Série des Documents de Travail

n° 2017-52
Preferences under ignorance

O. GOSSNER[1]
C.KUZMICS[2]

[1] CNRS; École Polytechnique; LSE. E-mail : olivier.gossner@polytechnique.edu
[2] University of Graz. E-mail : christoph.kuzmics@uni-graz.at

PREFERENCES UNDER IGNORANCE

OLIVIER GOSSNER AND CHRISTOPH KUZMICS

ABSTRACT. A decision maker (DM) makes choices from different sets of alternatives. The DM is initially ignorant of the payoff associated to each alternative, and learns these payoffs only after a large number of choices have been made. We show that, in the presence of an outside option once payoffs are learned, the optimal choice rule from sets of alternatives is one that is *as if* the DM had strict preferences over all alternatives. Under this model, the DM has preferences for preferences while being ignorant of what preferences are "right".

Keywords: consistency, rationality, weak axiom of revealed preferences, strict preference
Journal of Economic Literature classification numbers: C73, D01, D11

1. INTRODUCTION

Consider the classical economic problem of an agent who has to choose between different alternatives while being uncertain about their consequences. A cornerstone approach is the expected payoff approach, which originates in the work of Pascal (1670), and according to which the agent should rank alternatives according to the expected payoff brought by each of them. Albeit uncertain about which alternatives fare better than others, the agent still forms a ranking over them. In this case, preferences stem from beliefs; at the extreme, an agent who

We are very grateful to Michael Greinecker, Ran Spiegler, Tymon Tatur, Matthew White, conference participants at the conference on the Biological Basis of Economic Behavior, Vancouver, and at the workshop on Strategic Information Acquisition and Transmission, Munich, seminar participants at Northwestern University, London School of Economics, University of Warwick, Paris Game Theory Seminar, Roy Seminar, European University Institute, Center for Rationality at Jerusalem, Tel-Aviv University, Johns Hopkins University, and the University of British Columbia as well as Rakesh Vohra and two anonymous referees for helpful comments and suggestions.

is completely ignorant about payoffs is indifferent between all alternatives.[1] In this paper, we propose a setup in which optimal behavior commands even a fully ignorant agent to act *as if* she had strict preferences over alternatives. In our model, the ranking over alternatives is not driven by beliefs, but by the principle of maximization of an option value.

To fix ideas consider the following scenario. Two friends decide to go on a diet (with the purpose to lose weight, to feel better, to feel less tired, to combat an illness, or for some such goal) which they commit to following for a pre-specified length of time. Both dieters are offered, each day, food from different menus, but they are ignorant as to what choices are good to achieve their objective. Independently of each other, they both choose a choice rule, i.e., a rule that specifies what choice to make depending on each possible menu. After a while they meet and exchange their experience. The least successful dieter can then decide to adopt the most successful dieter's choice rule. How should they choose their diet to begin with?[2]

A model that is suitable to tackle this question must have the following ingredients. First, we must have individuals making choices in a variety of decision problems. Second, there must be some potentially attractive alternatives that individuals are ex-ante uncertain about whether they are "good" or "bad" choices. In fact we can ignore all clearly inferior choices such as eating stones or drinking salt-water, and focus only on those choices that are potentially "good". Third, realistically individuals will not have to stick to one rule of behavior throughout their whole life. They could learn something about the likely success of certain rules of behavior as they go along. However, and fourth, this learning is incomplete.

The model we provide to tackle this question can be sketched as follows. A decision maker (DM) will be asked to make repeated choices from subsets of a grand set of alternatives. The DM is asked to chose a choice rule that specifies what choice she would make for every possible subset of the set of all alternatives. A choice rule can be strictly

[1]With "complete ignorance" we mean that the agent has a symmetric belief about how good each alternative is for her: each alternative is just as likely to be good as any other.

[2]This example as well as others are developed in Section 1.1.

consistent, i.e., derived from a strict rational preference relation, it can also be non-consistent in the sense of exhibiting cycles or other non-transitivities. We allow all choice rules. At the time of this choice, the DM acts under a veil of ignorance and knows nothing about the value of the various alternatives to her. Nature then randomly chooses a gain function that attaches material gains to each alternative. After some time the DM learns how well her choice rule is doing on average without learning how each alternative contributes to the overall material gain, i.e., without learning the gain function itself. The DM can then stick to her chosen rule and obtain the resulting average material payoff or adopt an outside option, the value of which is chosen randomly and is possibly correlated with the gain function. The outside option captures any form of outside opportunity to the DM. In particular, it encompasses a reduced form model of the possibility of (incomplete) social learning.

We show that, in order to maximize expected gains, a choice rule must be strictly consistent. Moreover we identify conditions under which all strictly consistent rules, and only those rules, are equally optimal.

The argument for this claim is as follows. Using ignorance, it is easy to show that in such an environment all choice rules produce the same expected material gain. We show, and this is the crucial result, that strictly consistent rules are in some sense the most risky rules. To be more precise we note that any choice rule induces a probability distribution over material gains. We prove that for any non-strictly consistent choice rule there is a distribution over strictly consistent choice rules that induces a distribution over material gains, for any realized gain function, that is a strict mean preserving spread over the distribution of material gains induced by the given choice rule.

The DM will then strictly prefer this distribution over strictly consistent rules over the given choice rule, because increasing risk increases the value of the outside option.[3] Thus, for any non-strictly consistent choice rule, the DM will find a strictly consistent choice rule that she prefers strictly over the given choice rule.

[3]This is the same logic as in finance, e.g. in Merton (1973) and Rasmusen (2007), where increasing risk increases option value.

The paper proceeds as follows. In Section 1.1 we discuss a series of applications of our model. In Section 2 we provide the model. In Section 3 we state the main theorem and sketch its proof, in the course of which we establish two additional results that are of interest in their own right. Section 4 provides a discussion of the exact role the assumptions play for the various results. In section 5 we use a simple example with two purposes. First, it should help the reader to understand both the model and its workings. Second, it demonstrates the boundaries of our results by highlighting what is not true in this model. Section 6 finally discusses a few possible extensions of the model.

1.1. **Applications.** Our model relies on several major assumptions. The first of them is the decision maker's ex-ante ignorance as to what payoffs are associated with a list of potential items of choice. The second one is that the decision maker faces a number of decisions problems. For convenience, this number is assumed to be infinite in our main model, but it can as well be taken as large yet finite or, if the DM perfectly anticipates all decision problems, this number can even be relatively small, see Section 6. As a third assumption, the decision maker doesn't learn the payoffs associated to the different items available as different choices are made, or if she does learn then she does not or cannot use this information immediately. This assumption is in fact realistic in set-ups where all payoff realizations and information occur at a stage following decision making as in two of the four applications below. It is also realistic in set-ups where changing the choice rule is more costly than the value of information from the feedback from a single decision problem as in the other two applications below. And finally, the decision maker may have access to an outside option which can, for instance, take the form of a switch of rule or some insurance policy.

1.1.1. *Dieting.* According to the Boston Medical Group "[a]n estimated 45 million US Americans go on a diet each year", mostly with the desire to look better, to lose weight, or to be more healthy.

Although there is a growing consensus among nutritionists on the combination of diet, exercise and lifestyle that is best for goals of losing weight and being healthy, the number of diets people have tried and are still trying is almost endless.[4] Dieting is also very profitable for the dieting industry, with estimated revenues of $ 64 billion in 2014.

How is the observed diversity of diets consistent with our model assumptions?

First, scientific evidence on what diets work best is sometimes inconclusive, and poorly disseminated to the public. Hence the veil of ignorance. Second, the number of food decisions faced by an individual is large, typically between 3 and 5 per day. Third, the variance in weight measurements due, for instance, to different levels of hydration or medical conditions render the appreciation of a diet's efficiency difficult to assess in the short-run, and other diet-related medical conditions such as muscle loss may take a long time to detect. Overall, the long-run effect of one's diet on weight and health is difficult to assess in the short-run. Finally, dieters have outside options. If at the end of their diet they are not very successful they can adopt another diet, the diet of a more successful friend, or their old diet.

Our results suggest that the diversity of diets, all of which recommend consistent choices over foods, may actually be driven by these features. They predict that chosen diets may well be imperfect or maybe even detrimental for their goal. Thus, if we were to observe a dieters food choices and if we were trying to infer her "preferences" from her food choices, we would not recover her true preferences.

1.1.2. *Farming.* How does a farmer choose what seeds to plant and how to cultivate her seeds? Farmers have a huge variety of seeds they could grow and each piece of land may be suitable only for a subset of them. Before the beginning of the year, many uncertainties are unresolved, in particular it is unknown what yields each type of crop would generate as this depends on future climatic conditions. Also the future price of each crop is still uncertain. Insurance, when available, opens the possibility of an outside option to the farmer. Absent such insurance, a risk averse farmer would prefer planting a variety of seeds to minimize risk. However, when insurance is available, and even under uncertain future conditions, our results show that a not too risk averse farmer should plant as if having a ranking over crops: She will choose the "best" crop according to this ranking in all fields that allow it and

[4] see e.g. `https://en.wikipedia.org/wiki/List_of_diets`

choose the "second best" crop in all fields that do not allow the "best" crop but allow the "second best" and so on.

1.1.3. *Education.* What career should a young person pursue? How should this person best prepare herself for the future job market she will be competing in? Many years before this person will have a job and career, this person has to make a large array of decisions about which (learning) activities to choose from a large set of possible activities. Her goal is to maximize her eventual satisfaction in life with an appropriate career path. But at the time of decision making she is very uncertain about many aspects of this problem. She is uncertain about her own skills. Perhaps she is an apparently talented musician, but is it really her comparative advantage? Also future market conditions are uncertain. She could specialize in internet technology, but perhaps when she is finished with her education market conditions are such that this doesn't commands a particularly high salary. So her question is this. Should she try to specialize as much as possible (that is, behave as if she had consistent preferences over learning activities) or should she try to become a generalist?

Among the many decisions the agent has to make are which courses to follow, where the payoff associated to these is realized only at a later stage, sometime after the curriculum is completed. The outside option could be to start a new course from scratch or to work in the parents' business.

Our main result implies that, even if she does not know what career is ultimately the most rewarding, she should consistently pursue one specific career path (if her outside option is reasonable and she is not too risk averse) and make consistent choices. Even if the realized outcomes may be bad in some cases, on average, she will maximize expected payoffs this way.

1.1.4. *Medical Treatment.* There are many (smaller) ailments for which the medical profession has not yet found the perfect cure. For instance, back problems apparently can come in many forms and for many possible reasons. Furthermore a wide variety of treatment options is available and it is still very unclear which treatment option is best in which situation. Should it be surgery, osteopathic maneuvers, acupuncture, physical exercise, or a combination of all of these? And suppose the answer is that it should by physical exercise, what kind of physical exercise should it be? Walking, running, swimming, yoga?

As we mentioned, knowledge of the best medical treatment for each condition is not known. Furthermore, learning from the success or failure of treatment from each single patient is surely severely limited, the strong variance on the effect of each treatment on each patient tends to render learning difficult. Good feedback is sometimes also limited if the doctor does not see every patient again after prescribing treatment. One outside option that a doctor can adopt after following her choice rule of treatments for some time, is to adopt another choice rule of treatments.

Our results suggest that it may be optimal for a doctor to behave as if she has strict preferences over treatment options, even if her preferences are (or turn out to be) wrong. This could in this context also be interpreted as the doctor having a firm belief, correct or not, over the efficacy of the various treatments.

2. Model

2.1. **Choice.** Our set-up is based on the classical model of choice from choice sets. Let $K = \{1, ..., |K|\}$, $|K| > 1$ be the set of all possible **alternatives**. Let $\mathscr{L} = \mathscr{P}(K) \backslash \emptyset$ denote the set of all non-empty subsets of K. We call an element in $\mathscr{L}$ a **choice set**. A decision maker is repeatedly asked to make a choice from different choice sets.

Definition 1. *A* **choice rule** *is a function $R : \mathscr{L} \to \mathscr{L}$ such that $R(L) \subseteq L$ for all $L \in \mathscr{L}$. Let $\mathscr{R}$ denote the set of all such choice rules.*

Following Uzawa (1956) and Arrow (1959) (see also Chapter 1.B in Mas-Collel, Whinston, and Green (1995)), let $\succeq$ denote a binary (preference) relation over elements in K with the interpretation that when $i \succeq j$ an individual holding this preference relation weakly prefers i over j. The relation $\succeq$ is complete if for any two $i, j \in K$, $i \succeq j$ or $j \succeq i$ (or both), it is transitive if $i \succeq j$ and $j \succeq k$ imply $i \succeq k$. A complete and transitive relation is called **consistent** (often also termed "rational", see e.g. Definition 1.B.1 in Mas-Collel, Whinston, and Green (1995)). In this paper a special case of consistent preferences plays a prominent role, namely, strict preferences.

A relation $\succeq$ is anti-symmetric if whenever $i \succeq j$ and $j \succeq i$ then $i = j$. We call a preference relation **strictly consistent** if it satisfies completeness, transitivity, and anti-symmetry.

These definitions extend from preference relations to the corresponding individual's behavior.

Definition 2. *A choice rule $R \in \mathscr{R}$ is* **consistent** *if there exists a complete and transitive preference relation $\succeq$ such that, for every L, $R(L)$ is the set of maximal elements in L for $\succeq$. It is* **strictly consistent** *if it is consistent and $R(L)$ is a singleton for all $L \in \mathscr{L}$. Let $\mathscr{R}^s$ denote the set of strictly consistent rules.*

It is easily verified that a strictly consistent rule is one based on a strictly consistent preference relation.

2.2. **The environment.** An environment consists of two components. First, nature chooses a (material) gain function that associates gain levels to possible choices. It is useful to consider a fixed finite set of gain levels $G \subset \mathbb{R}_+$. A **gain function** $g \colon K \to G$ is then a function from the set of all possible choices to this set of possible gain levels, with the interpretation that $g(k) \in G$ is the gain an individual receives when choosing $k \in K$.

We extend any gain function to the set $\mathscr{L}$ of choice sets by setting

$$g(L) = \frac{1}{|L|} \sum_{k \in L} g(k)$$

for $L \in \mathscr{L}$, with the natural interpretation that $g(L)$ is the expected gain for the decision maker when L is the set of accepted alternatives, thus assuming that each element in L is then chosen by the decision maker with equal probability.[5]

Second, a **distribution over choice sets** $p \in \Delta(\mathscr{L})$ describes the frequency with which choice sets are presented to the decision maker. We assume that enough choice sets are available with positive frequency, thus making the assumption that p has full support over the non-singleton subsets of $\mathscr{L}$.

[5]This is an innocuous assumption, which, however, provides us with the property that the set of all decision rules is finite. The key lemma below, Lemma 2, extends to all stochastic choice models.

In some cases, it is useful to consider **neutral** distributions, for which all alternatives play the same role.

Definition 3. *A distribution* p *over choice sets is* **neutral** *if, for every permutation* π *of* K*, and every choice set* $L \subseteq K$*,* $p(L) = p(\pi(L))$*.*

Obviously, the uniform distribution is neutral. Other examples of neutral distributions over choice sets are the uniform distributions over choice sets of fixed size l, for $1 \leq l \leq |K|$.

Given a gain function g and a distribution of choice sets p, the (average) material gain of any rule $R \in \mathscr{R}$ is computed as:

$$g_p(R) = \mathbb{E}_p g(R(L)) = \sum_{L \in \mathscr{L}} p(L) g(R(L)).$$

Let $\mathscr{G}$ be a finite set of gain functions, and let $q \in \Delta(\mathscr{G})$ be a distribution over gain functions. For a permutation $\pi : K \to K$ and a gain function $g : K \to G$, we let $g^\pi : K \to G$ be the *permutation of* g defined by $g^\pi(k) = g(\pi(k))$ for all $k \in K$.

Definition 4. *A distribution over gain functions,* $q \in \Delta(\mathscr{G})$*, is* **symmetric** *if* $g^\pi \in \mathscr{G}$ *and* $q(g) = q(g^\pi)$ *for every gain function* $g \in \mathscr{G}$ *and for every permutation* $\pi : K \to K$*.*

The interpretation of the distribution q is that it is the decision maker's belief as to the likelihood of different gain functions. In what follows, we assume that the distribution q over gain functions is symmetric and that its support contains at least one non-constant gain function.

2.3. **Outside options.** After observing the "average" material payoff corresponding to the rule R, the decision maker may either stick to the induced material payoff, or switch to an outside option with material gain $\mathcal{g}$. The value $\mathcal{g}$ is random and its distribution can depend on the realized gain function g. The realized value of $\mathcal{g}$ is observed by the decision maker after she learns the average material payoff induced by her chosen rule. We assume that $\mathcal{g}$, conditional on any gain function g, has a positive density in the interval $[\min G, \max G]$. This assumption excludes the trivial cases in which $\mathcal{g}$ is either smaller than $\min G$ with probability 1 and the outside option is never chosen, as well as the case in which it is larger than $\max G$ with probability 1 and the outside option is always selected. Note however that it encompasses situations in which the outside option is available with positive probability only, as they are captured by distributions of $\mathscr{g}$ that put positive probability on values less than $\min G$. For some results we require the additional assumption that $\mathscr{g}$ is statistically independent of the distribution of the gain function. We indicate this when this is the case.

2.4. **The decision maker's problem.** The decision maker (DM) knows the set of alternatives K, the distribution p of choice sets, the distribution q of gain functions, as well as the distribution of the outside option $\mathscr{g}$ conditional on any gain function g.The timing of the decision problem is as follows. First, the DM chooses a rule in $\mathscr{R}$. Then nature chooses a gain function according to q. This gain function is not known to the DM at this time. The DM makes choices according to her chosen rule in every choice set L which she faces with frequency $p(L)$. The DM then learns the average realized gain $g_p(R)$. The outside option value $\mathscr{g}$ is realized and is observed by the DM, who can then choose the maximum of this average realized gain and $\mathscr{g}$.[6] In short, the DM chooses a rule $R \in \mathscr{R}$ in order to maximize her ex ante expected gain

$$\mathbb{E}_{q,\mathscr{g}}\left[\max\left\{g_p(R), \mathscr{g}\right\} \mid g\right].$$

The timing of events in the model is described in Table 1.

0	DM chooses rule R
1	gain function g and average gain $g_p(R)$ realizes
2	outside option $\mathscr{g}$ realizes
3	DM receives material gain $\hat{g} = \max\{g_p(R), \mathscr{g}\}$

TABLE 1. Timeline of events

3. Results

In this section we first state the main result, Theorem 1, and then sketch its proof by providing an intermediate result that is of interest in its own right, Theorem 2. The full proofs of all results are given in the appendix.

[6]For simplification, we abstain from considering Bayesian inferences on the gain function drawn from the observation of $g_p(R)$. We show how such inferences can be included to the analysis in section 6.1.

3.1. **Optimal choice.** The main result of this paper is the following theorem.

Theorem 1.

(1) *For every p, every optimal rule is strictly consistent.*
(2) *If p is neutral and the outside option is statistically independent of the distribution of the gain function then every strictly consistent rule is optimal.*

First we note that, given the assumption that q is symmetric, all choice rules yield the exact same ex ante expected gain. In other words, absent an outside option, all rules are equally good.

Lemma 1. *Let $R, R' \in \mathscr{R}$ be arbitrary decision rules. Then*

$$\mathbb{E}_q g_p(R) = \mathbb{E}_q g_p(R').$$

If all the rules give the same expected gain, they can still differ in the level of risk they provide.

Let R, R' be two rules. We say that R is strictly riskier than R' if the distribution of $g_p(R)$ under q is a strict mean-preserving spread of the distribution $g_p(R')$ under q. One distribution is a strict mean-preserving spread of another if it is a mean-preserving spread of and not identical to the other. If μ is a distribution over rules and R' is a rule, we say that μ is strictly riskier than R' if the distribution of $g_p(R)$ under q and μ is a strict mean-preserving spread of the distribution $g_p(R)$ under q.

The following result shows that the strictly consistent rules maximize risk in an unambiguous sense.

Theorem 2. *Let R be any non strictly consistent rule. There exists a distribution μ over strictly consistent rules such that μ is strictly riskier than R. If p is neutral, then every strictly consistent rule is strictly riskier than any non strictly consistent rule.*

By force of this theorem, the DM, when considering a non strictly consistent rule, will always find a distribution over strictly consistent rules (a mixed strategy putting weight only on strictly consistent rules) that she strictly prefers over the given rule. To complete the argument of point 1) of Theorem 1, we note that, as the DM strictly prefers this distribution over strictly consistent rules over the given rule, she must also strictly prefer one of these strictly consistent rules over the given rule.

To show point 2) of Theorem 1 we use the fact that, under the given assumptions, all strictly consistent rules are equivalent.

We have thus explained how Theorem 2 can be used to prove the main result, Theorem 1. The proof of Theorem 2, identifying how rules can be partially ordered by the mean-preserving spread order, rests on a key lemma, which we establish in the next subsection.

3.2. **Choice rules and choice distributions.** A key to a better understanding a choice rule's performance in the decision maker's problem is to consider the probability distribution over choices in K induced by this choice rule and by the distribution over choice sets. Given the distribution p over choice sets and a choice rule R, let $\lambda_p(R)(k)$ denote the overall probability with which an element $k \in K$ is selected under the rule R. it is given by:

$$\lambda_p(R)(k) = \sum_{L:k \in R(L)} \frac{p(L)}{|R(L)|}.$$

We call $\lambda_p(R)$ the **choice distribution** associated to R. This choice distribution summarizes the frequency with which each item in K is selected by R. This distribution is known to the agent. For a fixed g, a rule's average payoff is entirely determined by its choice distribution, through the following relation:

$$g_p(R) = \sum_k \lambda_p(R)(k) g(k).$$

For g unknown, the distribution of payoffs induced by R and g is entirely determined by $\lambda_p(R)$ and by the distribution of g. As we shall see, it is useful to think of the choice distribution induced by her rule as the object of choice for the agent.

For a given distribution p over choice sets, let Λ_p denote the set of all choice distributions available to the agent, i.e.,

$$\Lambda_p = \{\lambda_p(R), R \in \mathscr{R}\}.$$

Similarly, denote by Λ_p^s the subset of Λ_p consisting of distributions induced by strictly consistent rules, i.e.,

$$\Lambda_p^s = \{\lambda_p(R), R \in \mathscr{R}^s\}.$$

The following result locates the choice distributions induced by consistent rules as extreme points in the set of choice distributions. It shows that the extreme points of the convex hull of Λ_p consists of points in Λ_p^s only.

Lemma 2. *Every choice distribution in Λ_p is a convex combination of choice distributions in Λ_p^s.*

This lemma provides the key insight needed to prove Theorem 2 by establishing that the strictly consistent rules are, in the sense of the statement of the Theorem, the most risky. This lemma is proven in appendix A. We here provide an intuition for this result.

Consider a strictly consistent rule R^s that ranks alternatives in decreasing order $k_1, \ldots, k_K$. Such a rule maximizes the frequency of its preferred item k_1 among all rules. But this is not necessarily the only one with this property, since every rule, strictly consistent or not, that chooses k_1 whenever it is available, does the same. But, among all rules maximizing the probability of choosing k_1, R^s maximizes the frequency of k_2, and so on. This argument shows that every strictly consistent rule induces an extreme point in the set of achievable choice distributions. In order to show the converse property, i.e., that every extreme choice distribution is induced by a strictly consistent rule, it is important to remember that these extreme points are those which, among all in Λ_p, are the most extreme according to some direction, i.e., maximize some linear functional of the form $\sum_{k \in K} \alpha_k \lambda_p(k)$. Why is it that maximizing a linear functional is always achieved by a strictly consistent rule? Let us consider an agent who associates utility α_k with choice k. For this agent, a choice of rule R carries an expected utility $\sum_{k \in K} \alpha_k \lambda_p(R)(k)$. It should be quite intuitive that a rule that maximizes this expected payoff is one that chooses items in decreasing order of utilities (coefficients α_k) where ties in these utilities can be broken in any arbitrary way. Hence, an extreme point in the direction of the coefficients α_k can be achieved by a strictly consistent rule. Since this is true of all possible coefficients, all the extreme points are achieved by strictly consistent rules.

To visualize the sets Λ_p and Λ_p^s in an example, consider the example depicted in Figure 1. In this example the choice distributions given by strictly consistent rules are depicted by solid dots, whereas choice distributions of other non-strictly consistent singleton rules are depicted by hollow circles and squares. Note the extreme position of the choice distributions of strictly consistent rules within the set of all choice distributions.

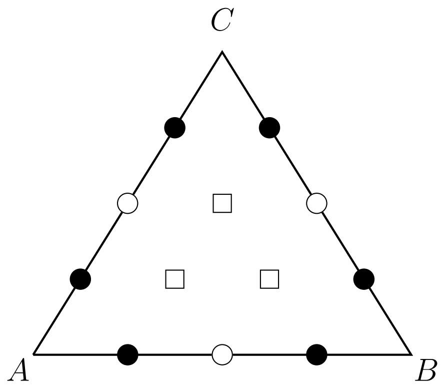

FIGURE 1. The set of choice distributions Λ_p for $K = \{A, B, C\}$ and p (neutral) such that $p(\{A, B, C\}) = p(\{A, B\}) = p(\{A, C\}) = p(\{B, C\}) = \frac{1}{4}$. All singleton rules are depicted. Solid circles represent the choice distributions that correspond to the strictly consistent rules.

4. Discussion of the assumptions

Here we briefly discuss the role played by the different assumptions in our main results. We first argue that the assumptions that p and q have full supports are not important and relaxing these changes the results only slightly. We then discuss why some results require the assumption that p is neutral and how the results change if q is not symmetric.

4.1. **The full support assumption for the distribution over choice sets.** Section 2 assumes that p has full support over non-singleton choice sets. Now suppose that p does not have full support. Note first that the conclusions of Lemmas 1 and 2 still hold. The conclusions of Theorem 2 are slightly modified: it is true that for any non strictly consistent rule there is a distribution over strictly consistent rules that yields a mean-preserving spread in terms of distributions of gains, but this spread does not have to be strict. The statement of Theorem 1 needs to be adapted. There still exists an optimal rule that is strictly consistent. This rule, however, is not unique when p does not have full support, since choices outside the support of p do not affect payoffs, thus are irrelevant. In this case, it can be shown that all optimal rules must coincide with a strictly consistent rule on the support of p.

4.2. **The full support assumption for the distribution of the outside option.** We also assumed that the outside option $\not{g}$, conditional on any gain function g, has full support over a sufficiently large interval. Note that this assumption is only relevant for Theorem 1. Relaxing this assumption changes the conclusion of Theorem 1 in the same way as relaxing the assumption that p has full support does: there exists a strictly consistent optimal rule, but not only strictly consistent rules may be optimal. To see this, observe for instance that if $\not{g}$ takes only values outside of the range of $g_p(R)$, all rules yield the same payoff hence are optimal.

4.3. **Statistically independent outside option.** The second part of Theorem 1 relies on the assumption that the outside option is statistically independent of the realized gain function. To see that this assumption is needed for this result consider the simple example with $K = \{a, b\}$ and two equally likely gain functions g^a, g^b such that $g^a(a) = g^b(b) = 1$ and $g^a(b) = g^b(a) = 0$. Let the outside option conditional on g^a have a distribution with probability close to one for values close to 0. Let the outside option conditional on g^b have a distribution with probability close to one for values close to 1. Let finally p be such that $p(\{a, b\}) = 1$.

In this case, the strictly consistent rule R^a that ranks $a \succ b$ is superior to the strictly consistent rule R^b that ranks $b \succ a$. The rule R^a achieves a payoff of 1 if g^a realizes (with the outside option not taken) and a payoff close to 1 also if g^b realizes (because of the outside option). The rule R^b, on the other hand, achieves a payoff of 1 when g^b realizes and a payoff of close to 0 when g^a realizes. Note that this is essentially the situation a DM would be in in our dieting example if she knows her friend chose rule R^b (which then serves as the outside option). Then she should choose rule R^a. Note also, that if she believes her friend chose one of the two rules with equal probability, then both rules are equally good for her.

4.4. **Non-neutral distributions over choice sets.** The most interesting implication of p non-neutral is the role p plays in Theorem 2 and in Theorem 1. The example in section 5 below shows that, for p non-neutral, yet q symmetric, it is not the case that all strictly consistent rules are most risky and that all strictly consistent rules are equally good and optimal. Also, different p's imply different most risky rules (even keeping q the same).

4.5. **Non-symmetric distribution over gain functions.** We finally turn to the two assumptions made on q. Assuming that there is at least one non-constant gain function in the support of q only avoids that the model is trivial. The second assumption, that q is symmetric, makes the model interesting by assuming the decision maker has a veil of ignorance. We believe that it is under this condition that results showing the optimality of strictly consistent rules are the most striking. Nevertheless, it is still interesting to examine the implications of an asymmetric q. The first observation in this case is that the conclusion of Lemma 1 does generally not hold if q is not symmetric. In this case (for instance in the trivial case in which q is supported by one payoff function only), some rules can provide a higher expected gain than others. Interestingly, however, under the presence of an outside option, the optimal rule is not generally the rule that maximizes the expected gain under the most likely gain function under q, as we show in section 5 below.

It is still true, however, that even if q is non-symmetric, if p and q have full support, the optimal rule (as in Theorem 1) is strictly consistent. The proof requires little adaptation. The key argument is the following. By Lemma 2, for every non strictly consistent rule, there exists a distribution over strictly consistent rules (as in Theorem 2) that produces a strict mean preserving spread in terms of choice distributions. This distribution also provides a strict mean preserving spread of payoffs for every q. Thus, the DM will, for any q, prefer this distribution of strictly consistent rules over the given non strictly consistent rule. Hence, at least one of these strictly consistent rules provides a higher expected payoff than the non strictly consistent rule. Which of the strictly consistent rules is optimal can then depend on q and the distribution of the outside option $\underline{q}$.

5. An example

We study an example in detail, showing in particular how the optimal choice rules can depend on the data of the problem when p is not neutral. We assume here that the distribution of the outside option $\underline{q}$ is statistically independent of the gain function.

Let $K = \{a, b, c\}$, and p be given by: $p(\{a, b\}) = p(\{a, c\}) = \frac{1}{4}$, and $p(\{b, c\}) = p(\{b\}) = p(\{c\}) = \frac{1}{8}$, $p(\{a, b, c\}) = p(\{a\}) = \frac{1}{16}$. Note that b and c are symmetrically treated in p, but that p is not neutral.

Given the symmetries in the setup, there are, without loss of generality, only three strictly consistent rules with potentially different payoff distributions. The strict preferences corresponding to these rules are:

$$R^a \quad a \succ b \succ c$$
$$R^b \quad b \succ c \succ a$$
$$R^c \quad c \succ a \succ b$$

Their corresponding choice distributions are $\lambda(R^a) = \frac{5}{8}a + \frac{1}{4}b + \frac{1}{8}c$, $\lambda(R^b) = \frac{1}{16}a + \frac{9}{16}b + \frac{3}{8}c$, and $\lambda(R^c) = \frac{5}{16}a + \frac{1}{8}b + \frac{9}{16}c$.

Let us consider gain functions that attach gain 1 to one element in K and 0 to the other two, and q the uniform distribution over these three gain functions. The payoff distributions of the strictly consistent rules under q are given in the following table (one • represents a probability weight of $\frac{1}{3}$).

R	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{3}{16}$	$\frac{4}{16}$	$\frac{5}{16}$	$\frac{6}{16}$	$\frac{7}{16}$	$\frac{8}{16}$	$\frac{9}{16}$	$\frac{10}{16}$
R^a		•		•						•
R^b	•					•			•	
R^c		•			•				•	

It can be seen that the payoff distributions of R^a and R^b are mean preserving spreads of the payoff distribution of R^c, but that neither the payoff distribution of R^a nor R^b is a mean preserving spread of the other. It follows that it is always the case that one of the two rules R^a or R^b is optimal.

We now show that which of R^a or R^b is optimal depends on the distribution of outside options. First consider a distribution of $\mathscr{q}$ with full support that puts high probability on some value $x \in \left(\frac{9}{16}, \frac{10}{16}\right)$, and for simplification think of the limit case in which the distribution puts probability 1 on x. Under R^b, the option is always chosen, hence the expected payoff is x, while under R^a the option is chosen with probability $\frac{2}{3}$ and the expected payoff is $\frac{1}{3}\frac{10}{16} + \frac{2}{3}x > x$. The option value is maximal under R^a which is then the only optimal rule. On the other hand, if the distribution of $\mathscr{q}$ puts high probability (think of it as being 1) on some value $x \in \left(\frac{1}{16}, \frac{2}{16}\right)$, the option is never chosen under R^a, which then yields an expected payoff of $\frac{1}{3}$, while it is chosen with probability $\frac{1}{3}$ under R^b which yields an expected payoff of $\frac{1}{3}x + \frac{1}{3}\frac{6}{16} + \frac{1}{3}\frac{9}{16} > \frac{1}{3}$. Hence in this second case the option value is maximal under R^b which is now the only optimal rule.

Note that ex ante all elements of K have the same chance of being the best choice. Nevertheless it is not true that all (strictly consistent) rules are equally good. Together with our result for p neutral and q symmetric, this implies that p has a subtle effect on which rules are good and which are bad. The optimal rule depends on p (just as much) as on q.

Now consider the same example but with a slightly different distribution over gain functions, denoted q'. Let q' be such that it is derived from q by taking a small $\epsilon > 0$ probability weight from all gain functions other than the $g^a \in \mathscr{G}$ with $g^a(a) = 1$ and $g^a(b) = g^a(c) = 0$ and move that total probability mass to that gain function g^a. Thus, g^a is the most likely gain function under q'. Let $\mathscr{q}$ put high probability on some value $x \in \left(\frac{1}{16}, \frac{2}{16}\right)$. Then for sufficiently small ϵ, rule R^b is strictly better than R^a, even though R^a is the unique optimal rule for gain function g^a. Thus, even if one gain function is more likely than all others, the strictly consistent rule associated with this gain function may not be optimal.

6. Extensions

We point out general robustness properties of our results. We first discuss Bayesian learning at the outside option stage. Then, we show that Theorems 1 and 2 are robust to several variations of the model.

6.1. **Bayesian updating.** Within our model an individual could, in principle, use Bayesian updating given her prior belief about the distribution over all environments and her average material payoff, in order to form a new and more informative belief about the environment she is facing. The main model of the paper rules out this possibility, and we show how the analysis extends to a model that permits such updating.

Assume the material payoff a gain function produces is the sum of two components. The first component is as modeled in Section 2. The second is a constant additive term, added to all payoffs in all decision problems irrespective of the chosen choice rule. This second term is highly uncertain (at least in the mind of the DM), and follows a uniform distribution in the interval $[-x, x]$ for $x > 0$. As x tends to infinity, Bayesian updating from the observed average material payoff provides no information about the realized gain function. Therefore, the extended model with Bayesian updating yields the same analysis and results as our original model.

6.2. **Costly experimentation and impatience.** The model studied so far considers that if the outside option is chosen, then the resulting utility is the one corresponding to the outside option's gain. This means that experimentation of a rule R is costless in the sense that when the outside option is chosen, the payoff generated by R is irrelevant. We can instead consider that experimentation is costly in the following sense. The payoff from R materializes in a first stage, and the agent obtains this payoff. Then in a second stage the agent may decide to switch to the outside option, or not. The agent has a discount factor of $0 < \delta < 1$, meaning that the objective is to maximize $(1-\delta)$ times the expected payoff in the first period plus δ times the expected payoff in the second period. The agent's problem then becomes to maximize over all rules R the total expected payoff:

$$\mathbb{E}_q \left[\mathbb{E}_{\mathscr{g}} \left[\max\{g(R), (1-\delta)g(R) + \delta \mathscr{g}\} \mid g \right] \right].$$

Note that $\max\{g(R),(1-\delta)g(R)+\delta\not{g}\} = (1-\delta)g(R)+\delta\max\{g(R),\not{g}\}$. Thus, the DM's objective is to maximize

$$(1-\delta)\mathbb{E}_q\left[g(R)\right] + \delta\mathbb{E}_q\left[\mathbb{E}_{\not{g}}\left[\max\{g(R),\not{g}\} \mid g\right]\right].$$

This new objective function differs from the one before only by the additional first term.

By Lemma 1 all rules yield the same expected gain. The first term in the objective function is thus irrelevant. Optimality is decided solely by the second term. This second term, however, coincides with the original objective function. Therefore, all conclusions of Theorem 1 remain valid.

6.3. **Finite sampling.** In our main model, we consider that the agent observes the expected payoff $g_p(R) = \mathbb{E}_p g(R(L))$ before deciding whether to use the rule R or take the outside option. The payoff $g_p(R)$ can be understood as the average of $g(R(L))$ over an infinite sequence of realizations of the choice set L according to p. Now consider a variation of the model in which the agent gets to observe the average payoff $\frac{1}{n}\sum_t g(R(L_t))$ over a finite and iid. sequence with law p of choice sets $L_1,\ldots,L_n$ before deciding to take the outside option or not.

In the modified model, the choice as to whether to switch to an outside option or not depends on a subtle Bayesian updating after observing $\frac{1}{n}\sum_t g(R(L_t))$. Still, the DM can use the following rule: switch to $\not{g}$ if and only if $\not{g} > \frac{1}{n}\sum_t g(R(L_t))$. Since by the law of large numbers, $\frac{1}{n}\sum_t g(R(L_t))$ converges almost surely to $\mathbb{E}_p g(R(L))$ when n becomes large, this switching rule yields an expected payoff going to $\max\{g_q(R),\not{g}\}$ when n becomes large. This implies that the choice of a rule in the modified problem gives an expected payoff that becomes arbitrarily close to the payoff in the original problem. Therefore, whenever all optimal rules are strictly consistent in the original model, the same remains true with finite sampling, for n large enough.

Note finally that the result of this section extends to any model in which $g_p(R)$ is observed with noise as long as the noise is small enough.

6.4. **Finite number of decision problems.** Another interpretation of our model is that the DM faces a finite series of choice sets $L \in \mathscr{L}$, where $p \in \Delta(\mathscr{L})$ is simply the empirical frequency distribution of choice sets, and this empirical frequency distribution is ex-ante known to the

DM. For instance, the DM could know that she is facing only two choice sets, simultaneously (as in the farming example of Section 1.1) or one after the other without (sufficient) feedback as in the dieting example. Then p is simply the empirical frequency that attaches a probability of one half to each of these two choice sets with the understanding that there is actually no randomness. That is, we do not need to appeal to a law of large numbers as in the previous subsection and the DM receives exactly the p-weighted sum of payoffs that accrue from her choices from the two choice sets. If the DM then has an outside option, all our results apply.

6.5. **Risk aversion.** Theorem 2 states that for every non-strictly consistent rule there is a distribution over strictly consistent rules that is more risky than the given rule. Does this imply that a risk averse DM, in the absence of an outside option, and under p neutral such that Lemma 1 holds, would prefer the non-strictly consistent rule? The answer to this question is: it depends. It depends on how risk aversion enters the DM's objective function. Suppose now that the DM does not care about material gains directly, but the utility that material gains give her. Then one way of modeling this would be to simply transform all gain levels $g \in G$ to utility levels $u(g)$, where u is an increasing and strictly concave map from $\mathbb{R}$ to $\mathbb{R}$. This, however, is simply a rescaling of gain levels and does not change the results in any way. In the dieting example this may well be the appropriate way to capture risk aversion.

But this would not be a good model of how risk enters the problem for the game show example. Suppose, in fact, we perform a lab experiment (using the game show as our guide). Suppose we ask the DM to make many choices, but all within an hour or so, and we pay her only a total amount at the end. The DM, if she is risk averse, would probably evaluate that total final payment with an increasing concave utility function and not each of the individual payments. In this case, such a risk averse DM would indeed prefer, under p neutral and no outside option, a non-strictly consistent rule. For instance, in the example of Figure 1 any cyclical rule that chooses A out of $\{A, B\}$, C out of $\{A, C\}$, and B out of $\{B, C\}$ then provides a higher ex ante expected utility than any strictly consistent rule. Note, however, that if risk aversion is small enough, all strict inequalities in the comparison of rules remain strict, so that the results of Theorem 1 still hold. In other words, if risk aversion is sufficiently small the DM prefers a strictly consistent rule. But if risk aversion is sufficiently high strictly consistent rules cease to be optimal.

6.6. **Heterogeneous preferences.** Recent empirical evidence shows that although individual agents' decisions are at large consistent with a theory of preferences, these preferences vary wildly across agents. For instance, using scanner data of household purchases, Echenique, Lee, and Shum (2011) and Dean and Martin (2015) find that individual households make consistent choices.[7] On the other hand Dean and Martin (2011) and Crawford and Pendakur (2013), show that households exhibit significant heterogeneity in preferences over consumption bundles.[8]

These findings are consistent with our main results. In fact, our model is able to account for heterogenous preferences for the two following reasons. First, we show there are conditions on the distribution over choice sets for the agents, under which even if all agents face the same such distribution over choice sets, all strictly consistent rules are equally good and all other rules suboptimal. Different agents may thus adopt different strictly consistent rules with each of them being optimal.[9] Second, we show that while all agents have the same utility function, different distribution over choice sets or different distributions over outside options lead to different optimal strictly consistent rules. Hence, a strictly consistent rule that is optimal for one agent may be suboptimal for another even though they both share the same utility function and face the same uncertainty as to which alternatives are good for them.

This interpretation of the apparent heterogeneity of preferences, thus, leaves room for educating people about what preferences they "should"

[7]To quote Echenique, Lee, and Shum (2011, p. 1205), "[i]t is fair to say that most of the empirical literature, using both field and experimental data, finds relatively few violations of GARP"

[8]Both findings of consistent and heterogenous behavior are confirmed by Choi, Fisman, Gale, and Kariv (2007) in the context of risk-preferences, see also Dean and Martin (2010, section 5.2.3).

[9]While each individual chooses one strictly consistent choice rule, the aggregate behavior will look like that of a random utility model as in Luce (1959) with implications as in Block and Marschak (1960), see also Gul and Pesendorfer (2013).

have if scientist identify which preferences would be "objectively better" and leaves room for paternalistic "nudging" people into the right direction.

Appendix A. Proofs

A.1. **Proof of Lemma 1.** Recall that, for a given rule $R \in \mathscr{R}$, the ex ante expected payoff is given by

$$\begin{aligned}
\mathbb{E}_q g_p(R) &= \mathbb{E}_q \mathbb{E}_p g(R(L)) \\
&= \sum_{g \in \mathscr{G}} q(g) \sum_{L \in \mathscr{L}} p(L) g(R(L)) \\
&= \sum_{g \in \mathscr{G}} q(g) \sum_{L \in \mathscr{L}} p(L) \frac{1}{|R(L)|} \sum_{k \in R(L)} g(k) \\
&= \sum_{L \in \mathscr{L}} \frac{p(L)}{|R(L)|} \sum_{k \in R(L)} \sum_{g \in \mathscr{G}} q(g) g(k),
\end{aligned}$$

where the last equality follows from a simple change in the order of summation.

We complete the proof of Lemma 1 by showing that $\sum_{g \in \mathscr{G}} q(g) g(k)$ does not depend on k. Since q is symmetric, for every permutation π of K we have

$$\sum_{g \in \mathscr{G}} q(g) g\,(k) = \sum_{g \in \mathscr{G}} q(g) g(\pi(k)).$$

By averaging over all permutations π we obtain:

$$\begin{aligned}
\sum_{g \in \mathscr{G}} q(g) g(k) &= \frac{1}{|K|!} \sum_{\pi} \sum_{g \in \mathscr{G}} q(g) g(\pi(k)) \\
&= \frac{1}{|K|!} \sum_{g \in \mathscr{G}} q(g) \sum_{\pi} g(\pi(k)) \\
&= \frac{1}{|K|!} \sum_{g \in \mathscr{G}} q(g) \sum_{k'} \frac{|K|!}{|K|} g(k') \\
&= \frac{1}{|K|} \sum_{k'} g(k').
\end{aligned}$$

QED

A.2. **Proof of Lemma 2.** We prove that Λ_p^s contains the extreme points of the convex hull of Λ_p in $\mathbb{R}^{|K|}$. By the supporting hyperplane theorem, it suffices to prove that, for any vector $v = (v(k))_k \in \mathbb{R}^{|K|}$, $\max_{\lambda_p \in \Lambda_p} \sum_k \lambda_p(k)v(k)$ is attained at some $\lambda_p \in \Lambda_p^s$. Interpret $v(k)$ as a "fictitious utility" for the choice k. For $L \subseteq K$, let $v(L) = \frac{1}{|L|}\sum_{k \in L} v(k)$. Let π be a permutation of K that orders the coordinates of v such that $v(\pi(1)) \geq v(\pi(2)) \geq \ldots \geq v(\pi(k))$. Maximizing $\sum_k \lambda_p(k)v(k)$ over $\lambda_p \in \Lambda_p$ is equivalent to maximizing the expected "fictitious utility" $\sum_{L \in \mathscr{L}} p(L)v(R(L))$ over all rules.

The rule R^π that selects the least element according to π in every choice set, $R(L) = \min\{k, \pi(k) \in L\}$, maximizes each term of the sum $\sum_{L \in \mathscr{L}} p(L)v(R(L))$, so it maximizes the sum. Also, R^π is strictly consistent, since it is the rule that corresponds to the preference relation $\pi(1) \succ \pi(2) \succ \ldots \succ \pi(k)$. Hence, $\lambda_p(R^\pi)$ belongs to Λ_p^s, and achieves $\max_{\lambda_p \in \Lambda_p} \sum_k \lambda_p(k)v_k$.

QED

A.3. **Proof of Theorem 2.** In order to prove Theorem 2 the following two Lemmas are useful.

Lemma 3. *Let $R \in \mathscr{R}^s$ and $R' \in \mathscr{R}$. If $\lambda_p(R') = \lambda_p(R)$, then $R' = R$.*

Proof: Consider w.l.o.g. the strictly consistent rule R corresponding to the preference relation $1 \succ 2 \succ 3 \succ ... \succ |K|$, and let R' be a rule such that $\lambda_p(R') = \lambda_p(R)$. Since $R(L) = \{1\}$ whenever $1 \in L$,

$$\lambda_p(R)(1) = \sum_{1 \in L} p(L) \geq \sum_{1 \in L, 1 \in R'(L)} \frac{p(L)}{|R'(L)|} = \lambda_p(R')(1).$$

Since p has full support, the inequality above is an equality if and only if $R'(L) = \{1\}$ whenever $1 \in L$. Now we have

$$\lambda_p(R)(2) = \sum_{2 \in L, 1 \notin L} p(L) \geq \sum_{2 \in L, 1 \notin L, 2 \in R'(L)} \frac{p(L)}{|R'(L)|} = \lambda_p(R')(2).$$

Here again, equality holds only if $R'(L) = \{2\}$ whenever $2 \in L$ and $1 \notin L$.

By induction on k, we obtain that $R'(L) = \{k\}$ whenever $k \in L$ and $1, \ldots, k-1 \notin L$, i.e., $R' = R$.

QED

Lemma 4. *For every non-constant vector* $(a_k)_{k\in K} \in \mathbb{R}^{|K|}$ *and every non-constant gain function* g*, there exists a permutation* g^π *of* g *such that* $\sum_k a_k g^\pi(k) \neq 0$*.*

Proof: Consider a vector $(a_k)_{k\in K} \in \mathbb{R}^{|K|}$ such that for all permutations g^π of a non-constant gain function g we have $\sum_k a_k g^\pi(k) = 0$. Consider the permutation π that only exchanges two indexes, $i, j \in K$. Then we have both

$$\sum_{k\neq i,j} a_k g(k) + a_i g(i) + a_j g(j) = 0$$

and

$$\sum_{k\neq i,j} a_k g(k) + a_i g(j) + a_j g(i) = 0.$$

The difference of these two expressions gives

$$a_i g(i) + a_j g(j) = a_i g(j) + a_j g(i),$$

or, equivalently,

$$(a_i - a_j)(g(i) - g(j)) = 0.$$

Thus, for every $i, j \in K$ we have $a_i = a_j$ or $g(i) = g(j)$. By assumption there exist $i, j \in K$ such that $g(i) \neq g(j)$, and thus for these we have $a_i = a_j$. Let $a = a_i = a_j$.

For every $k \neq i, j$, since we cannot have both $g(k) = g(i)$ and $g(k) = g(j)$ we have either $a_k = a_i = a$ or $a_k = a_j = a$. Therefore the vector $(a_k)_{k\in K}$ is constant.

QED

Proof of Theorem 2: Let $R \in \mathscr{R} \setminus \mathscr{R}^s$. By Lemma 2, $\lambda_p(R)$ is a convex combination of choice distributions in Λ_p^s. That is, there exists a distribution μ over $\mathscr{R}^s$ such that

$$\lambda_p(R) = \sum_{R'\in\mathscr{R}^s} \mu(R')\lambda_p(R').$$

We now have for every g:

$$g_p(R) = \sum_{R'\in\mathscr{R}^s} \mu(R') g_p(R') = \mathbb{E}_\mu g_p(R'). \tag{A.1}$$

Therefore, for every g, the distribution of $g_p(R')$ under μ is a mean preserving spread of the constant $g_p(R)$. This remains true when g is taken at random according to q: the distribution of $g_p(R')$ under q and μ is a mean preserving spread of the distribution of $g_p(R)$ under q.

We now show that this mean-preserving spread is strict. To show that it suffices to show that the mean preserving spread of equation (A.1) is strict for one g in the support of q. I.e., we need to prove that there exists g in the support of q and R' in the support of μ such that $g_p(R') \neq g_p(R)$.

By Lemma 3 there exists a rule $R' \in \mathscr{R}^s$ such that $\alpha(R') > 0$ and $\lambda_p(R') \neq \lambda_p(R)$. Let $a_k = \lambda_p(R')(k) - \lambda_p(R)(k)$. Since $\lambda_p(R') \neq \lambda_p(R)$, there exists k s.t. $a_k \neq 0$. But then since $\sum_k a_k = 0$, a is non-constant. Then, by q symmetric and Lemma 4, there exist g in the support of q s.t. $\sum_{k \in K} a_k g(k) \neq 0$. The results follows since $g_p(R) - g_p(R') = \sum_k a_k g(k)$.

To prove the final statement of Theorem 2 we note that under p neutral for any two strictly consistent rules $R', R'' \in \mathscr{R}^s$ their choice distributions $\lambda_p(R')$ and $\lambda_p(R'')$ are permutations of each other. But then under q symmetric the induced distribution over material gains g is the same for both rules. QED

A.4. **Proof of Theorem 1.** We need to prove that, for any p and for any $R' \in \mathscr{R} \setminus \mathscr{R}^s$ there is a $R^* \in \mathscr{R}^s$ such that the DM strictly prefers R^* over R', i.e., such that

$$\mathbb{E}_q \left[\mathbb{E}_{\not g} \left[\max \left\{ g_p(R'), \not g \right\} \mid g \right] \right] < \mathbb{E}_q \left[\mathbb{E}_{\not g} \left[\max \left\{ g_p(R^*), \not g \right\} \mid g \right] \right].$$

By Theorem 2 for any $R' \in \mathscr{R} \setminus \mathscr{R}^s$ there is a distribution μ over strictly consistent rule that is strictly riskier than rule R'. Note that this distribution μ does not depend on the realized gain function.

As the maximum is a convex function and as $\not g$, conditional on any gain function, has positive density in the whole range of possible gain levels (in particular it has support where the gains distributions induced by rule R' and distribution μ differ) we have

$$\mathbb{E}_q \left[\mathbb{E}_{\not g} \left[\max\{ g_p(R'), \not g \} \mid g \right] \right] < \mathbb{E}_q [\mathbb{E}_{\not g} [\sum_{R \in \mathscr{R}^s} \mu(R) \max\{ g_p(R), \not g \} \mid g]].$$

Interchanging the order of summation we have

$$\mathbb{E}_q \left[\mathbb{E}_{\not g} \left[\max\{ g_p(R'), \not g \} \mid g \right] \right] < \sum_{R \in \mathscr{R}^s} \mu(R) \mathbb{E}_q \left[\mathbb{E}_{\not g} \left[\max\{ g_p(R), \not g \} \mid g \right] \right].$$

Thus, there must be at least one $R^* \in \mathscr{R}^s$ such that

$$\mathbb{E}_q \left[\mathbb{E}_{\not{g}} \left[\max\{g_p(R'), \not{g}\} \mid g\right]\right] < \mathbb{E}_q \left[\mathbb{E}_{\not{g}} \left[\max\{g_p(R^*), \not{g}\} \mid g\right]\right].$$

To finish the proof note that under p neutral all strictly consistent rules induce the same distribution over gains. If the outside option is independent of the gain function we then have the desired result. QED

References

Arrow, K. J. (1959): "Rational choice functions and orderings," *Economica*, 26, 121–27.

Block, H. D., and J. Marschak (1960): "Random orderings and stochastic theories of responses," in *Contributions to probability and statistics*, ed. by I. Olkin, S. Ghrye, W. Hoeffding, W. Madow, and H. Mann, pp. 97–132, Stanford, CA. Stanford University Press.

Choi, S., R. Fisman, D. Gale, and S. Kariv (2007): "Consistency and heterogeneity of individual behavior under uncertainty," *The American Economic Review*, 97(5), 1921–1938.

Crawford, I., and K. Pendakur (2013): "How many types are there? A simple, theory-consistent approach to unobserved heterogeneity," *The Economic Journal*, 123(567), 77–95.

Dean, M., and D. Martin (2010): "How rational are your choice data?," mimeo.

——— (2011): "Testing for rationality with consumption data: Demographics and heterogeneity," *Available at SSRN 1894785.*

——— (2015): "Measuring rationality with the minimum cost of revealed preference violations," *Review of Economics and Statistics*, forthcoming.

Echenique, F., S. Lee, and M. Shum (2011): "The money pump as a measure of revealed preference violations," *Journal of Political Economy*, 119(6), 1201–1223.

Gul, F., and W. Pesendorfer (2013): "Random Utility Maximization with Indifference," mimeo, Princeton University.

Luce, R. D. (1959): *Individual Choice Behavior: A Theoretical Analysis.* Wiley, New York.

Mas-Collel, A., M. D. Whinston, and J. R. Green (1995): *Microeconomic Theory.* Oxford University Press, Oxford, UK.

MERTON, R. (1973): "Theory of Rational Option Pricing," *The Bell Journal of Economics and Management Science*, 4, 141–183.

PASCAL, B. (1670): *Pensées de M. Pascal sur la religion, et sur quelques autres sujets, qui ont esté trouvées après sa mort parmy ses papiers*. G. Desprez, Paris.

RASMUSEN, E. (2007): "When Does Extra Risk Strictly Increase the Value of Options?," *The Review of Financial Studies*, 20(5), 1647–1667.

UZAWA, H. (1956): "Note on preference and axioms of choice," *Annals of the Insitute of Statistical Mathematics*, 8, 35–40.

CNRS – ÉCOLE POLYTECHNIQUE AND LONDON SCHOOL OF ECONOMICS, LONDON, olivier.gossner@poly.polytechnique.edu

DEPARTMENT OF ECONOMICS, UNIVERSITY OF GRAZ, christoph.kuzmics@uni-graz.at