INFLUENCE AND CORRELATED CHOICE

CHRISTOPHER P. CHAMBERS, YUSUFCAN MASATLIOGLU,
CHRISTOPHER TURANSICK

Abstract. As a means for testing for the presence of influence across individuals, we discuss the concept of a stochastic choice function for a group of agents which allows observation of correlation structure (correlated choice rule). We ask when choice behavior is consistent with an underlying unobserved (latent) variable or signal which jointly governs preferences: this hypothesis represents absence of influence. Key is the property of *marginality*, which demands the independence of any given agents' budgetary choices from the budgets faced by the remaining agents. Marginality permits the construction of well-defined marginal stochastic choice functions. Marginality and non-negativity of an analogue of the Block-Marshack polynomials are equivalent to joint stochastic rationality for small environments. For larger environments, we offer an example of a correlated choice function establishing that each of the marginal stochastic choice functions may be stochastically rational while the correlated choice function is not. Thus, the detection of influence can be aided by studying correlated choice data.

1. Introduction

People are subject to social influence. This is not a controversial statement and the empirical literature has shown this to be true in many different environments.[1]

Date: March 9, 2021.
Chambers: Department of Economics, Georgetown University, ICC 580 37th and O Streets NW, Washington DC 20057. E-mail: Christopher.Chambers@georgetown.edu. Masatlioglu: University of Maryland, 3147E Tydings Hall, 7343 Preinkert Dr., College Park, MD 20742. E-mail: yusufcan@umd.edu Turansick: Department of Economics, Georgetown University, ICC 580 37th and O Streets NW, Washington DC 20057. E-mail: cmt152@georgetown.edu.

[1]To name a few, Topa (2001) finds that social connections aid in job search. Calvó-Armengol et al. (2009) find that higher levels of centrality in a social network lead to higher educational outcomes of students. Finally, Mas and Moretti (2009) find that there are productivity spillovers from high productivity workers to their coworkers.

Beyond this empirical literature, there is a rich econometric literature which concerns itself with the identification of social influence parameters in various models.[2] However, many of these empirical and econometric studies rely on heavily parametric models to find evidence of social influence in environments where we already expect to find social influence. One may then ask, absent parametric restrictions, what indicates the presence of social influence? We take a decision theoretic approach to answer this question.

The difficulty in identifying and detecting social influence is well described in Manski (1993). Manski terms this problem the *reflection problem*. It arises when researchers aim to infer the effect that the average behavior of a group has on members of that group. In a broader context, the reflection problem arises when researchers attempt to disentangle endogenous effects (social influence) from contextual effects (states of the world and preference draws). The problem lies in the simultaneity of choice and simultaneous actualization of preferences. Manski proposes two solutions to the reflection problem: tighter theory and richer data.

The study of social influence is a relatively new topic of investigation in the field of decision theory. Much of the work done takes the first approach suggested by Manski in developing tighter theory. Cuhadaroglu (2017) proposes a deterministic two stage model of choice where social influence is modeled by the (partial) completion of incomplete preferences. In the first stage, agents independently maximize a transitive but not necessarily complete preference. In the second stage, agents use the preferences of their peers to render comparisons that they were unable to make in the first stage. Borah et al. (2018) consider a deterministic two stage model of choice where social influence is modeled by consideration sets. In the first stage, agents refine their choice set by only considering options which are frequently chosen by others. In the second stage, agents maximize their preferences over their refined choice set. More recently, Chambers et al. (2020) consider a stochastic and parametric model of choice where social influence is modeled as one agent's choice (probability) altering the "utility" of

[2]See Bramoullé et al. (2009), Blume et al. (2011), and Blume et al. (2015) as examples. De Paula (2017) provides a selective review of recent studies.

that choice for other agents. Notably, this work studies joint stochastic choice behavior but does not consider the correlation structure of choices across agents.

We distinguish our paper from these past works in two important ways. The first is that we explore Manski's second proposed solution by considering a richer type of data. We consider correlated stochastic choice data, a novel type of data, which allow for researchers to observe the joint choices of multiple agents. Further distinguishing our paper from prior work is how we model influence. We model the *absence of influence*. We consider agents who face correlated random draws of preferences and think of these draws as the contextual effects of our model. This is the entirety of our model. Under the assumption of rational decision making, any falsification of our model can be thought of as induced by influence, the endogenous effect *potentially* in our environment.

In our model, agents are fully defined by three objects; their preferences, their choice sets, and their choices. Thus, if one agent were to influence another agent, it must be through one of these three objects. Perhaps unsurprisingly, influence along each of these channels is not equally detectable. Consider the following example. Suppose that we have one decision maker, who we will call the leader, whose preferences vary with an underlying state of the world. Now suppose that we have a second decision maker who, upon observing the preference realization of the leader, chooses their preference to match the preference of the leader. In the context of this story, there is obviously a strong type of influence present. However, note that this type of influence is observationally equivalent to perfectly correlated draws of preferences. With this in mind, we give the following (potentially) restrictive definition of *detectable influence*. We say that choices exhibit detectable influence if the choices cannot be rationalized by a common underlying state driving preferences. We term the model in which there is a common underlying state driving preferences the *correlated random utility model*, or CRUM for short.

CRUM is a multi-agent extension of the similarly named *random utility model* (RUM) of Block et al. (1959). The random utility model aims to explain stochastic choice via a process in which an agent randomly draws a preference and, upon realization of the preference, maximizes their preference.[3] Under this explanation, the random preference can be understood as depending on variables unobservable to an economist. For example, economic data may not include the state of the weather at a given moment in time. However, the weather clearly influences an individual's preference between hot or iced coffee. So, the underlying variables unobserved by the economist may vary in a random way, leading to the appearance of random utility.

Since Falmagne (1978) and McFadden and Richter (1990), we have understood the empirical content of the random utility model.[4] Thus, we can test whether a single decision maker behaves *as if* there is some underlying unobservable state driving preference draws.

As mentioned previously, we suggest that the absence of influence should be characterized by a form of joint stochastic rationality. If two decision makers' preferences are both determined by the same (unobservable to the economist) state, their choices will generally exhibit correlation. But this correlation should not be taken as direct evidence for the presence of influence. Both individuals may choose hot over iced coffee simply because it is cold. The same external state influences each of them, but they do not influence each other.

In order to distinguish between correlation and influence, we introduce a novel type of data into the stochastic choice model and presume observability of correlated choice behavior. We refer to this object as a *correlated choice rule*. One of our key contributions is the observation that correlated choice rules are much better equipped to detect influence than standard stochastic choice data. To see this, consider the following example. Suppose that two people, Alice and Bob, each decide whether to carry an umbrella with them for the day. When faced with this decision, Alice and Bob

[3]For other explanations of stochastic choice, see Machina (1985), Manzini and Mariotti (2014), Cerreia-Vioglio et al. (2019), and Allen et al. (2021).
[4]See also (Barberá and Pattanaik, 1986).

both decide to take an umbrella a quarter of the time. With these data, it would be reasonable to say that there is a chance of rain a quarter of the time. Now suppose that we observe that Alice and Bob never both choose to carry an umbrella at the same time. These data tell us that half of the time, at least one of Alice or Bob will choose to carry an umbrella. This observation refutes the claim that it rains a quarter of the time. An "underlying mechanism" which neatly explains these data is that Alice and Bob are a couple, it rains half of the time, and they choose to share a single umbrella. Table 1 describes the correlated choice rule for Alice and Bob.

	U	NU	
U	0	0.25	0.25
NU	0.25	0.5	0.75
	0.25	0.75	1

Table 1. Correlated choice rule which describes a couple's choice to carry an umbrella. U denotes the choice to carry an umbrella while NU denotes the choice to not carry an umbrella.

As we are interested in detecting influence, we are, by extension, interested in understanding whether variation of a common underlying state of the world could induce a given correlated choice rule. To better understand this question, once again consider Alice and Bob from the above example. Suppose that Bob has lost his umbrella and now we observe that Alice chooses to carry an umbrella half the time. If Alice's choices were independent of Bob's, the *marginal probability of her choice* would not depend on Bob's available choices. We call this property *marginality.* Marginality is a specific form of absence of influence—it allows us to construct a "marginal" stochastic choice function for each individual, independently of the budgets of remaining agents.

All the same, in this example, even the correlation between Alice's and Bob's choices is not enough confirm the presence of influence. Rather, the change in Bob's budget is what allowed the inference. The reason correlation alone does not tell us anything is that this correlation may come from external sources. This is a general principle: observation from a single budget is not enough to detect the presence of influence.

Perhaps Alice and Bob get their weather news from conflicting sources. Each of these sources will never falsely predict rain, but each of these sources will only predict rain half the time it actually does rain. We would not consider the agents' behavior to influence each other in this case. Instead, we would think it to be influenced by some external factor.

Our key finding is that correlated choice data can refute CRUM and, by extension, detect influence. We develop three finite tests which are able to detect to influence; marginality, *non-negativity* of an analogue of the Block-Marschak polynomials, and a condition we term *full capacity*. Our first main result shows that if any of these three conditions fails to hold, then we are able to refute CRUM. Our second main result shows that for "small" environments (those in which one of the agents has a global choice set containing at most three alternatives), marginality and non-negativity are both necessary and sufficient for a correlated choice rule to have a CRUM representation.

Our final contribution is that there are examples of correlated choice rules for pairs of agents which, if each agent is treated in isolation, appear to conform to the random utility model.[5] However, when the correlation structure is taken into account, we find that there can actually be no underlying state of the world driving preferences. Recall that non-negativity of an analogue of the Block-Marschak polynomials must be satisfied. Our example demonstrates that more conditions are necessary: it satisfies both our marginality condition and non-negativity of the Block-Marshack polynomials. We interpret the preceding phenomenon as a claim that influence can be "invisible" as a behavioral phenomenon without access to rich data—specifically correlated choice data, thus confirming Manski's claim. We conjecture that full capacity, in addition to marginality and non-negativity, is both necessary and sufficient for a correlated choice rule to be a CRUM.

There remains the question of what a full behavioral characterization of CRUM looks like. We actually offer one, which is based on standard techniques and an analogue of

[5] "Treated in isolation" means we observe the marginal probabilities of each agent's choices, while ignoring the underlying correlation structure. For this to be meaningful, marginality needs to be satisfied in the first place.

the *axiom of revealed stochastic preference* (ARSP) (see McFadden and Richter (1990) and McFadden (2005)) adapted to our environment. This axiom as stated can be viewed as an infinite collection of statements in propositional logic, and so differs from our previous tests. Each of those tests consists of a finite and bounded list of statements in propositional logic. Thus, verification that those tests are passed is straightforward, in contrast to ARSP.[6] We actually know simply by the form of the CRUM model that there are methods of characterizing the model which postulate only tests relying on finite lists of statements from propositional logic, finding these tests should be a goal of future research.

A refutation of CRUM might occur because of either "irrationality" at the individual level, or "influence" of the two agents in some form. We do not commit to specific interpretations of these two terms: we group them all under the rubric of *influence* for this paper. Clearly two agents, each expressing deterministic choices independently of each other, and exhibiting cyclic behavior would probably be more reasonably called irrational. It is not our purpose here to take a specific stance on this distinction, but instead describe how correlated data can lead to new insights.

The paper is organized as follows. Section 2, describes our model, tests, and main characterization theorem for small environments. Section 3 demonstrates that there are additional requirements implied by the model in larger environments. We offer a direction for future research. Finally, Section 4 concludes. Proofs are in an appendix.

2. The Model

Given is a (pair) of nonempty, finite sets of *alternatives* X and Y. Let $\mathcal{X}$ and $\mathcal{Y}$ be the sets of all non-empty subsets of X and Y, respectively.

In this paper, we consider a novel data set. In this world, (A, B) represents a decision problem where $A \in \mathcal{X}$ and $B \in \mathcal{Y}$. We assume that the outside observer has a technology for observing the probability of (a, b) being chosen from (A, B), where $a \in A$

[6]While linear programming algorithms can determine whetherARSP is satisfied, this is more evident from the structural form of the model than from ARSP itself.

and $b \in B$. For each $(A, B) \in \mathcal{X} \times \mathcal{Y}$, $p(A, B)$ is a probability measure.[7] Formally, we define a correlated choice rule now.

Definition. A *correlated choice rule* is a map $p : \mathcal{X} \times \mathcal{Y} \to \bigcup_{(A,B) \in \mathcal{X} \times \mathcal{Y}} \Delta(A \times B)$ for which for each $(A, B) \in \mathcal{X} \times \mathcal{Y}$, we have $p(A, B) \in \Delta(A \times B)$.

We now offer a few interpretations of this data set, though there are many more. Our focus is on the first type of data mentioned here. The key is that data consist of observations of choices of two entities, whether these be cross-sectional data or time series data.

We interpret X as the global choice set from which agent 1 may choose, and Y as the set from which agent 2 may choose. Correlated choice data represents the joint probability of agent 1's choices and agent 2's choices from two sets, respectively.

A reasonable example here might be two peers, one of whom is older. In line with our main hypothesis, we want to whether the choice of one agent directly "influences" the choice of the other. But there are other obvious examples. Gardening choices of two neighbors, food choices of wife and husband, and sport activity choices of two siblings are some examples of correlated choice data of two distinct agents.

There is at least one other, perhaps useful, interpretation of the model. With this interpretation, stochastic choice data are the outcome of repeated choices by the *same agent* (intrapersonal). This single agent has preferences on $X \times Y$, which might depend on unobservable random factors. We want to test whether the preferences over bundles in X and Y are *independent*, in the sense that consumption of a member of X does not influence the preferences over Y.[8]

Since choices are repeated, choices appear stochastic, where the randomness is induced by the unobserved external factor.

[7]It should be understood that the notation $p(\cdot|A, B)$ does not refer to conditional probability.

[8]Formally, we want to test whether preferences over $X \times Y$ can be written with a utility representation of the form: $U(a, b) = H(f(a), g(b))$, where H is increasing in each coordinate.

Under this interpretation, $p(a, b|A, B)$ represents the probability of the agent choosing a and b from A and B, respectively; when offered a budget of the form $A \times B$. Conventional data identify which specific product each customer chose and its price but do not describe the other products and their prices at the time of purchase. On the other hand, scanner data of purchase histories allow researchers to match the choice and price of multiple products. The repeated purchasing decisions of a single agent can be recovered by matching purchases made with the same credit card or debit card. Digital platforms offer another avenue to recover the types of data that we are considering. For many digital platforms, consumers are either required or suggested to make an account in order to make purchases. Tracking repeated purchasing decisions of a single agent then amounts to tracking the purchasing decisions of a single account. Correlated choice data could be i) two distinct products choices of the single agent overtime, or ii) the same product purchased on two different specific times (for example, purchases on weekday and weekend); for more on stochastic choice over time, see (Frick et al., 2019), (Lu and Saito, 2018), and (Duraj, 2018)).[9]

Alternatively, we may think of this single individual instead randomly choosing contingent choices or contracts (random choice of acts, see *e.g.* Lu (forthcoming)). X and Y now represent the available actions in two states of the world, state 1 and state 2, respectively. The fact that the individual chooses from $X \times Y$ means that there are no interstate constraints. The hypothesis we test in this environment is a very weak form of monotonicity across states, allowing for state depenendence. That is, it is weaker than either Savage's P3 (Savage, 1972) or the notion of monotonicity of (Anscombe et al., 1963).

Our data allow us to study the correlation structure of stochastic choice across a pair of agents.[10] Such a framework allows a richer language for discussing stochastic choice, and also allows for more restrictive testing. Table 2 illustrates our data for two decision problems. For each decision problem, we also provide marginal choice distributions.

[9]Dynamic stochastic choice has also been studied in the context of Luce's model or logit (Luce, 2005), see in particular Rust (1987), Fudenberg and Strzalecki (2015), or Pennesi (forthcoming).
[10]The theoretical framework can easily be generalized to accommodate any finite number of agents.

Here, the marginal distribution of $\{a_1, a_2, a_3\}$ is the same across two choice problems: $(\{a_1, a_2, a_3\}, \{b_1, b_2, b_3\})$ and $(\{a_1, a_2, a_3\}, \{b_1, b_2\})$. This is not true in general.

	b_1	b_2	b_3	
a_1	0.2	0	0.3	0.5
a_2	0.1	0.3	0	0.4
a_3	0	0	0.1	0.1
	0.3	0.3	0.4	1

	b_1	b_2	
a_1	0.3	0.2	0.5
a_2	0.1	0.3	0.4
a_3	0	0.1	0.1
	0.4	0.6	1

TABLE 2. Correlated choice rule for $(\{a_1, a_2, a_3\}, \{b_1, b_2, b_3\})$ and $(\{a_1, a_2, a_3\}, \{b_1, b_2\})$.

We first define the set of linear orders on any arbitrary set A by $\mathcal{L}(A)$.[11] For any set A and linear order $\succ$ over $\varnothing \neq B \subseteq A$, denote by $M(B, \succ)$ the unique maximal element of B according to $\succ$. A correlated choice rule has a random utility representation (RUM) if there is a probability distribution $\pi \in \Delta(\mathcal{L}(X \times Y))$ such that for any $(A, B) \in \mathcal{X} \times \mathcal{Y}$, and for every $(a, b) \in A \times B$, we have $p(a, b|A, B) = \pi\{\succ \in \mathcal{L}(X \times Y) : (a, b) = M(A \times B, \succ)\}$. This notion of RUM is the classical one. The only distinction is the form of the "budgets" on which the choice function operates: they must always take the form $A \times B$. Here, preferences can depend nontrivially on *pairs* of alternatives.

In terms of the story of peer influence, obviously we want to hypothesize that each set, X and Y belongs to a *different* agent, and to test whether they have individual preferences that act independently of each other.[12]

Definition. A correlated choice rule p has a correlated random utility representation (CRUM) if there is a probability distribution $\pi \in \Delta(\mathcal{L}(X) \times \mathcal{L}(Y))$ such that for any $(A, B) \in \mathcal{X} \times \mathcal{Y}$, and for every $(a, b) \in A \times B$, we have

$$p(a, b|A, B) = \pi\{(\succ, \succ') \in \mathcal{L}(X) \times \mathcal{L}(Y) : (a, b) = (M(A, \succ), M(B, \succ'))\}.$$

[11] A linear order is a binary relation which is complete, transitive, and antisymmetric.

[12] Of course we cannot rule out the possibility that the individuals "aggregate" their preferences in a Paretian fashion. The fact that there are no constraints across the individuals does not allow us to distinguish between individuals acting independently and agents using Paretian aggregation. They are observationally equivalent.

Observe that the probability π should be understood as randomly determining "tastes," perhaps as a function of some parameter observable to the agents in question, but unobservable to the analyst.

Not all correlated choice rules are consistent with CRUM. We now provide some necessary conditions for the CRUM model. Given the realization of the unobserved characteristic, each individual simply maximizes their preference, without referring to the other individual's choice. This motivates the following condition.

Axiom 1 (Marginality)**.** *For all $A \in \mathcal{X}$, $a \in A$, and all $B, B' \in \mathcal{Y}$, $\sum_{b \in B} p(a, b|A, B) = \sum_{b' \in B'} p(a, b'|A, B')$, with a similar statement for $b \in B$.*

Observe that any correlated choice rule satisfying marginality defines two classical stochastic choice functions, one over X and the other over Y, which can be termed *marginal choice rules*.

For example, given a a correlated choice rule p satisfying marginality, we can define the marginal $p_X : \mathcal{X} \to \bigcup_{\varnothing \neq A \subseteq X} \Delta(A)$, where $p_X(A) \in \Delta(A)$, via:

$$p_X(x|A) = \sum_{b \in B} p(a, b|A, B).$$

By marginality, p_X is well-defined independently of B. We define p_Y similarly. Further, we need an analogue of the Block-Marshak polynomials from Block et al. (1959). Define, for each $a \in A \subseteq X$ and $b \in B \subseteq Y$,

$$q(a, b|A, B) = \sum_{A' : A \subseteq A'} \sum_{B' : B \subseteq B'} (-1)^{|A' \setminus A| + |B' \setminus B|} p(a, b|A, B). \tag{1}$$

This expression looks complicated, but it really is a straightforward generalization of the classical single-agent concept. These numbers are set up so that if a correlated choice rule is a CRUM, then $q(a,b|A,B)$ would be exactly the probability of realizing a pair $(\succ_1,\succ_2) \in \mathcal{L}(X) \times \mathcal{L}(Y)$ for which for all $b \notin A$, $b \succ_1 a$ and for all $c \in A$, $a \succ_1 c$; with a similar statement for $\succ_2$. This can be seen easily by the formula:

$$p(a,b|A,B) = \sum_{A':A \subseteq A'} \sum_{B':B \subseteq B'} q(a,b|A',B'). \tag{2}$$

and applying a standard Mobius inversion formula, see *e.g.* Rota (1964). So, while the formula behind this expression is new (though bears an obvious resemblance to the classical formula), the idea behind it is not novel.

Axiom 2 (Non-negativity)**.** *For each $a \in A \subseteq X$ and $b \in B \subseteq Y$, we have $q(a,b|A,B) \geq 0$.*

Non-negativity of the Block-Marshack polynomials has been discussed before in the single-agent case, and the intuition is much the same here. Roughly, one should think of equation (2) as providing an analogue of a "cumulative distribution function" (cdf) in probability theory. The inverse equation (1) derives the "density" from the cdf. As probabilities must be non-negative, so must be the Block-Marshack polynomials.[13]

For a concrete example that speaks to the necessity of non-negativity, we turn our attention to Table 3. Table 3 provides an example of a correlated choice rule that does not admit a CRUM representation with each marginal choice rule admitting a RUM representation.[14] To see this, note that choice from $\{a_1,a_2,a_3\} \times \{b_1,b_2,b_3\}$ tells us that order pairs that rank a_1 and b_1 as best are drawn with a probability of one-half. Now, choice from $\{a_1,a_2,a_3\} \times \{b_1,b_2\}$ tells us that order pairs that rank a_1 and b_1 as best

[13]The curious form of the polynomials follows from properties of finite order differences. For example, given a probability measure μ on $\mathbb{Z}^2$, and a cdf $F(n,m) = \mu\{(a,b) : a \leq n, b \leq m\}$, it is easy to see that $\mu(\{(n,m)\}) = F(n,m) - F(n,m-1) - F(n-1,m) + F(n-1,m-1)$. This expression claims that there are non-negative differences of order 2; for higher dimensions, analogous expressions for non-negative differences of any order must be provided. These are what the Block-Marshack polynomials require.

[14]A marginal choice rule over X has a random utility representation (RUM) if there is a probability distribution $\pi \in \Delta(\mathcal{L}(X))$ such that for any $A \in \mathcal{X}$, and for every $a \in A$, we have $p_X(a|A) = \pi\{\succ \in \mathcal{L}(X) : a = M(A,\succ)\}$.

are never drawn. Putting this observation in terms of Block-Marschak polynomials, it turns out that $q(a_1, b_1|\{a_1, a_2, a_3\}, \{b_1, b_2\}) = -0.5$, a direct violation of non-negativity. To rationalize each marginal choice rule, we consider the following two distributions over linear orders.

$$\pi_1(\succ) = \begin{cases} 0.5 & a_1 \succ a_2 \succ a_3 \\ 0.25 & a_2 \succ a_1 \succ a_3 \\ 0.25 & a_3 \succ a_2 \succ a_3 \\ 0 & \text{otherwise} \end{cases} \qquad \pi_2(\succ) = \begin{cases} 0.5 & b_1 \succ b_2 \succ b_3 \\ 0.25 & b_2 \succ b_1 \succ b_3 \\ 0.25 & b_3 \succ b_2 \succ b_3 \\ 0 & \text{otherwise} \end{cases}$$

The distribution π_1 is a RUM representation of the marginal choice rule over different a. Similarly, the distribution π_2 is a RUM representation of the marginal choice rule over different b.

	b_1	b_2	b_3	
a_1	0.5	0	0	0.5
a_2	0	0.25	0	0.25
a_3	0	0	0.25	0.25
	0.5	0.25	0.25	1

	b_1	b_2	
a_1	0	0.5	0.5
a_2	0.25	0	0.25
a_3	0.25	0	0.25
	0.5	0.5	1

TABLE 3. Correlated choice rule for $(\{a_1, a_2, a_3\}, \{b_1, b_2, b_3\})$ and $(\{a_1, a_2, a_3\}, \{b_1, b_2\})$. The correlated choice rule does not admit a CRUM representation. The marginal choice rules admit RUM representations.

Our final axiom requires a few preliminary definitions. For an order $\succ$ and alternative a, we write the strict upper contour set of a as follows.[15]

$$U_\succ(a) = \{b | b \succ a\}$$

Definition. We call the tuple (a, b, A, B) a *choice tuple* if $a \in A \subseteq X$ and $b \in B \subseteq Y$.

Definition. We say that the choice tuple (a, b, A, B) is *associated* with the order pair $(\succ_1, \succ_2)$ if $U_{\succ_1}(a) = X \setminus A$ and $U_{\succ_2}(b) = Y \setminus B$.

[15] The strict upper contour set of a a is the set of all elements b satisfying $b \succ a$.

For example, let a correlated choice rule be generated by a degenerate distribution over the order pair $(\succ_1, \succ_2)$. Then $q(a, b|A, B) = 1$ if (a, b, A, B) is associated with $(\succ_1, \succ_2)$ and $q(a, b|A, B) = 0$ otherwise.

Definition. We say that the a set Γ of choice tuples *spans* $\mathcal{L}(X) \times \mathcal{L}(Y)$ if for all $(\succ_1, \succ_2) \in \mathcal{L}(X) \times \mathcal{L}(Y)$, there exists some $(a, b, A, B) \in \Gamma$ associated with $(\succ_1, \succ_2)$.

In simple terms, Γ is spanning if it "takes into account" every order pair. Let Θ denote the collection of all sets Γ of choice tuples (a, b, A, B) which span $\mathcal{L}(X) \times \mathcal{L}(Y)$.

Definition. We define the *capacity* of a a correlated choice rule as

$$C(p) = \min_{\Gamma \in \Theta} \sum_{(a,b,A,B) \in \Gamma} q(a, b|A, B)$$

The capacity of a correlated choice rule is meant to capture the portion of the correlated choice rule that can be generated by choice according to order pairs. Intuitively, for a correlated choice rule to have a CRUM representation, it must be that the entire correlated choice rule can be generated by choice according to order pairs. This motivates the following condition.

Axiom 3 (Full Capacity)**.** *For a correlated choice rule p, $C(p) = 1$.*

Alternatively, full capacity can be stated in terms of a set of linear inequalities. Since the capacity of a system is never greater than one,[16] full capacity is equivalent to asking that for all $\Gamma \in \Theta$ we have the following.

$$\sum_{(a,b,A,B) \in \Gamma} q(a, b|A, B) \geq 1.$$

We are interested in developing a finite test for the existence of a CRUM representation. With this alternate formulation of full capacity, checking for each of our axioms amounts to checking that a finite set of linear inequalities hold. With these three axioms, we are ready to present our first result.

[16]To see this, note that the set $\Gamma = \{(a, b, X, Y)\}_{a \in X, b \in Y}$ spans $\mathcal{L}(X) \times \mathcal{L}(Y)$ and $\sum_{a \in X} \sum_{b \in Y} q(a, b|X, Y) = \sum_{a \in X} \sum_{b \in Y} p(a, b|X, Y) = 1$.

Theorem 1. *Suppose that a correlated choice rule has a CRUM representation. Then it satisfies marginality, non-negativity, and full capacity.*

We leave all proofs to the appendix, but we discuss the intuition behind our result here. Upon drawing a pair of linear orders, each agent is able to maximize their own utility. This behavior directly implies marginality. As for non-negativity, recall that if a correlated choice rule has a CRUM representation, then each $q(a,b|A,B)$ is equal to the probability of some event and thus must be non-negative. Finally, every spanning Γ takes in to account each linear order pair at least once. This means that when we find the capacity of a correlated choice rule with a CRUM representation, it must be greater than or equal to the probability of drawing some order pair. Further, we can always find a spanning Γ which guarantees that the capacity of a system is no more than one. This leaves us to conclude that full capacity must be satisfied.

We now discuss the recovery of a CRUM representation from a correlated choice rule. For small choice environments, marginality and non-negativity are sufficient for the existence of a CRUM representation.

Theorem 2. *Suppose that $|X| \leq 3$ or $|Y| \leq 3$. If a correlated choice rule satisfies non-negativity and marginality, then it has a CRUM representation.*

Our proof is constructive and provides an algorithm that takes in a correlated choice rule and returns a CRUM representation of the correlated choice rule if one exists. At an intuitive level, we use the small choice environment and non-negativity to uniquely pin down the marginal distribution of preferences of the agent with the small choice environment. From there, we use marginality and non-negativity to ensure that the behavior of the agent with the large choice environment is consistent with utility maximization. We then condition on the preferences of the agent with the small choice environment and treat the other agent as if they were in the single agent case.

Interpretation-wise, observe that our Theorem 1 and Theorem 2 establish a complete characterization of correlated choice rules which admit a CRUM representation in the case where the number of alternatives for one of the agents is small (at most three).

As it turns out, in larger choice environments, margianlity and non-negativity cease to be sufficient for the existence of a CRUM representation. We discuss these larger choice environments in the next section.

3. Larger Choice Environments

In this section, we discuss the power of our axioms in choice environments larger than the environments considered in Theorem 2. Non-negativity and marginality fail to be sufficient conditions for the existence of a CRUM representation when both agents' choice sets have four or more elements. We first provide a counterexample that shows this. After discussing our counterexample, we discuss the power of full capacity in large choice environments.

3.1. Counterexample in the 4 by 4 Case.

We begin with a formal statement of our observation.

Observation 1. *Suppose that $|X| \geq 4$ and $|Y| \geq 4$. There are correlated choice rules which satisfy marginality and non-negativity but do not have a CRUM representation.*

Example 1 (Counterexample in a Large Choice Environment)**.** *Let $X = Y = \{a, b, c, d\}$. We use the table below to describe the behavior of each agent.*

	1	*2*
3	$a \succ b \succ c \succ d$	$b \succ a \succ d \succ c$
4	$a \succ b \succ d \succ c$	$b \succ a \succ c \succ d$

With probability $\frac{1}{2}$, each agent chooses according to a linear order from column 1. With probability $\frac{1}{2}$, each agent chooses according to a linear order from column 2. If the second agent's choice set contains a or b, then the first agent chooses according to row 3. If the second agent's choice set does not contain a or b, then the first agent chooses according to row 4. The behavior of the second agent is symmetric to that of the first agent.

We provide a more thorough description of the correlated choice rule generated by the above example in the appendix. We first note that the behavior described in the above example leads to a well defined correlated choice rule that satisfies both marginality and non-negativity. In this example, there is correlation between draws of linear orders as both agents choose their linear orders from the same column. However, each agent's choice set influences the other agent's order draw. This can be seen from the fact that the inclusion of a and b in an agent's choice set directly determines from which row the other agent draws their order. As the correlated choice rule of this example satisfies marginality, this type of influence is not observable by marginal choice data. This shows that further restrictions are needed in order to test for the existence of a CRUM representation. While this example showcases an extreme form of influence, this type of influence can show up in a much weaker form in various other correlated choice rules.

Our counterexample is closely related to the example that Fishburn (1998) gives to show that the standard random utility model is unidentified. We review this example and discuss how our counterexample relates to the example of Fishburn (1998).

Example 2 (Fishburn's Counterexample)**.** *Let* $X = \{a, b, c, d\}$*. Consider the following probability distributions over linear orders on* X*.*

$$\pi_1(\succ) = \begin{cases} \frac{1}{2} & \text{if } \succ \in \{a \succ b \succ c \succ d, b \succ a \succ d \succ c\} \\ 0 & \text{otherwise} \end{cases}$$

$$\pi_2(\succ) = \begin{cases} \frac{1}{2} & \text{if } \succ \in \{a \succ b \succ d \succ c, b \succ a \succ c \succ d\} \\ 0 & \text{otherwise} \end{cases}$$

These two probability distributions induce the same system of choice probabilities.

To begin, note that the four linear orders considered in Fishburn's counterexample are the same four linear orders used in our counterexample. Further, the marginal choice probabilities of the correlated choice rule from our example are exactly equal to the choice probabilities from Fishburn's counterexample. Consider the following formulation of our counterexample. Let an agent's marginal choice probabilities be the same as those in Fishburn's counterexample. If the other agent's choice set contains a or b, then an agent will choose according to π_1. If not, that agent will choose according to π_2. This formulation shows us that whenever the single agent random utility model fails to have a unique representation, the type of influence shown in our counterexample may be present.

3.2. **Full Capacity.** Theorem 1 and Theorem 2 provide a characterization of CRUM for small choice environments. The prior counterexample tells us that this characterization fails to hold in larger choice environments. As we note in Theorem 1, full capacity is a necessary condition for rationalizability. In this section, we conjecture that full capacity, along with marginality and non-negativity, is sufficient for rationalizability.

Conjecture 3. *If a correlated choice rule satisfies non-negativity, marginality, and full capacity, then the correlated choice rule has a CRUM representation.*

We leave the proof of this conjecture for future research. However, we explain the intuition behind why we believe full capacity is sufficient for the existence of a CRUM representation. We also explain why we are unable to provide a proof of our conjecture.

Our problem is closely related to the max-flow min-cut theorem of graph theory (Ford and Fulkerson, 1956). This theorem starts by supposing that we have some directed graph. This graph has two special nodes, the source, the node from which all maximal paths start, and the sink, the node at which all maximal paths end. Each edge has associated with it a positive real number which is interpreted as the capacity constraint of that edge. A flow assignment is an assignment of real numbers to maximal paths such that, for each edge, summing over the flows assigned to paths that pass through that edge is no more than the capacity constraint of that edge. The max-flow min-cut theorem characterizes the maximum possible total flow for a given directed graph. As the name of the theorem suggests, the maximum flow for a given directed graph is equal to the minimum value of all cuts of that graph. A cut of a graph is a set of nodes of the graph containing the source and not containing the sink. The edges associated with a cut are the set of edges connecting a node in the cut and a node not in the cut. The value of a cut is equal to the sum of the capacity constraints of the edges associated with the cut.

Now we draw a parallel between the setup for the max-flow min-cut theorem and our full capacity condition. In the appendix, we present a graphical representation of our choice environment.[17] In our representation, every order pair $(\succ_1, \succ_2)$ can be thought of as a path along the graph. Each choice tuple (a, b, A, B) can be thought of as an edge with corresponding capacity constraint $q(a, b|A, B)$. The edges of a cut are analogous to the choice tuples of a spanning Γ. Finally, the minimum value of a cut is analogous to the capacity of a correlated choice rule. In asking for the capacity of a correlated choice rule to be one, our aim is to find a flow assignment that assigns a total flow of one to order pairs. If such an assignment exists, it then also defines a CRUM representation of the correlated choice rule. However, in our graphical representation, there are more paths than there are order pairs. This is where the difficulty of our problem lies. The max-flow min-cut theorem ensures that there is a flow assignment to the entire set of paths that achieves a total flow equal to the minimum cut. In our case, we are considering a restricted set of paths on our graph.

To the best of our knowledge, this type of restricted max-flow min-cut result is an open or yet unasked question in the graph theory literature. We also note that the standard techniques used to prove the max-flow min-cut theorem no longer work when applied to the restricted case.

3.3. **The Axiom of Revealed Stochastic Preference.** We turn our attention to the Axiom of Revealed Stochastic Preference (ARSP) of McFadden and Richter (1990). We restate ARSP here.

[17]Here we consider the marginal graph representation.

Definition. A stochastic choice rule p satisfies *ARSP* if for every finite sequence $\{(a_i, A_i)\}_{i=1}^n$ where, for each i, $a_i \in A_i$, we have the following.

$$\sum_{i=1}^{n} p(a_i|A_i) \leq \max_{\succ \in \mathcal{L}(X)} \sum_{i=1}^{n} \mathbf{1}\{a_i = M(A_i, \succ)\}$$

Note that in order to verify ARSP, one must *in principle* check an infinite number of sequences; in contrast to testing non-negativity of the Block-Marschak polynomials. However, so long as a researcher is able to find a single sequence which fails to satisfy the inequality described in ARSP, then ARSP fails to hold. McFadden (2005) makes the observation that ARSP can be used to characterize the stochastic generalization of any set of finite choice rules. We record the straightforward observation that this extends to multi-agent and multi-dimensional choice rules. Consider the following restatement of ARSP in terms of correlated choice rules.

Definition. A correlated choice rule p satisfies ARSP if for every finite sequence $\{((a_i, b_i), A_i \times B_i)\}_{i=1}^n$ where, for each i, $A_i \times B_i \in \mathcal{X} \times \mathcal{Y}$ and $(a_i, b_i) \in A_i \times B_i$, we have the following.

$$\Sigma_{i=1}^n P_{A_i \times B_i}(a_i, b_i) \leq \max_{(\succ, \succ') \in \mathcal{L}(X) \times \mathcal{L}(Y)} \Sigma_{i=1}^n \mathbf{1}\{(a_i, b_i) = (M(A_i, \succ), M(B_i, \succ'))\}$$

As CRUM is the stochastic generalization of choice according to linear order pairs, CRUM is also characterized by the above formulation of ARSP.

Theorem 4. *A correlated choice rule p has a CRUM representation if and only p satisfies ARSP.*

The proof of Theorem 4 is not particularly novel. In fact, we base our proof heavily off of the proof of the equivalence of ARSP to stochastic rationality in Border (2007); it is mostly a matter of simply replacing the notation.

Up until now, we have assumed that we observe choice on every element of $\mathcal{X} \times \mathcal{Y}$. One of the main benefits of using ARSP is that it characterizes CRUM even on smaller choice domains.

4. Conclusion

In this paper, we develop three finite tests which are able to detect the presence of influence without the need for parametric restrictions. Our key findings are twofold. First, we show that correlated choice rules are better suited for detecting influence than standard stochastic choice data. Further, we find that menu variation is necessary in detecting influence.

One obvious shortcoming remains in this work: we have not been able to provide a full characterization of jointly rational correlated choice rules in terms of a finite system of linear inequalities analogous to Falmagne (1978). Indeed, we only provide a finite list of linear inequalities (marginality, non-negativity, and full capacity) which are *necessary* for the representation to hold. That said, the duality theory teaches us that there must necessarily exist a finite list of such linear inequalities, and so understanding what these are remains an open question.

While we discuss another dual characterization of jointly rational correlated choice rules (ARSP), the characterization is less appealing in two important ways. First off, ARSP does not provide intuitive insight into the choice behavior of jointly rational agents. Unlike marginality, which tells us that agents can maximize their preferences separately, ARSP has no intuitive interpretation in terms of choice behavior. In addition, ARSP has nothing to do with the theory of correlated choice. It would apply to the stochastic extension of any choice theory. The difficulty is in finding a *finite* list of linear inequalities which are characteristic of the model. This finite list (coming from the extreme rays of a dual polyhedron) depend on the particular model under consideration.

Appendix A. Mathematical Construction

A.1. **Choice Multigraph.** The proof of Theorem 2 is involved, and relies on the notion of a particular directed weighted multigraph which is derived from the given correlated choice rule. More formally, a multigraph is an object consisting of a collection of *nodes* and, between any pair of nodes, a potentially (empty) set of *directed edges*. Each edge is associated with a number, termed its *weight*.

First, an *order pair* is a pair of linear orders, or a member of $\mathcal{L}(X) \times \mathcal{L}(Y)$.

We construct an object we call a *choice multigraph*. In a formal sense, this will be a (directed) multigraph whose nodes will consist of pairs of subsets of X and Y, respectively.

What makes it a multigraph rather than a graph is that between any pair of nodes, there will in general be multiple edges.

Further, each edge in the multigraph will carry a *weight*, which is a real number. The real numbers, in an intuitive sense, will represent a kind of "flow." The meaning of this should become clearer as we proceed.

To begin, let $\mathcal{X}$ be the collection of nonempty subsets of X. The set of *nodes* for the choice multigraph of (X, Y) is a collection of nodes indexed by the collection of sets $(\mathcal{X}) \times (\mathcal{Y})$.

We now proceed to define the edges of the multigraph. Without loss, we will often refer to the nodes of the graph via their indices. Let $A \times B$ and $A' \times B'$ be two nodes.

Then, there exists a directed edge from the node $A \times B$ to the node $A' \times B'$ if one of the following is true:

(1) $A = A'$, $B' \subseteq B$ and $|B \setminus B'| = 1$
(2) $B = B'$, $A' \subseteq A$ and $|A \setminus A'| = 1$

Assuming that $X \cap Y = \varnothing$, then edges from the multigraph arise precisely from the covering relation of $\subseteq$ applied to $A \cup B$ and $A' \cup B'$.

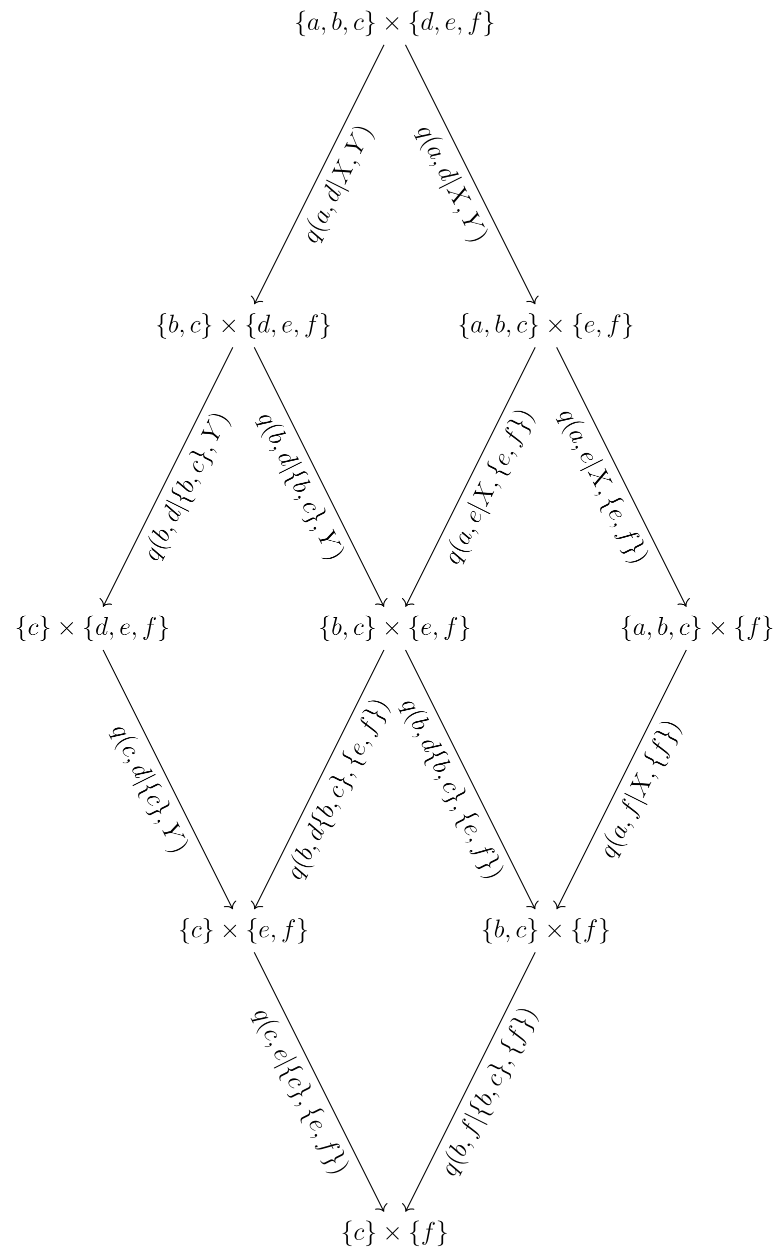


FIGURE 1. The face corresponding to the order pair $(a \succ_1 b \succ_1 c, d \succ_2 e \succ_2 f)$ in the choice multigraph corresponding to the collection of sets $\{\{a, b, c\}, \{d, e, f\}\}$.

Further, suppose that the pair of nodes under consideration takes the form (A, B) and (A, B'), where $B' \subseteq B$. Let $\{b\} = B \backslash B'$. Then for each such pair of nodes, we associate

$|A|$ edges, one for each member $a \in A$. For all $a \in A$, associate $q(a, b|A, B) \in \mathbb{R}$ with the a edge connecting the nodes associated with $A \times B$ and $A \times B'$.

We are interested in order pairs and their corresponding representation in the choice multigraph previously described. Each order pair will be associated with a submultigraph of the choice multigraph, which we will term a *face*.[18]

Formally, we define $F(\succ_1, \succ_2)$ to be the submultigraph of the choice multigraph defined as follows. First, it contains the node that corresponds to $X \times Y$. For the face corresponding to the linear order tuple $(\succ_1, \succ_2)$, a node indexed by $A \times B$ is on the face when there is $a \in X$ for which $A = \{b \in X : a \succ_1 b\}$ and there is $b \in Y$ for which $B = \{a \in Y : b \succ_2 a\}$. Any edge directed away from node (A, B) is given weight $q(a, b|A, B)$, where a is $\succ_1$ maximal for A and b is $\succ_2$ maximal for B.

See Figure 1, for an example of a three alternative environment and $F(\succ_1, \succ_2)$ where $a \succ_1 b \succ_1 c$ and $d \succ_2 e \succ_2 f$.

A.2. **Marginal Graph System.** In this section, we describe a graphical representation of a correlated choice rule, which we call the *marginal graph system*. For the pair of sets (X, Y), there are two marginal graph systems, one for each of the sets. For the pair of sets (X, Y), the marginal graph system for X is a "nested" collection of directed weighted graphs. As a first point, we construct a graph taking as nodes all subsets of X, and an edge from any set $A \subseteq X$ to any subset $B \subseteq A$ with $|B| = |A| - 1$; that is, edges coincide with the covering relation associated with $\subseteq$.

Since we have assumed marginality, we may define $p_X(a, A) = \sum_{b \in B} p(a, b|A, B)$ independently of $B \subseteq Y$. Then p_X can be interpreted as a classical single-agent stochastic choice function. Associated with p_X are its Block-Marshack polynomials, which we define via, for each $a \in A \subseteq X$:

$$q_X(a, A) = \sum_{A' : A \subseteq A'} (-1)^{|A' \setminus A|} p_X(a, A).$$

[18]Formally, by a submultigraph, we mean a multigraph whose nodes and edges are subsets of the original multigraph.

Assign $q_X(a, A)$ to the edge connecting the nodes corresponding to A and $A \setminus \{a\}$. This assignment both indexes the edges as well as records the "flow" along each edge. We call this first component of the marginal graph system the *marginal component* of the marginal graph system. Thus far, the setup is almost identical to the single agent case considered in Fiorini (2004).

For the second portion of the marginal graph system, for each edge with weight $q_X(a, A)$ of the marginal component, associate another directed, weighted graph. Each such graph is similar in structure to the marginal component, except it deals with the set Y. So, each such graph will have as nodes the subsets of Y, and an edge from A to B if $B \subseteq A$ and $|B| = |A| - 1$.

For the directed graph associated with the edge $q_X(a, A)$, assign $q(a, b|A, B)$ to the edge connecting B and $B \setminus \{b\}$. Once again, this assignment both indexes the edges as well as records the flow along each edge. We call these second types of graphs of the marginal graph system a *conditional component* of the marginal graph system. Observe that a conditional component depends on its associated edge in the marginal component. The terminology should cause no confusion.

When discussing full capacity, we consider an alternate and equivalent formulation of the marginal graph system. The above construction of the marginal graph system is equivalent to taking the marginal component of the marginal graph system and replacing each edge with the conditional component associated with that edge. When replacing an edge that connects node A with node $A \setminus \{a\}$ with its associated conditional component, we replace the node Y of the conditional component with the node A from the marginal component and the node $\varnothing$ from the conditional component with the node $A \setminus \{a\}$ from the marginal component. We call this formulation the *marginal graph*.

For future reference, we also note that we can define the marginal graph system corresponding to set Y in an analogous way. Below, we do so in some of our constructions.

Appendix B. Proofs

B.1. **Proof of Theorem 1.**

Proof. The correlated choice rule is rationalizable. This means that there exists some $\pi \in \Delta(\mathcal{L}(X) \times \mathcal{L}(Y))$ that induces the correlated choice rule.

Fix $A \times B \subseteq X \times Y$ with $(a, b) \in A \times B$. Let

$$M_{(a,b),A\times B} = \{(\succ_1, \succ_2) : \forall (x, x', y, y') \in A^c \times A \times B^c \times B, x \succ_1 a \succ_1 x' \text{ and } y \succ_2 b \succ_2 y'\}.$$

Then $M_{(a,b),A\times B}$ consists of the set of order pairs $(\succ_1, \succ_2)$ for which:

(1) Every member of A is $\succ_1$ ranked below every member outside of A, where x is at the top of A

(2) Every member of B is $\succ_2$ ranked below every member outside of B, where y is at the top of B.

The proof of the following claim is exactly analogous to the classical proofs in the single agent case, see *e.g.* Falmagne (1978).

Claim 1. *The correlated choice rule p is rationalized by $\pi \in \Delta(\mathcal{L}(X) \times \mathcal{L}(Y))$ if and only if $q(a, b|A, B) = \pi(M_{(a,b),A\times B})$ for all (a, b, A, B).*

Proof. For all $A \subseteq X$, all $B \subseteq Y$, all $a \in A$, and all $b \in B$, π rationalizes the correlated choice rule if and only if $p(a, b|A, B)$ is the π-probability of realizing a pair $(\succ_1, \succ_2)$ for which $A \subseteq \{z \in X : a \succ_1 z\}$ and $B \subseteq \{w \in Y : b \succ_2 w\}$. Namely, it is the π-probability of $\bigcup_{A\subseteq A'} \bigcup_{B\subseteq B'} M_{(a,b),A'\times B'}$. For a fixed (a, b), the sets $M_{(a,b),A'\times B'}$ are disjoint as $A \subseteq A'$ and $B \subseteq B'$, conclude:

$$p(a, b, A, B) = \sum_{A':A\subseteq A'} \sum_{B':B\subseteq B'} \pi(M_{(a,b),A'\times B'}).$$

The result now follows from the Mobius inversion formula. □

Non-negativity now follows as π is a probability, and hence non-negative.

The proof of marginality is obvious: $\sum_{b\in B} p(a,b,A,B) = \pi(\{(\succ_1,\succ_2) : x = M(A,\succ_1)\})$, which is independent of B.

To begin, consider the set of linear order pairs with which the choice tuple (a,b,A,B) is associated. This set is exactly $M_{(a,b),A\times B}$. We know that $q(a,b|A,B) = \pi(M_{(a,b),A\times B})$. Now we can decompose $M_{(a,b),A\times B}$ into its elements. This gives us the following.

$$q(a,b|A,B) = \pi(M_{(a,b),A\times B}) = \sum_{(\succ_1,\succ_2)\in M_{(a,b),A\times B}} \pi((\succ_1,\succ_2))$$

Consider some $\Gamma \in \Theta$.

$$\begin{aligned}\sum_{(a,b,A,B)\in\Gamma} q(a,b|A,B) &= \sum_{(a,b,A,B)\in\Gamma}\ \sum_{(\succ_1,\succ_2)\in M_{(a,b),A\times B}} \pi((\succ_1,\succ_2)) \\ &\geq \sum_{(\succ_1,\succ_2)\in\mathcal{L}(X)\times\mathcal{L}(Y)} \pi((\succ_1,\succ_2)) \\ &= 1\end{aligned}$$

The first line follows from the logic explained prior. The second line follows from the fact that Γ spans $\mathcal{L}(X)\times\mathcal{L}(Y)$ and thus considers each order pair at least once. The last line follows from the fact that π is a probability distribution. By this logic, whenever a correlated choice rule has a CRUM representation, each spanning Γ must have value of at least one. This means that the capacity of the correlated choice rule can be no less than one.

Now recall that every finite linear order has a maximal element. This means that the set $\{(a,b,X,Y)\}_{a\in X, b\in Y}$ spans $\mathcal{L}(X)\times\mathcal{L}(Y)$. From the definition of the BM polynomials we know the following.

$$\sum_{a\in X}\sum_{b\in Y} q(a,b|X,Y) = 1$$

This means that for each correlated choice rule, there is some spanning Γ which has a value of one. This, combined with the prior logic, gives us the following.

$$1 \leq C \leq 1$$

This gives us that $C = 1$. Thus rationalizability implies full capacity. This gives us that rationalizability implies marginality, non-negativity, and full capacity. □

B.2. **Proof of Theorem 2.**

B.2.1. *Preliminary Lemma.* Prior to moving to our proof of Theorem 2, we have one major lemma to state and prove.

Definition. We say that a correlated choice rule satisfies *recursivity* if for every $A \neq X$, every $B \subseteq Y$ and every $b \in B$, the following is satisfied (with a similar statement for sums across B).

$$\sum_{a \in A} q(a, b|A, B) = \sum_{c \in X \setminus A} q(c, b|A \cup \{c\}, B)$$

Lemma 1. *The correlated choice rule p satisfies marginality if and only if the corresponding Block-Marschak polynomials satisfy recursivity.*

Proof. First we show that marginality implies recursivity. Fix A, B and $b \in B$. Let us write the equations:

$$\sum_{a \in A} q(a, b|A, B) = \sum_{a \in A} \sum_{A \subseteq A'} \sum_{B \subseteq B'} (-1)^{|A' \setminus A| + |B' \setminus B|} p(a, b|A', B'). \tag{3}$$

Likewise,

$$\sum_{c \in X \setminus A} q(c, b|A \cup \{c\}, B) = \sum_{c \in X \setminus A} \sum_{A \cup \{c\} \subseteq A'} \sum_{B \subseteq B'} (-1)^{|A' \setminus (A \cup \{c\})| + |B' \setminus B|} p(c, b|A' \cup \{c\}, B'). \tag{4}$$

Let us subtract Equation (4) from equation (3). We will do a simple counting argument. In particular, for every set A', we will count the number of times it appears in the difference of the two equations. In equation (4), no term of the type $p(a, b|A, B')$ ever appears. Consequently, the difference of equation (3) and equation (4) has a term of $\sum_{a \in A} \sum_{B \subseteq B'} (-1)^{|B' \setminus B|} p(a, b|A, B')$.

Now consider any set A' for which $A \subseteq A'$ and $|A| < |A'|$. The total coefficient coming from equation (3) is obviously

$$\sum_{a \in A} \sum_{B \subseteq B'} (-1)^{|A' \setminus A| + |B' \setminus B|} p(a, b|A', B').$$

Likewise, the total coefficient coming from the negation of equation (4) is $(-1) \sum_{c \in A' \setminus A} \sum_{B \subseteq B'} (-1)^{|A' \setminus A| - 1 + |B' \setminus B|} p(c, b|A', B')$, or

$$\sum_{c \in A' \setminus A} \sum_{B \subseteq B'} (-1)^{|A' \setminus A| + |B' \setminus B|} p(c, b|A', B').$$

Overall, then, the difference of equation (3) and equation (4) is

$$\sum_{A \subseteq A'} \sum_{B \subseteq B'} (-1)^{|A' \setminus A| + |B' \setminus B|} \sum_{a \in A'} p(a, b|A', B').$$

Reverse the order of the sums according to A and B, and obtain:

$$\sum_{B \subseteq B'} \sum_{A \subseteq A'} (-1)^{|A' \setminus A| + |B' \setminus B|} \sum_{a \in A'} p(a, b|A', B').$$

Now, by marginality, $\sum_{a \in A'} p(a, b|A', B')$ is independent of A'. Since $b \in B$ is fixed, we can call this term $\iota(B')$. Therefore the expression becomes:

$$\sum_{B \subseteq B'} \sum_{A \subseteq A'} (-1)^{|A' \setminus A| + |B' \setminus B|} \iota(B').$$

One more rearrangement:

$$\sum_{B \subseteq B'} (-1)^{|B' \setminus B|} \iota(B') \sum_{A \subseteq A'} (-1)^{|A' \setminus A|}.$$

Obviously, though, since $A \neq X$, we know that $\sum_{A \subseteq A'} (-1)^{|A' \setminus A|} = 0$. This follows, as in the notation of Rota (1964), what we have is $\sum_{A \subseteq A'} \mu(A, A') \zeta(A', X)$ (here μ is the Mobius function of set inclusion), so that the entire expression is $\delta(A, X)$, where again this is the "Kronecker delta" referred to in Rota (1964) as being the identity element of the incidence algebra. This identity element is 0 if $A \neq X$ (otherwise is 1). So we are done.

Now we show that recursivity implies marginality. We begin with a claim.

Claim 2. *Suppose q satisfies recursivity. Then for each $n \geq 0$, we get*

$$\sum_{a\in A}\sum_{A':A\subseteq A':|A'|\leq|A|+n} q(a,b|A',B) = \sum_{A':A\subseteq A':|A'|=|A|+n}\sum_{a\in A'} q(a,b|A',B).$$

Further,

$$\sum_{a\in A}\sum_{A':A\subseteq A'} q(a,b|A',B) = \sum_{c\in X} q(c,b|X,B).$$

Proof of Claim 2. The proof is by induction on n. For $n = 0$ it is trivial. Suppose it is true for n and we will show for $n+1$. This gives us

$$\sum_{a\in A}\sum_{A':A\subseteq A':|A'|\leq|A|+n+1} q(a,b|A',B) = \sum_{a\in A}\sum_{A':A\subseteq A':|A'|\leq|A|+n} q(a,b|A',B) + \sum_{a\in A}\sum_{A':A\subseteq A':|A'|=|A|+n+1} q(a,b|A',B).$$

By the induction hypothesis, this gives

$$\sum_{a\in A}\sum_{A':A\subseteq A':|A'|\leq|A|+n+1} q(a,b|A',B) = \sum_{A':A\subseteq A':|A'|=|A|+n}\sum_{a\in A'} q(a,b|A',B) + \sum_{a\in A}\sum_{A':A\subseteq A':|A'|=|A|+n+1} q(a,b|A',B)$$

By applying recursivity to the first part of this sum, we know that

$$\sum_{a\in A'} q(a,b|A',B) = \sum_{c\in X\setminus A'} q(c,b|A'\cup\{c\},B).$$

So

$$\sum_{A':A\subseteq A':|A'|=|A|+n}\sum_{a\in A'} q(a,b|A',B) = \sum_{A':A\subseteq A':|A'|=|A|+n}\sum_{c\in X\setminus A'} q(c,b|A'\cup\{c\},B).$$

Now let us do a simple combinatorics argument.

Let us now consider a set A^* of cardinality $|A|+n+1$, which contains A. This set appears in the form of $A'\cup\{c\}$ in the above summation exactly $n+1$ times, one for each members $c \in A^* \setminus A$. And each time, it adds a value of $q(c,b|A^*,B)$. So, overall,

$$\sum_{A':A\subseteq A':|A'|=|A|+n}\ \sum_{c\in X\setminus A'} q(c,b|A'\cup\{c\},B) = \sum_{A':A\subseteq A':|A'|=|A|+n+1}\ \sum_{a\in A'\setminus A} q(a,b|A',B).$$

So the expression:

$$\sum_{A':A\subseteq A':|A'|=|A|+n}\ \sum_{a\in A'} q(a,b|A',B) + \sum_{a\in A}\ \sum_{A':A\subseteq A':|A'|=|A|+n+1} q(a,b|A',B)$$

gives exactly

$$\sum_{A':A\subseteq A':|A'|=|A|+n+1} q(a,b|A',B),$$

which is what we wanted to prove in regards to the first part of the lemma. The second part of the lemma follows from setting $n = |X| - |A|$. □

Now observe that $\sum_{a\in A} p(a,b|A,B) = \sum_{B\subseteq B'} \sum_{a\in A} \sum_{A\subseteq A'} q(a,b|A,B)$, which by the above claim is the same as $\sum_{B\subseteq B'} \sum_{a\in X} q(a,b|X,B)$, which is independent of A. So we are done.

□

B.2.2. *Main Theorem.*

Proof. To begin, let suppose that marginality and non-negativity hold for the correlated choice rule p. We describe an algorithm that takes in the choice multigraph (see Appendix A) formed by p and returns a CRUM representation of p. The algorithm is as follows.

(1) Initialize $i = 0$ (a counter variable) and associate with each order pair a $P_{(\succ^1,\succ^2)} = 0$. Define $q_0(a,b|A,B) = q(a,b|A,B)$.
(2) Consider the set of faces with strictly positive $q_i(\cdot)$ on each edge. Call this set F_i. If $F_i = \varnothing$, terminate the algorithm. Otherwise, consider the set of edges, e_i, associated with each face in F_i. Call this set E_i. Consider the set of $(q_i(\cdot), e_i)$ that are associated with edges in E_i. Call this set Q_i. Choose a $(q_i(\cdot), e_i)$ such that $q_i(\cdot)$ is minimal among $q \in Q_i$. Call this value q_i^*. By definition of F_i,

$q_i^* > 0$. Choose some face in F_i such that (q_i^*, e_i) is associated with an edge of that face. Call this face f_i^*. Consider the set of $(q_i(\cdot), e_i)$ associated with f_i^*. Call this set Q_i^*. For all $(q_i(\cdot), e_i) \in Q_i^*$, let $(q_{i+1}(\cdot), e_{i+1}) = (q_i(\cdot) - q_i^*, e_i)$. For all $(q_i(\cdot), e_i) \notin Q_i^*$, let $(q_{i+1}(\cdot), e_{i+1}) = (q_i(\cdot), e_i)$. Let $\succeq_i$ denote the order pair that is associated with f_i^*. Let $P_{(\succeq_i^1, \succeq_i^2)} = q_i^*$.

(3) Let $i = i + 1$ and return to step 2.

The algorithm can terminate in two cases. The first case is that there is zero flow along the choice multigraph at termination and the second case is that there is positive flow along some edge in the choice multigraph. Suppose that we are in the first case. In this case, let $\pi((\succ_1, \succ_2)) = P_{(\succ^1, \succ^2)}$. Recall that Claim 2 states that the correlated choice rule p is rationalized by $\pi \in \Delta(\mathcal{L}(X) \times \mathcal{L}(Y))$ if and only if $q(a, b|A, B) = \pi(M_{(a,b),A\times B})$ for all (a, b, A, B). In the case that the algorithm terminates with zero flow along the choice multigraph, exactly $q(a, b|A, B)$ amount of probability has been assigned to order pairs that rank a exactly at the top of A and b exactly at the top of B. This means that the probability distribution constructed by the algorithm satisfies $q(a, b|A, B) = \pi(M_{(a,b),A\times B})$ for all (a, b, A, B). Thus in the case where the algorithm terminates with zero flow on the choice multigraph, the correlated choice rule is rationalizable.

Now we show that positivity of the Block-Marschak polynomials, marginality holding, and termination of the algorithm with positive flow somewhere on the choice multigraph leads to a contradiction. The contradiction will establish the existence of a face on the choice multigraph having strictly positive flow along each of its edges. The existence of such a face contradicts the termination of the algorithm. To begin we need the following lemma.

Lemma 2. *If $q_i(\cdot)$ satisfy recursivity, then $q_{i+1}(\cdot)$ satisfy recursivity.*

Proof. Suppose $q_i(\cdot)$ satisfy recursivity. This means that $\sum_{a\in A} q_i(a, b|A, B) = \sum_{c\in X\setminus A} q_i(c, b|A \cup \{c\}, B)$. There are two cases to consider.

(1) There exists $c \in X \setminus A$ such that $(X \setminus (A \cup \{c\}) \succeq_i^1 c \succ_i^1 A)$ and $(Y \setminus B \succ_i^2 b \succeq_i^2 B)$. This means that $q_i(c, b|A \cup \{c\}, B)$ shows up on f_i^*. This implies that $q_{i+1}(c, b|A \cup \{c\}, B) = q_i(c, b|A \cup \{c\}, B) - q_i^*$. Without loss of generality, let a be such that $X \setminus A \succ_i^1 a \succeq_i^1 A$. This means that $q(a, b|A, B)$ shows up on f_i^*. Further this means that for all $a' \in A$ with $a' \neq a$, $q(a', b|A, B)$ does not show up on f_i^*. These two facts together mean that $q_{i+1}(a, b|A, B) = q_i(a, b|A, B) - q_i^*$ and $q_{i+1}(a', b|A, B) = q_i(a', b|A, B)$. Further, for all $c' \in X \setminus A$ with $d \neq c$, either $d \succ_i^1 c$ or $a \succ_i^1 d$. Both of these cases mean that $A \cup \{d\}$ does now appear on f_i^*. This means that $q_{i+1}(d, b|A \cup \{d\}, B) = q_i(d, b|A \cup \{d\}, B)$. Putting all these pieces together gets us the following.

$$\begin{aligned}\sum_{a \in A} q_{i+1}(a, b|A, B) &= \sum_{a \in A} q_i(a, b|A, B) - q_i^* \\ &= \sum_{c \in X \setminus A} q_i(c, b|A \cup \{c\}, B) - q_i^* \\ &= \sum_{c \in X \setminus A} q_{i+1}(c, b|A \cup \{c\}, B)\end{aligned}$$

Thus, in this case, recursivity holds.

(2) There does not exist $c \in X \setminus A$ such that $(X \setminus (A \cup \{c\}) \succeq_i^1 c \succ_i^1 A)$ or it is not the case that $(Y \setminus B \succ_i^2 b \succeq_i^2 B)$.
This means that, for all $a \in A$ and $c \in X \setminus A$, neither $q_i(a, b|A, B)$ or $q_i(c, b|A \cup \{c\}, B)$ appear on f_i^*. This means that $q_{i+1}(a, b|A, B) = q_i(a, b|A, B)$ and $q_{i+1}(c, b|A \cup \{c\}, B) = q_i(c, b|A \cup \{c\}, B)$. Putting this all together gets us the following.

$$\begin{aligned}\sum_{a \in A} q_{i+1}(a, b|A, B) &= \sum_{a \in A} q_i(a, b|A, B) \\ &= \sum_{c \in X \setminus A} q_i(c, b|A \cup \{c\}, B) \\ &= \sum_{c \in X \setminus A} q_{i+1}(c, b|A \cup \{c\}, B)\end{aligned}$$

Thus the lemma holds. □

Recall that we assume that the correlated choice rule satisfies marginality. This along with Lemma 1 and Lemma 2 implies that the choice multigraph satisfies recursivity at every step of the algorithm, including when the algorithm terminates. Without loss of generality, let $X = \{x_1, \ldots, x_N\}$ and $Y = \{a, b, c\}$. We are in the case where there is positive flow along some edge in the choice multigraph. Without loss, let it be that $q_i(a, b|A, B) > 0$. Suppose $B \neq Y$. Then there exists some $y' \in B^c$ such that $q_i(a, b'|A, B \cup \{y'\}) > 0$. This follows from recursivity and positivity of $q_i(a, b|A, B) > 0$. Apply this logic once more to get that there exists y'' such that $q_i(a, b''|A, Y) > 0$. We can apply this logic, with more iterations, to X as well. This means that without loss, $q_i(x_1, a|X, Y) > 0$. Now we will begin construction of the face with strictly positive entries.

(1) Let $X = X_1$. Without loss of generality, there exists $x_1 \in X_1$ such that $q_i(x_1, a|X_1, Y) > 0$. This follows from the previous logic.
(2) By recursivity, without loss of generality, $q_i(x_1, b|X_1, \{b, c\}) > 0$.
(3) Let $X \setminus \{x_1\} = X_2$. By recursivity, there exists $x_2 \in X_2$ such that $q_i(x_2, b|X_2, \{b, c\}) > 0$.
(4) Let $X_j = X \setminus \{x_k | k \in \mathbb{N}, k < j\}$. Let $j = 3$.
(5) By recursivity, there exists $x_j \in X_j$ such that $q_i(x_j, b|X_j, \{b, c\}) > 0$.
(6) If $j = N$, proceed to step 7. If $j < N$, let $j = j + 1$ and return to step 5.
(7) Note that for all $k \in \{1, \ldots, N\}$, $q_i(x_k, b|X_k, \{b, c\}) > 0$. This means that for all $k \in \{1, \ldots, N\}$, $q_i(x_k, c|X_k, \{c\}) > 0$. This follows from recursivity.
(8) Note that $q_i(x_k, b|X_k, \{b, c\}) > 0$ for all $k \in \{1, \ldots, N\}$ further implies that for all $k \in \{1, \ldots, N\}$, $q_i(x_k, a|X_k, Y) > 0$. This too follows from recursivity.

The above steps have shown that the face corresponding to $(x_1 \succ_1 \cdots \succ_1 x_N, a \succ_2 b \succ_2 c)$ has strictly positive flow along all of its edges. This contradicts the the algorithm terminating. This means that the algorithm must only terminate in case one, which we have shown leads to a rationalization of the system of choice probabilities. So we are done.

□

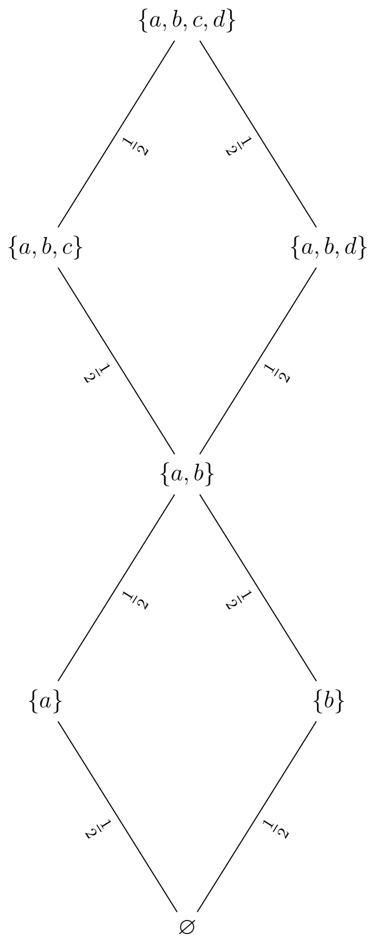
B.3. **Observation 1.** During this example, let $X = \{a, b, c, d\}$ and $Y = \{w, x, y, z\}$. The table below shows the non-zero BM-polynomials of the correlated choice rule. All the non-zero BM-polynomials are equal to 0.5.

Set	$\{w, x, y, z\}$	$\{x, y, z\}$	$\{w, y, z\}$	$\{y, z\}$	$\{y\}$	$\{z\}$
$\{a, b, c, d\}$	$q(d, w)$ $q(c, x)$	$q(d, x)$	$q(c, w)$	$q(c, z)$ $q(d, y)$	$q(c, y)$	$q(c, z)$
$\{a, b, c\}$	$q(c, w)$	$q(c, x)$	-	$q(c, y)$	-	$q(c, z)$
$\{a, b, d\}$	$q(d, x)$	-	$q(d, w)$	$q(d, z)$	$q(d, y)$	-
$\{a, b\}$	$q(a, x)$ $q(b, w)$	$q(b, x)$	$q(a, w)$	$q(a, y)$ $q(b, z)$	$q(b, y)$	$q(a, z)$
$\{a\}$	$q(a, w)$	$q(a, x)$	-	$q(a, z)$	$q(a, y)$	-
$\{b\}$	$q(b, x)$	-	$q(b, w)$	$q(b, y)$	-	$q(b, z)$

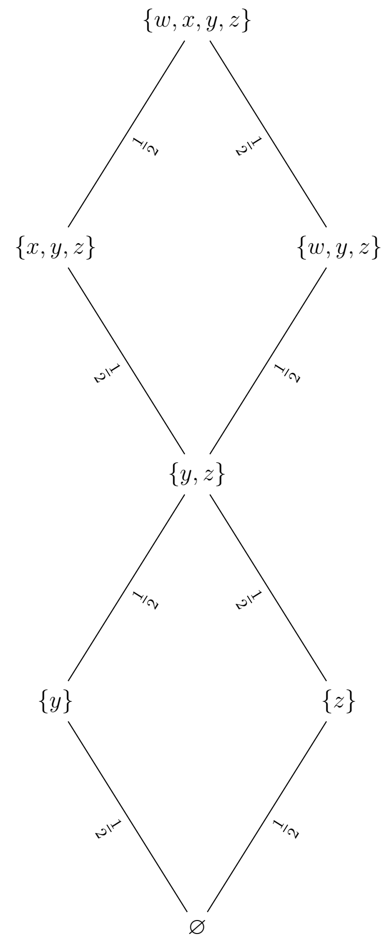
TABLE 4. Block-Marschak polynomial values for the 4 by 4 counterexample. The corresponding sets for the first agent are given by the set column and the corresponding sets for the second agent are given by the set row. Each cell indicates the non-zero BM-polynomials of the correlated choice rule. All the non-zero BM-polynomials are equal to 0.5

Figure 2 shows the marginal component of the X-Marginal graph system and the marginal component of the Y-Marginal graph system. Now we will go about describing the conditional components of each of these graph systems. The conditional component of the X-Marginal graph system can be described as follows.

(1) For the edge connecting $\{a, b, c, d\}$ to $\{a, b, c\}$ and the edge connecting $\{a, b, c\}$ to $\{a, b\}$, only the path corresponding to $w \succ x \succ y \succ z$ has positive flow along the conditional component. That flow is equal to $\frac{1}{2}$.
(2) For the edge connecting $\{a, b, c, d\}$ to $\{a, b, d\}$ and the edge connecting $\{a, b, d\}$ to $\{a, b\}$, only the path corresponding to $x \succ w \succ z \succ y$ has positive flow along the conditional component. That flow is equal to $\frac{1}{2}$.
(3) For the edge connecting $\{a, b\}$ to $\{a\}$ and the edge connecting $\{a\}$ to $\varnothing$, only the path corresponding to $w \succ x \succ z \succ y$ has positive flow along the conditional component. That flow is equal to $\frac{1}{2}$.



(A) Marginal component of the X-Marginal graph system.



(B) Marginal component of the Y-Marginal graph system.

FIGURE 2. Above are the marginal components for the two marginal graph systems from the four by four counterexample. All edges with zero flow are left out of the diagram.

(4) For the edge connecting $\{a, b\}$ to $\{b\}$ and the edge connecting $\{b\}$ to $\varnothing$, only the path corresponding to $x \succ w \succ y \succ z$ has positive flow along the conditional component. That flow is equal to $\frac{1}{2}$.

The conditional component of the Y-Marginal graph system is symmetric to the conditional component of the X-Marginal graph system. For clarity, we describe it as follows.

(1) For the edge connecting $\{w,x,y,z\}$ to $\{x,y,z\}$ and the edge connecting $\{x,y,z\}$ to $\{y,z\}$, only the path corresponding to $d \succ c \succ b \succ a$ has positive flow along the conditional component. That flow is equal to $\frac{1}{2}$.
(2) For the edge connecting $\{w,x,y,z\}$ to $\{w,y,z\}$ and the edge connecting $\{w,y,z\}$ to $\{y,z\}$, only the path corresponding to $c \succ d \succ a \succ b$ has positive flow along the conditional component. That flow is equal to $\frac{1}{2}$.
(3) For the edge connecting $\{y,z\}$ to $\{y\}$ and the edge connecting $\{y\}$ to $\varnothing$, only the path corresponding to $c \succ d \succ b \succ a$ has positive flow along the conditional component. That flow is equal to $\frac{1}{2}$.
(4) For the edge connecting $\{y,z\}$ to $\{z\}$ and the edge connecting $\{z\}$ to $\varnothing$, only the path corresponding to $d \succ c \succ a \succ b$ has positive flow along the conditional component. That flow is equal to $\frac{1}{2}$.

The structure of this counterexample can be found in the marginal graph system of correlated choice rules in larger choice environments. Whenever the marginal graph system of a correlated choice rule contains the above structure, the correlated choice rule fails to have a CRUM representation.

B.4. **Proof of Theorem 4.** Consider two agents with corresponding sets X and Y. Let $\Sigma \subseteq \mathcal{X} \times \mathcal{Y}$ be the choice domain with corresponding correlated choice rule p.

Proof. We begin by proving that having a CRUM representation implies ARSP. Suppose that the correlated choice rule has a CRUM representation. This implies that there exists some $\pi \in \Delta(\mathcal{L}(X) \times \mathcal{L}(Y))$ such that for all $A \times B \in \Sigma$, for all $(a,b) \in A \times B$, $P_{A\times B}(a,b) = \pi(\{(\succ,\succ')|(a,b) = (M(A_i,\succ), M(B_i,\succ'))\})$. This further means that for all finite sequences of $\{((a_i,b_i),A_i,B_i)\}_{i=1}^n$ with $(a_i,b_i) \in A_i \times B_i$, we have

$$\begin{aligned}\Sigma_{i=1}^n P_{A_i\times B_i}(a_i,b_i) &= \Sigma_{i=1}^n \pi(\{(\succ,\succ')|(a_i,b_i) = (M(A_i,\succ), M(B_i,\succ'))\}) \\ &\leq \max_{\pi' \in \Delta(\mathcal{L}(X)\times\mathcal{L}(Y))} \Sigma_{i=1}^n \pi'(\{(\succ,\succ')|(a_i,b_i) = (M(A_i,\succ), M(B_i,\succ'))\})\end{aligned}$$

The right hand side of the above equation is a maximization problem of a linear function on a compact and convex domain. This means that the maximizer is an extreme point.

Formally, this means that

$$\Sigma_{i=1}^{n} P_{A_i \times B_i}(a_i, b_i) \leq \max_{\pi' \in \Delta(\Pi_X \times \Pi_Y)} \Sigma_{i=1}^{n} \pi'(\{(\pi_X, \pi_Y) | (a_i, b_i) = (M(A_i, \succ), M(B_i, \succ'))\})$$

$$= \max_{(\pi_X, \pi_Y) \in \Pi_X \times \Pi_Y} \Sigma_{i=1}^{n} \mathbf{1}\{(a_i, b_i) = (M(A_i, \succ), M(B_i, \succ'))\}$$

Thus having a CRUM representation implies ARSP.

We will now show that ARSP holding implies that the correlated choice rule has a CRUM representation. We proceed by contraposition. Consider a matrix M with columns indexed by $(\succ, \succ') \in \mathcal{L}(X) \times \mathcal{L}(Y)$, the rows indexed by $((a, b), A \times B)$ for $A \times B \in \Sigma$ and $(a, b) \in A \times B$, and entry $((a, b), A \times B), (\succ, \succ'))$ given by $\mathbf{1}\{(a_i, b_i) = (M(A_i, \succ), M(B_i, \succ'))\}$. Let $W = \begin{bmatrix} M \\ 1 \end{bmatrix}$. Consider a column vector Q indexed by $((a, b), A \times B)$ with entry $((a, b), A \times B)$ given by $P_{A \times B}(a, b)$. Call $P = \begin{bmatrix} Q \\ 1 \end{bmatrix}$. If the correlated choice rule does not have a CRUM representation, that means there does not exist $\pi \in \mathbb{R}_+^{|\mathcal{L}(X) \times \mathcal{L}(Y)|}$ such that $W \cdot \pi = P$. By Farkas' Lemma, this means that there exists $j \in R^{|\mathcal{L}(X) \times \mathcal{L}(Y)|}$ such that $j \cdot W \leq 0$ and $j \cdot P > 0$. Writing out each line of the above equation gives us

$$\Sigma_{((a,b), A \times B)} j((a, b), A \times B) \mathbf{1}\{(a, b) = (M(A, \succ), M(B, \succ'))\} + j_{|P|} \leq 0$$

and

$$\Sigma_{((a,b), A \times B)} j((a, b), A \times B) P_{A \times B}(a, b) + j_{|P|} > 0$$

for all $(\succ, \succ') \in \mathcal{L}(X) \times \mathcal{L}(Y)$. Combining the above two equations gives us the following.

$$\Sigma_{((a,b), A \times B)} j((a, b), A \times B) P_{A \times B}(a, b) + j_{|P|} >$$
$$\Sigma_{((a,b), A \times B)} j((a, b), A \times B) \mathbf{1}\{(a, b) = (M(A, \succ), M(B, \succ'))\} + j_{|P|}$$

Let $J^+ = \{((a, b,), A \times B) | j((a, b,), A \times B) \geq 0\}$. Similarly, let
$J^- = \{((a, b,), A \times B) | j((a, b,), A \times B) < 0\}$. Using this notation to write the prior

equation gives us

$$\Sigma_{((a,b),A\times B)\in J^+} j((a,b),A\times B)P_{A\times B}(a,b) - \Sigma_{((a,b),A\times B)\in J^-}|j((a,b),A\times B)|P_{A\times B}(a,b) > \Sigma_{((a,b),A\times B)\in J^+} j((a,b),A\times B)\mathbf{1}\{(a,b)=(M(A,\succ),M(B,\succ'))\} - \Sigma_{((a,b),A\times B)\in J^-}|j((a,b),A\times B)|\mathbf{1}\{(a,b)=(M(A,\succ),M(B,\succ'))\}$$

Recall that

$$P_{A\times B}(a,b) = 1 - \Sigma_{(a,b)\in(A\times B)\setminus\{(a,b)\}} P_{A\times B}(a,b)$$

and that

$$\mathbf{1}\{(a,b)=(M(A,\succ),M(B,\succ'))\} = 1 - \Sigma_{(a,b)\in(A\times B)\setminus\{(a,b)\}}\mathbf{1}\{(a,b)=(M(A,\succ),M(B,\succ'))\}.$$

This implies that

$$\Sigma_{((a,b),A\times B)\in J^+} j((a,b),A\times B)P_{A\times B}(a,b) + \Sigma_{((a,b),A\times B)\in J^-}\Sigma_{(a,b)\in(A\times B)\setminus\{(a,b)\}}|j((a,b),A\times B)|P_{A\times B}(a,b) > \Sigma_{((a,b),A\times B)\in J^+} j((a,b),A\times B)\mathbf{1}\{(a,b)=(M(A,\succ),M(B,\succ'))\} + \Sigma_{((a,b),A\times B)\in J^-}\Sigma_{(a,b)\in(A\times B)\setminus\{(a,b)\}}|j((a,b),A\times B)|\mathbf{1}\{(a,b)=(M(A,\succ),M(B,\succ'))\}.$$

Since this is a finite system of strict linear inequalities, if there is a solution j, then there is a solution j where each component is rational. Let CD_j be the common denominator of this rational j. Let be K be defined as follows.

$$K((a,b),A\times B) = \begin{cases} CD_j j((a,b),A\times B) & \text{if } ((a,b),A\times B)\in J^+ \\ CD_j|j((a,b),A\times B)| & \text{if } ((a,b),A\times B)\in J^- \end{cases}$$

Using this new notation, the we can rewrite the last inequality as follows.

$$\Sigma_{((a,b),A\times B)\in J^+}K((a,b),A\times B)P_{A\times B}(a,b)+$$
$$\Sigma_{((a,b),A\times B)\in J^-}\Sigma_{(a,b)\in(A\times B)\setminus\{(a,b)\}}K((a,b),A\times B)P_{A\times B}(a,b) >$$
$$\Sigma_{((a,b),A\times B)\in J^+}K((a,b),A\times B)\mathbf{1}\{(a,b)=(M(A,\succ),M(B,\succ'))\}+$$
$$\Sigma_{((a,b),A\times B)\in J^-}\Sigma_{(a,b)\in(A\times B)\setminus\{(a,b)\}}K((a,b),A\times B)\mathbf{1}\{(a,b)=(M(A,\succ),M(B,\succ'))\}.$$

These $K((a,b),A\times B)$ correspond to a new finite sequence of $\{((a_i,b_i),A_i\times B_i)\}_{i=1}^n$. This along with the above inequality implies

$$\Sigma_{i=1}^n P_{A_i\times B_i}(a_i,b_i) > \Sigma_{i=1}^n \mathbf{1}\{(a_i,b_i)=(M(A_i,\succ),M(B_i,\succ'))\}.$$

The above inequality holds for all $(\succ,\succ')\in\mathcal{L}(X)\times\mathcal{L}(Y)$. This means that there exists some $\{(a_i,b_i),A_i\times B_i\}_{i=1}^n$ such that

$$\Sigma_{i=1}^n P_{A_i\times B_i}(a_i,b_i) > \max_{(\succ,\succ')\in\mathcal{L}(X)\times\mathcal{L}(Y)} \Sigma_{i=1}^n \mathbf{1}\{(a_i,b_i)=(M(A_i,\succ),M(B_i,\succ'))\}.$$

This is the negation of the inequality shown in ARSP. Thus by contraposition, ARSP holding implies that the correlated choice rule has a CRUM representation. Thus Theorem 4 holds.

□

References

Allen, R., P. Dziewulski, and J. Rehbeck (2021): "Revealed Statistical Consumer Theory," Unpublished.

Anscombe, F. J., R. J. Aumann, et al. (1963): "A Definition of Subjective Probability," *Annals of Mathematical Statistics*, 34, 199–205.

Barberá, S. and P. K. Pattanaik (1986): "Falmagne and the Rationalizability of Stochastic Choices in Terms of Random Orderings," *Econometrica*, 54, 707–715.

Block, H. D., J. Marschak, et al. (1959): "Random Orderings and Stochastic Theories of Response," Tech. rep., Cowles Foundation for Research in Economics, Yale University.

Blume, L. E., W. A. Brock, S. N. Durlauf, and Y. M. Ioannides (2011): "Identification of Social Interactions," in *Handbook of Social Economics*, Elsevier, vol. 1, 853–964.

Blume, L. E., W. A. Brock, S. N. Durlauf, and R. Jayaraman (2015): "Linear Social Interactions Models," *Journal of Political Economy*, 123, 444–496.

Borah, A., C. Kops, et al. (2018): "Choice via Social Influence," Unpublished.

Border, K. (2007): "Introductory notes on stochastic rationality," California Institute of Technology.

Bramoullé, Y., H. Djebbari, and B. Fortin (2009): "Identification of Peer Effects Through Social Networks," *Journal of Econometrics*, 150, 41–55.

Calvó-Armengol, A., E. Patacchini, and Y. Zenou (2009): "Peer Effects and Social Networks in Education," *The Review of Economic Studies*, 76, 1239–1267.

Cerreia-Vioglio, S., D. Dillenberger, P. Ortoleva, and G. Riella (2019): "Deliberately Stochastic," *American Economic Review*, 109, 2425–45.

Chambers, C. P., Y. Masatlioglu, and T. Cuhadaroglu (2020): "Behavioral Influence," Unpublished.

Cuhadaroglu, T. (2017): "Choosing on Influence," *Theoretical Economics*, 12, 477–492.

De Paula, A. (2017): "Econometrics of Network Models," in *Advances in Economics and Econometrics: Theory and Applications: Eleventh World Congress*, Cambridge University Press Cambridge, vol. 1, 268–323.

Duraj, J. (2018): "Dynamic Random Subjective Expected Utility," Unpublished.

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

Fiorini, S. (2004): "A Short Proof of a Theorem of Falmagne," *Journal of Mathematical Psychology*, 48, 80–82.

Fishburn, P. C. (1998): "Stochastic Utility," in *Handbook of Utility Theory*, ed. by S. Barbera, P. Hammond, and C. Seidl, Kluwer Dordrecht, 273–318.

Ford, L. R. and D. R. Fulkerson (1956): "Maximal Flow Through a Network," *Canadian journal of Mathematics*, 8, 399–404.

Frick, M., R. Iijima, and T. Strzalecki (2019): "Dynamic Random Utility," *Econometrica*, 87, 1941–2002.

Fudenberg, D. and T. Strzalecki (2015): "Dynamic Logit with Choice Aversion," *Econometrica*, 83, 651–691.

Lu, J. (forthcoming): "Random Ambiguity," *Theoretical Economics.*

Lu, J. and K. Saito (2018): "Random Intertemporal Choice," *Journal of Economic Theory*, 177, 780–815.

Luce, R. D. (2005): *Individual Choice Behavior: A Theoretical Analysis*, Dover.

Machina, M. J. (1985): "Stochastic Choice Functions Generated From Deterministic Preferences Over Lotteries," *The Economic Journal*, 95, 575–594.

Manski, C. F. (1993): "Identification of Endogenous Social Effects: The Reflection Problem," *The Review of Economic Studies*, 60, 531–542.

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

Mas, A. and E. Moretti (2009): "Peers at Work," *American Economic Review*, 99, 112–45.

McFadden, D. and M. K. Richter (1990): "Stochastic Rationality and Revealed Stochastic Preference," *Preferences, Uncertainty, and Optimality, Essays in Honor of Leo Hurwicz, Westview Press: Boulder, CO*, 161–186.

McFadden, D. L. (2005): "Revealed Stochastic Preference: A Synthesis," *Economic Theory*, 26, 245–264.

Pennesi, D. (forthcoming): "Intertemporal Discrete Choice," *Journal of Economic Behavior & Organization.*

Rota, G.-C. (1964): "On the Foundations of Combinatorial Theory I. Theory of Möbius Functions," *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 2, 340–368.

Rust, J. (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 57, 999–1033.

Savage, L. J. (1972): *The Foundations of Statistics*, Dover.

Topa, G. (2001): "Social Interactions, Local Spillovers and Unemployment," *The Review of Economic Studies*, 68, 261–295.