

Satisficing and Stochastic Choice

Victor H. Aguiar* María José Boccardi[†] and Mark Dean[‡]

August 31, 2015

Abstract

Satisficing is a hugely influential model of boundedly rational choice, yet it cannot be easily tested using standard choice data. We develop necessary and sufficient conditions for *stochastic* choice data to be consistent with satisficing, assuming that preferences are fixed, but search order may change randomly. The model predicts that stochastic choice can only occur amongst elements that are always chosen, while all other choices must be consistent with standard utility maximization. Adding the assumption that the probability distribution over search orders is the same for all choice sets makes the satisficing model a subset of the class of random utility models.

1 Introduction

People often do not pay attention to all the available alternatives before making a choice. This fact has led to an extensive recent literature aimed at understanding the observable implications of models in which the decision maker (DM) has limited attention.¹ In an important recent paper, Manzini & Mariotti (2014) characterize the stochastic choice data generated by a decision maker (DM) who has standard preferences, but only notices each alternative in their choice set with some probability. The chosen item is therefore the best alternative in the ‘consideration set’ of noticed items, which may be a strict subset of the items which are actually available.

The idea that a DM may not search exhaustively through all available alternatives is not new. Simon (1955) introduced the concept of satisficing: an intuitively plausible choice procedure by which the DM searches through alternatives until they find one that is ‘good

*Department of Economics, Brown University, Email: victor_aguiar@brown.edu

[†]Department of Economics, Brown University. Email: maria_jose_boccardi@brown.edu

[‡]Department of Economics, Columbia University. Email: mark.dean@columbia.edu

¹Notable examples include Masatlioglu et al. (2012), Caplin et al. (2011), Eliaz & Spiegler (2011), and Salant & Rubinstein (2008).

enough’, at which point they stop and choose that alternative.² This model has been hugely influential, both within economics, and in other fields such as psychology (Schwartz et al. (2002)) and ecology (Ward (1992)).

Despite the popularity of the satisficing model, testing its predictions can prove challenging. It has long been known that standard choice data, which records only the choices made from different choice sets, cannot be used to disentangle satisficing from utility maximization (see Caplin et al. (2011) for a discussion). Researchers have therefore typically resorted to richer data sets in order to test the satisficing model. For example, Caplin et al. (2011) make use of ‘choice process’ data, which records the evolution of choice with decision time, while Santos et al. (2012) use the order in which alternatives were searched as recorded from their internet browsing history.

In this paper, we characterize the observable implications of the satisficing choice procedure for *stochastic* choice data. Such data has been heavily studied in the economics literature.³ We assume that the DM has a fixed utility function and satisficing level. In any given choice set, they search sequentially until they find an alternative which has utility above their satisficing level, at which point they stop and choose that alternative. If they search the entire choice set and do not find a satisficing alternative then they choose the best available option. We assume that search order varies randomly, leading to stochasticity in choice. On the one hand, our paper is related to the work of Manzini & Mariotti (2014) (henceforth MM). It specifies a procedure by which attention is allocated, while MM is agnostic in this regard. On the other, it provides an alternative test of the satisficing model to that of Caplin et al. (2011) and Santos et al. (2012), using a data set which is readily available in many settings.

Our main observation is that the satisficing model implies that choice is stochastic only in choice sets where there are multiple alternatives above the satisficing level. If this is the case, then the order of search will affect the chosen alternative. If not, then either the choice set will be fully searched and the best option deterministically chosen, or the single satisficing alternative will always be chosen. This allows us to behaviorally identify the alternatives that are satisficing for the decision maker.

Without further restriction, any stochastic choice data set can trivially be made

²Caplin et al. (2011) show that satisficing behavior can be optimal under some circumstances

³See for example Block & Marschak (1960); Luce & Suppes (1965); Falmagne (1978); Gul & Pesendorfer (2006); Gul et al. (2014); Manzini & Mariotti (2014)

commensurate with the satisficing choice procedure by assuming that all alternatives are above the satisficing level, and the resulting distribution of choices reflects the distribution of search orders in that choice set. In order to generate meaningful behavioral implications, we must place further restrictions on the satisficing model. For our main theorem we make the assumption that the distribution of search orders has a full support property (i.e., each item has a positive probability of being searched first), and also rule out the possibility of indifference. This allows us to identify the set of above-reservation alternatives and characterize satisficing with two simple intuitive conditions. The first states that choice can be stochastic only amongst elements that are always chosen (with some probability) when available. The second says that revealed preference, defined via the support of the random choice rule in each set, must satisfy the Strong Axiom of Revealed Preference (SARP). Under these conditions, the data will admit a satisficing representation and the satisficing set, utility function and distribution over search orders can be identified to a high degree of precision.

Our baseline specification puts no restrictions on the relationship between the distribution of search orders across different choice sets. We next consider a refinement of the satisficing model in which the distribution of search order in each choice set is a manifestation of the same underlying search distribution. In order to guarantee such a representation we need a third axiom: the Total Monotonicity condition of Block & Marschak (1960). This condition on its own is necessary and sufficient for the data to be commensurate with the random utility model (RUM). Thus, the fixed distribution satisficing model is the precise intersection between satisficing and random utility.

We finish discussing three extensions to our results in which we relax the assumptions of full support, no indifference and the observation of a complete data set. We show that a satisficing model without full support, but with fixed distribution is equivalent to the random utility model. Allowing for indifference (but maintaining the full support assumption) is equivalent to dropping the requirement that stochasticity only take place amongst always chosen alternatives. If data is incomplete, our necessary and sufficient conditions are unchanged, but our ability to identify above satisficing elements is reduced.

Section 2 describes our set up. Section 3 characterizes the satisficing model. Section 4 considers the extensions described above, while section 5 discusses the related literature.

2 Set Up

2.1 Data

We consider a finite abstract choice set X , and let $\mathcal{D} \subseteq 2^X \setminus \emptyset$ be the set of menus in which behavior is observed. We assume that data comes in the form of a *random choice rule*, $p : X \times \mathcal{D} \mapsto [0, 1]$, which specifies for each menu $A \in \mathcal{D}$ the probability of choosing each element $a \in A$ (for example, if the DM has a one third probability of choosing x from $\{x, y, z\}$ then $p(x, \{x, y, z\}) = \frac{1}{3}$).

Definition 1 [*Data set*] A data set consists of a set of menus $\mathcal{D} \subseteq 2^X \setminus \emptyset$ and a random choice rule $p : X \times \mathcal{D} \mapsto [0, 1]$ such that $\sum_{a \in A} p(a, A) = 1 \forall A \in \mathcal{D}$. We say a data set is *complete* if $\mathcal{D} = 2^X \setminus \emptyset$.

Random choice rules have been heavily studied in the theoretical, as well as the applied literature.⁴ In practice, while a random choice rule is not directly observable, it can be estimated from observed choice frequencies, pooling either across repeated choices by the same individual, or by aggregating across the choices of different individuals.

2.2 The Satisficing Model

The satisficing choice procedure can be described as follows: when faced with a menu of options to choose from, the DM searches through the available alternatives one by one. If, at any point, they come across an alternative which is ‘good enough’, they stop searching and select that alternative from the menu. If they exhaustively search all alternatives without finding an element which satisfies their criteria, then they choose the best available alternative from the set. Note that the standard model of rational choice is a limiting case of the satisficing model in which no alternative is ‘good enough’.

As a concrete example, consider a DM searching for a book to buy in a bookshop prior to a flight. They examine the available books one by one, looking for one which satisfies their requirements (humorous, has good reviews, long enough to last the flight, not by Dan Brown). If they find such a book, they immediately go to the checkout and buy it. If they search the entire selection and don’t find a book which matches this criteria then they go back and choose the best of the books that they did see.

⁴Examples of early theoretical work include Block & Marschak (1960) and Luce & Suppes (1965). More recent work includes Gul & Pesendorfer (2006); Manzini & Mariotti (2014); Gul et al. (2014).

The satisficing choice procedure therefore has three building blocks. The first is a fixed utility function $u : X \rightarrow \mathbb{R}$, which describes the preferences of the DM. Following Manzini & Mariotti (2014), for our main results we rule out indifference, and therefore assume that u is injective. We discuss the implications of allowing indifference in section 4.2.

The second model element is a utility threshold u^* , which we will refer to as the *reservation utility*. This defines the concept of ‘good enough’: an alternative $x \in X$ is good enough if $u(x) \geq u^*$. We define $U^* = \{a \in X | u(x) \geq u^*\}$ as the set of satisficing elements according to u and u^* . For convenience, we will assume that there is at least one satisficing element: i.e. $u^* \leq \max_{x \in X} u(x)$. This assumption has no behavioral implication: a model in which only the best available alternative is above the reservation utility is indistinguishable from one in which there is no such alternative. However, it will streamline the statement of identification results in section 3.

The third element of the satisficing model is the order in which search occurs. A search order for a choice set A is defined by a linear order on that set.⁵ We use R_A to denote the set of linear orders on A , with r_A a typical element in R_A . Our key assumption is that the order of search is determined stochastically: we use $\gamma_A : R_A \rightarrow [0, 1]$ to denote the probability distribution over the set R_A , which we call a ‘stochastic search order’. Abusing notation slightly, we will use $\gamma_A(x \ r_A \ y)$ to denote the probability of all search orders in which x appears before y : i.e. $\gamma_A(r_A \in R_A | x \ r_A \ y)$.

We are agnostic about the source of this stochasticity. It could be that the DM randomly decides the order of search - in our example, sometimes the DM search through the books alphabetically, while sometimes they do so by genre. Alternatively, it could be that the random choice rule is generated by a DM who is faced by choice situations which are framed in different ways,⁶ with the framing unobservable to the researcher. For example, sometimes the bookstore puts the thrillers at the front of the store, while sometimes they put the romantic comedies at the front. These ‘frames’ affect the order in which the DM searches (though not their preferences), but are not known to the researcher.

A data set can be represented by the satisficing model if there exists a utility function, satisficing level and family of stochastic search orders which would generate the observed choice probabilities:

⁵i.e. a complete, transitive and antisymmetric binary relation on A with the interpretation ‘searched no later than’.

⁶In the sense of Salant & Rubinstein (2008).

Order	(1)	(2)	(3)	(4)	(5)	(6)
1st	a	a	b	b	c	c
2nd	b	c	a	c	a	b
3rd	c	b	c	a	b	a
γ_A^1	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{24}$	$\frac{1}{8}$	$\frac{1}{4}$

Table 1: An example of the satisficing model

Definition 2 [General Satisficing Model (GSM)] A data set (\mathcal{D}, p) has a Generalized Satisficing Model (GSM) representation if there exists an injective $u : X \rightarrow \mathbb{R}$, $u^* \in \mathbb{R}$ such that $u^* \leq \max_{s \in X} u(x)$, and $\{\gamma_A\}_{A \in \mathcal{D}}$ such that, for any $A \in \mathcal{D}$ and $a \in A$

$$p(a, A) = \begin{cases} \gamma_A(r_A | a r_A b \forall b \in A \setminus \{a\} \text{ s.t. } u(b) \geq u^*) & \text{if } u(a) \geq u^* \\ 1 & \text{if } a = \arg \max_{x \in A} u(x) < u^* \\ 0 & \text{Otherwise} \end{cases} \quad (1)$$

To illustrate how the model works consider the following example.

Example 1 Let $A = \{a, b, c\}$ and γ_A be as displayed in table 1. Consider first the case where a, b are satisficing alternatives, while c is not, that is $u(a), u(b) > u^* > u(c)$. Then, no matter the search order, c will never be chosen, and so $p(c, A) = 0$. However, as the DM will chose a if a is seen before b and b otherwise, their frequency depends on γ_A . In particular, $p(a, A) = \frac{3}{8}$ and $p(b, A) = \frac{5}{8}$. If instead all alternatives are below the satisficing level (i.e. $u^* > \max_{x \in A} u(x)$) then choice will be independent of search order: all alternatives will always be searched, and the best subsequently chosen.

3 Characterizing the Satisficing Model.

3.1 A Negative Result

The aim of this paper is to describe properties of a stochastic choice data set which are necessary and sufficient to guarantee a satisficing representation. However, our first result is negative: without further refinement, the GSM model provides no restriction on such a data set.

Proposition 1 Any data set (\mathcal{D}, p) has a GSM representation.

All proofs are relegated to the appendix

The GSM is flexible enough to match any data set because it places no restriction on the distribution over search orders in each decision problem. Thus one can always construct a distribution of search orders that will match the data by assuming that all alternatives are above the satisficing level.

To derive testable restrictions for the satisficing model, we introduce a ‘full support’ condition on the distribution of search orders. This restriction will allow us to identify satisficing alternatives as those which are chosen with positive probability in every choice set in which they appear. We can then utilize the underlying structure of the GSM model to derive behavioral restrictions. Intuitively, the stochastic nature of search generates stochastic behavior among satisficing alternatives. In contrast, we expect to observe deterministic utility maximizing behavior among choice sets which consist only of non-satisficing alternatives.

For the remainder of this section we concentrate on the simple case of full support, no indifference and complete data. We identify the behavioral conditions which characterize the resulting model. We also consider a special case of the model in which there is consistency in the distribution of search orders between choice sets. This additional restriction ensures that our model is behaviorally equivalent to a subset of the class of random utility models (RUMs). In section 4 we discuss extensions in which we drop the full support, no indifference, and complete data conditions.

3.2 Full Support Satisficing Models

Our model adds to the GSM the assumption that, in each choice set, any item will be searched first with some positive probability.

Assumption 1 [*Full Support*] For any $a \in A$ and all $A \in \mathcal{D}$: $\gamma_A(r_A \in R_A : a r_A b \quad \forall b \in A \setminus \{a\}) > 0$.

We describe a GSM which additionally satisfies this assumption as a *Full Support Satisficing Model (FSSM)*

Definition 3 [*Full Support Satisficing Model (FSSM)*] A data set (\mathcal{D}, p) has a *Full Support Satisficing Model (FSSM)* representation if it has a GSM representation in which the stochastic search order satisfies *Full Support*.

The assumption of Full Support has an important implication: we can identify above-reservation alternatives as those which are always chosen with positive probability in any choice set in which they appear. This is because Full Support implies that, for each such alternative, a search order in which it is searched first occurs with positive probability in each choice set, ensuring that it will be chosen. Furthermore, any alternatives that are *not* above reservation utility will be chosen with zero probability in any choice set which contains an above reservation utility alternative.

We define the set of alternatives which are always chosen:

Definition 4 [*Always Chosen Set*] For any data set (\mathcal{D}, p) , we define the **always chosen** set as $W^* = \{a \in X | p(a, A) > 0 \text{ for all } A \in \mathcal{D} \text{ such that } a \in A\}$.

For any complete data set generated by a FSSM, W^* must be equivalent to the set of above-reservation alternatives.

Lemma 1 Assume a complete data set (\mathcal{D}, p) admits an FSSM representation. Then, for any such representation $W^* = U^*$.

As we discuss in section 4.3, if the data set is not complete then W^* may be a strict superset of U^* : an below-satisficing alternative may always be chosen because it is only observed in choice sets containing other below-satisficing alternatives.⁷

Using the (observable) set W^* , we can define the first of two behavioral conditions which characterize the FSSM. It states that stochastic choice must only occur amongst elements of W^* . This follows from the fact that stochasticity in the satisficing model occurs only from stochasticity in search order.

Axiom 1 [*Deterministic no satisficing choice*] If $a \in X \setminus W^*$ then either $p(a, A) = 0$ or $p(a, A) = 1$ for all $A \in \mathcal{D}$.

The second condition ensures that the preference information revealed by a data set is well behaved. In order to state the condition, we introduce the following definitions.

Definition 5 [*Stochastic Revealed Preference*] Define $C(A) = \{a \in A | p(a, A) > 0\}$. a is **stochastically revealed directly preferred** to b if, for some $A \in \mathcal{D}$ $a, b \in A$ and $a \in C(A)$. a is **stochastically revealed preferred** to b if $\{a, b\}$ is in the transitive closure of the

⁷Completeness can be replaced for a weaker condition on the richness of the data set which requires observing choices from all two and three element sets.

stochastically revealed directly preferred relation. a is **stochastically strictly revealed preferred** to b if, for some $A \in \mathcal{D}$, $a \in C(A)$ and $b \notin C(A)$.

Notice that, for data generated by a FSSM, these revealed preference concepts align with the underlying utility function except in the case of two alternatives above that satisficing level. Such objects will be revealed indifferent to each other, yet may in fact be ranked according to the utility function. It is a defining feature of the satisficing model that utility differences above the threshold u^* are unimportant for behavior. Nevertheless, the FSSM implies that the stochastic revealed preference information must obey the Strong Axiom of Revealed Preference.

Axiom 2 [SARP] $C(A)$ must obey SARP: if a is stochastically revealed preferred to b then b must not be stochastically strictly revealed preferred to a .

Our first result is that axioms A1 and A2 are necessary and sufficient for a data set to have a FSSM representation

Theorem 1 *The following are equivalent:*

1. A stochastic choice dataset (\mathcal{D}, p) is generated by a FSSM.
2. A stochastic choice dataset (\mathcal{D}, p) satisfies A1 and A2.

To understand the sufficiency of the two axioms -A1 and A2-, note first that SARP allows us, through Afriat/Richter's theorem (1996), to construct a utility function which represents the stochastic revealed preference relation. Moreover, the elements of W^* will be maximal according to that utility representation, allowing for a u^* such that all elements of the always chosen set can be assigned a utility greater or equal than u^* ; while all the elements that are not always chosen are assigned an utility level below u^* . Axiom A1 guarantees deterministic choice in sets which contain at most one above-reservation alternatives, and SARP again ensures that such choices are utility maximizing. For all other choice sets, A1 ensures that alternatives with utility below u^* (and so outside W^*) are not chosen, and a suitable stochastic search order can be constructed from the random choice rules to explain the pattern of choice amongst above satisficing alternatives. .

Notice that, while Lemma 1 relies on the completeness of the data set, Theorem 1 does not. The behavioral content of the FSSM model is the same regardless of whether the data set is complete. However, the degree to which elements of the representation can be identified will be reduced in incomplete data sets, as we discuss in section 4.3.

The following examples illustrate the empirical content of the FSSM by presenting data sets which violate each of our axioms.

Example 2 [*Violation of Axiom 1*] Let $X = \{a, b, c\}$, and let $p(a, \{a, b\}) = 1$, $p(b, \{b, c\}) = \frac{1}{2}$, $p(a, \{a, c\}) = 1$ and $p(a, \{a, b, c\}) = 1$. This does not satisfy axiom 1 since $W^* = \{a\}$, but $p(b, \{b, c\}) = \frac{1}{2} \notin \{0, 1\}$. This behavior is incommensurate with the FSSM because the fact that b is chosen probabilistically from $\{b, c\}$ indicates that it must be above the satisficing level, yet this means that it should be chosen some of the time from $\{a, b, c\}$ due to the full support assumption.

Example 3 [*Violation of Axiom 2*] Let $X = \{a, b, c\}$, and let $p(a, \{a, b\}) = 1$, $p(b, \{b, c\}) = 1$, $p(a, \{a, c\}) = 0$ and $p(a, \{a, b, c\}) = 1$. This does not satisfy axiom 2 since $p(a, \{a, b\}) = 1$ means that a is stochastically strictly revealed preferred to b , $p(b, \{b, c\}) = 1$ means that b is stochastically revealed preferred to c , while $p(c, \{a, c\}) = 1$ means that c is stochastically strictly revealed preferred to a . Such behavior is also incommensurate with the FSSM as, in each case, the uniquely chosen object must have a utility strictly higher than those which are not chosen, either because all are below the satisficing level, in which case the best option is chosen, or because only the chosen object is above the satisficing level.

Theorem 1 shows the extent to which the FSSM can be tested and differentiated from other models. First, note that any data set in which $p(a, A) > 0$ for all $A \in \mathcal{D}$ will trivially satisfy both A1 and A2, and so admit an FSSM representation. This is because any data set in which the random choice rule has full support in every choice set can be rationalized by an FSSM in which every alternative is above the satisficing level, and the resulting pattern of choice is driven by the choice-set specific distribution over search orders. Second, note that the standard model of utility maximization (without indifference) is a limit case of the FSSM in which $|W^*| = 1$. Third, notice that an alternative interpretation of the FSSM is a model in which attention is complete and choices are governed by a preference relation which has indifference only amongst maximal elements. A model in which such indifference is resolved using a random tie breaking rule with full support amongst maximal elements is equivalent to the FSSM.

3.2.1 Recoverability in the FSSM

In the case of a complete data set which satisfies A1 and A2, many of the elements of the FSSM can be uniquely identified

Theorem 2 Let (\mathcal{D}, p) be a complete data generated by an FSSM $(u, u^*, \{\gamma_A\}_{A \in \mathcal{D}})$. For any FSSM representation of the data $(\bar{u}, \bar{u}^*, \{\bar{\gamma}_A\}_{A \in \mathcal{D}})$

1. $U^* = \bar{U}^*$
2. For all $a, b \notin U^*$, $u(a) > u(b) \Rightarrow \bar{u}(a) > \bar{u}(b)$
3. $\gamma_A(ar_A b \ \forall b \in \{A \cap U^*\} \setminus \{a\}) = \bar{\gamma}_A(ar_A b \ \forall b \in \{A \cap U^*\} \setminus \{a\})$ for all $A \in \mathcal{D}$,
 $a \in A \cap U^*$

Theorem 2 tells us that, in a complete data set we can uniquely identify the above-satisficing elements, the preference ordering over non-satisficing elements, and the probability that one satisficing element will be seen before another in any choice set.

3.3 Fixed Distribution Satisficing Models

So far, we have allowed stochastic search order to vary arbitrarily between choice sets: an alternative that is likely to be searched first in choice set A may be very unlikely to be searched first in choice set B . However, in some cases such an assumption may be inappropriate. For example, consider the case in which the probability of search is governed by the ‘salience’ of different alternatives: a book with a bright pink cover may be more likely to be looked at before one with a dark brown cover regardless of the set of available alternatives.

We now consider the implications of a satisficing model with full support in which the probability distribution over search orders is invariant to the set of available alternatives. We call this the ‘fixed distribution’ property.

Assumption 2 [*Fixed Distribution*] There exists a $\Gamma_X : R_X \rightarrow [0, 1]$ such that, for every $A \in \mathcal{D}$ and $r_A \in R_A$

$$\gamma_A(r_A) = \Gamma_X(r_X | r_A \subset r_X)$$

For every choice set A , it is as if the DM draws a search order from a distribution Γ_X over linear orders on the grand set of alternatives X . They then follow that search order, ignoring any alternatives that are not in fact available in A .

Definition 6 [*Fixed Distribution Satisficing Model (FDSM)*] A data set (\mathcal{D}, p) has a Fixed Distribution Satisficing Model (FDSM) representation if it has a FSSM representation in which the family of stochastic search orders $\{\gamma_A\}_{A \in \mathcal{D}}$ satisfy Fixed Distribution.

The conditions A1 and A2 are necessary for FDSM but not sufficient, implying that the FDSM is a strict subcase of FSSM. In order to obtain sufficiency, we make use of the

Total Monotonicity condition of Block & Marschak (1960). Total Monotonicity by itself is a sufficient and necessary condition for Random Utility Maximization in our environment. This implies that the FDSM is the exact intersection of the FDSM model with the Random Utility model of Block-Marschak and Falmagne (1978).

In order to define the total monotonicity condition, we first need to define the following function for each $A \in \mathcal{D}$ and $a \in A$:

$$f(a, A) = \sum_{D \in \mathbf{B}(A)} (-1)^{|D \setminus A|} p(a, D).$$

where $\mathbf{B}(A)$ is the class of supersets of A (i.e., $\mathbf{B}(A) \equiv \{D \in \mathcal{D} | A \subseteq D\}$).

Block & Marschak (1960) and Falmagne (1978) proved that the following behavioral axiom called Total Monotonicity (or Block-Marschak Monotonicity) is necessary and sufficient for a RUM representation:

Axiom 3 [*A3 Total Monotonicity*] $f(a, B) \geq 0$ for all $a \in X$, for all $B \in \mathcal{D}$.

Observe that f depends only on the data set so this axiom is testable.

Note that Total Monotonicity implies ‘standard’ monotonicity: the probability of choosing any given alternative falls as more alternatives are added to the choice set - that is $p(a, A) \geq p(a, B)$ when $A \subseteq B$.⁸ However, it is also stronger than this condition as we can see in the following example: Set $X = \{a, b, c, d\}$, let $p(a, \{a, b\}) = 0.2$, and $p(a, \{a, b, c\}) = p(a, \{a, b, d\}) = 0.19$ and $p(a, \{a, b, c, d\}) = 0.17$. We check $f(a, \{a, b\}) = p(a, \{a, b\}) + p(a, \{a, b, c, d\}) - [p(a, \{a, b, c\}) + p(a, \{a, b, d\})]$ and observe that $f(a, \{a, b\}) = -0.01$ negative and violating total monotonicity. However, we can see that standard monotonicity holds in this example.

Clearly, a FDSM cannot lead to a failure of regularity. If it did that would mean that a given satisficing item is more likely to be found first in a bigger menu than in an smaller one, which is not consistent with the idea that the probability of any search order is fixed across menus. The higher order monotonicity conditions implied by total monotonicity can be interpreted as saying that the likelihood of a satisficing item being found first decreases with its size of the menu, but the marginal effect of adding a new option to the menu

⁸To see this take the Mobius inverse representation of $p(a, A) = \sum_{D \in \mathbf{B}(A)} f(a, D)$ and $p(a, B) = \sum_{D' \in \mathbf{B}(B)} f(a, D')$ and note that if $D \in \mathbf{B}(B)$ then $D \in \mathbf{B}(A)$ and $f(a, D) \geq 0$ by total monotonicity, then we have that $p(a, A) \geq p(a, B)$.

decreases with its size. In the example above $f(a, \{a, b\}) \geq 0$ means that the impact of adding one additional item c in the menu $\{a, b\}$ on the probability of choosing a – i.e. $p(a, \{a, b\}) - p(a, \{a, b, c\})$ – is bigger than the impact of adding the same item c in the bigger menu $\{a, b, d\}$ on the probability of choosing a – i.e. $p(a, \{a, b, d\}) - p(a, \{a, b, c, d\})$.

We are ready to state the main result of this section.

Theorem 3 *The following are equivalent:*

1. A complete stochastic choice dataset (\mathcal{D}, p) is generated by a FDSM.
2. A complete stochastic choice dataset (\mathcal{D}, p) satisfies A1, A2 and Total Monotonicity (A3).

Note that, unlike Theorem 1, Theorem 3 requires a complete data set.

3.3.1 Recoverability in the FDSM

In the case of a complete data set which satisfies A1, A2 and Total Monotonicity (A3) several of the elements of the FDSM can be identified. In particular, the identification of the search orderings is improved upon the FSSM recoverability.

Theorem 4 *Let (\mathcal{D}, p) be a complete data generated by an FDSM (u, u^*, Γ_X) such that $X \setminus U^* \neq \emptyset$. For any FDSM representation of the data $(\bar{u}, \bar{u}^*, \bar{\Gamma}_X)$*

1. $U^* = \bar{U}^*$
2. For all $a, b \notin U^*$, $u(a) > u(b) \Rightarrow \bar{u}(a) > \bar{u}(b)$
3. $\Gamma_X(ar_X b) = \bar{\Gamma}_X(ar_X b)$ all $A \in D$ and $a, b \in U^*$

Theorem 4 tells us that, in a complete data set, we can identify the above-satisficing elements, granted that there is a positive probability of being seeing before some above-satisficing element. Moreover, we can identify the preference ordering over the alternatives that are surely non-satisficing elements, and the probability that one revealed satisficing element will be seen before another in any choice set, for those elements that are surely satisficing.

4 Extensions

In this section we extend our model by relaxing in turn the assumptions of full support, no indifference, and complete data

4.1 Satisficing Without Full Support

As discussed in section 3.1, the GSM model is vacuous without the full support assumption. Here we consider the empirical implication of dropping full support but maintaining the fixed distribution assumption. In such a case the identification of satisficing elements as those that are always chosen breaks down. A satisficing element a may not be chosen in some sets if it is always searched after another satisficing element b . To illustrate this point consider the following example.

Example 4 *Let $X = \{a, b, c, d\}$, $U^* = \{a, b\}$ and Γ a distribution over search orders on X with full support and where each possible search order is equally likely; moreover assume that $u(c) > u(d)$. Then, for any menu such that $U^* \subseteq A$, $p(a, A) = p(b, A) = \frac{1}{2}$, and if $A \cap U^* \neq \emptyset$, then $p(U^*, A) = 1$ and $p(c, A) = p(d, A) = 0$. Finally, $p(c, \{c, d\}) = 1$ and $p(d, \{c, d\}) = 0$. Note that this is the standard case describe in section 3.3. Now, notice that since we do not assume that Γ needs to have full support on the set of search orders on X , the same data set can be generated by the following fixed distribution satisficing model without full support: $\bar{U}^* = X$, $\Gamma((a, b, c, d)) = \Gamma((b, a, c, d)) = \frac{1}{2}$ and $\Gamma(r_X) = 0$ for all r_X linear order on X , such that $r_X \notin \{(a, b, c, d), (b, a, c, d)\}$. Furthermore, this alternative representation is not unique.*

Because it is not possible to identify the satisficing alternatives the only implication of the satisficing model without full support, but with fixed distribution is Total Monotonicity - in other words it is behaviorally indistinguishable from the random utility model. This can be seen by noting that a RUM can be reinterpreted as a satisficing model with fixed distribution by assuming that all alternatives are above the reservation level, and treating the preference orderings from the random utility model as search orders in the satisficing model.

Proposition 2 *The following are equivalent:*

1. *A complete stochastic choice dataset (\mathcal{D}, p) is generated by a FDSM without Full Support.*
2. *A complete stochastic choice dataset (\mathcal{D}, p) satisfies A3.*

4.2 Allowing for Indifference

Here we relax the no indifference assumption while keeping the Full Support conditions. Allowing for indifference potentially introduces stochasticity among non-satisficing alternatives due to the DM's rule to break ties. We assume that tie breaking works as follows, if

the DM is indifferent between two or more alternatives, and needs to choose one of them, she chooses at random from the set of indifferent alternatives with probabilities induced by the tie-breaking rule T .

Definition 7 [*Tie-breaking rule*] Let $T : X \times \mathcal{D} \rightarrow \mathbb{R}_{++}$ ⁹ be a function that assigns tie breaking weights to alternatives. In case of indifference between two or more alternatives in menu A , the DM applies the induced tie breaking rule as follows:

$$T(a|A^{\sim a}) = \frac{T(a, A)}{\sum_{b \in A^{\sim a}} T(b, A)} \quad (2)$$

where $T(a|A^{\sim a}) > 0$ is the always positive probability that a is chosen when a is indifferent to all the elements in the set $A^{\sim a} \equiv \{b \in A : u(b) = u(a)\}$, and superior to all other elements in A (i.e. $u(a) \geq u(b)$ for all $b \in A$).

Note that if $|A^{\sim a}| = 1$ then $T(a|A^{\sim a}) = 1$, and that $\sum_{b \in A^{\sim a}} T(a|A^{\sim a}) = 1$ in general.

We now extend the Full Support Search Model to allow for indifferences.

Definition 8 A data set (\mathcal{D}, p) has a Full Support Search Model with Indifferences (FSSMI) representation if there exists $u : X \rightarrow \mathbb{R}$, $u^* \in \mathbb{R}$ such that $u^* \leq \max_{a \in X} u(a)$, stochastic search orders $\{\gamma_A\}_{A \in \mathcal{D}}$ that satisfies Fixed Distribution and tie breaking rule $T : X \times \mathcal{D} \rightarrow \mathbb{R}_{++}$, such that, for any $a \in A$

$$p(a, A) = \begin{cases} \gamma_A(ar_A b \forall b \text{ s.t. } u(b) > u^*) & \text{if } u(a) \geq u^* \\ T(a|A^{\sim a}) & \text{if } a \in \arg \max_{x \in A} u(x) < u^* \\ 0 & \text{Otherwise} \end{cases} \quad (3)$$

The following example illustrates the FSSMI.

Example 5 Consider again example 1, and assume that DM's choices can be represented by a Full Support Search Model with Indifferences. Let $U^* = \{a\}$, and let $u(b) = u(c)$. Then, $p(a, A) = 1$, $p(b, A) = p(c, A) = 0$ for all A such that $a \in A$. Let the tie breaking rule be generated by $T(b, \{\{b, c\}\}) = 1$ and $T(c, \{\{b, c\}\}) = 2$, then $p(b, \{\{b, c\}\}) = \frac{1}{3}$, and $p(c, \{\{b, c\}\}) = \frac{2}{3}$

Axiom 1 is no longer necessary for the FSSMI model: stochasticity can occur amongst alternatives that are not always chosen due to indifference. In fact, it turns out that the

⁹Note that we rule out deterministic tie breaking since, the behavior of a DM that is indifferent between two alternatives a and b and always chooses a over b is behaviorally indistinguishable from a DM that prefers a over b .

behavioral implication of allowing for indifference is precisely the removal of this axiom from our set of necessary and sufficient conditions.

Theorem 5 *The following are equivalent:*

1. A complete stochastic choice dataset (\mathcal{D}, p) is generated by a FSSMI.
2. A complete stochastic choice dataset (\mathcal{D}, p) satisfies A2.

Theorem 5 highlights that the satisficing model *without* indifference can be reinterpreted as a standard optimizing model with random tie breaking, but allowing for indifference only amongst the best alternatives.

4.2.1 Recoverability in the FSSMI

Given $u^* \leq \max_{a \in X} u(a)$, the extension of the model to allow for ties, does not obscure the identification result for the satisficing set as in 4, where the always chosen set coincides with the satisficing set, i.e. $W^* = U^*$. The following theorem describes the degree to which the other elements of the model cannot be identified.

Theorem 6 *Let (\mathcal{D}, p) be a complete data generated by an FSSMI $(u, u^*, \{\gamma_A\}_{A \in \mathcal{D}}, T)$. For any FDSMI representation of the data $(\bar{u}, \bar{u}^*, \{\bar{\gamma}_A\}_{A \in \mathcal{D}}, \bar{T})$*

1. $U^* = \bar{U}^*$
2. For all $a, b \notin U^*$, $u(a) \geq u(b) \Rightarrow \bar{u}(a) \geq \bar{u}(b)$
3. $\gamma_A(ar_{Ab} \ \forall b \in \{A \cap U^*\} \setminus \{a\}) = \bar{\gamma}_A(ar_{Ab} \ \forall b \in \{A \cap U^*\} \setminus \{a\})$ for all $A \in \mathcal{D}$, $a \in A \cap U^*$
4. $T(a|A^{\sim a}) = \bar{T}(a|A^{\sim a})$ for all $a \in A$, $A \in \mathcal{D}$.

Theorem 6 tell us that in a complete data set we can uniquely identify the above-satisficing elements, the preference ordering among non-satisficing elements, the tie breaking rule when used and the probability that one satisficing element is seen before another in any choice set. Its proof follows from Theorem 2 when the identification of the tie breaking rule, since when it is used follows from the empirical choice probability among elements that are not revealed to be satisficing.

4.3 Incomplete Datasets

Here we relax the complete data set assumption while keeping the full support distribution condition and assuming no indifference. We do not work with the fixed distribution assumption since Total Monotonicity is not well defined for incomplete data sets, and the literature on Random Utility Models has not dealt with this extension.

Notice that complete data is not a necessary assumption for Theorem 1. Thus, if we drop completeness, the implications of the model are not affected, but identification becomes weaker. To see this note W^* may be a strict superset of U^* since a below-satisficing alternative may be always chosen because it is only observed in choice sets containing below-satisficing alternatives. The accuracy with which we can identify the primitives of the model given observed data depends on the richness of the data set.

Theorem 7 *Let (\mathcal{D}, p) be a data set (that needs not to be complete) that satisfies A1-A2. Then for a FSSM $(u, u^*, \{\gamma_A\}_{A \in \mathcal{D}})$ represents the data if and only if*

1. $\widetilde{W} \subseteq U^* \subseteq W^*$
2. u must represent the stochastic revealed preference relation on $X \setminus U^*$: that is if a is stochastically strictly revealed preferred to b then $u(a) > u(b)$, and if a is revealed preferred to b then $u(a) \geq u(b)$ for any $a, b \in X \setminus U^*$
3. $\gamma_A(ar_Ab \quad \forall b \in (U^* \cap A) \setminus \{a\}) = p(a, A)$ for all $A \in \mathcal{D}$, $a \in U^* \cap A$.

where $\widetilde{W} \equiv \{a \in X \mid \exists A \in \mathcal{D}, \text{ s.t } p(a, A) \in (0, 1)\}$.

Theorem 7 tell us that, when dealing with incomplete data sets one can only certainty identify satisficing choices if these have been observed chosen when other satisficing alternatives were available as well; that is, if we see them being chosen stochastically. As we established before, the satisficing set is a subset of the set of always chosen alternatives, but with incomplete data it may be a strict subset. Furthermore, one can only partially recover the preference order for the revealed not satisficing alternatives as well as a search order that only coincides with the one that generated the data for the relative probabilities for the elements that surely are in U^* .

5 Relation to Existing Literature

The paper closest in spirit to ours is Manzini & Mariotti (2014), which characterizes the random choice generated by a DM who makes choices by optimizing on a

stochastically generated consideration set. As in our model, preferences are deterministic, with randomness in choice coming from stochastic changes in attention. However, the behavioral implications of the two models are quite different, with the satisficing model being the more general. In the set up of MM, all alternatives are always chosen with positive probability in each set. In such a data set, axioms A1 and A2 are always satisfied, and so the FSSM is trivially more general than the stochastic consideration set model. Moreover, the FDSM also nests the stochastic consideration set model. This follows from the fact that, when restricted to the class of data in which all alternatives are chosen with positive probability, the FDSM model is equivalent to the class of all RUMs, and the model of MM is a strict subset of this class. Moreover, the FSSM and FDSM can accommodate data sets in which not all alternatives are chosen with positive probability.

Our work also contributes to the literature aimed at testing the satisficing model. It is well known that standard deterministic choice data cannot be used to distinguish rational choice from satisficing behavior, implying that richer data is needed. Caplin et al. (2011); Caplin & Dean (2011) showed how to test the satisficing model using ‘choice process’ data, which records not just final choice made by a decision maker, but also how choices change with contemplation time. Santos et al. (2012) utilize data in which the sequence of search is recorded to test the satisficing model. Our paper describes the implication of the satisficing model for stochastic choice data, which is arguably easier to collect than either choice process or search data.

Ours is not the first paper to characterize the behavior of random choice rules. Much of the previous work has focused on random utility models (RUMs), in which the DM chooses in order to maximize a utility function, drawn from some distribution (see for example Block & Marschak (1960); Falmagne (1978); Gul et al. (2014)). As discussed above, the FSSM is behaviorally distinct from the class of RUMs. It is easy to construct examples of FSSMs which violate regularity, and so cannot be modeled as the resulting from random utility maximization. Moreover, RUMs are not guaranteed to satisfy either axioms A1 or A2. In contrast the FDSM is behaviorally a subset of the class of RUMs. Total monotonicity A3 is necessary and sufficient for a RUM representation (Falmagne, 1978).

References

- Block, H. D. & Marschak, J. (1960). *Contributions to Probability and Statistics*, chapter Random Orderings and Stochastic Theories of Response. In Olkin (1960).
- Caplin, A. & Dean, M. (2011). Search, choice, and revealed preference. *Theoretical Economics*, 6(1), 19–48.
- Caplin, A., Dean, M., & Martin, D. (2011). Search and satisficing. *The American Economic Review*, (pp. 2899–2922).
- Eliaz, K. & Spiegel, R. (2011). Consideration sets and competitive marketing. *The Review of Economic Studies*, 78(1), 235–262.
- Falmagne, J. C. (1978). A representation theorem for finite random scale systems. *Journal of Mathematical Psychology*, 18(1), 52–72.
- Gul, F., Natenzon, P., & Pesendorfer, W. (2014). Random choice as behavioral optimization. *Econometrica*, 82(5), 1873–1912.
- Gul, F. & Pesendorfer, W. (2006). Random Expected Utility. *Econometrica*, 74(1), 121–146.
- Luce, R. & Suppes, P. (1965). *Handbook of Mathematical Psychology, Vol. III*, chapter Preference, Utility, and Subjective Probability. In R. D. Luce & Galanter (1965).
- Manzini, P. & Mariotti, M. (2014). Stochastic choice and consideration sets. *Econometrica*, 82(3), 1153–1176.
- Masatlioglu, Y., Nakajima, D., & Ozbay, E. Y. (2012). Revealed attention. *The American Economic Review*, 102(5), 2183–2205.
- McClellon, M. (2015). Unique Random Utility Representations.
- Olkin, I., Ed. (1960). *Contributions to Probability and Statistics*. Stanford: Stanford University Press.
- R. D. Luce, R. R. B. & Galanter, E., Eds. (1965). *Handbook of Mathematical Psychology, Vol. III*. New York: Wiley.
- Reutskaja, E., Nagel, R., Camerer, C. F., & Rangel, A. (2011). Search dynamics in consumer choice under time pressure: An eye-tracking study. *The American Economic Review*, (pp. 900–926).

- Richter, M. K. (1996). Revealed preference theory. *Econometrica: Journal of the Econometric Society*, (pp. 635–645).
- Salant, Y. & Rubinstein, A. (2008). (a, f): Choice with frames. *The Review of Economic Studies*, 75(4), 1287–1296.
- Santos, B. D. L., Hortacsu, A., & Wildenbeest, M. R. (2012). Testing models of consumer search using data on web browsing and purchasing behavior. *American Economic Review*, 102(6), 2955–80.
- Schwartz, B., Ward, A., Monterosso, J., Lyubomirsky, S., White, K., & Lehman, D. R. (2002). Maximizing versus satisficing: happiness is a matter of choice. *Journal of personality and social psychology*, 83(5), 1178.
- Simon, H. A. (1955). A behavioral model of rational choice. *The quarterly journal of economics*, (pp. 99–118).
- Ward, D. (1992). The role of satisficing in foraging theory. *Oikos*, (pp. 312–317).

A Proofs

A.1 Proof of Proposition 1

Proof. For any data set (\mathcal{D}, p) , set $U : X \rightarrow [0, 1]$ to be any arbitrary one to one real valued function and set $u^* = -2$; then $U^* = X$. For any menu $A \in \mathcal{D}$, let r_A^a be the set of linear orders on A such that $a r b$ for all $b \in A$. $\{r_A^a\}_{a \in A}$ therefore defines partition on R_A . Define

$$\gamma_A(r_A) = \frac{p(a, A)}{|r_A^a|} \text{ where } r_A \in r_A^a$$

Such a representation will generate p as, for any a , $u(a) \geq u^*$

$$\gamma_A(r_A | a r_A^a b \forall b \text{ s.t. } u(b) \geq u^*) = \gamma_A(r_A \in r_A^a) = p(a, A)$$

■

A.2 Proof of Lemma 1

Proof. ($W^* \subseteq U^*$) Let $a \in W^*$ then for any given $A \in \mathcal{D}$ we have $p(a, A) > 0$ then either: (i) $a \in U^*$ or (exclusive) (ii) $u(a) > u(b)$ for all $b \in A$ and $U^* \cap A = \emptyset$. Assume $a \notin U^*$ then

(since $U^* \neq \emptyset$) there exists $b \in U^*$ and, by completeness of the data, there is a menu $A' \in \mathcal{D}$ such that $a, b \in A'$, and therefore $p(a, A') = 0$, so we have a contradiction.

($U^* \subseteq W^*$) Let $a \in U^* \Rightarrow u(a) > u^*$, by FSSM $p(a, A) = \gamma_A(r_A|ar_A b \forall b \in A \text{ s.t. } u(b) > u^*) > 0$ where the last inequality follows from Full Support assumption. ■

A.3 Proof of Theorem 1

Proof. First we prove that (1) implies (2). To prove that a data set (\mathcal{D}, p) that admits a FSSM representation satisfies A1 first notice that $U^* \subseteq W^*$ even for incomplete data sets. To see this note that if $a \in U^*$ then $u(a) \geq u^*$ and since the data has a FSSM representation then $p(a, A) = \gamma_A(r_A|a r_A b \forall b \in A \cap U^* \setminus \{a\})$ for all $A \in \mathcal{D}$. Given the full support assumption, $\gamma_A(r_A|a r_A b \forall b \in A \cap U^* \setminus \{a\}) > 0$ for all $A \in \mathcal{D}$, therefore $p(a, A) > 0$ for all $A \in \mathcal{D}$ which in turn implies that $a \in W^*$

Then if $a \notin W^*$ we have that $a \notin U^*$, which in turn implies that $u(a) < u^*$. Since u is injective, there exists a $a_A^* = \text{argmax}_{a \in A} u(a)$. Then either (i) $a = \text{argmax}_{b \in A} u(b)$ or (ii) $u(a) < \max_{b \in A} u(b)$. If (i) since the data has a FSSM representation $p(a, A) = 1$; while if (ii) $p(a, A) = 0$. In either case A1 follows.

To show that A2 holds, assume, by the way of contradiction, that (i) a is stochastically revealed preferred to b and (ii) b is stochastically strictly revealed preferred to a . From (ii), given that data admits a FSSM we must have (by the full support assumption) that $u(a) < u^*$ and $u(a) < u(b)$. If a is stochastically revealed preferred to b then there must exist a sequence of alternatives c_1, \dots, c_N and choice sets A_1, \dots, A_{N-1} , such that $c_1 = a$, $c_N = b$, $c_n, c_{n+1} \in A_n$ and $c_n \in C(A_n)$. If $u(b) > u^*$ then it must be the case that $u(c_{N-1}) > u^*$ (otherwise c_{N-1} could not be chosen with positive probability when c_N was available). Iterating on this argument implies that $u(a) > u^*$. If $u(b) < u^*$ then this implies that $u(c_{N-1}) > u(b)$. If $u(c_n) < u^*$ for all n then iterating on this argument implies that $u(a) > u(b)$. Otherwise, the previous argument implies that $u(a) > u^*$. Either provides a contradiction.

Now we prove that (2) implies (1). For $A \subseteq X \setminus W^*$, $C(A) \equiv \{a \in A | p(a, A) = 1\}$ from A1. Given A2, we can generate an injective utility function using Afriat/Richter's theorem, such that $u : X \setminus W^* \rightarrow [0, \bar{u}]$ and $C(A) = \text{argmax}_{a \in A} u(a)$ for all $A \in \mathcal{D}$ such that $A \cap W^* = \emptyset$. Fix $u^* > \bar{u}$, enumerate the elements in W^* as a_1, \dots, a_N and let $u(a_n) = u^* + n$

for all $a_n \in W^*$. Thus $u : X \rightarrow \mathbb{R}$ is injective and $U^* \equiv W^*$. Furthermore, notice that A2 implies that W^* is non-empty, so that $u^* \leq \max_{x \in X} u(x)$.

For every A such that $A \cap W^*$ is non-empty, let $A^* \equiv A \cap W^*$ and define R_{A^*} the set of all linear orders on A^* , and let $R_A(r_{A^*})$ be the set of all linear orders $r_A \in R_A$ that induce the linear order $r_{A^*} \in R_{A^*}$. Then, set the probability of the set of linear orders that generate each r_{A^*} as

$$\gamma_A(r_A \in R_A(r_{A^*}) \mid a_i r_{A^*} a_j \ \forall a_j \in A^* \setminus \{a_i\}) = \frac{p(a_i, A)}{|\{r_{A^*} \in R_{A^*} : a_i r_{A^*} a_j \ \forall a_j \in A^* \setminus \{a_i\}\}|}$$

for all $a_i \in A^*$. Finally, distribute the probability mass above uniformly across the elements in $R_A(r_{A^*})$. For any $r_A \in R_A(r_{A^*})$, for a given r_{A^*} :

$$\gamma_A(r_A) = \gamma_A(r_A \in R_A \mid a_i r_{A^*} a_j \ \forall a_j \in A^* \setminus \{a_i\}) / |R_A(r_{A^*})|.$$

Where $|\cdot|$ stands for the cardinality map. Note that, as $r_A \in R_A(r_{A^*}) \mid a_i r_{A^*} a_j \ \forall a_j \in A^* \setminus \{a_i\}$ for some $a_i \in A^*$, and as $p(a_i, A) > 0$ by construction of W^* this distribution will have full support on R_A .

If $A \cap W^* = \emptyset$ then for all $r_A \in R_A$ define

$$\gamma(r_A) = \frac{1}{|R_A|}$$

Thus, between them u , u^* and $\{\gamma_A\}_{A \in \mathcal{D}}$ satisfy the requirements of an FSSM. To verify that we can generate (\mathcal{D}, p) notice that if we face a menu A we have the following cases:

(i) if $A \cap W^* = A$ then $u(a) > u^*$ for all $a \in A$ and, for each a , $p(a, A) = \gamma_A(r_A \in R_A \mid a r_A b \ \forall b \in A \setminus \{a\})$, to see that this is true observe that $\gamma_A(r_A \in R_A \mid a r_A b \ \forall b \in A \setminus \{a\}) = \sum_{r_{A^*} \in R_{A^*}} \mathbb{1}[r_{A^*} : a_i r_{A^*} a_j \ \forall a_j \in A^*] \gamma_A(r_A) = p(a, A)$.

(ii) If $A \cap W^* = \emptyset$ then $p(a, A) = 1$ if $u(a) > u(b)$ for all $b \in A \setminus \{a\}$ and zero otherwise. This follows directly from the fact that u was constructed to represent choice on such sets.

(iii) If $A \cap W^* \subset A$ and $A \cap W^* \neq \emptyset$ then we have $p(a, A) = \gamma_A(r_A \mid a r_A b \ \forall b \in (A \cap W^*) \setminus \{a\})$ if $u(a) \geq u^*$ and $p(a, A) = 0$ if $u(a) < u^*$. To see that this is true observe that by definition of W^* and A1 $p(A \cap W^*, A) = 1$. To see that this is true, observe that if we assume that $p(A \cap W^*, A) < 1$ we must have that $p(a, A) > 0$ for some $a \notin W^*$ but

that means by A1 that $p(a, A) = 1$ which is a contradiction of the fact that $p(A \cap W^*, A) > 0$.

Then $p(a, A) = 0$ if $u(a) < u^*$. For $a, b \in A \cap W^*$, the result follows as in (i). ■

A.4 Proof of Theorem 2

Proof. To prove (1) assume, by contradiction, that $U^* \neq \bar{U}^*$ and let $v \in \bar{U}^* \setminus U^*$. Then, it must be the case that $u(v) < u^*$ and $\bar{u}(v) \geq \bar{u}^*$. By completeness and the fact that $U^* \neq \emptyset$, $\exists a \in U^*$ and a $A \in \mathcal{D}$ such that $A = \{a, v\}$. Given that the data is represented by $(u, u^*, \{\gamma_A\}_{A \in \mathcal{D}})$, $p(a, A) = 1$. On the other hand, since $(\bar{u}, \bar{u}^*, \{\bar{\gamma}_A\}_{A \in \mathcal{D}})$ also represents the data it is the case that $p(a, A) \in (0, 1)$, which establishes a contradiction.

To prove (2) notice that from (1) $U^* = \bar{U}^*$. Since both, $(u, u^*, \{\gamma_A\}_{A \in \mathcal{D}})$ and $(\bar{u}, \bar{u}^*, \{\bar{\gamma}_A\}_{A \in \mathcal{D}})$ are generated by the same FSSM, then u and u^* represent the preferences given by definition 5 for all $a \notin U^*$. Therefore, it must be the case that u is a strictly increasing transformation of \bar{u} . on X/U^*

To prove (3) assume, by contradiction, that for some $A \in \mathcal{D}$, $a \in A \cap U^*$

$$\gamma_A(ar_Ab \quad \forall b \in \{A \cap U^*\} \setminus \{a\}) \neq \bar{\gamma}_A(ar_Ab \quad \forall b \in \{A \cap U^*\} \setminus \{a\})$$

then $p(a, A|\gamma_A) \neq p(a, A|\bar{\gamma}_A)$ which in turn implies that both, $(u, u^*, \{\gamma_A\}_{A \in \mathcal{D}})$ and $(\bar{u}, \bar{u}^*, \{\bar{\gamma}_A\}_{A \in \mathcal{D}})$ cannot represent the same data. ■

A.5 Proof of Theorem 3

Proof. First we prove (1) implies (2). If the complete data is generated by a FDSM then there is a triple (u, u^*, Γ_X) , take a realization of Γ_X with support on R_X and call it r_X , then define the linear ordering on $X \succ_X$: (1) $a \succ_X b$ if ar_Xb and $a, b \in U^*$, (2) $a \succ_X b$ if $u(a) > u(b)$ and $a, b \notin U^*$ and (3) $a \succ_X b$ if $a \in U^*, b \notin U^*$. Now, assign this linear ordering \succ_X the probability $\Gamma_X(r_X)$. It is direct to see that there is a Random Utility Maximization model without indifference with realizations \succ_X with probability $\Gamma_X(r_X)$. By Block & Marschak (1960) it follows that the generated data set (\mathcal{D}, p) satisfies Total Monotonicity (A3). The fact that A1 and A2 hold follows from Theorem 1.

Second we prove (2) implies (1). Because FDSM is a subcase of FSSM and A1 and A2 hold we can build an utility $u : X \mapsto \mathbb{R}$ and a threshold $u^* \in \mathbb{R}$ such that $u(x) \geq u^*$ for all

$x \in W^*$ and $u(x) < u^*$ for all $x \in X \setminus W^*$. Finally, if a complete stochastic choice dataset (\mathcal{D}, p) satisfies A3 then the model is a Random Maximization Utility model so we can recover a distribution over linear orders on X $\Gamma_X^Q : \bar{R}_X \mapsto [0, 1]$ such that $p(a, A) = \Gamma_X^Q(a\bar{r}_X b \forall b \in A)$ thanks to Falmagne (1978). Call its support \bar{R}_X the set of quasi-search ordering, because they are the results of both deterministic (degenerate) utility maximization choice and fixed random searching and satisficing behavior. To construct the Γ_X we build each element of its support by taking an element of the support of the Quasi-search ordering $\bar{r}_X \in \bar{R}_X$ and we restrict it to W^* . Here we define an equivalence class on \bar{R}_X if they have the same restriction $\bar{r}_X|W^*$, if we have any two elements $\bar{r}_X, \bar{r}'_X \in \bar{R}_X$ that have the same restriction to W^* (i.e., $x\bar{r}_X y \iff x\bar{r}'_X y$ for $x, y \in W^*$) we say $\bar{r}_X \equiv_{W^*} \bar{r}'_X$, the equivalence class set is denoted as $[\bar{r}_X]_{\equiv_{W^*}} = \{\bar{r}'_X \in \bar{R}_X | \bar{r}_X \equiv_{W^*} \bar{r}'_X\}$ then we assign to the representative of the equivalence class or the restriction $r_X|W^*$ the probability corresponding to the sum $\sum_{\bar{r}'_X \in [\bar{r}_X]_{\equiv_{W^*}}} \Gamma_X^Q(\bar{r}'_X)$. For any given restricted ordering $\bar{r}_X|W^*$ we build its transitive closure or the set of transitive extensions to X and call this set $R_X(\bar{r}_X) \subset X \times X$. We assign each of the elements of this set $\hat{r}_X \in R_X(\bar{r}_X)$ the probability $\sum_{\bar{r}_X \in [\bar{r}_X]_{\equiv_{W^*}}} \Gamma_X^Q(\bar{r}_X) / |R_X(\bar{r}_X)|$ where the numerator is the probability of the restricted to W^* quasi-search ordering $\bar{r}_X|W^*$ and the denominator is the cardinality of the previously defined set. Doing this for all elements of \bar{R}_X we build a new support R_X with probabilities as indicated that provide us with Γ_X .

Note that Γ_X has full support due to how W^* is constructed. Because W^* is the always chosen set, we know that any element of the set of restrictions $\bar{r}_X|W^*$ such that for each $a \in W^*$, $a\bar{r}_X b$ for all $b \in W^* \setminus \{a\}$ has positive probability. It follows that by definition, for any $A \in \mathcal{D}$, and for any $x \in W^*$ we have $p(x, A) > 0$, this means that $\Gamma_X^Q(\bar{r}_X \in \bar{R}_X : x\bar{r}_X y \ \forall y \in X \setminus \{x\}) > 0$. This implies that all representatives of the equivalence class of restricted orderings $r_X|W^*$ where x is searched first in W^* have positive probability. Then we have extended them to X with the uniform distribution for each set $R_X(\bar{r}_X)$ thus preserving the full support for the whole X . The reason is that the transitive closure to X of any $r_X|W^*$ contains linear search orders that have each $x \in W^*$ as the first searched element, because it contains the ordering that preserves the elements in W^* in the top and the rest at the bottom. But also it contains search orders with each element of $X \setminus W^*$ at the top for each of such elements and W^* at the bottom all with positive probability.

We have extended each restriction $\bar{r}_X|W^*$ such that we can let Γ_X be the fixed distribution of search orders. We have built a FDSM or a triple (u, u^*, Γ_X) . To verify that this FDSM model generate (\mathcal{D}, p) notice that if we face a menu A we have the following cases:

(i) if $A \cap W^* = A$ then $p(a, A) = \Gamma_X(r_X \in R_X | ar_X b \quad \forall b \in A \setminus \{a\})$, to see that this is true observe that $\Gamma_X(r_X \in R_X | ar_X b \quad \forall b \in A \setminus \{a\}) = \Gamma_X^Q(\bar{r}_X \in \bar{R}_X | a\bar{r}_X b \quad \forall b \in A \setminus \{a\})$. In this case the first equality follows from the equivalence to random utility when restricted to W^* .

(ii) If $A \cap W^* = \emptyset$ then $p(a, A) = 1$ if $u(a) > u(b)$ for all $b \in A \setminus \{a\}$ and zero otherwise. This is direct from the fact that u represents C in such sets

(iii) If $A \cap W^* \subset A$ then we have $p(a, A) = \Gamma_X(r_X | ar_X b \quad \forall b \in (A \cap U^*) \setminus \{a\})$ if $u(a) \geq u^*$ and $p(a, A) = 0$ if $u(a) < u^*$. To see that this is true observe that by definition of W^* and A1 $p(A \cap W^*, A) = 1$ then $p(a, A) = 0$ if $u(a) < u^*$. For $a, b \in A \cap W^*$ observe that by construction $\Gamma_X(r_X \in R_X | ar_X b \quad \forall b \in (A \cap U^*) \setminus \{a\}) = \Gamma_X^Q(\bar{r}_X \in \bar{R}_X | a\bar{r}_X b \quad \forall b \in (A \cap W^*) \setminus \{a\})$ Finally, observe that $\Gamma_X^Q(\bar{r}_X \in \bar{R}_X | a\bar{r}_X b \quad \forall b \in (A \cap W^*) \setminus \{a\}) = \Gamma_X^Q(\bar{r}_X \in \bar{R}_X | a\bar{r}_X b \quad \forall b \in A \setminus \{a\}) = p(a, A)$ because the facts that $p(c, A) = 0$ for any $c \notin W^*$ and that Γ_X^Q represents choice implies that $\Gamma_X^Q(\bar{r}_X \in \bar{R}_X | c\bar{r}_X a) = 0$. ■

A.6 Proof of Theorem 4

Proof. For (1) and (2) we use the results of Theorem 2.

(3) follows from the fact that $W^* \equiv U^*$ and the FDSM behaves as Random Utility in this for the elements in W^* . Assume by contradiction that $\Gamma_X(xr_X y) \neq \bar{\Gamma}_X(xr_X y)$ for some $x, y \in W^*$ but that means that in the menu $\{x, y\}$, $p(x, \{x, y\} | \Gamma_X) \neq p(x, \{x, y\} | \bar{\Gamma}_X)$ which is a contradiction.

■

A.7 Proof of Proposition 2

Proof. First we prove that (1) implies (2). If a complete data has a FDSM without full support representation then there is a triple (u, u^*, Γ_X) , take a realization of Γ_X with support on some subset or all R_X and call it r_X , then define the linear ordering on X \succ_X : (1) $a \succ_X b$ if $ar_X b$ and $a, b \in U^*$, (2) $a \succ_X b$ if $u(a) > u(b)$ and $a, b \notin U^*$ and (3) $a \succ_X b$ if $a \in U^*, b \notin U^*$. Now, assign this linear ordering \succ_X the probability $\Gamma_X(r_X)$. It is direct to see that there is a Random Utility Maximization model without indifference with realizations \succ_X with probability $\Gamma_X(r_X)$. By Block & Marschak (1960) it follows that the generated data set (\mathcal{D}, p) satisfies Total Monotonicity (A3).

Now we prove that (2) implies (1). If a complete stochastic choice dataset (\mathcal{D}, p) satisfies A3 then the model is a Random Maximization Utility model so we can recover a distribution over linear orders on X $\Gamma_X : \bar{R}_X \mapsto [0, 1]$ such that $p(a, A) = \Gamma_X^Q(a\bar{r}_X b \forall b \in A)$ thanks to Falmagne (1978). Interpret Γ_X as a fixed distribution search order, and select any injective $u : X \rightarrow \mathbb{R}$ and u^* such that $u(x) > u^*$ for all $x \in X$. By construction this triple (u, u^*, Γ_X) generates the observed data. ■

A.8 Proof of Theorem 5

Proof. First we prove that (1) implies (2). To show that A2 holds, assume, by the way of contradiction, that (i) a is stochastically revealed preferred to b and (ii) b is stochastically strictly revealed preferred to a . From (ii), given that data admits a FSSMI we must have (by the full support assumption) that $u(a) < u^*$ and $u(a) < u(b)$. If a is stochastically revealed preferred to b then there must exist a sequence of alternatives c_1, \dots, c_N and choice sets A_1, \dots, A_{N-1} , such that $c_1 = a$, $c_N = b$, $c_n, c_{n+1} \in A_n$ and $c_n \in C(A_n)$. If $u(b) > u^*$ then it must be the case that $u(c_{N-1}) > u^*$ (otherwise c_{N-1} could not be chosen with positive probability when c_N was available). Iterating on this argument implies that $u(a) > u^*$. If $u(b) < u^*$ then this implies that $u(c_{N-1}) > u(b)$. If $u(c_n) < u^*$ for all n then iterating on this argument implies that $u(a) \geq u(b)$. Otherwise, the previous argument implies that $u(a) > u^*$. Either provides a contradiction.

Now we prove that (2) implies (1). We say two items are revealed stochastically indifferent if a is revealed stochastically preferred to b and b is revealed stochastically preferred to a , in that case we denote aI^*b . We modify this relation $I^* \subseteq X \times X$ by removing its elements that have at least one item from always chosen set W^* that is non-empty by SARP. Formally we define the relation $I = \{(a, b) \in X \times X : (a, b) \in I^* \text{ and } a \in X \setminus W^* \text{ or } b \in X \setminus W^*\} \cup D(W^*)$ where $D(W^*)$ is the diagonal ordering in W^* (i.e., it contains only the elements $(a, a) \in W^* \times W^*$). I is an equivalence relation because it is reflexive, symmetric and transitive. Because A2 holds I^* is an equivalence relation, and I is still an equivalence relation because it only eliminates the indifference of the items in W^* except for reflexivity, namely the elements $(a, a) \in I^*$ for $a \in W^*$. By A2 and the definition of W^* no item in $X \setminus W^*$ is revealed indifferent to an item in W^* . The relation I induces an equivalence class that we denote as $[a]$. We concentrate on the quotient set $X^I = X/I$, we define the canonical projection $j : X \mapsto X/I$ and its inverse mapping $j^{-1} : X/I \mapsto X$. We let $\mathcal{D}^I \equiv \{j(A)\}_{A \in \mathcal{D}}$ be the indexed set by \mathcal{D} . In particular define $p_I : X^I \times \mathcal{D}^I \mapsto [0, 1]$ as $p(a^I, A^I) = \sum_{a \in j^{-1}(a^I) \cap A} p(a, A)$ for $A \in \mathcal{D}$ such that $j(A) = A^I$ and $a^I \in A^I$, this mapping is well defined. If A2 holds it follows that

the quotient dataset $\{p(a^I, A^I)\}_{a^I \in X^I, A^I \in \mathcal{D}^I}$ also satisfies SARP. Also observe, that in the quotient dataset the always chosen set $W^{I,*} = \{a^I \in X^I : p(a^I, A^I) > 0 \quad \forall A^I \in \mathcal{D}^I\}$ is such that $j^{-1}(a^I) \in W^*$ for all $a^I \in W^{I,*}$, this follows from the construction of I because the equivalence classes in W^* are singletons. Observe also that A1 holds in the quotient dataset $\{p(a^I, A^I)\}_{a^I \in X^I, A^I \in \mathcal{D}^I}$, because if $b^I \in X^I \setminus W^{I,*}$ then either $p(b^I, A^I) = 0$ or (exclusively) $p(b^I, A^I) = 1$. In fact, the set X^I is a finite choice set with elements associated with degenerated probabilities of choice $p([a] \cap A, A) = \sum_{a \in [a] \cap A} p(a, A) \in \{0, 1\}$ if $[a] \subseteq X \setminus W^*$. To see this is true assume that $p([a] \cap A, A) \in (0, 1)$ and $[a] \subseteq X \setminus W^*$, this means that there is a third element $c \in A$ such that $c \in A \setminus [a]$, that is stochastically revealed preferred to all $a \in [a]$ (i.e, $p(c, A) > 0$) and of course all elements $a \in [a]$ are stochastically revealed preferred to c , but that means that $a I c$ for all $a \in [a]$ which means that $c \in [a]$, this is a contradiction. By theorem (1) we conclude that the quotient dataset $\{p(a^I, A^I)\}_{a^I \in X^I, A^I \in \mathcal{D}^I}$ can be generated by a FSSM without indifference, thus we build a triple $(\bar{u}, u^* \{\bar{\gamma}\}_{A^I \in \mathcal{D}^I})$, that generates the quotient dataset.

With this in hand we build the FSSMI in the actual dataset $\{p(a, A)\}_{a \in X, A \in \mathcal{D}}$. (i) We build a utility function $u : X \mapsto \mathbb{R}$, by the composition $u = \bar{u} \circ j$ where j is the canonical projection defined above. By construction $u(a) > u^*$ for all $a \in W^*$ and $u(a) = u(b)$ if $b \in [a]$. Moreover, $u(b) < u^*$ for all $b \in X \setminus W^*$. (ii) The search probabilities are defined over the quotient set, we build search probabilities for the actual set X . $\bar{\gamma}$ defines a full support search distribution on each $A^I \in \mathcal{D}^I$. Now we define γ by the following algorithm: For each search ordering r_{A^I} defined in the quotient dataset, define the set of linear orders in the corresponding $A \in \mathcal{D}$ such that $j(A) = A^I$, and call it $R_A(r_{A^I})$ by $a r_A b$ if $j(a) r_{A^I} j(b)$, also for the elements of the top equivalence class in r_{A^I} (i.e., $a^{I,*} r_{A^I} b^I$ for all $b^I \in A^I \setminus \{a^I\}$). Obtain all possible linear search orders over $j^{-1}(a^{I,*}) \cap A$, for the rest of equivalence classes fix one arbitrary linear search order over it $j^{-1}(a^I) \cap A$ for each of the linear search orders for the top equivalence class. Now assign to each element $r_A \in R_A(r_{A^I})$ the probability $\gamma_A(r_A) = \bar{\gamma}(r_{A^I})/k!$ with $k = |j^{-1}(a^{I,*})|$, where $k!$ is the number of possible linear orders in the top equivalence class. This construction provides as with $\{\gamma_A\}_{A \in \mathcal{D}}$ that defines a FS random linear ordering on each \mathcal{D} . (iii) To build the menu dependent tie breaking rules we calibrate them as follows $T(a, A) = p(a, A)$ if $a \in X \setminus W^*$ and $a \in [a]$ such that $p([a] \cap A, A) = 1$. If $p([a] \cap A, A) = 0$ then we let $T(a|A \sim^a) = 1/|[a] \cap A|$. This guarantees a tie breaking rule that is always positive and that adds up to 1 as required.

We have generated a tuple $(u, u^*, \{\gamma_A\}_{A \in \mathcal{D}}, T)$ or a FSSMI representation that generates the complete dataset $\{p(a, A)\}_{a \in X, A \in \mathcal{D}}$. To verify this claim, notice that this follows imme-

diately from applying theorem (1) to generate the quotient dataset, and noticing that for non-satisficing elements we can generate the actual dataset using the calibrated tie breaking rule directly and observing that the elements in $W^{I,*}$ have a one to one correspondence to the elements in W^* . ■

A.9 Proof of Theorem 6

Proof. (1)-(3) follows from Theorem 4. To prove (4) notice that, given $U^* = \bar{U}^*$, for all $a \notin U^*$, if $T(a|A^{\sim a}) \neq \bar{T}(a|A^{\sim a})$ then, from the definition of the model $p(a, A) \neq \bar{p}(a, A)$. ■

A.10 Proof of Theorem 7

Proof. To prove (1) notice that, since we do not allow for ties, $p(a, A) \in (0, 1)$ if $a \in U^*$ then $\widetilde{W} \subseteq U^*$. Moreover if $a \in U^*$ then, given the full support assumption, $p(a, A) > 0$ for all $A \in \mathcal{D}$ then $a \in W^*$ as in Theorem 1.

(2) follows from A2. Notice that SARP does not requires complete data sets to guarantee the existence of a utility function that represents the revealed preference relation.

(3) follows from the definition of the model. Note that identification is only possible when surely revealed satisficing elements are available together in a menu. That is, we can identify the probability of a seen first than b in $A \in \mathcal{D}$, i.e. $\gamma_A(ar_A b \ \forall b \in (A \cap U^*) \setminus \{a\})$ if $a \in A \cap U^*$. Moreover notice that if this is the case, then $a, b \in \widetilde{W}$. ■