MANY INSPECTIONS ARE MANIPULABLE

ERAN SHMAYA

ABSTRACT. A self proclaimed expert uses past observations of a stochastic process to make probabilistic predictions about the process. An inspector applies a test function to the infinite sequence of predictions provided by the expert and the observed realization of the process in order to check the expert's reliability. If the test function is Borel and the inspection is such that a true expert will always pass it, then it is also manipulable by an ignorant expert. The proof uses Martin's theorem about determinacy of Blackwell games. Under the axiom of choice, there exist non-Borel test functions that are not manipulable.

1. INTRODUCTION

At every period $n = 0, 1, 2, \dots$ nature chooses an outcome s_n from a finite set S. An expert claims to know the underlying distribution behind nature's choices. To prove his claim, at each period n the expert provides a probabilistic prediction p_n about s_n before s_n is realized. The question addressed in this paper is whether the expert's reliability can be tested from the infinite sequence of predictions $(p_0, p_1, \dots)$ provided by the expert and the actual observed sequence $(s_0, s_1, \dots)$.

Date: January 27, 2008.
I am grateful to Federico Echenique for two rounds of detailed comments and suggestions which substantially improved the paper, and also to Chris Chambers, Eddie Dekel, Yossi Feinberg, Wojciech Olszewski, Alvaro Sandroni and Leeat Yariv.

Assume that an inspector decides on the reliability of the expert by applying some test function – a function whose arguments are the predictions $(p_0, p_1, \dots)$ made by the expert and the actual realization $(s_0, s_1, \dots)$, and whose value is either 'pass' if the predictions fit the realization and 'reject' otherwise. The calibration test is a well known example of such a test function, which turns out to be manipulable: an ignorant expert, who does not know the distribution of the process, can fake predictions that match the performance of a true expert, who predicts according to the correct distribution. In this paper I prove that every Borel test function is manipulable, without making any assumption about its form.

Olszewski and Sandroni [15], extending a previous theorem of Sandroni [17], have already proved a general manipulability result of the type sought in this paper. They consider an inspection in which the expert can be rejected only at some finite period: if he wasn't rejected at any finite period then he passes the test. This assumption is natural from an economic perspective, since a real world inspection is based on a finite data sequence. Topologically, this assumption translates to semi-continuity of the test function. Olszewski and Sandroni proved that such a test function is always manipulable[1]. My contribution is twofold: I extend Olszewski and Sandroni's result from semi-continuous test functions to arbitrary Borel functions, thus dispensing with the assumption that rejection is determined at a finite period, and I give an example of a non-Borel test function that is not manipulable. My proof

[1]In fact, their result is formulated in a wider framework, without the assumption that the predictions are only about the current period. See Section 3.3

uses Martin's theorem about determinacy of Blackwell games, which is a new tool in this literature.

The manipulability theorem of this paper is in sharp construct with results of Dekel and Feinberg [5] and Olszewksi and Sandroni [16]. These authors prove the existence of non-manipulable inspections that are based on a prediction in $\Delta(S^{\mathbb{N}})$ about the entire infinite realization of the process, which the expert announces before any data is realized. The manipulability theorem of this paper, on the other hand, relies on the fact that at every period the expert provides predictions about the current period, or, more generally, about a finite number of future periods, but not about events that are are only determined at infinity. To emphasize this point, I give an example of a non-manipulable sequential inspection, in which at every period the expert makes a prediction about a single event that is only determined at infinity.

Theorem 1 and Theorem 2 in Section 2 are the main results of this paper – every Borel test function that does not reject the truth with high probability is manipulable, and, under the axiom of choice, there exists a non-Borel test function that is not manipulable. Section 3 discusses related literature. Section 4 presents Martin's theorem . The proofs of the theorems are in Sections 5 and 6. In Section 7 I give an example of a Borel non-manipulable inspection that is based on repeated predictions about a single event. Section 8 concludes.

2. Manipulable and non-manipulable tests

Let S be a finite set. Elements of $\Delta(S)$ are called *predictions*. At every period $n = 0, 1, 2, \dots$ an *outcome* $s_n \in S$ is realized. At every period, before s_n is realized, an expert declares a prediction $p_n \in \Delta(S)$ about s_n, based on the past outcomes $s_0, \dots, s_{n-1}$. A *realization* is given by an infinite sequence $s \in S^{\mathbb{N}}$ of outcomes, where $\mathbb{N} = \{0, 1, 2, \dots\}$.

Let $S^{<\mathbb{N}} = \bigcup_{n \geq 0} S^n$ be the set of all finite sequences of elements of S, including the empty sequence e. For every realization $s = (s_0, s_1, \dots) \in S^{\mathbb{N}}$ and every $n \in \mathbb{N}$ let $s|_n = (s_0, \dots, s_{n-1})$ be the initial segment of s of length n. In particular, $s|_0 = e$.

Definition 1. A *prediction rule* is given by a function $f : S^{<\mathbb{N}} \to \Delta(S)$.

If the expert uses a prediction rule f then his prediction about s_n after observing $(s_0, \dots, s_{n-1})$ would be $f(s_0, \dots, s_{n-1})$.

Definition 2. A *test function* is a function $T : \Delta(S)^{\mathbb{N}} \times S^{\mathbb{N}} \to \{0, 1\}$.

A test function T dictates, for every infinite sequence of predictions and every realization, whether or not the expert passes the inspection:

Definition 3. Let T be a test function and $s \in S^{\mathbb{N}}$ a realization. A sequence $p \in \Delta(S)^{\mathbb{N}}$ of predictions *passes* T *over* s if $T(p, s) = 1$. A prediction rule f *passes* T *over* s if p passes T over s, where $p = (p_0, p_1, \dots)$ is the sequence of predictions of f along the realization, $p_n = f(s_0, \dots, s_{n-1})$.

Definition 4. A test function T *does not reject the truth with probability* $1 - \epsilon$ if

$$\mathbb{P}\big(f \text{ passes } T \text{ over } \Theta_0, \Theta_1, \dots\big) \geq 1 - \epsilon$$

for every sequence of random variables $\Theta_0, \Theta_1, \dots$ with values in S, where the prediction rule $f : S^{<\mathbb{N}} \to \Delta(S)$ is given by

$$f(s_0, \dots, s_{n-1})[s_n] = \mathbb{P}(\Theta_n = s_n | \Theta_0 = s_0, \dots, \Theta_{n-1} = s_{n-1}) \tag{1}$$

for every $s_0, \dots, s_n \in S$.[2]

Thus, T does not reject the truth if a true expert, who knows the distribution of the process will pass the test with high probability by predicting according to this distribution.

Definition 5. A test function T is ϵ-manipulable if there exists some probability measure ξ over prediction rules such that

$$\xi\big(\{f | f \text{ passes } T \text{ over } s\}\big) \geq 1 - \epsilon,$$

for every $s \in S^{\mathbb{N}}$.

If a test function is ϵ-manipulable then an ignorant expert can randomize his prediction rule according to ξ and pass the test with high probability, regardless of the actual realization. The following theorem shows that this situation is typical.

Theorem 1. *Let $T : \Delta(S)^{\mathbb{N}} \times S^{\mathbb{N}} \to \{0, 1\}$ be a Borel test function. If T does not reject the truth with probability $1 - \epsilon$, then T is $\epsilon + \delta$-manipulable for every $\delta > 0$.*

[2] The random variables $\Theta_0, \Theta_1, \dots$ are defined over some probability space $(\Omega, \mathcal{A}, \mathbb{P})$. I suppress the underlying probability space when using the language of random variables.

Remark 1. For the purpose of Theorem 1, the space $\Delta(S)^{\mathbb{N}} \times S^{\mathbb{N}}$ is a topological space with the product of discrete topologies over $\Delta(S)$ and S. The *sigma-algebra of Borel sets* of $\Delta(S)^{\mathbb{N}} \times S^{\mathbb{N}}$ is the sigma-algebra generated by the topology. The test function T is a Borel function if the set $\{(p,s)|T(p,s) = 1\}$ is a Borel subset of $\Delta(S)^{\mathbb{N}} \times S^{\mathbb{N}}$. Note that since the discrete topology is stronger than the natural Euclidean topology over $\Delta(S)$ it gives rise to a larger class of Borel functions.

Can a test functions $T : \Delta(S)^{\mathbb{N}} \times S^{\mathbb{N}} \to \{0, 1\}$ that is not Borel be manipulable? It follows from the proof of Theorem 1 that there is a model of set theory without the axiom of choice in which Theorem 1 is valid for an arbitrary test function (Remark 2). The next theorem shows that in ZFC there exists a test function that is not manipulable.

Theorem 2. *Let $S = \{0, 1\}$. Under the axiom of choice, there exists a test function $T : \Delta(S)^{\mathbb{N}} \times S^{\mathbb{N}} \to \{0, 1\}$ such that*

1. *T does not reject the truth with probability 1.*
2. *For every probability distribution ξ over prediction rules there exists some $s \in S^{\mathbb{N}}$ such that*

$$\xi\big(\{f|f \text{ passes } T \text{ over } s\}\big) = 0.$$

In particular, T is not ϵ-manipulable for any $\epsilon > 0$.

3. Related literature

3.1. **Calibration tests.** Calibration tests [4, 7, 9, 8, 10, 18] compare the observed frequency of events over a set of time with the average predictions over the same set of time. I follow Lehrer's paper [10].

For $p \in \Delta(S)$ and a subset V of S let $P[V] = \sum_{s \in V} p[s]$ and let $\mathbf{1}_V : S \to \{0,1\}$ be the indicator function of V.

A *simple calibration test* is given by a pair (U, C), where U and C are functions that assign for every observation $(s_0, \ldots, s_n) \in S^{<\mathbb{N}}$ subsets $C(s_0, \ldots, s_n)$ and $U(s_0, \ldots, s_n)$ of S, such that $C(s_0, \ldots, s_n) \subseteq U(s_0, \ldots, s_n)$. The interpretation is that $U(s_0, \ldots, s_n)$ is the local universe considered after $s_0, \ldots, s_n$ and, within this universe the event $C(s_0, \ldots, s_n)$ is checked. The simple calibration test $T^{U,C} : \Delta(S)^{\mathbb{N}} \times S^{\mathbb{N}} \to \{0,1\}$ induced by (U, C) checks whether the conditional probability attached by the expert to the events $C(s_0, \ldots, s_n)$ given $U(s_0, \ldots, s_n)$ matches the empirical relative frequency. Formally,

$$T^{U,C}(p,s) = 1 \text{ if } \sum_{n=0}^{\infty} \mathbf{1}_{U(s|_n)}(s_n) = \infty \text{ implies}$$

$$\lim_{n \to \infty} \frac{\sum_{i=0}^{n} \mathbf{1}_{U(s|_i)}(s_i) \cdot \left(p_i[U(s|_i)] \mathbf{1}_{C(s|_i)}(s_i) - p_i[C(s|_i)]\right)}{\sum_{i=0}^{n} \mathbf{1}_{U(s|_i)}(s_i)} = 0,$$

for every realization $s = (s_0, s_1, \ldots) \in S^{\mathbb{N}}$ and every infinite sequence of predictions $p = (p_0, p_1, \ldots) \in \Delta(S)^{\mathbb{N}}$. Then $T^{U,C}$ is a Borel test function that does not reject the truth with probability 1.

A more general calibration test is given by a mixture of simple calibration tests: Let $\mathcal{S}$ be the set of simple calibration tests, and let λ be a probability distribution over $\mathcal{S}$. Assume the inspector first chooses a simple calibration test $T \in \mathcal{S}$ using λ and then applies this test to the expert's predictions. For every such λ, Lehrer constructs a prediction rule that passes the inspection λ-almost surely. Lehrer's result relates to Theorem 1 of this paper in the following way: Let

$\Lambda : \Delta(S)^{\mathbb{N}} \times S^{\mathbb{N}} \to \{0,1\}$ be the test function that is given by

$$\Lambda(p,s) = 1 \text{ if and only if } \int T(p,s)\lambda(\mathrm{d}T) = 1.$$

Then Λ is a Borel test function that does not reject the truth with probability 1. By Theorem 1, it is ϵ-manipulable for every ϵ. Thus, there exists a manipulation scheme that passes the calibration test with high probability. For the case of calibration tests, Lehrer proves a stronger result: He shows that the test is 0-manipulable and that the manipulation scheme is *pure* – there exists a prediction rule f such that f passes the test Λ over s for every s[3]. Most importantly, Lehrer constructs the manipulating rule, while my proof is not constructive.

3.2. **Infinite-horizon predictions.** According to the setup of Section 2, at each period the expert has to provide a prediction about the outcome of that period. Dekel and Feinberg [5] and Olszewski and Sandroni [16] consider a different framework, in which at the start of the inspection the expert must inform the inspector his prediction about the entire realization of the process. A test function in this context is a function $t : \Delta(S^{\mathbb{N}}) \times S^{\mathbb{N}} \to \{0,1\}$, where $\Delta(S^{\mathbb{N}})$ is the set of all probability measures over the set $S^{\mathbb{N}}$ of realizations. I call elements of $\Delta(S^{\mathbb{N}})$ *infinite-horizon predictions*, denote tests that are based on infinite horizon predictions by lower case t, and call them *infinite-horizon tests*. When necessary to avoid confusion, I call elements of $\Delta(S)$ *one-period predictions*, and tests a la Definition 2 *one-period test*.

[3]Other calibration tests do not admit a pure manipulation [18, Example 2.1].

There is a natural correspondence $f \leftrightarrow \mu_f$ between prediction rules according to Definition 1 and probability measures over S^∞: For every prediction rule f, μ_f is the joint distribution of a sequence $\Theta_0, \Theta_1, \dots$ of random variables satisfying (1). Because of this correspondence, every one-period test function $T : \Delta(S)^{\mathbb{N}} \times S^{\mathbb{N}} \to \{0,1\}$ naturally induces[4] an infinite-horizon test function t such that $t(\mu_f, s) = 1$ iff f passes T over s, but the converse is not true. The papers [5] and **??** prove the existence of infinite horizon test functions that do not reject the truth with probability 1 and that are not manipulable. In light of Theorem 1 it should also be mentioned that the test function constructed by Olszewski and Sandroni is a Borel function, when $\Delta(S^\infty)$ is equipped with its standard Borel structure. In Section 7 I discuss in detail the difference between finite horizon and infinite horizon predictions.

3.3. **Future independent tests.** Olszewski and Sandroni [15] consider (infinite horizon) test functions $t : \Delta(S^{\mathbb{N}}) \times S^{\mathbb{N}} \to \{0,1\}$ of the form

$$t(\mu_f, s) = \begin{cases} 0, & \text{if } (f|_n, s|_n) \in R \text{ for some } n \in \mathbb{N}, \\ 1, & \text{otherwise.} \end{cases} \tag{2}$$

for some $R \subseteq \bigcup_{n \geq 0} \left(\Delta(S)^{\bigcup_{i=0}^{n} S^i} \times S^n \right)$, where $f|_n$ is the restriction of f to $\Delta(S)^{\bigcup_{i=0}^{n-1} S^i}$ for every prediction rule f and $s|_n = (s_0, \dots, s_{n-1})$ for every realization $s = (s_0, s_1, \dots) \in S^{\mathbb{N}}$. The underlying assumption is that rejection must occur at some period n, and that the decision

[4]There is a minor inaccuracy here because μ_f does not determine f uniquely. Cantankerous readers can assume that $T(p, s) = 0$ whenever $p_n[s_n] = 0$ for some $n \in \mathbb{N}$, so that T determines t uniquely.

of the inspector at that period depends only on the segment of the data $s|_n$ that was realized before that period and on the prediction rule before that period. The set R (the rejection set) consists of all finite segments of realizations and prediction rules that are considered to be inconsistent with each other.

Test functions of the form (2) are called *future independent*: The decision whether to reject an infinite-horizon prediction μ_f at some period does not depend on predictions made by f at later periods. For future independent tests, Olszewski and Sandroni prove an analogue of Theorem 1: A future independent test that does not reject the truth with high probability is manipulable.

It is interesting to compare the theorem of Olszewski and Sandroni with Theorem 1 of this paper. Consider first a one-period test function $T : \Delta(S)^{\mathbb{N}} \times S^{\mathbb{N}} \to \{0, 1\}$. If the infinite-horizon test function induced by T is future independent, then the set $\{(p, s)|T(p, s) = 1\}$ is closed (Recall Remark 1) and, in particular, Borel. Therefore in the framework of test function studies in the paper, the scope of Theorem 1 is wider than that of Olszewski and Sandroni's theorem. However, there are future independent, infinite-horizon tests that are not induced by one-period test functions. In fact, neither of the theorems is contained in the other. Olszewski and Sandroni's inspector is more restricted in that he must decide to reject using a finite number of predictions. On the other hand, he can use the entire prediction rule up to the rejection point, including predictions conditioned on observation outside the actual realization.

3.4. **Comparative testing of experts.** Two recent papers study a different setup, in which several experts are tested together by making simultaneous predictions about the same process. Feinberg and Stewart [6] construct a cross-calibration test that does not reject a true expert, and that, in the presence of a true expert, is non-manipulable in a strong sense by ignorant experts. Al-Najjar and Weinstein [1] show that an ignorant expert passes the likelihood ratio tests against a true expert only if his predictions happen to be close to those of the true expert. They apply their result in a Bayesian setting, where the true distribution of the process is chosen randomly according to some prior over $\Delta(S^{\mathbb{N}})$. When tested separately, an ignorant expert who predicts according to the induced distribution over $S^{\mathbb{N}}$ passes the inspection. On the other hand, if the prior over $\Delta(S^{\mathbb{N}})$ is diffusive enough, then the likelihood ratio against the true expert is non-manipulable.

4. Blackwell games

A *Blackwell game* is a two-player zero-sum game that is given by (A, B, r) where A and B are the sets of *actions* of player 1 (the maximizer) and 2 (the minimizer) respectively and $r : (A \times B)^{\mathbb{N}} \to [0, 1]$ is the *payoff function*.

The game is played as follows: At every stage $n = 0, 1, 2, \ldots$ both players, simultaneously and independently, choose an action. At the end of the stage, each player is informed of his opponent's action. Let a_n, b_n be the actions chosen by player 1 and 2 respectively at stage n. The payoff that player 2 pays player 1 is given by $r(a_0, b_0, a_1, b_1, \ldots)$.

Let $\mathcal{H} = (A \times B)^{<\mathbb{N}} = \bigcup_{n \geq 0}(A \times B)^n$ be the set of *finite histories* of the game, including the empty history e. A behavioral strategy σ of player 1 is given by $\sigma : \mathcal{H} \to \Delta(A)$. Behavioral strategies τ of player 2 are defined analogously. Every pair σ, τ of behavioral strategies induces a probability distribution $\mu_{\sigma,\tau}$ over the set $(A \times B)^{\mathbb{N}}$ of *infinite histories* or *plays*. Let $R(\sigma, \tau) = \int r \, \mathrm{d}\mu_{\sigma,\tau}$ be the expected payoff in the game if the players play according to σ, τ.

Determinacy. The *upper value* $\overline{V}(G)$ and the *lower value* $\underline{V}(G)$ of G of a Blackwell game G are given by

$$\overline{V}(G) = \inf_{\tau} \sup_{\sigma} R(\sigma, \tau).$$

$$\underline{V}(G) = \sup_{\sigma} \inf_{\tau} R(\sigma, \tau),$$

where the suprema range over all behavioral strategies σ of player 1 and the infima over all behavioral strategies τ of player 2. A strategy σ of player 1 is δ-*optimal* if $R(\sigma, \tau) \geq \underline{V}(G) - \delta$ for every behavioral strategy τ of player 2. The game G is *determined* if $\underline{V}(G) = \overline{V}(G)$. Blackwell [2, 3] proved the determinacy of Blackwell games (which he called infinite games with imperfect information) with a payoff function that is the indicator function of a G_δ set, and conjectured that every Blackwell game with Borel payoff function is determined. Vervoort [19] advanced higher in the Borel hierarchy, proving determinacy for indicators of $G_{\delta\sigma}$ sets. The conjecture was proved by Donald A. Martin in 1998 [13] (See also Maitra and Sudderth's paper [12] for applications to stochastic games).

Martin's Theorem. *Let A, B be two countable sets, at least one of which is finite, and let $r : (A \times B)^{\mathbb{N}} \to [0,1]$ be a Borel function. Then the Blackwell game (A, B, r) is determined.*

Random plays. Let (σ, τ) be a pair of behavioral strategies in a the Blackwell game (A, B, r). A (σ, τ)-*random play*, is a sequence

$$\alpha_0, \beta_0, \dots, \alpha_n, \beta_n, \dots$$

of random variables over some probability space, where the values of α_n (respectively β_n) are in A (respectively B) such that

$$(3) \quad \mathbb{P}\left(\alpha_n = a, \beta_n = b | \alpha_0, \beta_0, \dots, \alpha_{n-1}, \beta_{n-1}\right) = \sigma(\alpha_0, \beta_0, \dots, \alpha_{n-1}, \beta_{n-1})[a] \cdot \tau(\alpha_0, \beta_0, \dots, \alpha_{n-1}, \beta_{n-1})[b]$$

for every $a \in A$ and $b \in B$.

The measure $\mu_{\sigma,\tau}$ that is induced by (σ, τ) over $(A \times B)^{\mathbb{N}}$ is the joint distribution of some (σ, τ)-random play. The payoff function associated with a pair of behavioral strategies (σ, τ) can also be written in terms of random plays: $R(\sigma, \tau) = \mathbb{E}\left(r\left(\alpha_0, \beta_0, \alpha_1, \beta_1, \dots\right)\right)$.

Pure and mixed strategies. A *pure strategy* of player 1 in the Blackwell game (A, B, r) is given by a function $f : B^{<\mathbb{N}} \to A$: for every sequence $b_0, \dots, b_n$ of past actions of player 2, $f(b_0, \dots, b_n)$ player 1's action at stage $n+1$. Every pure strategy is in particular a Behavioral strategy. A *mixed strategy* of player 1 is a probability distribution over pure strategies. Kuhn's Theorem establishes the equivalence between behavioral and mixed strategies. In particular, for every $\delta > 0$ player

1 has a δ-optimal mixed strategy in every Blackwell game, i.e. a mixed strategy ξ such that $R(\xi, g) \geq \underline{V}(G) - \delta$ for every behavioral strategy τ of player 2, where $R(\xi, \tau) = \int R(f, g)\xi(\mathrm{d}f)$ is the expected payoff for player 1 under ξ, τ.

5. Proof of Theorem 1

Let $\Delta^Q(S) = \{p \in \Delta(S) | p[s] \in \mathbb{Q} \text{ for every } s \in S\}$ be the set of elements of $\Delta(S)$ with rational values. For a test function T let $G(T)$ be the Blackwell game in which the set of action of player 1 is $\Delta^Q(S)$, the set of actions of player 2 is S and the payoff function is the restriction of T to $(\Delta^Q(S) \times S)^{\mathbb{N}}$. Note that every pure strategy of player 1 in $G(T)$ is a prediction rule according to Definition 1. Roughly speaking, player 1 represents the expert and player 2 represents nature. However, in the game $G(T)$ player 2 is allowed to condition his actions on past actions of player 1 (as if nature picks the value of s_n depending on previous predictions made by the expert) and player 1 is only allowed to make predictions with rational values[5].

The game $G(T)$ satisfies the assumptions of Martin's theorem: the action set of player 1 is finite, the action set of player 2 is countable, and the payoff function is Borel. Therefore $\overline{V}(G(T)) = \underline{V}(G(T))$. The following two lemmas complete the proof of Theorem 1.

[5] Olszewski and Sandroni [16] use another game theoretic representation of the expert's problem. Their game is a normal form one-shot game in which nature chooses a realization and the expert chooses a prediction rule. They use topological properties of the test function to deduce the determinacy of the game using the classical minimax theorem.

Lemma 1. *If T does not reject the truth with probability $1-\epsilon$ then $\overline{V}(G(T)) \geq 1-\epsilon$.*

Lemma 2. *T is $1-\underline{V}(G(T))+\delta$ manipulable for every $\delta>0$*

The proof of Lemma 1 uses the following lemma. Recall that, for a finite set S and $p, p' \in \Delta(S)$ a *coupling of* (p, p') is a pair (Θ, Θ') of random variables such that $\mathbb{P}(\Theta = s) = p[s]$ and $\mathbb{P}(\Theta' = s) = p'[s]$ for every $s \in S$, i.e. the marginal distributions of Θ and Θ' are p and p' resp.

Coupling Lemma. [11, Chapter 1, Theorem 5.2] *Let S be a finite set and let $p, p' \in \Delta(S)$. Then there exists a coupling (Θ, Θ') of (p, p') such that $\mathbb{P}(\Theta \neq \Theta') = \|p - p'\|_1/2$.*[6]

Proof of Lemma 1. Let τ be a behavioral strategy for player 2 in $G(T)$. We have to construct a good response for player 1 against τ. The strategy will be such that at every stage player 1 predicts the action of player 2 for that stage. Note that since τ is given, at every stage player 1 knows the probability distribution according to which player 2 is going to choose an action. However, since in $G(T)$ player 1 is only allow to make predictions with rational values, his strategy will only approximate this distribution.

Let $\delta > 0$ and let $f, f' : S^{<\mathbb{N}} \to \Delta(S)$ be the prediction rules defined inductively as follows: for every $(s_0, \ldots, s_n) \in S^{<\mathbb{N}}$ let

$$f(s_0, \ldots, s_n) = \tau(p_0, \ldots, p_n, s_0, \ldots, s_n), \tag{4}$$

[6] I essentially use the coupling lemma to prove that prediction rules with rational image give rise to a set of probability measures over S^∞ that is dense in the norm topology. Cf. Lemma 3 in Olszewski and Sandroni's paper [15].

where $p_i = f'(s_0, \dots, s_{i-1})$ and let $f'(s_0, \dots, s_n) \in \Delta^Q(S)$ be such that

$$\|f'(s_0, \dots, s_n) - f(s_0, \dots, s_n)\|_1 < \delta/2^n. \tag{5}$$

Then f' is a pure strategy of player 1 in $G(T)$. I am going to construct a (f', τ)-random play $(\Pi_0, \Theta_0, \Pi_1, \Theta_2, \dots)$ and, on the same probability space, a stochastic process $(\Theta'_0, \Theta'_1, \dots)$ that equals $(\Theta_0, \Theta_1, \dots)$ with high probability, such that f' is the correct prediction rule for $(\Theta'_0, \Theta'_1, \dots)$.

Let $\Theta = (\Theta_n)_{n \in \mathbb{N}}$ and $\Theta' = (\Theta'_n)_{n \in \mathbb{N}}$ be random variables over some probability space defined inductively such that, for $s_0, s'_0, \dots, s_{n-1}, s'_{n-1} \in S$, the conditional joint distribution of Θ_n, Θ'_n given the event $\{\Theta_i = s_i, \Theta'_i = s'_i \text{ for } 1 \le i < n\}$ satisfies

$$\mathbb{P}\left(\Theta_n = s | \Theta_i = s_i, \Theta'_i = s'_i \text{ for } 0 \le i < n\right) = f(s_0, \dots, s_{n-1})[s], \text{ and} \tag{6}$$

$$\mathbb{P}\left(\Theta'_n = s' | \Theta_i = s_i, \Theta'_i = s'_i \text{ for } 0 \le i < n\right) = f'(s'_0, \dots, s'_{n-1})[s'], \tag{7}$$

and, if $s_i = s'_i$ for $0 \le i < n$ then also

$$\mathbb{P}\left(\Theta_n \ne \Theta'_n | \Theta_i = s_i, \Theta'_i = s'_i \text{ for } 0 \le i < n\right) \le \delta/2^n. \tag{8}$$

If $s_i = s'_i$ for $0 \le i < n$ then the existence of a pair of random variables that satisfy (6),(7),(8) follows from (5) and the coupling lemma. If $s_i \ne s'_i$ for some $0 \le i < n$ then the conditional joint distribution of Θ_n, Θ'_n given the event $\{\Theta_i = s_i, \Theta'_i = s'_i \text{ for } 0 \le i < n\}$ can be chosen arbitrarily with the marginals given by (6) and (7). Note that from (6) and (7) it follows that

$$\mathbb{P}\left(\Theta_n = s | \Theta_0 = s_0, \dots, \Theta_{n-1} = s_{n-1}\right) = f(s_0, \dots, s_{n-1})[s], \text{ and} \tag{9}$$

$$\mathbb{P}\left(\Theta'_n = s' | \Theta'_0 = s'_0, \dots, \Theta'_{n-1} = s'_{n-1}\right) = f'(s'_0, \dots, s'_{n-1})[s']. \tag{10}$$

Also, from (8) it follows that

$$\mathbb{P}(\Theta_n \neq \Theta'_n | \Theta_i = \Theta'_i \text{ for } 0 \leq i \leq n) \leq \delta/2^n,$$

and therefore

$$\mathbb{P}(\Theta \neq \Theta') \leq \sum_{n \in \mathbb{N}} \mathbb{P}(\Theta_n \neq \Theta'_n | \Theta_i = \Theta'_i \text{ for } 0 \leq i \leq n) \leq 2\delta. \tag{11}$$

Let $\Pi = (\Pi_n)_{n \in \mathbb{N}}$ be given by

$$\Pi_n = f'(\Theta_0, \dots, \Theta_{n-1}). \tag{12}$$

Then it follows from (4), (9) and (12) that

$$\mathbb{P}(\Theta_n = s | \Pi_0, \Theta_0, \dots, \Pi_{n-1}, \Theta_{n-1}) = \tau(\Pi_0, \Theta_0, \dots, \Pi_{n-1}, \Theta_{n-1})[s]. \tag{13}$$

From (12) and (13) it follows that

$$(\Pi_0, \Theta_0, \Pi_1, \Theta_1, \dots) \text{ is a } (f', \tau)\text{-random play in } G(T). \tag{14}$$

Therefore

$$\begin{aligned}R(f',\tau) =&\mathbb{E}\big(T(\Pi_0,\Theta_0,\Pi_1,\Theta_1,\dots)\big) = \\ &\mathbb{P}\big(\Pi \text{ passes } T \text{ over } \Theta\big) \geq \mathbb{P}\big(\Pi \text{ passes } T \text{ over } \Theta'\big) - \mathbb{P}(\Theta \neq \Theta') = \\ &\mathbb{P}\big(f' \text{ passes } T \text{ over } \Theta'\big) - \mathbb{P}(\Theta \neq \Theta') \geq \\ &1-\epsilon-\mathbb{P}(\Theta \neq \Theta') \geq 1-\epsilon-2\delta,\end{aligned}$$

where the first equality follows from (14), the second equality from Definition 3, the first inequality from the fact that

$$\{\Pi \text{ passes } T \text{ over } \Theta'\} \subseteq \{\Pi \text{ passes } T \text{ over } \Theta\} \cup \{\Theta' \neq \Theta\},$$

the third equality from (12) and Definition 3, the second inequality from (10) and the fact that T does not reject the truth with probability $1-\epsilon$, and the third inequality from (11). Thus, for every strategy τ of player 2 and every $\delta > 0$ we built a pure strategy f' of player 1 such that $R(f',\tau) \geq 1-\epsilon-2\delta$. Therefore $\overline{V}(G(T)) \leq 1-\epsilon$ as desired. □

Proof of Lemma 2. Let ξ be a mixed δ-optimal strategy for player 1 in $G(T)$. We claim that ξ, viewed as a distribution over prediction rules, $1-\underline{V}(G(T))+\delta$-manipulates T. Indeed, let s be a realization and let g be the pure strategy of player 2 in $G(T)$ that is given by $g(p_0,\dots,p_{n-1}) = s_n$ for every $p_0,\dots,p_{n-1} \in \Delta^Q(S)$.

Let f be a pure strategy of player 1 in $G(T)$. Then it follows from Definition 3 that

$$R(f,g) = \begin{cases} 1, & \text{if } f \text{ passes } T \text{ over } s, \\ 0, & \text{otherwise.}\end{cases}$$

It follows that

$$\underline{V}(G(T)) - \delta \leq R(\xi, g) = \int R(f, g)\xi(\mathrm{d}f) = \xi\left\{f | f \text{ passes } T \text{ over } s\right\},$$

as desired. □

Remark 2. There is a model of set theory without the axiom of choice in which every Blackwell game is determined [13, Theorem 13]. In this model every set is universally measurable, which makes Definitions 4 and 5 meaningful for an arbitrary test function T. It follows from the proof that in such a model Theorem 1 is valid for an arbitrary test function.

6. Proof of Theorem 2

The test is a modification of the non-manipulable (infinite-horizon) test of Dekel and Feinberg [5, Proposition 2]. A subset M of $\{0,1\}^{\mathbb{N}}$ is *universally null* if M is universally measurable and $\mu(M^c) = 1$ for every non-atomic probability measure μ over $\{0,1\}^{\mathbb{N}}$. It follows from the axiom of choice that there exist universally null sets in $\{0,1\}^{\mathbb{N}}$ of cardinality $\aleph_1$[14, Theorem 5.3][7]. Such a set cannot be a Borel set. Note that for a universally null set M and an arbitrary probability measure μ one has

$$\mu(M^c \cup A(\mu)) = 1, \tag{15}$$

where $A(\mu) = \{s \in \{0,1\}^{\mathbb{N}} | \mu(\{s\}) > 0\}$ is the set of atoms of μ.

[7]Dekel and Feinberg use the fact that every Lusin set is universally null. Existence of Lusin set follows from the continuum hypothesis

Let $S=\{0,1\}$ and Let M be an uncountable universally null subset of $S^{\mathbb{N}}$. Let $T:\Delta(S)^{\mathbb{N}}\times S^{\mathbb{N}}\rightarrow\{0,1\}$ be the test function that is given by

$$T(p,s)=\begin{cases}0, & \text{if } s\in M \text{ and } \prod_{n\in\mathbb{N}}p_n[s_n]=0,\\ 1, & \text{otherwise.}\end{cases}$$

I claim that T satisfies the requirements of Theorem 2.

Let $f:S^{<\mathbb{N}}\rightarrow\Delta(S)$ be a prediction rule, and let $s\in S^{\mathbb{N}}$. Let $\Theta_0,\Theta_1,\dots$ be a sequence of random variables satisfying (1), and let $\mu_f\in\Delta(S^{\mathbb{N}})$ be their joint distribution. Then

$$\mu_f(\{s\})=\mathbb{P}\big(\Theta_n=s_n \text{ for every } n\in\mathbb{N}\big)= \prod_{n\in\mathbb{N}}\mathbb{P}\big(\Theta_n=s_n|\Theta_i=s_i \text{ for } 0\leq i<n\big)=\prod_{n\in\mathbb{N}}p_n[s_n], \tag{16}$$

where $p_n=f(s_0,\dots,s_{n-1})$.

By the last equation and the definition of T it follows that f passes T on $s\in M$ if and only if $\mu_f(\{s\})>0$, i.e. $s\in A(\mu_f)$. The rest of the argument is the same as in Dekel and Feinberg's paper: For every prediction rule f

$$\mu_f\big(\{s|f \text{ passes } T \text{ over } s\}\big)=\mu_f(M^c\cup A(\mu_f))=1$$

(the last equality follows from (15)), and therefore T does not reject the truth with probability 1.

To prove the second assertion of Theorem 2, let ξ be a probability distribution over prediction rules and let $s \in M$. If

$$\xi\big(\{f|f \text{ passes } T \text{ over } s\}\big) = \xi\big(\{f|\mu_f(\{s\}) > 0\}\big) > 0,$$

then in particular $\bar{\xi}(\{s\}) > 0$ where $\bar{\xi} \in \Delta(S^{\mathbb{N}})$ is the barycenter of ξ, given by

$$\bar{\xi}(B) = \int \mu_f(B)\xi(\mathrm{d}f)$$

for every Borel subset B of $\Delta(S^{\mathbb{N}})$. Thus ξ will pass the test with some positive probability over $s \in M$ only if $s \in A(\bar{\xi})$. Since M is uncountable and $A(\bar{\xi})$ is countable, it follows that there are some $s \in M$ over which ξ passes the test with probability 0.

Remark 3. The test constructed by Olszewski and Sandroni [16] has the stronger property that for every randomly generated prediction rule ξ, the set of all realizations s over which ξ passes tests with positive probability is a set of first Baire category.

7. Finite horizon predictions and predictions about a finite set

Theorem 1, which deals with one-period tests, can be generalized to tests with finite horizon: Let $k \geq 1$ be a natural number. Consider an inspection in which, at every period n, the expert provides a probabilistic prediction about $(s_n, s_{n+1}, \ldots, s_{n+k-1})$. A *$k$-horizon prediction* is an element of $\Delta(S^k)$. A *k-horizon test function* is a function $T : \Delta(S^k)^{\mathbb{N}} \times S^{\mathbb{N}} \to \mathbb{R}$. A k-horizon prediction rule is a function $f : S^{<\mathbb{N}} \to \Delta(S^k)$. Definitions 4 and 5 extend in an obvious way to the case of k-horizon test functions, and the analogue of Theorem 1 is also true: If a k-horizon test function does not reject the truth with probability $1-\epsilon$ then it is $\epsilon+\delta$-manipulable for every $\delta > 0$. Moreover, the number k need not be constant or bounded, and can depend on past realizations and predictions.

On the other hand, as Dekel and Feinberg and Olszewski and Sandroni show, infinite horizon tests can be non-manipulable. In order to emphasize that the manipulability result of this paper relies on the fact that the predictions requested from the expert are about events in the finite horizon, and not just on the fact that the set S^k over which predictions are made at each period is finite, consider the following situation: Fix a Borel set $B \subseteq S^{\mathbb{N}}$. Assume that at every period n, given the partial realization $(s_0, \ldots, s_{n-1})$ observed at that period, the expert is asked to make a prediction $q_n \in [0,1]$ about whether or not the event B will occur, that is whether or not the infinite realization s is in B. Consider the test function $T_B : [0,1]^{\mathbb{N}} \times S^{\mathbb{N}} \to \{0,1\}$ that is given by

$$T_B(q,s) = \begin{cases} 1, & \text{if } \lim_{n\to\infty} q_n = \mathbf{1}_B(s) \\ 0, & \text{otherwise.} \end{cases}$$

A *prediction rule about* B is a function $f : S^{<\mathbb{N}} \to [0,1]$. A prediction rule passes the test T_B over realization $s \in S^{\mathbb{N}}$ if $T_B(q,s) = 1$ where $q \in [0,1]^{\mathbb{N}}$ is given by $q_n = f(s_0, \ldots, s_{n-1})$. It follows from the martingale convergence theorem that the test T_B does not reject the truth with probability 1, that is, for every stochastic process $\Theta_0, \Theta_1, \ldots$ with values in S one has

$$\mathbb{P}(f \text{ passes } T_B \text{ on } \Theta_0, \Theta_1, \dots) = 1$$

where f is the prediction rule that is given by

$$f(s_0, \dots, s_{n-1}) = \mathbb{P}\big((\Theta_0, \Theta_1, \dots) \in B | \Theta_0 = s_0, \dots, \Theta_{n-1} = s_{n-1}\big).$$

Note that in the inspection induced by T_B the expert is always asked to state a prediction about only two possibilities – either B occurs or B does not occur. Still, as I show in the following example, T_B need not be manipulable when B is an event in the infinite horizon.

Example 1. Let $B \subseteq S^{\mathbb{N}}$ be a Borel set that is not an F_σ set. Then the test T_B is not ϵ-manipulable for any $\epsilon < 1/2$.

Indeed, Let ξ be a probability measure over prediction rules. Let

$$\overline{B} = \left\{ s \in S^{\mathbb{N}} \middle| \xi \left(\left\{ f | \liminf_{n \to \infty} f(s|_n) > 1/2 \right\} \right) > 1/2 \right\}.$$

Since for every $s \in S^{\mathbb{N}}$ one has

$$\left\{ f | \liminf_{n \to \infty} f(s|_n) > 1/2 \right\} = \bigcup_{r=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} \{ f | f(s|_k) > 1/2 + 1/r \}$$

it follows that that

$$\begin{aligned} s \in \overline{B} &\leftrightarrow \exists r, n \quad \xi \left(\bigcap_{k=n}^{\infty} \{ f | f(s|_k) > 1/2 + 1/r \} \right) > 1/2 \\ &\leftrightarrow \exists r, n, t \, \forall m \geq n \quad \xi \left(\bigcap_{k=n}^{m} \{ f | f(s|_k) > 1/2 + 1/r \} \right) > 1/2 + 1/t. \end{aligned}$$

Therefore

$$\overline{B} = \bigcup_{r,n,t=1}^{\infty} \bigcap_{m=n}^{\infty} \overline{B}(r,n,t,m),$$

where

$$\overline{B}(r,n,t,m) = \left\{ s \middle| \xi \left(\bigcap_{k=n}^{m} \{f | f(s|_k) > 1/2 + 1/r\} \right) > 1/2 + 1/t \right\}$$

Since the sets $\overline{B}(r,n,t,m)$ are clopen (membership to $\overline{B}(r,n,t,m)$ depends only on finite number of coordinates) it follows that $\overline{B}$ is an F_σ set. By the choice of B it follows that $B \neq \overline{B}$. If s is any element in the symmetric difference of B and $\overline{B}$ then $\xi\left(\{f|f \text{ passes } T_B \text{ on } s\}\right) \leq 1/2$. In particular, T_B is not ϵ-manipulable for any $\epsilon < 1/2$.

Remark 4. Let $S = \{0,1\}$. As an example of a Borel subset B of $S^{\mathbb{N}}$ that is not an F_σ set one can take $B = \{s \in S^{\mathbb{N}} | \sum_{n\in\mathbb{N}} s_n = \infty\}$, the set of realizations with infinitely many 1's.

8. Conclusions

The inspections considered in this paper are *sequential*: they require the expert to announce at every period a probabilistic prediction in $\Delta(P)$ for some finite set P of possibilities. In one-period inspections $P = S$, the set of possible outcomes in the period. In k-horizon inspections (Section 7) $P = S^k$, the set of possible outcomes in the k-s next periods. In inspections T_B about an event B (Section 7), P is a set of two possibilities: 'B occurs' and 'B does not occur'. In contrast, the inspections studied by Dekel and Feinberg [5] and Olszewski and Sandroni [16] require the expert to provide one prediction in $\Delta(S^{\mathbb{N}})$ about the entire realization of the process.

The paper bears good news and bad news to the inspector: There exist sequential inspections that are not manipulable (Theorem 2 and Example 1). However, such inspections must be either non-Borel or rely on predictions about events that are not determined at any finite period.

References

[1] Nabil I. Al-Najjar and Jonathan Weinstein. Comparative testing of experts. mimeo, 2007.

[2] David Blackwell. Infinite G_δ games with imperfect information. *Matematwyki Applicationes Mathematicae*, 10, 1969.

[3] David Blackwell. Operator solution of infinite G_δ games of imperfect information. In D.L. Igelhart T.W. Anderson, K.B. Athreya, editor, *Probability, statistics, and mathematics: Papers in honor of S. Karlin*. Acamedmic Press, New York, 1989.

[4] A.P. Dawid. The well calibrated bayesian. *Journal of the American Statistical Association*, 77, 1982.

[5] Eddie Dekel and Yossi Feinberg. Non-bayesian testing of a stochastic prediction. *Review of Economic Studies*, 73, 2006.

[6] Yossi Feinberg and Colin Stewart. Testing multiple forecasters. mimeo, 2007.

[7] Dean P. Foster and Rakesh Vohra. Calibrated learning and correlated equilibrium. *Games and Economic Behavior*, 21, 1997.

[8] Drew Fudenberg and David K. Levine. An easier way to calibrate. *Games and Economic Behavior*, 29:131–137, 1999.

[9] Ehud Kalai, Ehud Lehrer, and Rann Smorodinsky. Calibrated forecasting and merging. *Games and Economic Behavior*, 29, 1999.

[10] Ehud Lehrer. Any inspection is manipulable. *Econometrica*, 69, 2001.

[11] Thomas Lindwall. *Lectures on the Coupling Method*. John Wiley & Sons, 1992.

[12] A. Maitra and W. sudderth. Finitely additive stochastic games with Borel measurable payoff. *International Journal of Game Theory*, 1998.

[13] Donald A. Martin. The determinacy of Blackwell games. *The Journal of Symbolic Logic*, 63, 1998.

[14] A. W. Miller. Special subsets of the real line. In Kenneth Kunen and Jerry E. Vaughan, editors, *Handbook of set-theoretic topology*. North-Holland, 1984.

[15] Wojciech Olszewski and Alvaro Sandroni. Future independent tests. mimeo, 2006.

[16] Wojciech Olszewski and Alvaro Sandroni. A nonmanipulable test. mimeo, 2007.

[17] Alvaro Sandroni. The reproducible properties of correct forecasts. *Internat. J. Game Theory*, 32(1):151–159, 2003.

[18] Alvaro Sandroni, Rann Smorodinsky, and Rakesh V. Vohra. Calibration with many checking rules. *Mathematics of Operation Research*, 28, 2003.

[19] Marco R. Vervoort. Blackwell games. In T. S. Fergurson L. S. Shapley and J. B. MacQueen, editors, *Statistics, probability and game theory: Papers in honor of David Blackwell*. Institute of Mathematical Statistics, 1996.

E-mail address: `eshmaya@ist.caltech.edu`

Information Science and Technology, California Institute of Technology