochastically dominates y, and, in stark contrast with the POS' property of Section 5.2, $p(y, \{x, y\}) > 0$.[12] First, the extent of the violation of the POS' property provides a natural quantification of the behavioral inconsistency. Second, and more interesting, our model can be easily extended to incorporate this inconsistency. Arguably, violations of stochastic dominance can be understood as mistakes on the part of the decision-maker. Then, in the spirit of the trembling hand approach in game

[12] The logic below applies trivially to larger sets of alternatives and to settings other than risk.

theory, we can regard our decision-maker as behaving according to the SCRUM with a large probability and with a small probability experiencing a tremble where, in this case, the dominated alternative is chosen.

Let us now consider the compromise and attraction effects. In the standard representation of the compromise effect, there are three ordered alternatives $x \prec y \prec z$ such that in the set $\{x, y, z\}$ the middle alternative y is seen as a compromise between the two others, and then $p(y, \{x, y\}) < p(y, \{x, y, z\})$, violating MON. In the attraction effect, $x \prec y \to z$ and $x \prec z$. Here, the presence of z is perceived as giving strength to the alternative that dominates it, y, and hence, again, $p(y, \{x, y\}) < p(y, \{x, y, z\})$ representing a violation of MON. In both cases, the inclusion of a new alternative in the menu, i.e. z, makes the decision-maker change her views of the choice situation. Again, MON can be used to quantify the relevance of the effects. More interesting, a natural way of introducing these considerations in our setting is by contemplating a menu-dependent SCRUM, where the probability function μ on the collection of utility functions $\{U\}_{t=1}^{T}$ depends on the menu of alternatives. In particular, one can entertain the possibility that introducing an alternative that is to the right in $\prec$ causes the decision-maker to shift her attention to the right of $\{U\}_{t=1}^{T}$, which would rationalize the two menu effects.

The two extensions of SCRUMs commented on above would smoothly accommodate the behavioral phenomena of interest, while retaining the spirit of the stochastic choice model.

References

[1] Agranov, M. and P. Ortoleva (2016), "Stochastic Choice and Preferences for Randomization," *Journal of Political Economy*, forthcoming.

[2] Apesteguia, J. and M.A. Ballester (2016), "Monotone Stochastic Choice Models: The Case of Risk and Time Preferences," mimeo.

[3] Athey, S. (2001), "Single Crossing Properties and the Existence of Pure Strategy Equilibria in Games of Incomplete Information," *Econometrica*, 69.4:861–889.

[4] Barberà, S. and P.K. Pattanaik (1986). "Falmagne and the Rationalizability of Stochastic Choices in Terms of Random Orderings," *Econometrica*, 54.3:707–715.

[5] Benoît, J.-P. and E.A. Ok (2007). "Delay Aversion," *Theoretical Economics*, 2.1:71–113.

[6] Block, H.D. and J. Marschak (1960). "Random Orderings and Stochastic Theories of Response," in I. Olkin et al, eds., Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling, 97–132. Standford, Calif.:Standford University Press.

[7] Caplin, A. and M. Dean (2016). "Revealed Preference, Rational Inattention, and Costly Information Acquisition," *American Economic Review*, forthcoming.

[8] Cox, J.C., D. Friedman, and V. Sadiraj (2008). "Revealed Altruism," *Econometrica*, 76.1:31–69

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

[10] Fishburn, P.C. (1998). "Stochastic Utility," in S. Barberà, P.J. Hammond, and C. Seidl, eds., Handbook of Utility Theory, 272–311. Heidelberg, Germany: Springer.

[11] Fudenberg D. and T. Strzalecki (2015). "Dynamic Logit with Choice Aversion," *Econometrica*, 83.2:651–691.

[12] Fudenberg D., R. Iijima and T. Strzalecki (2015). "Stochastic Choice and Revealed Perturbed Utility," *Econometrica*, 83.6:2371–2409.

[13] Gans, J. and M. Smart (1996). "Majority Voting with Single Crossing Preferences," *Journal of Public Economics*, 59.2:219–237.

[14] Gul, F. and W. Pesendorfer (2006). "Random Expected Utility," *Econometrica*, 74.1:121–146.

[15] Gul, F., P. Natenzon, and W. Pesendorfer (2014). "Random Choice as Behavioral Optimization," *Econometrica*, 82.5:1873–1912.

[16] Jewitt, I. (1987). "Risk Aversion and the Choice between Risky Prospects: The Preservation of Comparative Statics Results," *Review of Economic Studies*, 54.1:73–85.

[17] Lu, J. and K. Saito (2016), "Random Intertemporal Choice," mimeo.

[18] Luce, R.D. (1959). "Individual Choice Behavior: A Theoretical Analysis," New York, New York: Wiley.

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

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

[21] Milgrom, P. and C. Shannon (1994). "Monotone Comparative Statics," *Econometrica*, 62.1:157–180.

[22] Mirrlees, J. (1971). "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, 38.2:175–208.

[23] Persson, T. and G. Tabellini (2000). "Political Economics: Explaining Economic Policy". Cambridge, MA: MIT Press.

[24] Persico, N. (2000). "Information Acquisition in Auctions," *Econometrica*, 68.1:135–148.

[25] Spence, A.M. (1974). "Market Signaling: Information Transfer in Hiring and Related Screening Processes," Cambridge, MA: Harvard University Press.