Testing axioms of stochastic discrete choice using population choice probabilities

William J. McCausland*[1] and A.A.J. Marley †[2]

[1]Université de Montréal
[2]University of Victoria

April 14, 2017

Abstract

We test several axioms of stochastic discrete choice for population level choice probabilities, using choice data from a new experiment featuring 20 different choice domains, including fine art, travel itineraries, and pizza toppings. For each choice domain, there is a universe of five objects, and for each of the 26 doubleton and larger subsets of a given universe, we observe the choices of 40 subjects. Each subject faces exactly one choice subset from each domain, making this a strictly between-subject design; unlike similar studies with within-subject or mixed designs, the common assumption that choices are independent, and on each choice set, identically distributed, is innocuous. We report evidence, in the form of Bayes factors, for and against each of the axioms, using data for each domain. We find...

1 Introduction

Let $T = (x_1, \ldots, x_n)$ be a universe of choice objects. When faced with a non-empty choice set $A \subseteq T$, an agent chooses a single object from A. The probability that the agent chooses $x \in A$ is denoted $P_A(x)$. A *random choice structure* (RCS) is the complete specification of the $P_A(x)$, $x \in A \subseteq T$, and is denoted (T, P).

We distinguish between two different interpretations of a RCS. An *individual* RCS governs the choices of a single individual; a *population* RCS, those of a random individual selected from a population of individuals. Note that if each individual in a population is governed by an individual RCS then the population RCS will be a convex combination of individual RCSs: each population choice distribution P_A, $A \subseteq T$ is a mixture—with random sampling from the population being the common mixing distribution—of the various individual distributions P_A.

Where there are multiple trials, we complete the probabilistic model by assuming that the same $P_A(\cdot)$ governs every choice from A and that choices are statistically independent across trials. Following others, we refer to these two assumptions together as the iid assumption; note that "identically distributed" applies separately to each choice set while "independent" applies globally.

There are good reasons to be skeptical of these assumptions in experiments featuring within-subject or mixed designs; subjects may be learning or their attention may be waning during the

*william.j.mccausland@umontreal.ca, File: `among.tex`, Very preliminary.
†ajmarley@uvic.ca

course of the experiment. While steps can be taken to attenuate these problems by limiting the number of repetitions of each choice set and by using distractor trials to make it more difficult for subjects to recognize choice objects, and while we can test the assumptions to some extent, we can never be sure that the assumptions hold. In a strictly between-subject design, however, with no communication between subjects, the iid assumption is much more plausible.

1.1 Axioms of discrete stochastic choice

Various axioms, conditions, properties and hypotheses about probabilistic choice behaviour can be expressed as restrictions over the various choice probabilities of a RCS. Henceforth, we will use the term axiom as a generic term to include all of these. Examples, all defined in Appendix A, include weak, moderate and strong stochastic transitivity, regularity, the triangle inequality, the Block-Marshak inequalities and the multiplicative inequality. Falmagne (1978) showed that the Block-Marshak inequalities are necessary and sufficient for random utility. Tversky (1972a) and Sattath and Tversky (1976) establish that the Block-Marshak inequalities, moderate stochastic transitivity and the multiplicative inequality are all necessary conditions for, and thus testable implications of, the Elimination by Aspects model introduced by Tversky (1972b).

Whether or not an axiom applies to a population RCS and whether or not it applies to individual RCSs are both of interest, but they are quite distinct questions. For example, weak stochastic transitivity (WST) is considered by many to be a compelling consistency principle applying to individual choice. However, the Condorcet voting paradox shows us that we should not always expect population choice probabilities to satisfy WST.

Nonetheless, for some axioms, if the RCS for all individuals in the population satisfy the axiom, then the population RCS must as well. Expressed in contrapositive form to make plain the implications for testing, if the axiom is violated for a population, than it is violated by at least one individual. It is easy to see that the triangle inequality, regularity and random utility are examples of these axioms: each of these axioms consists of a set of linear inequalities, and so the region where the axiom is satisfied is a convex set; the population RCS, being a convex combination of individual RCSs, must satisfy the axiom whenever the individuals all do.

However, there are other axioms that can hold for every individual in a population but not the population itself. This is true of weak, moderate and strong stochastic transitivity, and the multiplicative inequality. Table 1 shows the relevant counterexamples: sets of individual RCS's satisfying an axiom aggregating to population RCS's that do not. In all cases, the RCS's are induced by random preferences. Thus, the non-aggregation result holds whether or not we require the RCSs to be induced by random preference. Lines 1, 2 and 3 give degenerate ranking and choice probabilities for three individuals (or types of individuals) in a population. The choice behaviour of each individual satisfies weak, moderate and strong stochastic transitivity. The convex combination of the three individual random choice structures, with equal weights, gives the population random choice structure of line 4. These probabilities violate even weak stochastic transitivity, since $p(x,y) = p(y,z) = 2/3$, but $p(x,z) = 1/3$. This first example is essentially the Condorcet voting paradox. A second counterexample pertains to the multiplicative inequality. Lines 5 and 6 in Table 1 give choice probabilities for two individuals, each satisfying the multiplicative inequality. The equal-weighted convex combination of the two gives the population RCS of line 7, which violates the multiplicative inequality since $P_{\{x,y,z\}}(y) = 1/2 < p(y,x)p(y,z) = 3/4$.

While some axioms and conditions involve binary choice probabilities only, (e.g. weak, moderate and strong stochastic transitivity, the triangle inequality) other axioms involve multiple choice probabilities. Regularity, the multiplicative inequality and the Block-Marshak inequalities jointly constrain all choice probabilities on doubleton and larger choice sets. It is easy to see from the

Line	p_i	π_{xyz}	π_{xzy}	π_{yxz}	π_{yzx}	π_{zxy}	π_{zyx}	$p(x,y)$	$p(y,z)$	$p(x,z)$	$P_T(x)$	$P_T(y)$
1	$\frac{1}{3}$	1						1	1	1	1	
2	$\frac{1}{3}$				1				1			1
3	$\frac{1}{3}$					1		1				
4		$\frac{1}{3}$			$\frac{1}{3}$	$\frac{1}{3}$		$\frac{2}{3}$	$\frac{2}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
5	$\frac{1}{2}$	$\frac{1}{2}$		$\frac{1}{2}$				$\frac{1}{2}$	1	1	$\frac{1}{2}$	$\frac{1}{2}$
6	$\frac{1}{2}$				$\frac{1}{2}$		$\frac{1}{2}$		$\frac{1}{2}$			$\frac{1}{2}$
7		$\frac{1}{4}$		$\frac{1}{4}$	$\frac{1}{4}$		$\frac{1}{4}$	$\frac{1}{4}$	$\frac{3}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{2}$

Table 1: Two aggregations of individual choice probabilities. Each line gives a random preference (columns π_{xyz} through π_{zyx}) and the induced RCS (columns $p(x,y)$ through $P_T(x)$). The individual random preferences and RCS's in lines 1, 2 and 3 aggregate (with mixing probabilities p_i) to the population random preference and RCS in line 4; likewise, lines 5 and 6 aggregate to line 7.)

definitions of these axioms that in each case, if we are given an RCS satisfying the axiom and an arbitrary $A \subseteq T$ with at least two elements, we can alway modify $P_A(\cdot)$ in such a way that the modified RCS violates the axiom.

1.2 Context effects

In Psychology, the term "context effect" refers broadly to an effect of the environment on the perception of stimuli. It also refers more narrowly, in Marketing, Economics and Psychology, to the effect on consumers' preferences of the set of available options. Three documented context effects are relevant to the subject of stochastic discrete choice: the asymmetric dominance, similarity and compromise effects.

The asymmetric dominance effect was introduced by Huber, Payne, and Puto (1982), and is defined up to a dominance relation on the universe of choice objects. In pertains to choice environments with three objects: a "target" x, a "competitor" y and a "decoy" z; x "dominates" z, but y does not; neither x nor y dominates the other. The effect occurs when $P_{\{x,y,z\}}(x) > p(x,y)$, and constitutes evidence against regularity and random utility.

Tversky (1972b) introduces the similarity effect, defined up to the definition of a similarity relation on the choice objects under consideration. The similarity effect occurs if there are x, y and z such that x and y are similar, x and z are not similar and y and z are not similar, and

$$p(x,z) > \frac{P_{\{x,y,z\}}(x)}{P_{\{x,y,z\}}(x) + P_{\{x,y,z\}}(z)} \quad \text{or} \quad p(y,z) > \frac{P_{\{x,y,z\}}(y)}{P_{\{x,y,z\}}(y) + P_{\{x,y,z\}}(z)}.$$

The inequality on the left holds when adding the object y, similar to the target object x, to the choice set $\{x,z\}$ lowers the *relative* probability of choosing x, compared to z. The inequality on the right is similar, except that y is the target object and x is the similar object. The effect constitutes evidence against simple scalability.

The compromise effect, introduced by Simonson (1989) and made more precise in Tversky and Simonson (1993) pertains to choice environments where objects can be described in terms of a set of real-valued attributes. An object y is *between* objects x and z, denoted $x|y|z$, if two conditions hold: first, for each attribute, y has a level of that attribute intermediate between the levels of x

and z; and second, that each of x and z have the highest level of at least one attribute. This effect is said to occur when $x|y|z$ and one of the following inequalities, from page 1183 in Tversky and Simonson (1993), hold:

$$p(y,x) < \frac{P_{\{x,y,z\}}(y)}{P_{\{x,y,z\}}(y) + P_{\{x,y,z\}}(x)} \quad \text{or} \quad p(y,z) < \frac{P_{\{x,y,z\}}(y)}{P_{\{x,y,z\}}(y) + P_{\{x,y,z\}}(z)}.$$

The inequality on the left holds when adding the object z to the choice set $\{x, y\}$, making y a "compromise" between x and z, raises the relative probability of choosing y, compared to x. The inequality on the right is similar, except that x is added to $\{y, z\}$, again making y a "compromise" between the two. The effect also constitutes evidence against simple scalability.

1.3 Our contribution

With these issues in mind, we have designed an experiment with the following features. First, for each choice domain, we collect choice data on all doubleton and larger sets of the universe of objects. In doing this, we expose *every* implication of regularity, the multiplicative inequality and random utility to possible falsification. Second, we use a strictly between-subject design, where each subject makes a single choice from each choice domain. This makes the iid assumption very reasonable. Third, rather than asking subjects to make repeated choices in a domain, we allocate their limited attention to a wide variety of choice domains. In this way, we hope to identify regularities in population choice that apply widely across domains. We have included domains where choice objects were chosen to elicit context effects as well as other domains where there is no obvious reason to expect context effects.

2 Experimental Design

Describe implementation of experiment, role of SSI and I4C here.

Describe experimental conditions here.

Describe demographic information on subjects here.

Our experiment has J choice domains, described in Section 3 below, each with a universe of size five. There are $I = 2^5 - 5 - 1 = 26$ choice sets (binary and larger subsets of the universe) for each domain, for a total of JI choice sets. A total of NI subjects completed the survey, where N is the number of subjects responding to each of the JI choice sets. Each subject sees exactly one choice set from each domain, and so faces J trials. Thus, there are a total of NIJ trials.

For each domain j, the NI subjects are partitioned into I groups of size N, with subjects in each group seeing the same choice set from that domain. Random partitions are statistically independent across domains.

Elements of a choice set are presented in random position (left, middle, right, for example) on the screen, independently across the N participants in the relevant group. The J trials assigned to each participant are presented in random order, with order independent across participants.

Subjects must choose exactly one object from each choice set.

3 Domains

Here, we describe a set of J choice domains.

We classify choice domains into four categories. In the first category, objects have no numerical attributes. In the second category, there is a single attribute, whose level is not explicitly given. In the third category, there are two numerical attributes and the experimental design is intended to elicit one or more context effects. In the fourth category, there are many attributes, and the objects are chosen to resemble objects in discrete choice experiments used in Marketing, although with fewer attributes.

The section titles, "Male stars" for example, give the names of the domains for the purpose of reporting results. The specification of a domain is shown in a grey box and consists of a question and five choice objects, or responses. In the actual experiment, for a given domain, all participants see the same question (such as "Which movie star would you most like to have lunch with?") but not the same set of responses; different participants will see from two to five of the possible responses, in random order, as described in Section 2. For the purpose of reporting results, the choice objects in each domain are labelled a, b, c, d and e, in the order presented in the grey boxes. The subjects do not see these labels, and accordingly, we do not use them within the boxes.

We adopt the convention that whenever there is an established order for the items in a list, we order them in decreasing order.

3.1 No numerical attributes

3.1.1 Male stars

The source for this domain is the IMDb list "Top 25 Biggest MOVIE STARS in the World!" on May 25th, 2015. The choices are the top five male actors in that list, in order.

Which movie star would you most like to have lunch with?

- Robert Downey Jr.
- Leonardo DiCaprio
- Tom Cruise
- Johnny Depp
- George Clooney

3.1.2 Female stars

The source for this domain is the IMDb list "Top 25 Biggest MOVIE STARS in the World!" on May 25th, 2015. These are the top five female actors in that list, in order.

Which movie star would you most like to have lunch with?

- Meryl Streep
- Jennifer Lawrence
- Emma Stone
- Kristen Stewart
- Anne Hathaway

3.1.3 Films

The source for this domain is the IMDb list "Most Popular Feature Films Released 1990 to 1999". The decade was chosen so that the films would not be easily recognizable by most participants.

> Judging from the following descriptions of films, which film would you most want to see?
>
> - Two imprisoned men bond over a number of years, finding solace and eventual redemption through acts of common decency.
> - Mathilda, a 12-year-old girl, is reluctantly taken in by Léon, a professional assassin, after her family is murdered. Léon and Mathilda form an unusual relationship, as she becomes his protégé and learns the assassin's trade.
> - The lives of two mob hit men, a boxer, a gangster's wife, and a pair of diner bandits intertwine in four tales of violence and redemption.
> - A sexually frustrated suburban father has a mid-life crisis after becoming infatuated with his daughter's best friend.
> - Identical twins, separated at birth and each raised by one of their biological parents, discover each other for the first time at summer camp and make a plan to bring their wayward parents back together.

3.1.4 Star pairs

Here, choice objects are pairs of movie stars from a set of four movie stars: Tom Hanks, Scarlett Johansson, Brad Pitt and Angelina Jolie. The only missing pair is Brad Pitt and Angelina Jolie. One possible measure of similarity is the number of actors in common between two pairs, with values 0 and 1. There are nine doubleton choice sets (i.e. pairs of actor pairs) with one star in common and one ($\{c, d\}$) without any stars in common. Thus there are three triples ($\{a, c, d\}$, $\{b, c, d\}$, $\{c, d, e\}$) where one might expect a similarity effect.

In this example, subjects' preferences may depend not only on their liking of particular actors but also on complementaries between actors.

> Knowing only who is starring, which new film would you most like to see?
>
> - Tom Hanks and Scarlett Johansson
> - Scarlett Johansson and Brad Pitt
> - Tom Hanks and Brad Pitt
> - Scarlett Johansson and Angelina Jolie
> - Tom Hanks and Angelina Jolie

3.1.5 Pizzas

The source for this domain is the FF Pizza menu at the Beaubien location in Montreal. All these pizzas are either 12 or 13 dollars.

Which of the following pizzas would you prefer?

- Mozzarella, tomato sauce, basil
- Pepperoni, mushrooms, green pepper, mozzarella, tomato sauce
- Red onion, tomato sauce, feta, mozzarella, olive oil, Greek spices, tomato sauce
- Bacon, white onion, mozzarella, parmesan, fresh cream, tomato sauce, ground pepper
- Mushrooms, green pepper, mozzarella, tomato sauce

3.1.6 Juices

Which of the following fresh juices would you prefer?

- Mango
- Orange
- Apple
- Grapefruit
- Pineapple

3.1.7 Colours

Which of these colours do you like best?

- Red
- Purple
- Pink
- Blue
- Green

3.1.8 Color combinations

The source for this domain is the website “The top tens”, page “Two colors that look good side by side.” The color combinations here are ranked 1, 4, 5, 13 and 14. We chose a selection of five high ranking combinations among which there were many colors in common. Using a similarity measure equal to the number of colours in common between two pairs, there are two doubleton choice sets where the two colour pairs have no colours in common ($\{a, e\}$ and $\{b, d\}$) and eight where the two colour pairs have one colour in common. This gives six tripleton pairs in which one might expect a similarity effect.

Which of these colour combinations do you like the best?

- Black and red
- Black and purple
- Black and blue
- Blue and red
- Blue and purple

3.1.9 Events

This domain involves comparisons of the probabilities of future events. Logically, the probability of event e must be as least as great as the probability of a, which must in turn be as least as great as the probability of d; also, the probability of b must be as great as the probability of c.

> Which of the following events is most likely to happen in the next twenty years?
>
> - Scotland becomes an independent country.
> - Either Catalonia or Quebec become independent countries.
> - Catalonia becomes an independent country.
> - Scotland and Quebec become independent countries.
> - Either Scotland or Quebec become independent countries.

3.1.10 Radio formats

The domain is radio formats, and the choice objects are the top 5 radio formats in Canada in 2015, according to

https://byrnesmedia.com/2015/10/05/the-6-best-performing-radio-formats-in-canada/

Descriptions of formats are from

http://www.newsgeneration.com/broadcast-resources/guide-to-radio-station-formats/

> Suppose you were on a two hour road trip and you had to choose among radio stations with the following formats. Which would you choose?
>
> - News
> - Hot Adult Contemporary, or Hot AC (A variety of classic and contemporary mainstream music geared towards adults.)
> - Classic Hits (Rock and pop, roughly 1964-1989)
> - Country Music
> - Adult Contemporary, or AC (Adult-oriented pop/rock with no hard rock.)

3.1.11 Musical artists

The choice objects in this domain are the top selling musical artists of all time, according to Wikipedia. They should be familiar to a large majority of subjects.

> Which of the following musical artists do you prefer?
>
> - The Beatles
> - Elvis Presley
> - Michael Jackson
> - Madonna
> - Elton John

3.1.12 Aboriginal art

Which of the following examples of Australian aboriginal art do you find most appealing?

3.1.13 Impressionist art

Which of the following examples of Impressionist art do you find most appealing?

3.1.14 Sentences

This domain consists of sentences that have been used in experiments of linguistic judgement. They are, respectively sentences 7j, 7h, 7p, 7a and 7e in Bard, Robertson, and Sorace (1996). Figure 1 of that paper shows acceptability scores for these and other sentences given by two individual linguists, an acceptability score aggregating the scores of four linguists and an acceptability score aggregating the scores of four "naive subjects", all undergraduate anatomy students. There is broad, but not perfect, agreement in terms of order, and in the following list, they are in decreasing order of acceptability according to the measure aggregating the judgements of four linguists.

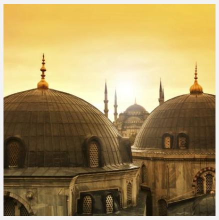
Which of the following sentences do you find the most grammatically acceptable?

- Who did Bill buy the car to please?
- This is a book which reading would be fun.
- Where did Bill buy the car to drive?
- Which man do you wonder when to meet?
- With which pen do you wonder what to write?

3.1.15 Travel

The source is Tripadvisor. These are the top five travel destinations, according to the results of an on-line contest where visitors to a Tripadvisor site could make pairwise choices between travel destinations.

Which of the following travel destinations would you prefer?

1. Marrakech, Morocco 
2. Istanbul, Turkey 
3. Hanoi, Vietnam 
4. Siem Reap, Cambodia 
5. Praque, Czech Republic 

3.2 Single attributes, not directly observed

The five domains in this section contain choice objects with an objective rank order.

3.2.1 Latitude

The five cities of this domain have a latitude close to 45 degrees north. In the following list, they are ordered from furthest north to furthest south. According to Wikipedia, their latitudes are, respectively, $46°12'N$, $45°46'N$, $44°59'N$, $44°25'N$ and $43°42'N$.

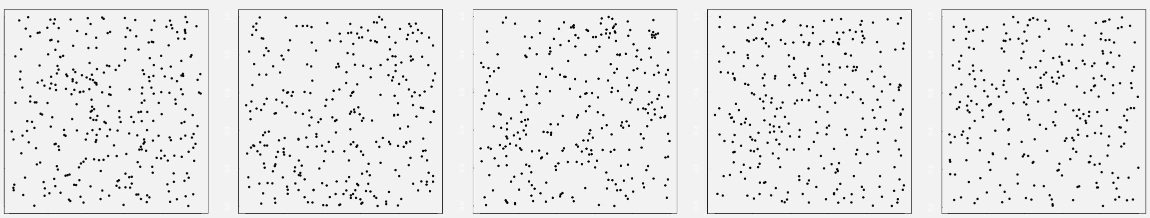
Which of the following cities do you think is furthest north?

- Geneva, Switzerland
- Lyon, France
- Minneapolis, United States
- Bucharest, Romania
- Toronto, Canada

3.2.2 Scatterplots

This domain is a perception example. The true numbers of points are, respectively, 320, 310, 300, 290 and 280. It is much clearer that there are more points in the first panel than in the fifth, than that there are more points in the first than in the second. There is an obvious similarity measure here that might be expected to lead to similarity effects.

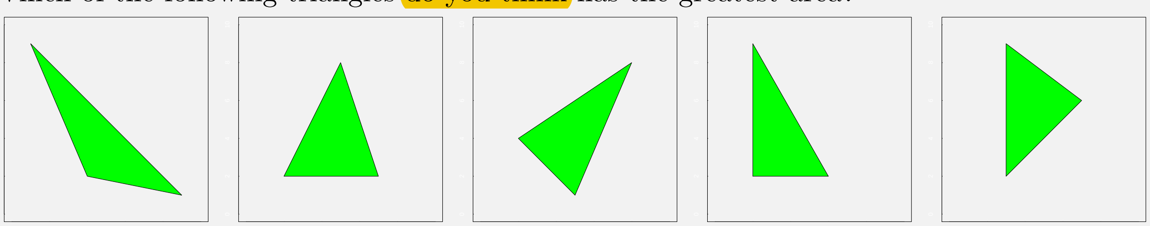
Which of the following boxes has the greatest number of points?



3.2.3 Triangles

This domain is another perception example. The true areas are, respectively, 16, 15, 15, 14 and 14 units.

Which of the following triangles do you think has the greatest area?



3.2.4 Population

These countries were ranked, respectively, 4th through 8th in terms of population in 2016, when their populations, in millions, were 258, 206, 202, 186 and 156.

Domain	Dominance	Similarity w/o dominance	Betweenness w/o similarity
Beer	$e > c$, $b > c$, $d > c$	$b \sim d$, $d \sim e$, $b \sim e$	$a\|b\|e$
Cars	$a > d$, $b > e$		$a\|c\|b$, $a\|c\|e$, $a\|c\|b$, $d\|c\|e$
Restaurants	$b > e$, $c > e$	$b \sim c$	$a\|b\|d$, $a\|c\|d$, $a\|e\|d$
Flights I		$b \sim c$, $c \sim d$, $b \sim d$	$a\|b\|e$, $a\|c\|e$, $a\|d\|e$

Table 2: Relations of dominance, similarity and betweenness in two-attribute domains

Which of the following countries do you think has the biggest population?

- Indonesia
- Brazil
- Pakistan
- Nigeria
- Bangladesh

3.2.5 Surface area

These countries are ranked, respectively, 2nd through 6th in terms of surface area, including inland bodies of water. In millions of square kilometres, those surface areas are 9.984, 9.573, 9.525, 8.516 and 7.692.

Which of the following countries do you think has the greatest surface area, including inland bodies of water?

- Canada
- United States of America
- China
- Brazil
- Australia

3.3 Objects with two attributes

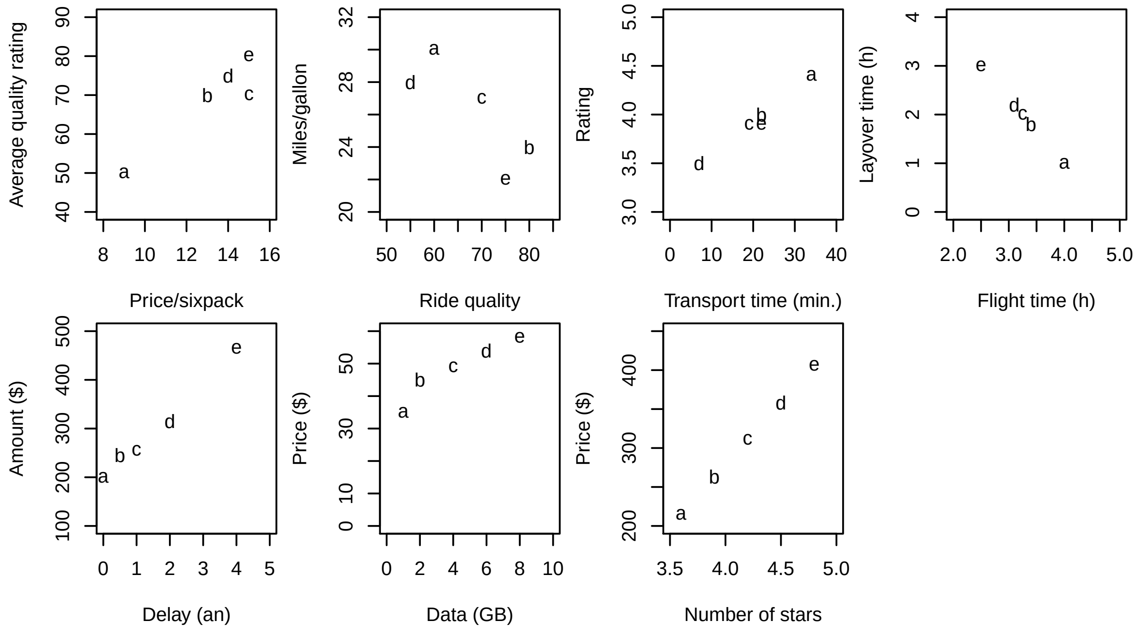
Here we describe domains whose choice objects have exactly two real-valued attributes. In all cases, the levels of these attributes are presented directly to the subject. Figure 1 plots the five objects of each of these domains in attribute space, revealing potential conditions for context effects.

Table 2 shows, for the first four of these domains, dominance, similarity and betweenness relations that we might expect to generate context effects. Following usual practice, we exclude similarity relations for two objects in which there is also a dominance relation between the two; and betweenness relations among three objects when two of them are similar and the third is not similar to either of the first two.

3.3.1 Beer

This domain is from an experiment reported in Huber, Payne, and Puto (1982) used to illustrate an asymmetric dominance effect. The prices are multiplied by 5 and we added choice objects d and e to allow for two more asymmetric dominance effects and two similarity effects. The first panel

Figure 1: Designs for context effects. Each panel plots the five objects of a domain in attribute space. The domains are, in row major order, Beer, Cars, Restaurants, Flights I, Delayed Choice, Phone Plans, Hotels



of Figure 1 show the choice objects in attribute space. Table 1 shows the relations of dominance, similarity and betweenness among objects associated with context effects. The most commonly used domains to illustrate the attraction effect involve Beer, Cars, Apartments, Computers, Restaurants and Televisions.

Below you will find three brands of beer. You know only the price per sixpack and the average quality ratings made by subjects in a blind taste test. Given that you had to choose one brand to buy on this information alone, which one would it be?

Price/sixpack	Average quality rating (100 = Best; 0 = Worst)
$9.00	50
$13.00	70
$15.00	70
$14.00	75
$15.00	80

3.3.2 Cars

This domain is based on an experiment from Wedell and Pettibone (1996). There are two experiments involving cars in that article, the experiment in question is numbered 18 in the appendix to that paper. Objects below have the same attributes as in that experiment and a similar range of levels. We adapted the objects to allow for compromise effects. The second panel of Figure 1 shows the choice objects in attribute space. Table 1 shows the relations of dominance, similarity and betweenness among objects associated with context effects.

Which of the following cars would you prefer to drive, all other features begin equal? Ride quality is a on a scale of 0 to 100.

Ride quality	Miles per gallon
60	30
80	24
70	27
55	28
75	22

3.3.3 Restaurants

This domain is based on another experiment in Wedell and Pettibone (1996), numbered 19 in the appendix to that paper. The third panel of Figure 1 shows the choice objects in attribute space. Table 2 shows the relations of dominance, similarity and betweenness among objects associated with context effects.

Which restaurant would you most like to go to for your next restaurant meal, based on transportation time (in minutes) and average customer ratings (from 1 to 5).

Transportation time	Rating
34	4.4
22	4.0
19	3.9
7	3.5
22	3.9

3.3.4 Flights I

The fourth panel of Figure 1 shows the choice objects in attribute space. Table 2 shows the relations of dominance, similarity and betweenness among objects associated with context effects.

Which of the following flight itineraries would you choose? All involve two flights, with one layover between them.

Total flight time	Layover time
4:00	1:00
3:24	1:48
3:15	2:00
3:06	2:12
2:30	3:00

3.3.5 Delayed Choice

This domain is loosely based on an experiment by Benzion, Rapoport, and Yagil (1989), in which subjects are asked to assign present values equivalent to the receipt of $200 at time horizons of 0.5, 1, 2 and 4 years. Based on implied discount factors at various terms, we constructed five choice objects designed to have approximately the same present value equivalent.

Which of the following would you rather have?

- $200 credited to your bank account immediately.
- $245 credited to your bank account in six months.
- $255 credited to your bank account in one year.
- $315 credited to your bank account in two years.
- $465 credited to your bank account in four years.

3.3.6 Phone plans

The source for this domain is the website of Fido Mobile, with rates quoted on March 1, 2017 in Canadian dollars.

If you had to choose among the following cell phone plans, which would you choose? In all cases, unlimited calling, text picture and video messages to Canadian and international mobile numbers are included. Excess data usage is billed at $10 per 500 MB.

- 1 GB data per month, $35 per month.
- 2 GB data per month, $45 per month.
- 4 GB data per month, $49 per month.
- 6 GB data per month, $54 per month.
- 8 GB data per month, $58 per month.

3.3.7 Hotel rooms

Using Expedia results, I did a linear regression of price per night on a constant and the Expedia rating, in numbers of stars, for a sample of available hotels. The five levels of numbers of stars correspond roughly to the mean, plus and minus one sample standard deviation and plus and minus two standard deviations. Prices are approximately equal to fitted values in the linear regression.

Suppose you are staying over two nights in New York city for the purposes of attending a meeting or conference. Which of the following hotels would you choose, based on customer ratings and price per night?

- 3.6/5 stars, 215$ per night
- 3.9/5 stars, 263$ per night
- 4.2/5 stars, 311$ per night
- 4.5/5 stars, 358$ per night
- 4.8/5 stars, 406$ per night

3.4 Objects with multiple attributes

3.4.1 Flights II

This domain illustrates three-way tradeoffs. The points form a constellation in the simplex that resembles the pattern of points on the “five” side of a die.

Which of the following flight itineraries would you choose? All involve two flights and have a total duration of six hours.

1st flight	Layover	2nd flight
1:30	1:15	3:15
3:15	1:15	1:30
2:15	1:30	2:15
1:30	1:45	2:45
2:45	1:45	1:30

3.4.2 Televisions

The source for this domain is the website of Best Buy Canada, with prices in Canadian dollars.

Which of the following televisions would you be most likely to buy if you were in the market for a television? All are LED televisions. Resolution refers to number of horizontal lines. Smart indicates internet connectivity.

Brand	Resolution	Smart	Price ($)	Screen Size (inches)
Sharp	1080	Yes	309	32
Insignia	720	No	209	32
Sony	720	Yes	439	32
Samsung	1080	Yes	459	40
Toshiba	1080	No	409	43

4 Results

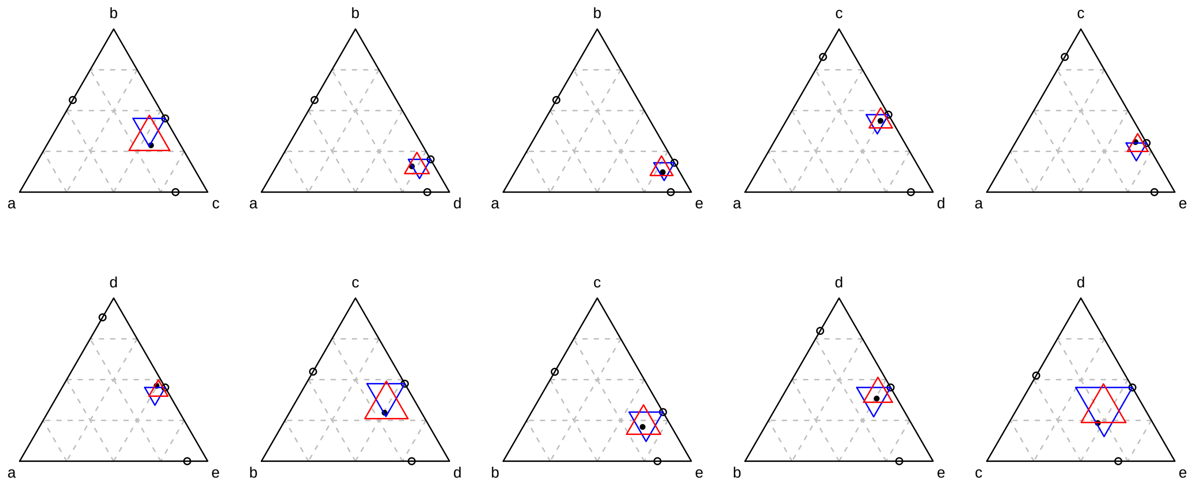
4.1 Graphical Illustrations

Figures 2 through 2 illustrate binary and ternary choice probabilities, as they relate to regularity and the multiplicative inequality, for the various domains. (Note to reader. The data from our experiment is not yet collected. Figure 2 shows data from a real experiment, where choice objects are lotteries, but not the experiment described here. We do this in order to give an idea of the figures we intend to show.) There is one figure for each domain, and each figure consists of 10 panels, with each panel corresponding to one of the ten tripleton choice sets from that domain.

Each panel features an equilateral triangle, with vertices labelled with three of the choice objects from the domain. Take the first panel as an example, where the vertices are labelled a (bottom left) b (top) and c (bottom right). Each point in the triangle, including the interior and the boundary, represents a triple $(P_A(a), P_A(b), P_A(c))$ of choice probabilities, in a Barycentric coordinate system. Thus, the distance of a point to the right side of the triangle gives $P_A(a)$, the distance to the base gives $P_A(b)$ and the distance to the left side gives $P_A(c)$. The point is also the convex combination of the vertices a, b and c—which have Barycentric coordinates $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$, respectively—with weights $P_A(a)$, $P_A(b)$ and $P_A(c)$.

The hollow dots on the left, right and bottom sides of the triangle represent the binary probabilities $p(a, b)$, $p(b, c)$ and $p(a, c)$, respectively. For example, the dot on the left side of the triangle is the convex combination of the vertices labelled a and b, with weights $p(a, b)$ and $p(b, a)$ respectively. Throughout, we reserve hollow dots for binary probabilities and solid dots for ternary probabilities, to avoid any ambiguity for points on the boundary of the triangle.

Figure 2: Binary and ternary probabilities for the lottery domain.



The blue triangle in each figure gives necessary conditions on the ternary probability, given the three binary probabilities, for regularity to hold: ternary probabilities within the blue triangle satisfy $P_A(a) \leq \min(p(a,b), p(a,c),\ P_A(b) \leq \min(p(b,a), p(b,c))$ and $P_A(c) \leq \min(p(c,a), p(c,b))$. Similarly, the red triangle gives necessary conditions on the ternary probability, give the binary probabilities, for the multiplicative inequality to hold: here the conditions are $P_A(a) \geq p(a,b)p(a,c)$, $P_A(b) \geq p(b,a), p(b,c)$ and $P_A(c) \geq p(c,a), p(c,b)$.

Figures for each domain here. Add information on WST, MST, SST, TI to diagrams.

Comment on graphical results here.

4.2 Evaluating axioms using formal model comparison tools

In this section we tests axiom by measuring how much the predictive performance of an unrestricted model of choice probabilities improves when we impose the restriction implied by the axiom. Given some observed choice data y, the Bayes factor in favour of the restricted model, relative to the unrestricted model, is the ratio of the probability of y occurring under the restricted model to the probability of y occurring under the unrestricted model. Letting Λ be the event that the axiom holds, we can express this Bayes factor as the ratio on the left hand side of equation (1).

$$\frac{\Pr[y|\Lambda]}{\Pr[y]} = \frac{\Pr[\Lambda|y]}{\Pr[y]}. \tag{1}$$

Now by Bayes rule, we obtain the equality in (1). The ratio on the right hand side is something we can estimate by simulation. It is the ratio of posterior to prior probability that the axiom holds, under the unrestricted model. High posterior probability in the numerator favours an axiom that is compatible with observed choice frequencies; low prior probability in the denominator favours an axiom that is restrictive, a kind of model parsimony.

In order to obtain prior and posterior probabilities of an axiom, we must decide on a prior distribution for P. We use a set of nine hierarchical prior distributions described in McCausland

Table 3: Log Bayes factor in favour of axioms, by domain and axiom

Domain	WST	MST	SST	TI	Reg	RU	Mul
Male stars	-	-	-	-	-	-	-
Female stars	-	-	-	-	-	-	-
Films	-	-	-	-	-	-	-
etc.							

and Marley (2013) in order to assess the robustness of our results to the prior specification. Each of these priors is symmetric with respect to choice objects and has full support on the space of random choice structures.

We compute prior and posterior probabilities of axioms numerically using the similation methods described in McCausland and Marley (2014). We compute numerical standard errors of prior and posterior probabilities using the overlapping batch mean method described in Flegal and Jones (2010) and combine these standard errors to compute standard errors for Bayes factors using the delta method.

Table 3 shows log Bayes factors in favour of the various axioms, by domain and axiom.

Comment on Bayes factors here.

5 Conclusions

Conclusions here.

A Definitions of axioms and basic results

A random choice structure satisfies

WST *weak stochastic transitivity* if and only if for all distinct x, y and z,

$$p(x,y) \geq \frac{1}{2} \text{ and } p(y,z) \geq \frac{1}{2} \Rightarrow p(x,z) \geq \frac{1}{2},$$

MST *moderate stochastic transitivity* if and only if for all distinct x, y and z,

$$p(x,y) \geq \frac{1}{2} \text{ and } p(y,z) \geq \frac{1}{2} \Rightarrow p(x,z) \geq \min[p(x,y), p(y,z)],$$

SST *strong stochastic transitivity* if and only if for all distinct x, y and z,

$$p(x,y) \geq \frac{1}{2} \text{ and } p(y,z) \geq \frac{1}{2} \Rightarrow p(x,z) \geq \max[p(x,y), p(y,z)],$$

TI the *triangle inequality* if and only if for all distinct x, y and z,

$$p(x,y) + p(y,z) + p(z,x) \geq 1,$$

Reg *regularity* if and only if for all $x \in A \subseteq B \subseteq T$,

$$P_A(x) \geq P_B(x),$$

MI the *multiplicative inequality* if and only if for all $x \in A, B \subseteq T$,

$$P_{A \cup B} \geq P_A(x) P_B(x),$$

FI the *Falmagne inequalities* if and only if for all non-empty $A \subseteq T$ and all $x \in A$,

$$\sum_{B\colon A \subseteq B \subseteq T} (-1)^{|B \setminus A|} P_B(x) \geq 0. \tag{2}$$

The multiplicative inequality is due to Sattath and Tversky (1976), who show that it is a testable implication of the Elimination by Aspects model described in Tversky (1972b) and Tversky (1972a) and the independent random utility model.

References

Bard, E. G., Robertson, D., and Sorace, A. (1996). 'Magnitude Estimation of Linguistic Acceptability', *Language*, 72(1): 32–68.

Benzion, U., Rapoport, A., and Yagil, J. (1989). 'Discount Rates Inferred from Decisions: An Experimental Study', *Management Science*, 35: 270–84.

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

Flegal, J. M., and Jones, G. L. (2010). 'Batch means and spectral variance estimators in Markov chain Monte Carlo', *Annals of Statistics*, 38(2): 1034–1070.

Huber, J., Payne, J. W., and Puto, C. (1982). 'Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis', *Journal of Consumer Research*, 9: 90–98.

McCausland, W. J., and Marley, A. A. J. (2013). 'Prior Distributions for Random Choice Structures', *Journal of Mathematical Psychology*, 57: 78–93.

——— (2014). 'Bayesian Inference and model comparison for random choice structures', *Journal of Mathematical Psychology*, 62-63: 33–46.

Sattath, S., and Tversky, A. (1976). 'Unite and Conquer: A Multiplicative Inequality for Choice Probabilities', *Econometrica*, 44: 79–89.

Simonson, I. (1989). 'Choice Based on Reasons: The Case of Attraction and Compromise Effects', *Journal of Consumer Research*, 16.

Tversky, A. (1972a). 'Choice by Elimination', *Journal of Mathematical Psychology*, 9: 341–367.

Tversky, A. (1972b). 'Elimination by Aspects: A Theory of Choice', *Psychological Review*, 79(4): 281–299.

Tversky, A., and Simonson, I. (1993). 'Context-Dependent Preferences', *Management Science*, 39: 1179–1189.

Wedell, D. H., and Pettibone, J. C. (1996). 'Using Judgments to Understand Decoy Effects in Choice', *Organization Behavior and Human Decision Processes*, 67(3): 326—344.