Working Paper Series

No. 03-2021

Roy Allen

Making sense of monkey business: Re-examining tests of animal rationality

Roy Allen
Department of Economics
University of Western Ontario
rallen46@uwo.ca

Pawel Dziewulski
Department of Economics

Since there is only a single choice in these environments, there is no di erence between examining a distribution of choices or a single choice.

While the above conceptual distinctions are important for interpretation, this paper budget

$$B(p) = \{x \in \mathbb{R}^2_+ \mid p_1x_1 + p_2x_2 \le 1\}:$$

In Figure 1, the set is represented by the budget line (i.e., the set of bundles $(x_1; x_2)$ for which the expenditure $p_1x_1 + p_2x_2$ is equal to 1) and all the points below it. Throughout the paper, we assume the two budget lines under consideration intersect.[3]

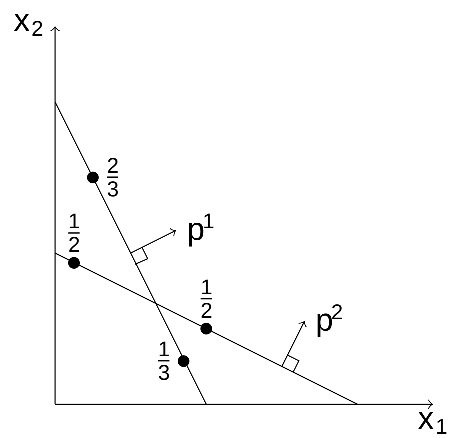


Figure 1: Two overlapping budget sets with two consumption goods (x_1 and x_2). The consumption goods for animals are often food (e.g. apples and grapes, water and flavored water, etc.). The prices (p^1 and p^2) define a trade-off rate between two goods. Choices made on each budget line are in black and the proportion each bundle is chosen is next to it.

Unlike human experiments which describe the task to the subject, nonhuman animals cannot have the task explained to them. This has led researchers to have nonhuman animals make choices from a given budget several times. For example, a nonhuman animal may be asked to trade tokens for apples and grapes several times until the choices of the animal do not vary too much. While this is an *ex-ante* sensible thing to do, it also introduces the conceptual problem of trying to discern whether an individual has a preference over the distribution of their choices.

We denote the primitive dataset of choices in the two budgets by $D = \{(p^1; x^{1;n})\}_{n=1}^{N_1} \cup \{(p^2; x^{2;m})\}_{m=1}^{N_2}$. Here p^1 is the vector of normalized prices from the first budget environment. The bundle $x^{1;n}$ is interpreted as the n-th repetition of the task with budgets defined by normalized prices p^1. The choices made from the second budget (under prices p^2) are indexed by $x^{2;m}$. The primitive dataset also admits a distributional dataset $D_D = \{(p^1; \;^1); (p^2; \;^2)\}$ where ^{1} is the sample distribution of all choices from the first budget set. For example, given the first budget with the price p ^{2}

3 Models of rationality

This section defines the three relevant models of rationality.

3.1 Standard deterministic utility maximization

Utility maximization supposes that an individual makes choices that maximize some locally nonsatiated utility function defined by $u : \mathbb{R}^2_+ \to \mathbb{R}$ over the linear budget constraint. A utility function is defined as "locally nonsatiated" when there is always a better bundle "nearby."[4] Thus, we assume that the choice of bundles is generated by

$$\max_{x \in \mathbb{R}^2_+} \quad u(x)$$
$$\text{s.t.} \quad p_1 x_1 + p_2 x_2 \leq 1.$$

This model does not involve any distribution of choices and thus we work directly with the primitive dataset D. For this model, one could only see two distinct bundles such as $x^{1;1} \neq x^{1;2}$ chosen from the same budget when the utility from the first bundle equals the

3.2 Average rationality

Average rationality is essentially deterministic utility rationalization over the average bundle chosen. This is the type of rationality that has been examined in Kagel et al. (1975), Battalio et al. (1981), Kagel et al. (1981), Battalio et al. (1985), Kagel et al. (1995), and Chen et al. (2006). Since the average is a property of the distribution of choices, we consider a model where an individual can have a preference for randomization. In

3.3 Random utility

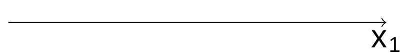
We now define the random utility maximization following Block and Marschak (1960), McFadden and Richter (1990), McFadden (2005), Hoderlein and Stoye (2015), and Kitamura and Stoye (2018). The idea behind random utility models is that an individual's choices might be governed by a distribution over different preferences. For example, one could interpret these preferences as different moods.

To represent random utility models mathematically, let $\mathcal{U}$ be the space of strictly quasiconcave locally nonsatiated utility functions $u : \mathbb{R}^2_+ \to \mathbb{R}$. A random utility model generates choices in the following manner. First, there is a probability measure μ over the space of functions $\mathcal{U}$ such that the distribution of choices $\pi \in \Delta$ satisfies

$$\pi(O) = \mu\left(\left\{u \in \mathcal{U} : \operatorname{argmax}_{x \in \mathbb{R}^2_+ | p_1 x_1 + p_2 x_2 \leq 1} u(x) \in O\right\}\right); \qquad (1)$$

for any measurable set $O \subseteq \mathbb{R}^2_+$. The argmax set is a singleton since $\mathcal{U}$ consists of strictly quasiconcave functions. In other words, the probability of choosing a bundle in the set O is equal to the probability of drawing a utility function that is maximized over the budget at some point in the set O.

The method to test random utility models significantly differs from the deterministic utility model and average rationality. The paper of Hoderlein and Stoye (2015) shows that for the random utility model, one only needs to check conditions on the probabilities of making choices in certain regions. The regions of choice are shown in Figure 2. Here $R^{1|1}$ is the region on the first budget line that lies above the second budget line in Figure 2. More generally, we refer to $R^{r|t}$ as the r-th region of the t-th budget as defined in Figure 2. Since we assume that preferences are locally nonsatiated, the only relevant regions are on the budget lines.



Hoderlein and Stoye (2015) show that random utility models can be checked by examining choice probabilities in these regions with a distributional data set D_D. A similar result holds for higher dimensional consumption with multiple budgets as shown in Kitamura and Stoye (2018). The condition essentially requires that there cannot be a large proportion of violations of deterministic rationality. For an appropriately measurable set $S \quad R^2_+$ and the probability measure $\; 2 \Delta$, we let $\;(S)$ be the probability a choice is in the set S. We record the result from Hoderlein and Stoye (2015) below.

Proposition 3. The distributional dataset $D_D = \;(p^1; \;^1); (p^2; \;^2)$ is consistent with a random utility model if and only if $\;^1(R^{1j1}) \quad \;^2(R^{1j2})$ and $\;^1(R^{3j1}) = \;^2(R^{3j2})$.

4 Statistical testing of average rationality

The method of inference we present is designed for application to the common datasets in nonhuman animal experiments mentioned throughout. Thus, we have two goods and two budget sets.[6] This empirical setting has several features that simplify statistical testing relative to a more general setup. First, prices and income are known exactly, so they do not need to be estimated. Second, there are several realized choices from each budget, which allows us to use the central limit theorem to justify a normal approximation of the means of the sampled distributions. Third, because we have two budgets, there is a

The alternative is the same as the statement in Proposition 2 and written as

$$H_a : p^1 \quad E[X^{2;m}] \quad 1 \quad \text{and}$$
$$p^2 \quad E[X^{1;n}] \quad 1 \quad \text{with one inequality strict.}$$

Recall that in the application, prices and income are measured without error, so we use lower case letters to indicate that they are nonrandom.

Thus, on the event

$$p_1 \quad \overline{X}^2 \qquad 1 \quad \hat{\ }^2 c^p \frac{}{1};$$

or equivalently

$$p_1 \quad \overline{X}^2 + \hat{\ }^2 c^p \frac{}{1} \qquad 1;$$

we have strong evidence against

$$p^1 \quad E[X^{2;m}] \qquad 1:$$

A similar argument holds when testing the other inequality. Thus, when the test function is 1, we have strong evidence against both inequalities

$$p^1 \quad E[X^{2;m}] \quad 1 \text{ and } p^2 \quad E[X^{1;n}] \quad 1;$$

i.e., strong evidence against H_0. The choice of $c^p \frac{}{1}$ comes from the fact that we have to reject two inequalities. In more detail, this threshold is motivated by

$$\text{Prob} \quad p^1 \quad E[X^{2;m}] \quad p^1 \quad \overline{X}^2 + \hat{\ }^2 c^p \frac{1c}{1}$$

$c^p_{1-} > c_{0:5} = 0$.

One may not nd the failure to reject an interesting way to di erentiate between datasets. We suggest one way to di erentiate between datasets in the spirit of the Afriat e ciency index (Afriat, 1973). Instead, one can cona8Glute the least nominal size under which the null is rejected. To do this, one can cona8Glute the largest c_{1-} under which the test rejects the data and then nd the corresponding . The construction of is related to a p-value but is distinct because there are con gurations consistent with the null hyothesis in which rejection probabilities are not asymptotically equal to nominal size .

High gives less evidence against the null. To see why this is the case, note that if is high ($> 0:75$), then c^p_{1-} is negative. Thus, this requires the data to not only satisfy the condition stated in Proosition 2 when failing to reject, but that the distance of the average bundles must also be far away from rejecting. The closer theis to one, the lower the evidence is against the null since this suggests that the mean bundles are far apart.

5 Application to Chen et al. (2006)

We apply the above procedure to the data on choices by caluchin monkeys that was previously analyzed in Chen et al. (2006). We examine whether the caluchin monkeys are rational according to deterministic utility maximization, average rationality, and random utility models. Thus, we check the conditions of the earlier proositions.

The exleriment waserformed on three di erent caluchin monkeys. We refer to the three subjects under their abbreviations: AG, FL, and NN. The consumption bundles consisted of two goods: slices of apples and gelatin cubes or grapes (depending on the subject). In the exleriment, caluchin monkeys traded tokens for food items under two di erent exchange rate regimes. In the rst regime, one token could be exchanged for one slice of apple or one gelatin cube/grape. In the second regime, two slices of apple could be exchanged for one gelatin cube/grape. Per each subject, the data generated in the exleriment consisted of multiple choices from the two budget sets. For additional information on the exleriment, we refer the reader to Chen et al. (2006).

The details and results of the test are contained in Table 1. We note that two of the three subjects from Chen et al. (2006) refute utility maximization, while all three subjects could be described by a random utility model and average rationality. We note that all monkeys satisfy average rationality so the statistical test prolosed earlier does not reject for any $< 0:75$. To try to better discern between the di erent datasets, we also report the least cona8Glu1

	Subject		
	AG	FL	NN
	Budget 1		
p^1	(1=12; 1=12)	(1=12; 1=12)	(1=12; 1=12)
N_1	12	11	6
X^1	(6:08; 5:92)	(5:64; 6:36)	(5; 7)
	Budget 2		
p^2	(1=18; 1=9)	(1=20; 1=10)	(1=20; 1=10)
N_2	22	14	10
X^2	(9; 4:5)	(13:86; 3:07)	(12:8; 3:6)
Deterministic	No	No	Yes
Random utility	Yes	Yes	Yes
Average choice	Yes		No

et al. (1981), Kagel et al. (1981), Battalio et al. (1985), and Kagel et al. (1995) evaluating average rationality can also be interpreted as evidence for a preference over distributions. This complements an approach in Natenzon (2019) that studies stochastic choice for nonhuman subjects without studying an explicit preference over distributions.

References

Afriat, S. N. (1967). The construction of utility functions from expenditure data. *International economic review 8*(1), 67{77.

Afriat, S. N. (1973). On a system of inequalities in demand analysis: An extension of the classical method. *International economic review*, 460{472.

Agranov, M., P. J. Healy, and K. Nielsen (2020). Stable randomization. Working Paper.

Agranov, M. and P. Ortoleva (2017). Stochastic choice and preferences for randomization. *Journal of Political Economy 125*(1), 40{68.

Agranov, M. and P. Ortoleva (2020). Ranges of preferences and randomization. Working Paper.

Allen, R., P. Dziewulski, and J. Rehbeck (2021). Revealed statistical consumer theory. Working paper.

Andreoni, J. and J. Miller (2002). Giving according to garp: An experimental test of the

Inte133.276 -11.955 Td[(19)1(81)0t8m7(an)1(d)-296(P)6((2)1 rg B7eldeS.276 -mleoni, J. A737{(ia
co(si)1(s:)entcory7and3t202iat1(.)-3s(g)-324(an)lt(r)1(eo,oa)1(c)27tm

Choi, S., S. Kariv, W. M

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(2), 245–264.