Dominance and Optimality¹

by Xienan Cheng ⊗ Tilman Börgers

August 29, 2022

Abstract

This paper proposes a general theory of dominance among choices that encompasses strict and weak dominance among strategies in games, Blackwell dominance among experiments, and first or second order stochastic dominance among monetary lotteries. One choice dominates another if in a variety of situations the former choice yields higher expected utility than the latter. We then investigate whether, in a finite set of possible choices, all undominated choices are optimal in some situation. We present a formal framework in which the answer to this question is positive, and we show that within this framework the set of undominated choices is the smallest set to which the decision maker can restrict attention ex ante without running the risk of not having an optimal choice in the particular situation in which she finds herself. For this result it is crucial that the dominating alternatives are allowed to be convex combinations (in games: mixed strategies). A detailed analysis of dominance in game theory, Blackwell dominance, and first or second order stochastic dominance in one common framework also allows us to compare the properties of these concepts, and to obtain insights into why certain versions of our result apply only to some, but not all of these concepts.

¹We are grateful to Yaron Azrieli, Jianpei Li, and Shuo Liu for comments. We would like to thank Duarte Gonçalves for bringing Weinstein (2020) to our attention.
1. INTRODUCTION

Economic theory has developed many notions of “dominance” of some choice over another. For example, in game theory the notions that one strategy “strictly dominates” another one, or that it “weakly dominates” another one, are fundamental concepts (see, for example, Pearce (1984)). In the theory of choice among monetary lotteries the concepts of “first order stochastic dominance” (Quirk and Saposnik, 1962) and “second order stochastic dominance” Rothschild and Stiglitz (1970) are frequently used. In the theory of information “Blackwell dominance” among experiments (Blackwell (1951), Blackwell (1953)) is an important concept.

All these concepts of dominance provide partial orders of sets of alternatives among which a decision maker chooses. One alternative dominates another alternative if it is the better choice regardless of certain aspects of the decision maker’s decision problem. For example, one strategy dominates another one if it is a better choice regardless of the player’s belief about the other players’ strategy choices. One monetary lottery first order stochastically dominates another lottery if it yields higher expected utility regardless of the decision maker’s utility function, provided that this utility function is increasing. One monetary lottery second order stochastically dominates another lottery if it yields higher expected utility regardless of the decision maker’s utility function, provided that this utility function is increasing and concave. One experiment Blackwell dominates another experiment if it allows the decision maker to achieve higher expected utility regardless of which decision problem the decision maker faces.\(^2\)

Many famous results of economic theory provide equivalent characterizations of dominance relations that are easier to check than the original definition. For example, to check whether one lottery first order stochastically dominates another lottery one may equivalently compare the two lotteries’ cumulative distribution functions (this is the main result in Quirk and Saposnik (1962)). To determine whether one experiment Blackwell dominates another experiment one may equivalently check whether the latter experiment can be obtained by “garbling” the former experiment (Theorem 5 in Blackwell (1953)). Also, to check whether one strategy dominates another one one can equivalently check whether for every pure strategy combination that the other players might choose the former strategy yields higher utility than the latter.\(^3\)

\(^2\)In this paragraph we have been deliberately vague about whether we refer to “strict” or “weak dominance” and whether we mean by “higher expected utility” that the utility is strictly or weakly larger. We shall, of course, be more precise about these issues later in the paper.

\(^3\)Unlike the other results mentioned in this paragraph, this last result is a trivial observation.
When we eliminate dominated choices in a decision problem we narrow down the options that a rational decision maker might choose without completely specifying all characteristics of this decision maker. In this paper we will ask in a variety of contexts whether ruling out dominated choices is the best we can do without specifying the decision maker’s characteristics further, or whether there are other, not dominated choices that a rational decision maker nonetheless will never choose. We prove a very general theorem that shows that nothing more than dominated choices can be ruled out. Our main result is of the following form: for every not dominated alternative there exists some specification of the decision maker’s problem in which a rational decision maker will choose this alternative.

The precise details matter for this result, however. In particular, the result is not true unless we consider the possibility that an alternative is dominated by a convex combination of the other alternatives. The potential relevance of convex combinations of alternatives is familiar from game theory: A strategy is a best response to some belief if and only if it is not strictly dominated by any of the other strategies nor by any convex combination of the other strategies (i.e. nor by any mixed strategy). It is well-known that this result would not be true if we had not included the possibility that the dominating strategy is a mixed strategy. In other words: there are well-known examples of games in which a strategy is not strictly dominated by any other pure strategy, yet it is not a best response to any belief of the player. Our main theorem is built on this insight, but applies to a much more general setting than just games.

One such setting is the choice among experiments where one might wish to use Blackwell dominance to rule out some choices. Suppose a decision maker can choose one experiment from a finite set of available experiments. Assume that all experiments are available at no cost. Consider an experiment that is not Blackwell dominated by any other experiment. Does there exist a decision problem in which it is optimal to choose this experiment from the set of available experiments? It turns out that the answer to this question is “yes” only if we allow for the possibility that the dominating experiment is an appropriately defined convex combination of the other experiments. We give in Section 5 a counterexample that demonstrates how the result otherwise fails. On the other hand, once convex combinations of experiments are considered, the result is an immediate implication of our main theorem.

Other details matter. Our most general result will characterize those alternatives that are, for some specification of the details of the decision problem, the only optimal
choice of the decision maker. We call such alternatives “uniquely optimal.” Thus, our main result establishes an equivalence between an alternative not being dominated by convex combinations of other alternatives, and an alternative being a uniquely optimal choice. We also establish equivalences between an alternative not being dominated and the alternative being one optimal choice, or the alternative being an optimal choice in some “non knife-edge” circumstances. These versions of our result are the ones that are familiar from game theory, but they are only applicable under a set of assumptions that are satisfied in finite strategic games, but that are not satisfied when, say, comparing experiments.

One might question our focus on uniquely optimal choices. Of course, sometimes decision makers will face situations in which they are indifferent between several optimal choices. But we shall provide sufficient conditions such that, when there are multiple optimal choices in a decision problem, then at least one of those choices is a uniquely optimal choice in some other decision problem. Therefore, the set of uniquely optimal choices is the minimal set of alternatives to which a decision maker may restrict attention such that this set includes an optimal choice regardless of the specifics of the decision maker’s decision problem.

To formalize the argument explained in the previous paragraph we introduce “minimally sufficient” sets of alternatives, that is, sets of alternatives that are sufficient in the sense that whatever the particulars of the decision problem the sets always contain at least one optimal choice, and that are minimal in the sense that they have no subset that is also sufficient. We motivate minimally sufficient sets as the sets of alternatives that a decision maker would restrict attention to if attention is costly for the decision maker, but the decision maker is not willing to give up any material payoff in the decision problem in return for lower attention cost. We prove that under some conditions the set of uniquely optimal choices is the only minimally sufficient set of alternatives.

The general result on which we build our analysis is presented in Section 2. We motivate a focus on uniquely optimal actions in Section 3. In the subsequent sections we apply our general analysis first to dominance relations in games, then to dominance relations among experiments, and finally to dominance relations among monetary lotteries. We mention related literature in each of the applications sections.

This paper builds to a significant extent on Peter Fishburn’s work on decision theory. Fishburn (1964) provided a general analysis of dominance, and emphasized the importance of dominance by convex combinations. However, Fishburn focused on
dominance among monetary lotteries and studied notions that are similar to those in Quirk and Saposnik (1962) and Rothschild and Stiglitz (1970). Fishburn emphasized (footnote 12 in Fishburn (1964)) that dominance, as used in his analysis “should not be confused with dominance in game theory or statistical decision theory.” Our argument is the opposite: dominance in games and dominance in statistical decision theory are closely related to the other dominance notions. Thus, the results for monetary lotteries that we discuss in Section 6 and that overlap with results due to Fishburn (1964) are derived from the same general result as the results for games that we discuss in Section 4 and that overlap with results in Pearce (1984) and Weinstein (2020). Our analysis of dominance among experiments are, as far as we are aware, new to this paper, as is the unification of all these results in a general framework.

One part of our general result in Section 2 relies on a separation theorem in Fishburn (1975). Moreover, the proof of all parts of the general result is inspired by the way in which Fishburn (1975) used separation theorems.

2. General Results

Let $X$ and $Y$ be two non-empty sets, and let $u : X \times Y \rightarrow \mathbb{R}$. Here, $X$ is the set of actions $x$ that a decision maker can choose from, while $Y$ is the set of “situations” $y$ that this decision maker might find herself in. $u(x, y)$ is the decision maker’s utility if choosing action $x$ in situation $y$. The decision maker first observes the situation $y \in Y$ and then chooses an action $x \in X$. Throughout this section we shall make the following assumption:

**Assumption 1.** $X$ is finite. $Y$ is a convex subset of a topological vector space. $u$ is linear and continuous in $y$.

The assumption that $X$ is finite greatly simplifies the analysis below. In the applications that we shall consider, $Y$ is, depending on the context, the set of beliefs the decision maker might hold, or the set of value functions corresponding to the decision problems the decision maker might face, or the set of utility functions the decision maker might have. This is why it is convenient to let $Y$ be a subset of a topological vector space. $Y$ will be convex in all our applications. The linearity of $u$ will reflect that decision makers in our applications are expected utility maximizers.
The convexity of $Y$ together with the linearity of $u$ will allow us to use separating hyperplane theorems in our proofs.

**Definition 1.** An action $x \in X$ is optimal if there exists a $y \in Y$ such that:

$$u(x, y) \geq u(x', y) \text{ for all } x' \in X.$$  

We denote by $X_O$ the set of optimal actions.

**Definition 2.** An action $x \in X$ is interior optimal if there exists a $y \in ri(Y)$ (where $ri(Y)$ denotes the relative interior of $Y$) such that:

$$u(x, y) \geq u(x', y) \text{ for all } x' \in X.$$  

We denote by $X_{IO}$ the set of interior optimal actions. It may not seem intuitively obvious why it is relevant whether the situation $y$ in which an action $x$ is a best response is interior or not. However, when $Y$ represents a set of beliefs, then the relative interior of $Y$ will represent the set of full support beliefs. Having beliefs with full support has been interpreted in the game theoretic literature as sign of caution by the decision maker. This is why interior situations $y$ will receive special attention in this section.

**Definition 3.** An action $x \in X$ is uniquely optimal if there exists a $y \in Y$ such that:

$$u(x, y) > u(x', y) \text{ for all } x' \in X \text{ such that } x \neq x'.$$  

We denote by $X_{UO}$ the set of uniquely optimal actions. It is not immediately obvious why the set of uniquely optimal actions should receive special attention. We address this issue therefore in detail in the next section.

Our objective in this section is to characterize the sets of optimal, interior optimal, and uniquely optimal actions in terms of dominance notions. We therefore next introduce the dominance notions that we are considering.

**Definition 4.** An action $x \in X$ is strictly dominated if there are a set $\{x_1, x_2, \ldots, x_n\} \subseteq X \setminus \{x\}$ and a vector $(\lambda_1, \lambda_2, \ldots, \lambda_n) \in \mathbb{R}_+^n$ with $\sum_{i=1}^n \lambda_i = 1$ such that:

$$\sum_{i=1}^n (\lambda_i u(x_i, y)) > u(x, y) \text{ for all } y \in Y.$$
We denote by $X_{NSD}$ the set of all actions that are not strictly dominated. It is essential that we consider the possibility here that an action is dominated not by a single action but by a convex combination of actions. We shall illustrate this point in the applications that we consider later in the paper. One may think of the convex combination of actions as a “mixed action” in analogy to mixed strategies in game theory.

**Definition 5.** An action $x \in X$ is weakly dominated if there are a set $\{x_1, x_2, \ldots, x_n\} \subseteq X \setminus \{x\}$ and a vector $(\lambda_1, \lambda_2, \ldots, \lambda_n) \in \mathbb{R}^n_+$ with $\sum_{i=1}^n \lambda_i = 1$ such that:

$$\sum_{i=1}^n (\lambda_i u(x_i, y)) \geq u(x, y) \quad \text{for all } y \in Y$$

with strict inequality for at least one $y \in Y$.

We denote by $X_{NWD}$ the set of all actions that are not weakly dominated.

Strict and weak dominance are standard notions that are familiar from game theory. We will use a third concept that is less familiar, but that will prove crucial for some of our results.

**Definition 6.** An action $x \in X$ is redundant if there are a set $\{x_1, x_2, \ldots, x_n\} \subseteq X \setminus \{x\}$ and a vector $(\lambda_1, \lambda_2, \ldots, \lambda_n) \in \mathbb{R}^n_+$ with $\sum_{i=1}^n \lambda_i = 1$ such that:

$$\sum_{i=1}^n (\lambda_i u(x_i, y)) = u(x, y) \quad \text{for all } y \in Y.$$

Thus, an action is redundant if it is equivalent in expected utility to a convex combination of the other actions. Denote by $X_{NR}$ the set of actions that are not redundant.

We are now ready to state our main result in this section. This result generalizes a number of results familiar from the literature on game theory.

**Theorem 1.** (i) If an action is optimal, then it is not strictly dominated: $X_O \subseteq X_{NSD}$.

(ii) If $Y$ is compact, then an action that is not strictly dominated is optimal: $X_{NSD} \subseteq X_O$.

(iii) If an action is interior optimal then it is not weakly dominated: $X_{IO} \subseteq X_{NWD}$. 
(iv) If $Y$ is finite dimensional, then an action that is not weakly dominated is interior optimal: $X_{NWD} \subseteq X_{IO}$.

(v) An action is uniquely optimal if and only if it is not weakly dominated and not redundant: $X_{UO} = X_{NWD} \cap X_{NR}$.

Note that results (ii) and (iv) are based on assumptions regarding $Y$ that go beyond those made in Assumption 1. Our proofs use these additional assumptions. Of course, the theorem does not claim that these additional assumptions are necessary.

Our proof of Theorem 1 below is based on elementary separating hyperplane theorems in finite dimensional Euclidean space. Thus, it is a simple and geometric proof. The ideas on which some parts of the proof are based are related to ideas in Fishburn (1975).

Proof. Step 1: We first prove that $X_O \subseteq X_{NSD}$. The proof is indirect. Suppose $x$ were optimal for some $\bar{y} \in Y$, but that $x$ were strictly dominated. Let $\{x_1, x_2, \ldots, x_n\} \subseteq X \setminus \{x\}$ be the actions in the support of the strictly dominating convex combination, and let $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ be the corresponding weights. We have: $\sum_{i=1}^{n} (\lambda_i u(x_i, y)) > u(x, y)$ for all $y \in Y$. But this implies that for some $i$ we have: $u(x_i, \bar{y}) > u(x, \bar{y})$, contradicting that $x$ is optimal in situation $\bar{y}$.

We next prove that $X_{IO} \subseteq X_{NWD}$. The proof is indirect. Suppose $x$ were optimal for some $\bar{y} \in ri(Y)$, but that $x$ were weakly dominated. Let $\{x_1, x_2, \ldots, x_n\} \subseteq X \setminus \{x\}$ be the actions in the support of the weakly dominating convex combination, and let $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ be the corresponding weights. We then have: $\sum_{i=1}^{n} (\lambda_i u(x_i, y)) \geq u(x, y)$ for all $y \in Y$ and $\sum_{i=1}^{n} (\lambda_i u(x_i, y^*)) > u(x, y^*)$ for some $y^* \in Y$. Because $x$ is optimal at $\bar{y}$, we must have: $\sum_{i=1}^{n} (\lambda_i u(x_i, \bar{y})) - u(x, \bar{y}) = 0$. Define: $\hat{y} \equiv (1+\varepsilon)\bar{y} - \varepsilon y^*$. Because $\bar{y}$ is in the relative interior of $Y$, we have $\hat{y} \in Y$ for sufficiently small $\varepsilon > 0$.

Now note that:

$$\sum_{i=1}^{n} (\lambda_i u(x_i, \hat{y})) - u(x, \hat{y})$$

$$= (1 + \varepsilon) \left( \sum_{i=1}^{n} (\lambda_i u(x_i, \bar{y})) - u(x, \bar{y}) \right) - \varepsilon \left( \sum_{i=1}^{n} (\lambda_i u(x_i, y^*)) - u(x, y^*) \right)$$

$$= -\varepsilon \left( \sum_{i=1}^{n} (\lambda_i u(x_i, y^*)) - u(x, y^*) \right) < 0,$$

This contradicts the assumption that the convex combination weakly dominates $x$. 

We finally prove that $X_{UO} \subseteq X_{NWD} \cap X_{NR}$. The proof is indirect. Suppose $x$ were uniquely optimal for some particular $\bar{y} \in Y$. but that $x$ were weakly dominated. Let $\{x_1, x_2, \ldots, x_n\} \subseteq X \setminus \{x\}$ be the actions in the support of the weakly dominating convex combination, and let $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ be the corresponding weights. We have: 

$$
\sum_{i=1}^{n} (\lambda_i u(x_i, y)) \geq u(x, y) \text{ for all } y \in Y.
$$

But this implies that for some $i$ we have: 

$$
u(x_i, \bar{y}) \geq u(x, \bar{y}),$$

contradicting that $x$ is uniquely optimal in situation $\bar{y}$. The same argument proves that a uniquely optimal action cannot be redundant.

**Step 2:** We now prove the converses of the statements in Step 1. We first prove that $X_{NSD} \subseteq X_{O}$ if $Y$ is compact. The proof is indirect. Suppose $x$ were not optimal for any $y \in Y$. We prove that then a convex combination of actions strictly dominates $x$. Consider the following set.

$$
C \equiv \{(u(x_1, y) - u(x, y), u(x_2, y) - u(x, y), \ldots, u(x_n, y) - u(x, y)) | y \in Y\},
$$

where we take $x_1, x_2, \ldots, x_n$ to be an enumeration of the set $X \setminus \{x\}$. If $x$ is not optimal in any situation $y \in Y$, then:

$$C \cap R_n^+ = \emptyset. \quad \text{(4)}$$

Observe that $R_n^+$ is a closed and convex set, and that $C$ is convex and compact (convex because of the linearity of $u$ and the convexity of $Y$ and compact because of the continuity of $u$ and the compactness of $Y$). We can then apply the following hyperplane theorem:

**Separating Hyperplane Theorem 1:** Suppose $C \subseteq R^n$ is convex and compact. If $C \cap R_n^+ = \emptyset$ then there exists $\lambda \in R_n^+$ with $\lambda \neq 0$ such that $\lambda \cdot x > 0$ for all $x \in C$.

For completeness, we briefly derive this result from a standard separating hyperplane theorem:

**Proof.** The strict separating hyperplane theorem (for example Theorem 3.7 in Vohra (2005)) implies that there are $\lambda \in R^n$ with $\lambda \neq 0$ and $\varepsilon \in R$ such that $\lambda \cdot x < \varepsilon$ for all $x \in R^n$, and $\lambda \cdot x > \varepsilon$ for all $x \in C$. It easily follows that $\lambda \in R_n^+$ and that $\varepsilon > 0$. Therefore, $\lambda \cdot x > \varepsilon$ for all $x \in C$ implies $\lambda \cdot x > 0$ for all $x \in C$. \qed

Obviously, we may normalize the vector $\lambda$ to which the theorem refers so that its components add up to 1. Applying the theorem to our setting, we therefore find that

---

4. We denote by $R_n^-$ the set of all vectors $x = (x_1, x_2, \ldots, x_n) \in R^n$ such that $x_i \leq 0$ for all $i$. 

there is a convex combination of actions such that:
\[
\sum_{i=1}^{n} \left[ \lambda_i (u(x_i, y) - u(x, y)) \right] > 0 \text{ for all } y \in Y.
\]
This means that the convex combination of the actions \(x_1, x_2, \ldots, x_n\) with weights \(\lambda_1, \lambda_2, \ldots, \lambda_n\) strictly dominates \(x\).

Next we prove that \(X_{NWD} \subseteq X_{IO}\) if \(Y\) is finite dimensional. If \(Y\) is finite dimensional then it is without loss of generality to assume that it is a subset of a finite dimensional Euclidean space. Our proof of the result is again indirect. Suppose \(x\) were not optimal for any \(y \in ri(Y)\). Define the set \(C\) as before. Because \(u\) is linear in \(y\), the relative interior of \(C\) is the image of the relative interior of \(Y\). Therefore we have:
\[
ri(C) \cap R^n_+ = \emptyset.
\]
We now apply the following separating hyperplane theorem:

**Separating Hyperplane Theorem 2:** Suppose \(C \subseteq R^n\) is convex. If \(ri(C) \cap R^n_+ = \emptyset\) then there exists \(\lambda \in R^n_+\) with \(\lambda \neq 0\) such that \(\lambda \cdot x \geq 0\) for all \(x \in C\) and \(\lambda \cdot x > 0\) for at least one \(x \in C\).

*Proof.* By Theorem 6.2 of Rockafellar (1970), \(ri(C)\) is non-empty and convex. By Lemma 5 in Fishburn (1975), there exists a \(\lambda \in R^n_+\) with \(\lambda \neq 0\) such that \(\lambda \cdot x \geq 0\) for all \(x \in ri(C)\) and \(\lambda \cdot x > 0\) for at least one \(x \in ri(C)\). By continuity, \(\lambda \cdot x \geq 0\) for all \(x\) in the topological closure of \(ri(C)\), and by Theorem 6.3 in Rockafellar (1970), the topological closure of \(ri(C)\) is a superset of \(C\). Therefore, \(\lambda \cdot x \geq 0\) for all \(x \in C\). Finally, because \(ri(C) \subseteq C\), we have \(\lambda \cdot x > 0\) for at least one \(x \in C\). \(\square\)

Normalizing again the vector \(\lambda\) to which the theorem refers so that its components add up to 1, we find that there is a convex combination of actions such that:
\[
\sum_{i=1}^{n} \left[ \lambda_i (u(x_i, y) - u(x, y)) \right] \geq 0 \text{ for all } y \in Y,
\]
with strict inequality for at least one \(y \in Y\). This means that the convex combination of the actions \(x_1, x_2, \ldots, x_n\) with weights \(\lambda_1, \lambda_2, \ldots, \lambda_n\) weakly dominates \(x\).

We finally prove that \(X_{NWD} \cap X_{NR} \subseteq X_{UO}\). The proof is indirect. Suppose \(x\) were not uniquely optimal for any \(y \in Y\). Defining the set \(C\) as before, this means that:
\[
C \cap R^n_{<0} = \emptyset. \quad (5)
\]
\(^5\)We denote by \(R^n_{<0}\) the set of all vectors \(x = (x_1, x_2, \ldots, x_n) \in R^n\) such that \(x_i < 0\) for all \(i\).
We now apply the following separating hyperplane theorem:

**Separating Hyperplane Theorem 3:** Suppose $C \subseteq \mathbb{R}^n$ is convex. If $C \cap \mathbb{R}_{\geq 0}^n = \emptyset$ then there exists $\lambda \in \mathbb{R}^n_+$ with $\lambda \neq 0$ such that $\lambda \cdot x \geq 0$ for all $x \in C$.

**Proof.** Minkowski’s separating hyperplane theorem (for example Theorem 3.6 in Vohra (2005) implies that there are $\lambda \in \mathbb{R}^n$ with $\lambda \neq 0$ and $\varepsilon \in \mathbb{R}$ such that $\lambda \cdot x \leq \varepsilon$ for all $x \in \mathbb{R}_{\leq 0}^n$ and $\lambda \cdot x \geq \varepsilon$ for all $x \in C$. It easily follows that $\lambda \in \mathbb{R}^n_+$ and $\varepsilon \geq 0$. Therefore, $\lambda \cdot x \geq \varepsilon$ for all $x \in C$ implies $\lambda \cdot x \geq 0$ for all $x \in C$. \qed

Normalizing again the vector $\lambda$ to which the theorem refers so that its components add up to 1, we find that there is a convex combination of actions such that:

$$\sum_{i=1}^{n} \lambda_i u(x_i, y) \geq u(x, y) \text{ for all } y \in Y.$$  

Thus, the convex combination of $\{x_1, \ldots, x_n\}$ with weights $\lambda_1, \ldots, \lambda_n$ either weakly dominates $x$ or is equivalent to $x$, which contradicts the assumption with which we began this indirect proof. \qed

3. **Limited Attention Without Loss of Optimality**

We now provide a rationale for focusing on uniquely optimal actions as defined in the previous section. Specifically, we state conditions under which the set of uniquely optimal actions is the smallest set that the decision maker can limit attention to if she wants to choose an optimal action in every situation. We thus envisage the following scenario: the decision maker first restricts attention to some subset $\hat{X}$ of the set of all actions. Then she observes the situation $y$, and then she picks an action $x$ from the subset of actions $\hat{X}$ to which she has restricted attention. We assume that attention is costly: the decision maker wants to restrict attention to a set that is small (in terms of set-inclusion). Finally, attention cost are of second order importance: the decision maker’s first priority is to take an optimal action in every situation.

To formalize this, we introduce some additional notation. For every $y \in Y$ the set of optimal actions is:

$$O(y) = \{x \in X \mid u(x, y) \geq u(x', y) \text{ for all } x' \in X\}.$$  

In the following definition, the key term that we define is “minimal sufficiency.”
Definition 7. A set $\hat{X} \subseteq X$ is sufficient if for every $y \in Y$:
\[ O(y) \cap \hat{X} \neq \emptyset. \]

A set $\hat{X} \subseteq X$ is minimally sufficient if it is sufficient and there is no sufficient set $\bar{X} \subseteq X$ such that:
\[ \bar{X} \subset \hat{X}. \]

A minimally sufficient set of actions is thus a smallest set of actions to which the decision maker may restrict attention if she wants to choose optimally in every situation $y$.

The following result provides sufficient conditions for the set of uniquely optimal actions to be the unique minimally sufficient subset of the set of all actions. We emphasize that this result does not rely on Assumption 1.

Theorem 2. If $X$ is finite, if for any $x, x' \in X$ with $x \neq x'$ there is a $y \in Y$ such that $u(x, y) \neq u(x', y)$, and if:
\[ X_{UO} = X_{NWD} \cap X_{NR}, \]
then $X_{UO}$ is the unique minimally sufficient subset $\hat{X}$ of $X$.

Among the three assumptions of this theorem, the condition $X_{UO} = X_{NWD} \cap X_{NR}$ is not formulated in terms of the primitives of our model. However, Theorem 1 shows assumptions for the primitives of our model that imply that $X_{UO} = X_{NWD} \cap X_{NR}$ holds.

The proof of Theorem 2 describes an algorithm for finding the set $X_{NWD} \cap X_{NR}$ that may be of independent interest.

Proof. It is obvious that every sufficient set must include $X_{UO}$. What remains to be shown is that $X_{UO}$ contains for every situation $y \in Y$ an optimal action. Because the proposition assumes that $X_{UO} = X_{NWD} \cap X_{NR}$, this is equivalent to the statement that $X_{NWD} \cap X_{NR}$ contains for every situation $y \in Y$ an optimal action.

To prove this, we first observe that $X_{NWD} \cap X_{NR}$ can be constructed by the following algorithm. Set $X^0 = X$. For $k = 1, 2, \ldots, |X|$ (where $|X|$ is the number of elements of $X$), if no action $x \in X^{k-1}$ is either weakly dominated by, or equivalent to, a convex combination of the actions in $X^{k-1} \setminus \{x\}$, then set $X^{k-1} = X^k$. Otherwise, pick
arbitrarily some such \( x^{k-1} \in X^{k-1} \), and set \( X^k = X^{k-1} \setminus \{x^{k-2}\} \). Note that after at most \(|X|\) steps no further actions are eliminated. We claim that the final set is \( X^{|X|} = X_{NWD} \cap X_{NR} \). It is clear that \( X^{|X|} \) includes all actions in \( X_{NWD} \cap X_{NR} \). It thus remains to show that any action \( x \notin X_{NWD} \cap X_{NR} \) is eliminated in some step of this algorithm.

To prove the claim we show that if \( x \) is weakly dominated by, or equivalent to, a convex combination of the actions in \( X^{k-1} \setminus \{x\} \), and if \( x \in X^k \), then \( x \) it is also weakly dominated by, or equivalent to, a convex combination of the actions in \( X^k \setminus \{x\} \). Let the elements of \( X^{k-1} \setminus \{x\} \) be \( \{x_1, x_2, \ldots, x_n\} \), and let the weights of the convex combination be \( \lambda_1, \lambda_2, \ldots, \lambda_n \). Without loss of generality assume that \( x_1 \) is eliminated in step \( k \), that is: \( x^{k-1} = x_1 \). This means that \( x_1 \) is either weakly dominated or equivalent to, a convex combination of the elements of the set \( \{x_2, \ldots, x_n\} \cup \{x\} \). Let the weights of the convex combination be: \( \hat{\lambda}_2, \ldots, \hat{\lambda}_n, \hat{\lambda}_x \). It is then obvious that \( x \) is also weakly dominated, or equivalent to, a convex combination of \( \{x_2, \ldots, x_n\} \cup \{x\} \) with weights: \( \lambda_2 + \lambda_1 \hat{\lambda}_2, \ldots, \lambda_n + \lambda_1 \hat{\lambda}_n, \lambda_1 \hat{\lambda}_x \).

We next show: \( \lambda_1 \hat{\lambda}_x < 1 \). Suppose the opposite: \( \lambda_1 \hat{\lambda}_x = 1 \), hence \( \lambda_1 = \hat{\lambda}_x = 1 \). This means that \( x_1 \) is either weakly dominated or equivalent to \( x \), and that \( x \) is either weakly dominated or equivalent to \( x_1 \). But this means that \( x_1 \) and \( x \) are duplicates, which is a case that we ruled out in the assumptions of Theorem 2.

Now observe that, if \( x \) is weakly dominated by, or equivalent to, a convex combination of \( \{x_2, \ldots, x_n\} \cup \{x\} \) with weights:

\[
\lambda_2 + \lambda_1 \hat{\lambda}_2, \ldots, \lambda_n + \lambda_1 \hat{\lambda}_n, \lambda_1 \hat{\lambda}_x
\]

it is also weakly dominated by, or equivalent to, a convex combination of \( \{x_2, \ldots, x_n\} \) with weights:

\[
\frac{\lambda_2 + \lambda_1 \hat{\lambda}_2}{1 - \lambda_1 \hat{\lambda}_x}, \ldots, \frac{\lambda_n + \lambda_1 \hat{\lambda}_n}{1 - \lambda_1 \hat{\lambda}_x}.
\]

The division by \( 1 - \lambda_1 \hat{\lambda}_x \) is well-defined because, as we showed in the previous paragraph, \( \lambda_1 \hat{\lambda}_x < 1 \). We have now concluded the proof that the algorithm that we have described terminates with the set \( X_{NWD} \cap X_{NR} \).

Now consider any situation \( y \in Y \) and suppose \( x \in X \) is optimal in situation \( y \). Obviously, \( x \in X^0 \). Also, either \( x \in X^1 \), or \( x \) is weakly dominated by, or equivalent to, a convex combination of the actions in \( X^1 \). Then one of the actions in \( X^1 \) must also be a best response to \( y \). Iterating this argument leads to the conclusion that one of the actions in \( X^{|X|} \) is optimal in situation \( y \). \( \square \)
We now provide an example in which \( X_{UO} \neq X_{\text{NW}D} \cap X_{\text{NR}} \), and in which \( X_{UO} \) is not the unique minimally sufficient subset of \( X \). In this example \( X = \{x_1, x_2, x_3, x_4\} \) and \( Y = \{y_1, y_2\} \). The utility function is given in the following table:

<table>
<thead>
<tr>
<th>( x )</th>
<th>( y_1 )</th>
<th>( y_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x_1 )</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>( x_2 )</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>( x_3 )</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>( x_4 )</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

**Table 1.** An example where \( X_{UO} \) is not minimally sufficient.

In this example, \( X_{UO} = \emptyset \) and \( X_{\text{NW}D} \cap X_{\text{NR}} = \{x_2, x_4\} \). Obviously, \( X_{UO} \) is not sufficient. The minimally sufficient sets of actions are all sets with two elements such that one of the elements is \( x_1 \) or \( x_2 \), and the other element is \( x_3 \) or \( x_4 \).

4. **Dominance and Optimality of Strategies in Games**

We now explain how to apply the results of Section 2 to games. We focus on a player \( i \) in a strategic game who has to choose one strategy from a finite set of strategies \( S_i \). There are finitely many other players \( j \neq i \) and the Cartesian product of their strategy sets is \( S_{-i} \). Player \( i \)'s utility function is \( u_i : S_i \times S_{-i} \rightarrow \mathbb{R} \). Player \( i \)'s belief about the other players’ choices is a probability measure \( \mu_i \) on \( S_{-i} \). Denote the set of all such probability measures by \( \Delta(S_{-i}) \).\(^6\) Player \( i \)'s expected utility when she has belief \( \mu_i \) and chooses strategy \( s_i \) is:

\[
\sum_{s_{-i} \in S_{-i}} u_i(s_i, s_{-i})\mu_i(s_{-i}).
\]

In the notation of Section 2 in this application the set \( X \) of actions is the set \( S_i \) of strategies, and the set \( Y \) of situations is the set \( \Delta(S_{-i}) \) of beliefs. Note that the set of beliefs is a compact and convex subset of \( \mathbb{R}^{|S_{-i}|} \), and that the utility function \( u_i \) is linear (and therefore continuous) in a player’s beliefs.

The notions of strict and weak dominance introduced in Section 2 correspond to the thus named notions in game theory. Note that when checking dominance relations

\(^{6}\)From now on, for any finite or compact set \( A \), we denote by \( \Delta(A) \) the set of all (Borel-) probability distributions on \( A \).
among two strategies $s_i$ and $s_j$ in games we typically only compare expected utility for any given pure strategy combination of the other players, that is, we only consider beliefs that are Dirac measures on $S_{-i}$. This is sufficient because $\Delta(S_{-i})$ the convex hull of the set of Dirac measures on $S_{-i}$ and because of the linearity of expected utility. This is a general point: The properties of actions $x \in X$ defined in Definitions 4, 5, and 6 would not change if one replaced the expression “for all $y \in Y$” by the expression “for all $y \in Y^*$” where $Y^*$ is a subset of $Y$ such that the convex hull of $Y^*$ equals $Y$.

All parts of Theorem 1 hold in this setting. In particular, note that part (ii) applies because the set $\Delta(S_{-i})$ is compact. Parts (i) and (ii) of Theorem 1, if applied to finite strategic games, are the same as Lemma 3 in Pearce (1984) whose proof was different from ours, however. Pearce’s proof was built on the existence of Nash equilibria in zero sum games. As regards parts (iii) and (iv) of Theorem 1 note that an element of the relative interior of $\Delta(S_{-i})$ is a full support belief, and therefore parts (iii) and (iv) of Theorem 1 correspond to Lemma 4 in Pearce (1984). Again, the proof in Pearce (1984) is different from ours. Finally, part (v) of Theorem 1, if applied to strategic games, is a special case of Proposition 3 in Weinstein (2020), who, like Pearce, presents a proof that is built on the theory of zero-sum games.

It is well-known that the results listed above for strategic games would not be true if one considered dominance by pure strategies only, not by mixed strategies. Mixed strategies are the equivalent of the “convex combinations” in Section 2. As an example, consider the two player game in Table 2, where only player 1’s utility is shown. Player 1’s strategy $B$ is strictly dominated, but not by any pure strategy. It is strictly dominated by, for example, the mixed strategy that places probability 0.5 on $T$ and $M$.

<table>
<thead>
<tr>
<th></th>
<th>$L$</th>
<th>$R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T$</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>$M$</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>$B$</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 2. A strategy that is not strictly dominated by any pure strategy may be strictly dominated by a mixed strategy.
5. Dominance and Optimality of Experiments

In this section we apply the results of Section 2 to experiments. Let $\Omega$ be a finite set of states of the world, and let $\mu \in \Delta(\Omega)$ be a decision maker’s prior belief about the state. The decision maker can observe a signal about the state of the world before making a decision. Here, we mean by a signal a mapping: $s : \Omega \rightarrow \Delta(M_s)$ a signal, where $M_s$ is a finite set of signal realizations and $s(\omega) \in \Delta(M_s)$ is the distribution of signal realizations conditional on $\omega$. There is a finite set $\mathcal{S}$ of such signals from which the decision maker must choose one. Signals are costless.

The decision maker faces a decision problem $(A, u)$. Here, $A$ is a finite set of actions and $u : A \times \Omega \rightarrow \mathbb{R}$ is a von Neumann Morgenstern utility function. We denote by $\mathcal{A}$ the set of all such decision problems.

For every signal $s \in \mathcal{S}$ we denote by $\mu_s \in \Delta(\Delta(\Omega))$ the corresponding distribution of posterior beliefs. For every decision problem $(A, u)$ we denote by $v_{A,u} : \Delta(\Omega) \rightarrow \mathbb{R}$ the value function:

$$v_{A,u}(\nu) = \max_{a \in A} \sum_{\omega \in \Omega} (u(a, \omega)\nu(\omega)).$$

Here, $\nu$ stands for an arbitrary posterior belief of the decision maker. If the decision maker faces decision problem $(A, u)$, has access to signal $s$ before choosing an action, and chooses an action that maximizes her expected utility, she obtains ex ante expected utility:

$$\int_{\Delta(\Omega)} v_{A,u}(\nu) d\mu_s.$$

Blackwell introduced a partial order over signals. Blackwell (1951) and Blackwell (1953) showed various conditions all to be equivalent to the original definition of the Blackwell order. In Definition 8, we don’t present Blackwell’s original definition of the order, but we use one of the conditions that Blackwell showed to be equivalent to the original definition to define the Blackwell order. This is more in line with the way in which dominance orders in other areas of economics are conventionally defined.

**Definition 8.** Signal $s$ Blackwell dominates signal $\hat{s}$ if for every decision problem $(A, u)$ in $\mathcal{A}$:

$$\int_{\Delta(\Omega)} v_{A,u}(\nu) d\mu_s \geq \int_{\Delta(\Omega)} v_{A,u}(\nu) d\mu_{\hat{s}}.$$
We now explain how to fit Blackwell dominance into our framework. The set $X$ is the set of signals among which the decision maker can choose. The set $Y$ is the set of all value functions that correspond to a decision problem in $\mathcal{A}$. We endow this set with the standard vector space structure and with the topology of uniform convergence. Observe that the set $Y$ is convex. To see this note that the convex combination of two value functions corresponding to decision problem $(A, u)$ and decision problem $(A', u')$ with weights $\lambda$ and $1 - \lambda$ is the value function for the decision problem in which the decision maker chooses from $A \times A'$ and with probability $\lambda$ the first choice matters, and utility is given by $u$, and with probability $1 - \lambda$ only the second choice matters, and utility is given by $u'$.

The utility function $u(x, y)$ from Section 2 is in the setting of this section the expected utility $\int_{\Delta(\Omega)} v_{A, u}(\nu) d\mu_s$. Note that this utility function is linear and continuous in the value function.

We explain next how in this setting convex combinations of signals can be interpreted as signals in themselves. For the purposes of this discussion we assume that no two signals in $\mathcal{S}$ have overlapping message sets: $M_s \cap M\hat{s} = \emptyset$ for every $s, \hat{s} \in \mathcal{S}$ with $s \neq \hat{s}$. This is not a substantial assumption. Rather, this assumption allows us to simplify the notation in the following definition.

**Definition 9.** Suppose $\{s_1, s_2, \ldots, s_n\} \subseteq \mathcal{S}$ and assume that the vector $(\lambda_1, \lambda_2, \ldots, \lambda_n) \in \mathbb{R}^n_+$ satisfies $\sum_{i=1}^{n} \lambda_i = 1$. The convex combination of the signals $\{s_1, s_2, \ldots, s_n\}$ with weights $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ is the signal

$$s : \Omega \to \Delta \left( \bigcup_{i=1}^{n} M_{s_i} \right)$$

such that for every $\omega \in \Omega$, every $i \in \{1, 2, \ldots, n\}$, and every $m_{s_i} \in M_{s_i}$, we have:

$$s(m_{s_i} | \omega) = \lambda_i s_i(m_{s_i} | \omega).$$

Intuitively, a convex combination of the signals in set $\{s_1, s_2, \ldots, s_n\}$ with weights $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ is the following signal: The decision maker observes the message of signal $s_i$ with probability $\alpha_i$. This signal yields for the decision maker the expected utility that we attribute to the convex combination of actions in Section 2. Therefore, in our context, we can interpret a convex combination of signals as another signal.
Observe that, in the terminology of Section 2 a signal is Blackwell dominated if and only if it is either weakly dominated by a convex combination of other signals or is redundant.

Items (i), (iii) and (v) of Theorem 1 apply directly to our setting. By contrast, items (ii) and (iv) do not apply. This is because the set of all value functions generated by finite action problems is neither compact nor finite dimensional. Observe that claims (i) and (ii) are, however, vacuously true for the Blackwell order. This is because no signal ever strictly dominates another signal. This is because we have not ruled out from consideration those decision problems \((A, u) \in \mathcal{A}\) in which the utility function \(u\) does not depend on the state \(\omega\). All signals are useless in such decision problems, and therefore, no signal is ever strictly better than another signal in all decision problems. Parts (iii) and (iv) are in our setting in this section of no interest because in the setting of this section it is easily checked that the relative interior of the set \(Y\) with respect to the topology of uniform convergence is empty. We therefore focus now on part (v) of Theorem 1.

For maximum clarity we translate the definition of unique optimality, and the assertion of part (v) of Theorem 1 to our setting.

**Definition 10.** A signal \(s \in S\) is a uniquely optimal choice in decision problem \((A, u)\) if:
\[
\int_{\Delta(\Omega)} v_{A,u}(\nu) d\mu_s > \int_{\Delta(\Omega)} v_{A,u}(\nu) d\mu_{\hat{s}} \text{ for all } \hat{s} \in S \text{ with } \hat{s} \neq s.
\]

Part (v) of Theorem 1 says in our setting:

**Proposition 1.** Signal \(s \in S\) is a uniquely optimal choice in some decision problem \((A, u)\) if and only if it is not Blackwell dominated by any convex combination of signals in \(S \setminus \{s\}\).

We now show by means of an example that the result would not be true if we replace did not allow the Blackwell dominating signal to be a convex combination of the other signals, but required the Blackwell dominating signal to be one of the other signals. Consider the following example: \(\Omega = \{\omega_1, \omega_2, \omega_3\}, \mu(\omega) = \frac{1}{3}\) for all \(\omega \in \Omega\), \(S = \{s_1, s_2, s_3, s_4\}\). \(M_i = \{m_{11}^i, m_{12}^i\}\) for \(i = 1, 2, 3\), and \(M_4 = \{m_{11}^4, m_{12}^4, \ldots, m_{16}^4\}\). For each of the signals \(s_1, s_2,\) and \(s_3,\) and for each state of the world, the corresponding
row in Table 3 below indicates the conditional probability of observing each signal realization.

\[
\begin{array}{c|cc}
   & s_1 & s_2 & s_3 \\
\hline
\omega_1 & 1 & 0 & 0 \\
\omega_2 & 0 & 1 & 1 \\
\omega_3 & 0 & 1 & 1 \\
\end{array}
\]

Table 3. Conditional distributions of \(s_1, s_2, \) and \(s_3\)

Table 4 provides the same information for signal 4.

\[
\begin{array}{c|cccccc}
   & s_4 & m_4^1 & m_4^2 & m_4^3 & m_4^4 & m_4^5 & m_4^6 \\
\hline
\omega_1 & 1/4 & 0 & 0 & 3/8 & 3/8 & 3/8 & 0 \\
\omega_2 & 0 & 1/4 & 0 & 3/8 & 0 & 3/8 \\
\omega_3 & 0 & 0 & 1/4 & 0 & 3/8 & 3/8 \\
\end{array}
\]

Table 4. Conditional distributions of \(s_4\)

We claim that \(s_4\) is not Blackwell dominated by any of \(s_1, s_2, s_3,\) but that it is Blackwell dominated by the convex combination of these three signals that places weight 1/3 on each of these signals. To see that \(s_4\) is not Blackwell dominated by \(s_1\) note that signal \(s_4\) has a realization \((m_2^2)\) which reveals that the true state is \(\omega_2,\) whereas \(s_1\) has no such realization. This implies that \(s_4\) cannot be Blackwell dominated by \(s_1.\) Analogous arguments show that \(s_4\) is not Blackwell dominated by \(s_2\) or \(s_3.\)

To see that \(s_4\) is Blackwell dominated by the convex combination of signals \(s_1, s_2, s_3\) that places probability 1/3 on each of these signals we consider the distribution of
posterior beliefs generated by \( s_4 \) and compare it to the distribution of posterior beliefs generated by the convex combination. For each state in \( \Omega \) signal 4 generates with probability \( 1/12 \) a posterior belief that is a Dirac measure on this state. For each state in \( \Omega \) the convex combination of signals \( s_1, s_2, s_3 \) generates with probability \( 1/9 \) (\( > 1/12 \)) a posterior belief that is a Dirac measure on this state. Also, for each pair of states in \( \Omega \), signal 4 generates with probability \( 1/4 \) a posterior belief that places probability \( 1/2 \) on each of the two states in this pair. For each pair of states in \( \Omega \) the convex combination of signals \( s_1, s_2, s_3 \) generates with probability \( 2/9 \) (\( < 1/8 \)) a posterior belief that places probability \( 1/2 \) on two of the three states. One can now easily show that the distribution of posterior beliefs under the convex combination of signals is a mean-preserving spread of the distribution of posterior beliefs that is generated by signal \( s_4 \). By standard results, this implies that \( s_4 \) is Blackwell dominated by the convex combination of signals \( s_1, s_2, s_3 \) that places probability \( 1/3 \) on each of these signals.

Because \( s_4 \) is Blackwell dominated by a convex combination of \( s_1, s_2 \) and \( s_3 \) in every decision problem one of signals \( s_1, s_2, \) or \( s_3 \) yields at least as high expected utility as \( s_4 \). Yet \( s_4 \) is not Blackwell dominated by any single of the signals \( s_1, s_2 \) and \( s_3 \). The example in Tables 3 and 4 is therefore the analogue for signals of the example in Table 2.

We conclude this section by briefly considering briefly some alternatives to the Blackwell order and the applicability of the results in Theorem 1 to these orders. Suppose that the set of states \( \Omega \) is a lattice. We can then restrict attention to decision problems such that set of actions \( A \) is a finite lattice and the utility function \( u \) is supermodular in the product lattice on \( \Omega \times A \). Let us call such decision problems “monotone,” and let us denote the set of all such decision problems by \( \mathcal{M} \).

**Definition 11.** Signal \( s \) monotonically dominates signal \( \hat{s} \) if for every monotone decision problem \((A, u) \in \mathcal{M} \):

\[
\int_{\Delta(\Omega)} v_{A,u}(\nu) d\mu_s \geq \int_{\Delta(\Omega)} v_{A,u}(\nu) d\mu_{\hat{s}}.
\]

We can apply part (v) of Theorem 1 to conclude:

**Proposition 2.** Signal \( s \in S \) is a uniquely optimal choice in some monotone decision problem \((A, u) \in \mathcal{M} \) if and only if it is not monotonically dominated by any convex combination of signals in \( S \setminus \{s\} \).
The order that we have introduced in Definition 11 is closely related to the orders introduced in Lehmann (1988) and Athey and Levin (2018) but it does not coincide with either of these. Lehmann, and also Athey and Levin, assume the action set to be a subset of the set of real numbers, and thus they assume the action set to be completely ordered. With this assumption the argument that we used above to show that the set of value functions is convex no longer applies. This is because that argument involved creating a new decision problem from two given decision problems in which the decision maker’s action set was the Cartesian product of the original action sets. It was important that this new decision problem was included in the set of admissible decision problems. But if action sets have to be one-dimensional, this argument no longer holds. Kim (2022) allows multi-dimensional action sets but imposes joint conditions on signals and decision problems. A detailed consideration of his order is outside of the scope of this paper.

One might also modify the Blackwell order by considering not only the value functions that are generated by finite decision problems, but instead all value functions that are convex in the posterior, and in addition normalize value functions, so that every value function must assume 0 as the minimum expected utility and 1 as the maximum expected utility. Note that this rules out constant value functions. With this construction, the concept of strict dominance among signals is no longer vacuous. For example, a perfectly informative signal will strictly dominate a completely uninformative signal. A statistical characterization of this dominance relation among signals is left for future work.

6. Dominance and Optimality of Monetary Lotteries

Consider an expected utility maximizer who chooses one lottery from a finite set of lotteries. Here, a lottery is a probability distribution over \( \mathbb{R} \). One lottery first order stochastically dominates another lottery if the former lottery yields at least as high expected utility as the latter provided that the decision maker’s utility is non-decreasing in money. One lottery second order stochastically dominates another lottery if the former lottery yields at least as high expected utility as the latter provided that the decision maker’s utility is non-decreasing and concave in money.\(^7\)

\(^7\)Our definitions of the stochastic dominance orders are taken from Mas-Colell et al. (1995), where they are Definitions 6.D.1 and 6.D.2. The definitions differ slightly from the original definition of first order stochastic dominance in Quirk and Saposnik (1962) and second order stochastic dominance in Rothschild and Stiglitz (1970).
To apply the results of Section 2 we let the set $X$ of actions be the set of monetary lotteries that the decision maker can choose from, and we let the set $Y$ be the set of non-decreasing (in the case of first order stochastic dominance), or the set of non-decreasing and concave (in the case of second order stochastic dominance) utility functions with domain $\mathbb{R}$. The utility function $u$ from Section 2 is then the expected utility, that is, if $x$ is the lottery with cumulative distribution function $F$, and $y$ is the utility function $u : \mathbb{R} \to \mathbb{R}$, then

$$u(x, y) = \int u(z) dF.$$

In this specification, the set $Y$ is a convex set, although it is not compact, and has an empty relative interior. The utility function $u$ is linear in $u$ and continuous in the topology of uniform convergence.

As was the case with signals, convex combinations of lotteries have again an intuitive interpretation. The convex combination that attaches weight $\lambda$ to the lottery with cumulative distribution function $F_1$ and weight $1 - \lambda$ to the lottery with cumulative distribution function $F_2$ is the lottery with cumulative distribution function $\lambda F_1 + (1 - \lambda)F_2$. This new lottery yields exactly the expected utility specified in Section 2 for convex combinations.

A lottery is first or second-order stochastically dominated by a convex combination of other lotteries if it is either weakly dominated by the convex combination, or if it is made redundant by the convex combination of the other lotteries.

We can apply parts (i), (iii) and (v) of Theorem 1. As in the previous section, strict dominance among lotteries is a vacuous notion because we have allowed the constant utility function. Interior optimality is vacuous because the relative interior of $Y$ is empty. Therefore, we focus on part (v) of the Theorem.

We illustrate in Table 5 why part (v) of the Theorem would not hold if we had not considered convex combinations. We display three lotteries, $\ell_1$, $\ell_2$ and $\ell_3$. Each row corresponds to a Dollar amount, each column corresponds to a lottery, and the entry in the table indicates the probability of the given dollar amount.

Note that neither lottery 1 nor lottery 2 first order stochastically dominates lottery 3. However, the convex combination of lotteries 1 and 2 that attaches weight $2/3$ to lottery 1 and weight $1/3$ to lottery 2 attaches probability $1/3$ each to 0 Dollars, 1 Dollar, and 2 Dollars. This convex combination first order stochastically dominates lottery 3. Indeed, a decision maker with a strictly increasing utility function will
choose either lottery 1 or lottery 2, depending on the particular utility function, but will never choose lottery 3.

Our discussion suggests that it would be interesting to modify the framework by considering only lotteries with some given compact support, and then to allow only strictly increasing utility functions that assume 0 as the minimum value and 1 as the maximum value. The notion of strict dominance would then not be vacuous, and in addition one might try to adapt the framework so that part (ii) of Theorem 1 applies. We leave this to future work.

Table 5. Payoffs of three lotteries

<table>
<thead>
<tr>
<th></th>
<th>$\ell_1$</th>
<th>$\ell_2$</th>
<th>$\ell_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0$</td>
<td>$\frac{1}{2}$</td>
<td>0</td>
<td>$\frac{4}{9}$</td>
</tr>
<tr>
<td>1$</td>
<td>0</td>
<td>1</td>
<td>$\frac{7}{9}$</td>
</tr>
<tr>
<td>2$</td>
<td>$\frac{1}{2}$</td>
<td>0</td>
<td>$\frac{2}{9}$</td>
</tr>
</tbody>
</table>

7. Conclusion

We have investigated for various dominance orders in economics the relation between an action not being dominated and an action being optimal in some situation. Our investigation has uncovered shared properties of, as well as differences among, dominance orders in economics. Our paper has emphasized the importance of considering convex combinations of actions in dominance relations.

Our investigation has sometimes suggested modifications of the definition of dominance relations. For example, for dominance relations among signals, we have suggested a modification of the definition of monotone dominance. It remains an open problem whether one can provide simple characterizations of these modified dominance relations.

In some contexts, previous literature has perhaps not sufficiently emphasized the importance of convex combinations. In the theory of the optimal choice of investment portfolios it might be of interest to investigate “efficient” portfolios where such portfolios are defined as those that are not dominated by convex combinations of other portfolios. In the theory of information acquisition, when discussing the choice of signals of an expected utility maximizing decision maker, it seems worthwhile to
investigate which signals are undominated within particular sets of signals. In a forthcoming companion paper we shall tackle this latter problem in a setting in which a signal is two-dimensional, and the marginal distributions of signals are given and fixed. We characterize joint distributions are not Blackwell dominated by convex combinations of other signals. Results of this type yield insight into optimal choices of decision makers without relying on specific assumptions about their environment or their preferences. This seems of particular interest in the context of theoretical statistics. Our results for two-dimensional signal distributions, for example, have implications for the theory of efficient sampling methods.

References


