Theory and Decision (2020) 88:257–268
https://doi.org/10.1007/s11238-019-09723-6

Stochastic choice over menus

Pedram Heydari[1]

Published online: 12 October 2019

Abstract
Models of choice over menus aim at capturing the effect of some behavioral or non-standard element of decision-making on the behavior of a single decision-maker. These models are usually compared with the standard model of choice over menus, in which the decision-maker chooses a menu whose best item is better than that of all other available ones. However, in many empirical settings such as experimental studies, choice data come from a population of decision-makers with possibly heterogeneous attitudes and tastes. This heterogeneity can make the observed choices over menus stochastic. This fact calls for a stochastic characterization of models of choice over menus to be able to better compare and contrast different models empirically. In this paper, I do this task for the standard model, which would be an extension of the random utility model to the realm of choice over menus. In particular, I provide the necessary and sufficient conditions, i.e., axioms on (stochastic) choice data over menus for it to be consistent with a population of decision-makers each of whom behaves according to the standard model. The axioms that characterize the model are the axiom of revealed stochastic preferences over singletons and three rationality axioms.

Keywords Stochastic choice · Random utility · Dynamic choice · Menu

1 Introduction

This paper is an attempt to extend the random utility model (RUM) to the realm of dynamic choice. The framework that I am going to use is the one introduced by Kreps (1979). In this framework, the decision is made in two stages. In the first stage, the decision-maker chooses a menu from a set of available menus, out of which he/she would choose an individual item in the second stage. For example, he/she might first

This paper is the second chapter of my Ph.D. dissertation. I thank Christopher Chambers, Mark Machina, Joel Sobel, Simone Galperti, and seminar participants at UCSD for their helpful comments.

✉ Pedram Heydari
pedramh68@gmail.com

[1] Geisinger ADMI, 120 Hamm Dr, Lewisburg, PA 17837, USA

choose a restaurant to then choose a meal from its menu once in the restaurant. In this framework, menus derive all their values from what they contain. In other words, we abstract away from situations in which menus have intrinsic values.

Following Kreps (1979), choices over menus have been employed by numerous papers to model various issues and behavioral biases that are unidentifiable within the classical framework when we only look at choices from menus, i.e., choices in the second stage.[1] All of these models characterize the behavior of a single decision-maker. These models are often compared with the standard model where the decision-maker chooses a menu that offers the best item according to her preferences over individual items.[2] However, in many empirical studies, such as experimental settings, choice data come in an aggregate form from a population of decision-makers who might differ in their preferences and attitudes. In such scenario, choice data will appear stochastic in the sense that different menus might be observed to be chosen from the same set of available menus. This fact calls for modeling choices over menus in a stochastic fashion to better test, calibrate, compare, and contrast different models of behavior.

This paper does precisely the above task when the underlying model of individual behavior is the standard model. To do so, I assume that choices are generated via random utility maximization. Specifically, I assume that there exists a probability distribution over the set of all possible preferences over individual items, which in turn leads to a probability distribution over the set of preferences over menus of items, according to the standard model. For example, restaurant A is better than B if the best meal of A is better than the best meal of B, according to the realized preference relation over meals.[3] I provide the necessary and sufficient conditions on stochastic choices over menus for them to be compatible with the described model.

2 Related literature

On one hand, this paper is related to the literature on stochastic choice and in particular the random utility model. To understand what the random utility model entails, let X be the finite universal set of items. Menus, denoted by m, m', m_1, etc. are non-empty subsets of X. Denote the probability of choosing an item x in a menu m by $\rho(x|m)$. $\rho : X \times 2^X \setminus \emptyset \to [0, 1]$ is called the *random choice rule* and satisfies $\sum_{x \in m} \rho(x|m) = 1$. This condition is a feasibility condition requiring the chosen option to be available, i.e., to belong to m. Also, let $O(X)$ be the set of all strict orderings of X and denote a generic element of it by u. According to RUM, there exists $\mu \in \Delta(O(X))$, such that

[1] Some examples include Gul and Pesendorfer (2001), Ortoleva (2013), Kopylov (2012). These models are discussed in more detail in Sect. 2.

[2] This setting implies that the decision-maker does not perceive any uncertainty regarding future mood, nor any temptation, cost of thinking, or any other behavioral biases in the second stage and her preferences are stable over time.

[3] Like the standard random utility model, one can attribute the stochasticity of choice alternatively to a single decision-maker with random preferences or a population of decision-makers who have different preferences in the first stage.

$$\rho(x|m) = \mu(u \in O(X)|\forall y \in m, u(x) \geq u(y)).$$

In words, $\rho(x|m)$ is equal to the likelihood of having an strict ordering that places x above all other items in m.[4]

Not all random choices are consistent with RUM. McFadden and Richter (1990) provided the condition that characterizes RUM.[5] The condition is called the *axiom of revealed stochastic preferences* (ARSP), which goes as the following:
Axiom of Revealed Stochastic Preferences: For any list $x_1 \in m_1, \ldots, x_n \in m_n$, the following is true:

$$\sum_{i=1}^{n} \rho(x_i|m_i) \leq \max_{u \in O(X)} \sum_{i=1}^{n} 1_{\{\forall \quad x_j \in m_i, u(x_i) \geq u(x_j)\}}.$$

The intuition for ARSP is the following: For every $u \in O(X)$, let $T(u)$ be the number of items in the list that are ranked above all other items in their respective menus, according to u. Also, let the right-hand side of the inequality above be equal to T^*. Therefore, for any $u \in O(X)$, $T(u) \leq T^*$. On the other hand, the left-hand side of the inequality above is some weighted average of $(T(u))_{u \in O(X)}$ and, therefore, is less than T^*, hence the inequality above.

On the other hand, this paper is closely related to the literature that studies choices over menus. Menus are meant to represent the first stage of a dynamic decision process that determine the available options in future stages. Apart from their relevance to everyday decisions, in the theoretical literature, menus have been used to model and identify some anticipated behavioral aspect or irregularity in decision-making that cannot be fully identified in the classical static environments.

In the first paper on the subject, Kreps (1979) first characterizes deterministic preferences over menus that are compatible with a single standard decision-maker. It turns out a single axiom is all that is needed, which requires the union of any two menus to be indifferent to the more preferred one. Furthermore, Kreps (1979) studies the effect of subjective uncertainty regarding future mood (i.e., second-stage preferences) on preferences over menus and contrasts it with preferences of a standard decision-maker, who does not perceive such uncertainty. He shows that while a standard decision-maker does not find the combination of two menus more valuable than both, a decision-maker who perceives uncertainty regarding his/her future preferences might do so. In other words, such uncertainty leads to a preference for flexibility.

Gul and Pesendorfer (2001), henceforth GP (2001), models the effect of temptation and costly self-control in the second stage on preferences over menus. They consider two scenarios. In one scenario, the decision-maker has temptation, but has the ability to abstain by exerting costly self-control. In the other scenario, the decision-maker falls to his/her temptation, no matter how much will they exert their will. The second scenario is equivalent to the dual-self-model of intertemporal decision-making

[4] RUM only focuses on strict orderings of X, because when we allow for arbitrary indifferences among items, the model loses its empirical content, and any random choice rule will be rationalizable by the model. This can be done if we put all the probability mass on the ordering that is indifferent between all items in X and break the ties in a way that matches our random choice rule.

[5] Block and Marschak (1960) also gave an alternative axiomatizations of the random utility model.

proposed by Strotz (1955). In contrast to Kreps (1979), the decision-makers modeled in GP (2001) might prefer commitment rather than flexibility. In particular, for these decision-makers, the combination of any two menus will always be ranked in between the two.

Ortoleva (2013) studies the effect of costly contemplation of choice or thinking in the second stage on preferences over menus. In this model, such costly contemplation might also lead to a preference for commitment. Also, Kopylov (2012) studies the effect of perfectionism and guilt aversion on choices over menus. In all these models, in the first stage, the decision-maker anticipates these irregularities and adjusts her choices over menus accordingly to maximize her own utility. In some cases, the preferences of the first- and second-stage selves are aligned, such as in Kreps (1979), Ortoleva (2013) and temptation with self-control model of GP (2001). In some other cases, there might be a conflict of interest between the two, such as the temptation without self-control model of GP (2001).

There are a number of papers that incorporate random preferences in modeling choices over menus. Stovall (2010) models a decision-maker who has temptation *with* self-control as defined in GP (2001), but is uncertain about the temptation preference which he/she would have in the second stage. In other words, the decision-maker expects to have random temptations in the second stage when he/she is to choose an item from the menu chosen in the first stage. Stovall (2010) assumes that the random temptation preferences come from a finite set. The decision-maker anticipates this randomness in the first stage and attaches an “expected utility” to each available menu with the expectation being taken with respect to the probabilities of different temptation preferences. In contrast to my work, since the randomness is resolved in the second stage in this model, preferences and choices over menus will be deterministic.

In a similar model, Dekel and Lipman (2012), henceforth DL (2012), characterize a Strotzian decision-maker—one who suffers from temptation but without self-control as defined in GP (2001) [6]—who has a single normative preference over individual items, but is uncertain about which temptation preference which he/she is going to have in the second stage. They call such decision-maker random Strotzian. In this paper, as in Stovall (2010), since the randomness is resolved in the second stage, it will lead to deterministic preferences and choices in the first stage. DL (2012) show that preferences that are modeled in Stovall (2010)—what they call random GP preferences—have also a random Strotzian representation, assuming that the probability measure over the set of temptation preferences that the decision-maker considers is nontrivial. They define a measure as nontrivial if it assigns zero measure to the set of utility functions (preferences) that are indifferent between all individual items.

Chatterjee and Krishna (2009) consider a decision-maker that has two selves, a long run self and a virtual self who have deterministic but different preferences over the individual items. At stage one, the decision-maker—possessed by his/her long run self—anticipates that his/her virtual self will take over in the second stage with some privately known probability. Responding to this anticipation, the decision-maker evaluates each menu by the expected utility of the chosen item in the second stage evaluated by the long run self’s utility function and the expectation being taken with

[6] The Strotzian model is inspired by Strotz (1955).

respect to the probability that the virtual self will take over. Once again, this model leads to deterministic preferences and choices in the first stage.

Another close paper to mine is a more recent work by Fudenberg and Strzalecki (2015), henceforth FS (2015). This paper also aims at importing stochastic properties of choice to a dynamic environment. In some aspects, FS (2015) Fudenberg and Strzalecki (2015) is more general than my approach. In particular, this paper allows for environments with an arbitrary finite number of stages, rather than only two stages. Moreover, at each stage, they allow the menus to not only determine the future available options, but also to have intrinsic values of their own. In contrast to my approach, FS (2015) assumes that we observe choices in all stages. Moreover, it focuses on specific stochastic choice models, such as logit, and does not characterize the general effect of having random utilities on choices over menus.

In the next section, I introduce the stochastic version of the standard model and provide a characterization of it.

3 Standard model

3.1 Preliminaries

X is the universal set of items and is finite. *Menus* are subsets of X and are denoted by m, m', m_1 and so on.[7] $M \subseteq 2^X \setminus \{\{\emptyset\}\}$ denotes a generic collection of menus that contain more than just the empty menu. S_M denotes the union of all menus in M. For every M and m, let $M - m = \{m' - m | m' \in M\}$. Call m a singleton if $|m| = 1$. The set of all strict orderings of X is denoted by $O(X)$, and the set of weak orderings of X is denoted by $WO(X)$. Without loss of generality, we can represent each ordering of X, weak or strict, by an ordinal utility function. For the sake of simplicity, I use orderings and utility functions interchangeably.

I assume that observable data consist of the probability of choosing an arbitrary menu m in an arbitrary collection of menus M that contains m. Such probabilities are represented via a function called the random menu choice rule, (RMCR), denoted by $\rho : 2^X \times 2^{2^X} \setminus \{\{\emptyset\}\} \to [0, 1]$, such that $\sum_{m \in M} \rho(m|M) = 1$. Also, for all menus m and collection of menus M, ρ satisfies $\rho(m|M \cup \{\emptyset\}) = \rho(m|M)$. The latter means that empty menus do not affect the choice probability of other available menus.

3.2 Representation

Before the first stage, $u \in O(X)$ is realized according to a probability distribution that is denoted by $\mu \in \Delta(O(X))$. In the first stage, one of the available menus in M whose best item is better than that of all others according to u is chosen. In this model, indifferences can arise whenever the intersection of m with some other menu in M is not empty. To be more precise, whenever the best item according to the realized utility function u lies in the intersection of m and some other menu in M, and then,

[7] This treatment is a bit different from the standard treatment of menus in the literature such as the one introduced in the previous section. Here, for the ease of exposition, I allow menus to be empty as well.

the decision-maker becomes indifferent between m and those menus containing that best item. This model stays silent about how these ties are broken. In other words, conditional on a realized utility function, u, I only assume that the chosen menu is one of the best available ones according to V_u.[8] Formally, this translates to the following: every $u \in O(X)$ induces a weak ordering $V_u \in WO(2^X \setminus \{\emptyset\})$ of menus in the following way:

$$V_u(m) = \max_{x \in m} u(x).$$

As such, the standard model is equivalent to:

$$\begin{aligned}\mu[u|\forall m' \in M, V_u(m) > V_u(m')] &\leq \rho(m|\{m\} \cup M) \\ &\leq \mu[u|\forall m' \in M, V_u(m) \geq V_u(m')].\end{aligned} \tag{1}$$

The inequality on the left requires that the probability of choosing m against the menus in M to be weakly greater than the probability of having a utility function u, such that V_u ranks m strictly above all menus in M. The inequality on the right requires the probability of choosing m against the menus in M not to be greater than the probability of having a utility function u, such that V_u ranks m weakly above all menus M.

3.3 Axioms

In this section, I provide the axioms that characterize the standard model. Notice that in this model, the restriction of observed choices to the collection of singletons is similar to choices that are generated by a random utility maximizer in the realm of static choice. Thus, the restriction of ρ to the set of singletons must satisfy ARSP. This is stated formally in the following axiom.

1. *ARSP Over Singletons* Let n be a an arbitrary natural number. Also, for every $i \in \{1, 2, \ldots, N\}$, let M_i be a collection of singletons and $\{x_i\} \in M_i$. Then, the following holds:

$$\sum_{i=1}^{n} \rho(\{x_i\}|M_i) \leq \max_{u \in O(X)} \sum_{i=1}^{n} 1_{\{\forall \{x_j\} \in M_i, u(x_i) \geq u(x_j)\}}.$$

Now, consider three available restaurants with mutually exclusive menus m, m', m''. Suppose that after a while, m and m' merge into a single restaurant that serves the menu $m \cup m'$. After this change, the probability of choosing the new restaurant, $m \cup m'$ would be the same as the probability of choosing either m or m', before the merging took place. This observation leads us to the next merging additivity axiom.

[8] This is because it is not clear how such ties should be broken. Moreover, any tie breaking rule will only make the model more complicated, without adding much to its content.

2. *Merging Additivity* For all mutually exclusive menus m, m', and m'', we have:

$$\rho(m \cup m'|\{m \cup m', m''\}) = \rho(m|\{m, m', m''\}) + \rho(m'|\{m, m', m''\}).$$

To see how the next axiom works, consider a set of available restaurants M whose menus do not overlap with restaurant m. The probability of choosing m against the restaurants in M is only sensitive to the union of the items offered by restaurants in M and not to how these items are distributed across different restaurants. This property is formalized in the following reorganization invariance axiom.

3. *Reorganization invariance* For every menu m and sets of menus M and M', $S_M = S_{M'}$ and $m \cap S_M = \emptyset$ imply:

$$\rho(m|\{m\} \cup M) = \rho(m|\{m\} \cup M').$$

The next axiom, disjointing inequalities, consists of two inequalities. To understand the axiom, consider an arbitrary set of available restaurants. If one of the restaurants discards all the items on their menu that are found on the menu of the other available restaurants, its choice probability would weakly decrease. Moreover, the level of decrease in its choice probability is not greater than its choice probability when it offers only the discarded items, and those items are removed from the menus of all other available restaurants. These inequalities are formally stated below.

4. *Disjointing inequalities* For every menu m and set of menus M, we have:

(a)

$$\rho(m - S_M|\{m - S_M\} \cup M\}) \leq \rho(m|\{m\} \cup M).$$

(b)

$$\begin{aligned}\rho(m|\{m\} \cup M) - \rho(m - S_M|\{m - S_M\} \cup M\}) \\ \leq \rho(m \cap S_M|\{m \cap S_M\} \cup (M - m)).\end{aligned}$$

Now, we are ready to state the theorem.

Theorem *An RMCR ρ satisfies axioms 1–4 if and only if there is a probability distribution over O(X). such that the ρ satisfies (1), i.e.. ρ is rationalizable by the standard model.*

Proof The formal proof is in the appendix. Here, I provide an informal sketch for the "only if" part. First, by ARSP over singletons and the fact that ARSP characterizes RUM, it is guaranteed that a probability measure μ over preferences exists, such that the restriction of ρ to singletons can be represented by (1). Next, using merging additivity and reorganization invariance, for any such μ, I prove the validity of (1) for bigger menus and when the menu in consideration is disjoint with other available menus. Finally, using the disjointing inequalities, I prove the validity of (1) for cases where the available menus might overlap. □

Identification Note that when we confine ourselves to menus of deterministic outcomes in X, the inferred probability measure $\mu \in O(X)$ above may not be unique.[9] However, in an alternative framework, we can allow menus to contain lotteries over X and assume that the decision-maker has a random cardinal utility function that evaluates lotteries in $\Delta(X)$. In this framework, if we require the decision-maker to be an expected utility maximizer in the realm of lotteries, then we can replace the axioms of random expected utility model of Gul and Pesendorfer (2006) with our ARSP axiom over singletons. Since the random expected utility model of Gul and Pesendorfer (2006) uniquely identifies the probability distribution of the cardinal utility function, our model would do so, as well. Other than that, the rest of our axioms remain unchanged.

One important feature of the model is that adding an item x to a menu m will always increase the chance of choosing that menu against any collection of menus M that are disjoint with $m \cup \{x\}$. This is because all the preferences that rank m above all menus in M still do so for $m \cup \{x\}$. However, there are extra new preferences that rank $m \cup \{x\}$ above all menus in M, those that place x above all items that are found in m and the menus in M. Therefore, all such preferences also choose $m \cup \{x\}$ over all menus in M. One can think of this property as a form of stochastic monotonicity.

4 Conclusion

This paper is an attempt to characterize the implications of random preferences over a set of items, when the actual choice objects are menus of such items. I assumed that menus are not intrinsically valuable, and they draw all their values from what they contain. I studied choices over menus generated by a population of decision-makers, each of whom behaves according to the standard model. I showed that such choices are characterized by four axioms, namely ARSP over singletons, merging additivity, reorganization invariance, and disjointing inequalities.

One natural step forward is to investigate stochastic choices of a population of Strotzian decision-makers who have different pairs of time-inconsistent preferences over the set of individual items. Another natural future step is to incorporate the stochasticity of choice in other famous models of choice over menus, such as the temptation with self-control model of GP (2001) and the model of preference for flexibility in Kreps (1979), among others.

A Proof of the Theorem

To prove the theorem, the following lemma is useful.

Lemma 1 *For all disjoint m and m', the following is true:*

$$\mu[u|V_u(m - m') > V_u(m')] + \mu[u|V_u(m \cap m') > V_u(m' - m)]$$

[9] It is a well-known property of RUM that the inferred probability measure is not always uniquely identified. For an example of the multiplicity of the probability measure, see McFadden (1978).

$$= \mu[u|V_u(m) \geq V_u(m')].$$

Proof Consider a utility function u for which $V_u(m) \geq V_u(m')$ and let $x_u^* = \underset{x\in m\cup m'}{\operatorname{argmax}} \; u(x)$. Therefore

$$x_u^* \in m.$$

This implies that $x_u^* \in m - m'$ or $x_u^* \in m \cap m'$. If the former is true, then $V_u(m - m') > V_u(m')$, and if the latter is true, then $V_u(m \cap m') > V_u(m' - m)$. Therefore

$$\begin{aligned}&\mu[u|V_u(m - m') > V_u(m')] + \mu[u|V_u(m \cap m') > V_u(m' - m)] \\ &\quad \geq \mu[u|V_u(m) \geq V_u(m')].\end{aligned} \tag{*}$$

Similarly, we can show that if for a utility function u if $V_u(m - m') > V_u(m')$ or $V_u(m \cap m') > V_u(m' - m)$, then $V_u(m) \geq V_u(m')$. This implies that

$$\begin{aligned}&\mu[u|V_u(m - m') > V_u(m')] + \mu[u|V_u(m \cap m') > V_u(m' - m)] \\ &\quad \geq \mu[u|V_u(m) \leq V_u(m')].\end{aligned} \tag{**}$$

$(*)$ and $(**)$ together will prove the lemma. □

Now, we are ready to state the proof of the theorem.
If part It is easy to check that if ρ satisfies (1), that is:

$$\begin{aligned}\mu[u|\forall m' \in M, V_u(m) > V_u(m')] &\leq \rho(m|\{m\} \cup M) \\ &\leq \mu[u|\forall m' \in M, V_u(m) \geq V_u(m')],\end{aligned} \tag{1}$$

then it satisfies ARSP over singletons, merging additivity, and reorganization invariance. It is also easy to check that if ρ satisfies (1) and $m \cap S_M = \emptyset$, then ρ satisfies the disjointing inequalities. Therefore, we only need to show the disjointing inequalities follow by assuming (1) and $m \cap S_M \neq \emptyset$. Since $(m - S_M) \cap S_M = (S_M - m) \cap (m \cap S_M) = \emptyset$, we have the following:

$$\begin{aligned}&\rho(m - S_M|\{m - S_M\} \cup M\}) = \rho(m - S_M|\{m - S_M, S_M\}), \\ &\quad = \mu[u|V_u(m - S_M) > V_u(S_M)] \qquad (2) \\ &\rho(m \cap S_M|\{m \cap S_M\} \cup (M - m)) = \rho(m \cap S_M|\{m \cap S_M, S_M - m\}) \\ &\quad = \mu[u|V_u(m \cap S_M) > V_u(S_M - m)]. \qquad (3)\end{aligned}$$

The first equalities in both lines above follow from merging additivity, which is a consequence of (1). The second equalities follow directly from (1). Now, (2) and (3) imply the following:

$$\rho(m - S_M|\{m - S_M, S_M\}) + \rho(m \cap S_M|\{m \cap S_M, S_M - m\})$$

$$= \mu[u|V_u(m - S_M) > V_u(S_M)] + \mu[u|V_u(m \cap S_M) > V_u(S_M - m)]$$
$$= \mu[u|V_u(m) \geq V_u(S_M)], \tag{4}$$

where the last equality follows from lemma 1. On the other hand, it is easy to see that

$$\mu[u|V_u(m) \geq V_u(S_M)] = \mu[u|\forall m' \in M, V_u(m) \geq V_u(m')].$$

The latter together with (1) and (4) imply

$$\rho(m|\{m\} \cup M) \leq \rho(m - S_M|\{m - S_M\} \cup M) + \rho(m \cap S_M|\{m \cap S_M\} \cup (M - m)).$$

Also, (1) and (2) imply:

$$\rho(m - S_M|\{m - S_M\} \cup M) \leq \rho(m|\{m\} \cup M).$$

The last two inequalities establish the disjointing inequalities. □

Only if part Assume that ρ satisfies axioms ARSP, merging additivity, reorganization invariance, and the disjointing inequalities. ARSP over singletons guarantees that there exists a probability measure, μ over the set of strict orderings of singletons or equivalently X, such that for any set of singletons $\{x_1\}, \{x_2\}, \ldots, \{x_n\}$, we have:

$$\rho(\{x_1\}|\{\{x_1\}, \{x_2\}, \ldots, \{x_n\}\}) = \mu[u|\forall i \neq 1, u(x_1) > u(x_i)]. \tag{5}$$

Now, we show that ρ together with μ satisfies (1). We break down the proof into two separate, exhausting cases:

1. $\rho(m|\{m\} \cup M)$ where m is disjoint with all menus in M.
2. $\rho(m|\{m\} \cup M)$ where m is allowed to overlap with menus in M.

- *Case 1*
  For any menu m, denote the set of all its subsets of cardinality one, i.e.. the finest partition of it by FP_m. Since m is disjoint with all menus in M, and it is the union of all singletons created by its elements, merging additivity through induction implies:

$$\rho(m|\{m\} \cup M) = \sum_{x \in m} \rho(\{x\}|FP_m \cup M).$$

  By reorganization invariance, we have:

$$\rho(\{x\}|FP_m \cup M) = \rho(\{x\}|FP_m \cup FP_{S_M}) = \rho(\{x\}|FP_{S_M \cup m}).$$

  Now, since all menus in FP_{S_M} are singletons, (5) implies:

$$\rho(\{x\}|FP_{S_M \cup m}) = \mu[u|\forall x' \in FP_{S_M \cup m} - \{x\}, u(x) > u(x')].$$

Therefore

$$\sum_{x\in m}\rho(x|FP_m\cup M)=\sum_{x\in m}\mu[u|\forall x'\in FP_{S_M\cup m}-\{x\},u(x)>u(x')].$$

And finally

$$\sum_{x\in m}\mu[u|\forall x'\in FP_{S_M\cup m}-\{x\},u(x)>u(x')]=\mu[u|\forall m'\in M,V_u(m)>V_u(m')].$$

Therefore

$$\rho(m|\{m\}\cup M)=\mu[u|\forall m'\in M,V_u(m)>V_u(m')],$$

which is equivalent to (1), when m does not intersect with any of the menus in M.

– *Case 2*

In this case, m is allowed to have a non-empty intersection with S_M. According to the disjointing inequalities and reorganization invariance, we have:

$$0\leq\rho(m|\{m\}\cup M)-\rho(m-S_M|\{m-S_M,S_M\})\leq\rho(m\cap S_M|\{m\cap S_M,S_M-m\}).$$

Now, since $(m-S_M)\cap S_M=(m\cap S_M)\cap(S_M-m)=\emptyset$, according to case 1:

$$\begin{aligned}\rho(m-S_M|\{m-S_M,S_M\})&=\mu[u|V_u(m-S_M)>V_u(S_M)]\\ \rho(m\cap S_M|\{m\cap S_M,S_M-m\})&=\mu[u|V_u(m\cap S_M)>V_u(S_M-m)].\end{aligned}$$

Therefore

$$\begin{aligned}&\rho(m-S_M|\{m-S_M,S_M\})+\rho(m\cap S_M|\{m\cap S_M,S_M-m\})\\ &\quad=\mu[u|V_u(m-S_M)>V_u(S_M)]+\mu[u|V_u(m\cap S_M)>V_u(S_M-m)]\\ &\quad=\mu[u|\forall m'\in M,V_u(m)\geq V_u(m')],\end{aligned}$$

where the last equality follows from Lemma 1. On the other hand, the following is true:

$$\rho(m-S_M|\{m-S_M,S_M\})=\mu[u|\forall m'\in M,V_u(m)>V_u(m')].$$

The last two equations together with the disjointing inequalities imply

$$\begin{aligned}&\mu[u|V_u(m)>V_u(m')\quad\forall m'\in M]\leq\rho(m|\{m\}\cup M)\leq\mu[u|V_u(m)\\ &\quad\geq V_u(m')\quad\forall m'\in M].\end{aligned}$$

$\square$

References

Block, H. D., & Marschak, J. (1960). Random orderings and stochastic theories of responses. *Contributions to Probability and Statistics*, *2*, 97–132.

Chatterjee, K., & Krishna, R. V. (2009). A "dual self" representation for stochastic temptation. *American Economic Journal: Microeconomics*, *1*(2), 148–67.

Dekel, E., & Lipman, B. L. (2012). Costly self control and random self indulgence. *Econometrica*, *80*(3), 1271–1302.

Fudenberg, D., & Strzalecki, T. (2015). Dynamic logit with choice aversion. *Econometrica*, *83*(2), 651–691.

Gul, F., & Pesendorfer, W. (2006). Random expected utility. *Econometrica*, *74*(1), 121–146.

Gul, F., & Pesendorfer, W. (2001). Temptation and self control. *Econometrica*, *69*(6), 1403–1435.

Kopylov, I. (2012). Perfectionism and choice. *Econometrica*, *80*(5), 1819–1843.

Kreps, D. M. (1979). A representation theorem for "preference for flexibility". *Econometrica*, *47*(3), 565–577.

McFadden, D. (1978). Modeling the choice of residential location. *Transportation Research Record*, *673*, 72–77.

McFadden, D., & Richter, M.K. (1990). Stochastic rationality and revealed stochastic preference. *Preferences, uncertainty, and optimality, essays in honor of Leo Hurwicz* (pp. 161–186). Westview Press: Boulder, CO

Ortoleva, P. (2013). The price of flexibility: Towards a theory of thinking aversion. *Journal of Economic Theory*, *148*(3), 903–934.

Stovall, J. E. (2010). Multiple temptations. *Econometrica*, *78*(1), 349–376.

Strotz, R. H. (1955). Myopia and inconsistency in dynamic utility maximization. *The Review of Economic Studies*, *23*(3), 165–180.