ANY INSPECTION IS MANIPULABLE

by

Ehud Lehrer*
School of Mathematical Sciences
Sackler Faculty of Exact Sciences
Tel Aviv University, Ramat Aviv
Tel Aviv 69978, Israel
e-mail: lehrer@math.tau.ac.il

First version: January 1997
March 11, 2001

Abstract. A forecaster provides a probabilistic prediction regarding the following day's state of nature. To examine the forecaster, an inspector employs calibration tests that compare the average prediction and the empirical frequency of pre-specified events. This paper shows that any mixed test can be manipulated in the sense that, independently of the state realizations, the difference between the average prediction and the past empirical frequency that corresponds to almost any test employed diminishes to zero. In other words, a forecaster has a prediction scheme that passes almost any test. In particular, a forecaster can pass all the tests in a countable set **simultaneously**.

* I am grateful to Rann Smorodinsky, Sylvain Sorin and two anonymous referees for their helpful comments. This research was partially supported by the US-Israel Binational Science Foundation, Grant no. 9600043.

1. Introduction

Consider a sequential forecast which may vary with time and history. A calibration test (see Dawid (1982), Foster and Vohra (1997) and Kalai, Lehrer and Smorodinsky (1999)) compares the observed empirical frequency over a set of times with the average prediction, over the same set of times. Each test checks different events at different times. A forecasting scheme passes a calibration test at a sequence of state realizations if the gap between these two figures diminishes to zero. In other words, each test defines what is consistent with the observed facts differently. A forecast is consistent with the observed facts according to a given test if the empirical frequency of the events checked coincides asymptotically with the average prediction of these events. When the forecast and the true distribution (according to which the states are selected) are the same, the forecast passes any calibration test with probability one. A calibration test is, therefore, Type I-error free.

The calibration tests treated in the literature are pure: After any history it is known what specific event (for instance, "rain", "foggy") is being considered. Here, we introduce mixed calibration tests, where, after each history, a randomly chosen event is checked. A mixed test checks whether the average prediction of these randomly chosen events is asymptotically equal to the frequency of their realizations. Like a deterministic test, a mixed calibration test is also Type I-error free.

The main result of this paper shows that there is no mixed calibration test which rejects any inaccurate forecast. That is, for any test of this kind, there exists a forecasting scheme which is always consistent with the realized events. A constructive method to produce this scheme is provided as well.

Another interpretation of the main result involves a forecaster and an inspector. The inspector randomly chooses a calibration test according to a known distribution. It is shown that a forecaster, without knowing anything regarding the way states are selected can pass almost every test chosen by the inspector, independently of the sequence of realizations. In other words, the forecaster has a forecasting method which is consistent with the observed facts in the meaning induced by almost all tests.

Instead of randomly choosing one test, as discussed previously, the inspector can employ a set of tests simultaneously. The main result implies that if the inspector employs countably many calibration tests, for instance all those computable by a Turing machine, a forecaster can, likewise, pass them all[1] simultaneously. It means that a consultant or

[1] The forecasting scheme will be non-computable in this case.

a scientist (e.g., a weather forecaster) who knows the (countable) set of tests his model needs to pass, can manipulate. That is, no matter what the sequence of realizations is, his manipulative forecast will pass all the tests.

Foster and Vohra (1997), who were the first to introduce the subject of calibration into game theory, deal with a combined calibration test. The forecast passes a combined test if over those times, where the forecaster says that the next state will be ω with probability p, the asymptotic empirical frequency of ω is indeed p. Foster and Vohra showed that there always exists a forecast that is consistent with the observed realizations in the sense that it passes the combined test for every sequence of state-realizations. Fudenberg and Levine (1999a) proved this result by the minmax theorem. Hart and Mas-Colell (1996) did the same using Blackwell's approachability theorem (Blackwell (1956), Foster and Vohra (1999)).

The proof of the main result relies on an extension of Blackwell's approachability theorem (Blackwell (1956)) to games with payoffs in infinite dimensional spaces (Lehrer (1997)).

2. Forecasts and Checking Rules

Let Ω be a finite state space. At time n an outcome $\omega_n \in \Omega$ is realized. For each infinite stream of outcomes, $\omega^\infty = (\omega_1, \omega_2, ...) \in \Omega^\infty$, let $\omega^n = (\omega_1, \omega_2, ..., \omega_n)$ denote the history of length n, where ω^0 denotes the null history. At any time a probabilistic forecaster, after observing the past history ω^n, assigns the distribution $\mu(\cdot|\omega^n)$ to the possible events in the subsequent period, $n+1$. Formally, $\mu(\cdot|\omega^n)$ is a probability over Ω.

Remark 1. By the Kolmogorov extension theorem (see Shiryayev (1984)) the set of all probabilities $\mu(\cdot|\omega^n)$ induces a unique distribution, μ, over Ω^∞. For any subset A of Ω, $\mu(A|\omega^n)$ can be therefore interpreted as the conditional probability assigned by μ, given ω^n, that ω_{n+1} will be in A.

Definition. A *checking rule* is a pair, $D = (C, U)$, where C and U are functions defined on the set of all histories, $\cup_{n=0}^{\infty}\Omega^n$, with Ω^0 being a singleton. For every history ω^n, $C(\omega^n)$ and $U(\omega^n)$ are events (i.e., measurable subsets) in Ω. Moreover, $C(\omega^n) \subseteq U(\omega^n)$.

The checking rule D is interpreted as follows. After the history ω^n, D determines what (local) "universe" to consider (i.e., $U(\omega^n)$) and, within this "universe", which event to check (i.e., $C(\omega^n)$).

To empirically test the reliability of a forecast, Dawid (1982) introduced the notion of calibration. A belief is declared as being calibrated with the truth if observed frequencies of events match their forecasted probabilities. The checking rule D is used to empirically test the reliability of μ. This is done by comparing the relative frequency of the C's among the realizations of U to the conditional probability (according to μ) of C given U. Relatedly, Kalai, Lehrer and Smorodinsky (1999) considered a special kind of checking rules, where $U(\omega^n)$ is either Ω or $\emptyset$.

For any checking rule define, $I^n(D,\omega^n) = 1$ if $\omega_n \in U(\omega^{n-1})$. Otherwise, $I^n(D,\omega^n) = 0$. Thus, I^n is equal to 1 if the realized state is in the "universe" checked. In this case, we say that D is active. Let $T^n(D,\omega^n) = \sum_{t=1}^n I^t(D,\omega^t)$. T^n is the number of times, up to stage n, when the checking rule was active.

For any checking rule D, a forecast μ, and ω^∞, denote,

(1)
$$f^n(D,\omega^n,\mu) = \frac{\sum_{t=1}^n \mathbb{1}_{\{\omega_t \in U(\omega^{t-1})\}}\Big(\mu\big(U(\omega^{t-1})|\omega^{t-1}\big)\mathbb{1}_{\{\omega_t \in C(\omega^{t-1})\}} - \mu\big(C(\omega^{t-1})|\omega^{t-1})\big)\Big)}{T^n(D,\omega^n)},$$

where $0/0$ is defined as 0 and ω^0 denotes the null history.

Remark 2. Conditional on ω^{t-1}, the expectation, with respect to μ, of the t-th summand of (1), $\mathbb{1}_{\{\omega_t \in U(\omega^{t-1})\}}\Big(\mu\big(U(\omega^{t-1})|\omega^{t-1}\big)\mathbb{1}_{\{\omega_t \in C(\omega^{t-1})\}} - \mu\big(C(\omega^{t-1})|\omega^{t-1}\big)\Big)$, is zero. Therefore, by the strong law of large numbers, $f^n(D,\omega^n,\mu)$ converges to zero μ-almost surely.

Remark 3. In case $I^t = 0$, the t-th summand of (1) is equal to zero. That is, when D is inactive at time t, $f^{t-1}(D,\omega^{t-1},\mu) = f^t(D,\omega^t,\mu)$.

Remark 4. An alternative definition of $f^n(D,\omega^n,\mu)$ can be given using the terminology of conditional expectation[2]. Let $\mathcal{F}_t$ be the field generated by the histories of length t. Define two random variables C_t and U_t, both being $\mathcal{F}_t$-measurable, which take only two values, 0 or 1. $C_t = 1$ only if $\omega_t \in C(\omega^{t-1})$ and $U_t = 1$ only if $\omega_t \in U(\omega^{t-1})$. In other words, C_t and U_t are the characteristic functions of $C(\omega^{t-1})$ and $U(\omega^{t-1})$, respectively. Now, $f^n(D,\omega^n,\mu) = \frac{\sum_{t=1}^n U_t\left(E\left(U_t|\mathcal{F}_{t-1}\right)C_t - E\left(C_t|\mathcal{F}_{t-1}\right)\right)}{\sum_{t=1}^n U_t}$. Since $C_t = 1$ implies $U_t = 1$, the right side of this equality can also be written as $\frac{\sum_{t=1}^n \left(E(U_t|\mathcal{F}_{t-1})C_t - E(C_t|U_t,\mathcal{F}_{t-1})U_t\right)}{\sum_{t=1}^n U_t}$ or as $\frac{\sum_{t=1}^n U_t E(U_t|\mathcal{F}_{t-1})\left(C_t - E(C_t|U_t,\mathcal{F}_{t-1})\right)}{\sum_{t=1}^n U_t}$.

Definition. The forecast μ *passes the calibration test induced by* D *at* $\omega^\infty = (\omega_1,\omega_2,...)$, if $T^n(D,\omega^n) \to \infty$ implies

$$\lim_{n\to\infty} f^n(D,\omega^n,\mu) = 0.$$

[2] This formulation was suggested by one of the referees.

The fact that $\lim_{n\to\infty} f^n(D,\omega^n,\mu)=0$ means that the forecaster managed to have a good track record along ω^∞, according to D.

Definition. A *test* is a function, ψ, which attaches to any measure μ and $\omega^\infty \in \Omega^\infty$ an element in the set $\{PASS, FAIL\}$. A test ψ is *Type I-error free* if for any measure μ, $\psi(\mu,\omega^\infty)$ is $PASS$ with μ-probability one.

In other words, a test is Type I-error free if it is immunized against committing a Type I-error of rejecting a true forecasting scheme.

Remark 5. Remark 2 implies that for every calibration test D μ passes the calibration test induced by D at μ-almost every ω^∞ (see also Dawid (1982) and Kalai, Lehrer and Smorodinsky (1999)). Stated differently, the calibration tests are Type I-error free.

The combined calibration test introduced by Foster and Vohra (1997) is not Type I-error free. For every p, a distribution over Ω, denote, $N(p,\omega^n)=\sum_{t=1}^n \mathbb{1}_{\{\mu(\cdot|\omega^{t-1})=p\}}$. Thus, $N(p,\omega^n)$ is the number of times, up to time n, that the forecast is p. Let $\rho(p,\omega,\omega^n)$ be the fraction of those times for which the realization is $\omega\in\Omega$. Formally,

$$\rho(p,\omega,\omega^n)=\frac{\sum_{t=1}^n \mathbf{1}_{\{\mu(\cdot|\omega^{t-1})=p\}}\mathbf{1}_{\{\omega^t=\omega\}}}{N(p,\omega^n)}$$

if $N(p,\omega^n)>0$ and zero otherwise. The forecast μ *passes the combined calibration test at* ω^∞ if

$$\lim_{n\to\infty}\sum_p |\rho(p,\omega,\omega^n)-p(\omega)|\frac{N(p,\omega^n)}{n}=0$$

for every $\omega\in\Omega$, where the summation is taken over all the distributions, p.

Consider a forecasting scheme that satisfies the following two properties: (a) at any time, the forecast along the sequence of realizations ω^∞ differs from all previous forecasts, and (b) at any time, the probability assigned to each $\omega\in\Omega$ is bounded away from zero. Such a forecasting scheme does not pass the combined calibration test at ω^∞. This is so because up to any time n, there are n different p's for which $N(p,\omega^n)=1$ and, moreover, for each such p, the difference $|\rho(p,\omega,\omega^n)-p(\omega)|$ is bounded away from zero. Note that the forecast does not pass the combined calibration test even if it is true, namely, even if the sequence of realizations is randomly chosen according to the forecasting scheme. Thus, the combined calibration test is not Type I-error free.

3. The Main Result

Denote by $\mathcal{D}$ the set of all checking rules. This set is naturally endowed with a topology (an open set includes all the rules that agree on a finite set of histories), and a σ-algebra (the algebra generated by all the open sets).

Suppose that the inspector chooses a checking rule out of $\mathcal{D}$ according to the distribution λ. The distribution λ will be referred to also as a mixed checking rule.

Theorem 1. *Given λ, there is a forecasting μ s.t. for any ω^∞, the sequence $f^n(D, \omega^n, \mu)$, of random variables defined over $\mathcal{D}$, converges λ-almost surely to zero.*

The theorem establishes the existence of a forecast which passes the calibration test induced by almost every checking rule, at any realized sequence of states.

Corollary 1. *For any given countable set $\mathcal{D}' \subseteq \mathcal{D}$, there is a forecast μ s.t. μ passes every calibration test induced by any checking rule in $\mathcal{D}'$ at any ω^∞ (i.e., $f^n(D, \omega^n, \mu)$ converges to zero for every $D \in \mathcal{D}'$ at any ω^∞.)*

In other words, Corollary 1 states that for any given countable set $\mathcal{D}'$, there is a forecast μ which passes the calibration test induced by any $D \in \mathcal{D}'$ over any sequence of realizations, ω^∞. This means that any countable set of tests can be manipulated by a smart forecaster.

Proof of Corollary 1. Let $\mathcal{D}'$ be a countable set of checking rules and let λ be a distribution which assigns to any $D \in \mathcal{D}'$ a positive probability. Since, by Theorem 1, f^n converges λ-almost surely to zero, $f^n(D)$ converges to zero for every $D \in \mathcal{D}'$. ∎

The proof of the main theorem, Theorem 1, requires some results regarding the approachability of random variables. These results are of interest in themselves, and are presented in the following sections.

4. The Game Played by the Forecaster and the Inspector

Fix $\omega^\infty \in \Omega^\infty$ and consider the following sequential game. At time n, after the history ω^{n-1} of states, the forecaster chooses a distribution $\mu(\cdot|\omega^{n-1})$ over Ω and the inspector chooses a pair $(C(\omega^{n-1}), U(\omega^{n-1}))$ such that $C(\omega^{n-1}) \subset U(\omega^{n-1}) \subset \Omega$. In other words, the forecaster's strategy is a complete forecasting scheme, μ, and the inspector's pure strategy is a checking rule, D. The forecaster wins the game if μ passes the calibration test induced by D at ω^∞, and looses otherwise. We then say that $U_{\omega^\infty}(\mu, D)$ is equal to 1 if the forecaster wins and equal to 0 otherwise. Denote this game by Γ_{ω^∞}.

Note that in Γ_{ω^∞}, both players cannot condition their strategies on the previous choices made by their opponents. In subsection 10.1, we refer to the game where the inspector can condition his choices on the previous forecaster's predictions and claim that the same results hold. In any case, the forecaster may rely solely on previously realized states.

Returning to this game, the inspector is allowed to randomly choose a pure strategy. Let λ be an inspector's mixed strategy. Thus, λ is a distribution over $\mathcal{D}$. We now extend the range of U_{ω^∞} to accommodate also the inspector's mixed strategies as follows: $U_{\omega^\infty}(\mu, \lambda) = E_\lambda(U_{\omega^\infty}(\mu, D))$.

Since the players are not informed of ω^∞ before the game starts, the players actually play an unknown game selected from the set $\{\Gamma_{\omega^\infty}\}_{\omega^\infty \in \Omega^\infty}$. This is an incomplete information game without a prior distribution over the set of possible games. Non-Bayesian games of this kind were previously treated also in Banõs (1968) and in Megiddo (1980).

The following rephrases Theorem 1.

Theorem 1*. *Given an inspector's mixed strategy λ, there is a forecasting scheme μ s.t. $U_{\omega^\infty}(\mu, \lambda) = 1$ for any ω^∞.*

The fact that the forecaster can win the game for **any** ω^∞ means that whatever the true distribution over Ω^∞ may be, say, π, the following holds: $\min_\lambda \max_\mu \; E_\pi(U_{\omega^\infty}(\mu, \lambda)) = 1$. Thus, if the strategy of the inspector is known, the forecaster has a strategy that ensures him a win without knowing the true distribution that governs the evolution of states.

5. The Principle of Approachability

This section is devoted to the geometric principle behind the proof. The simplest version of this principle called the principle of approachability, is the following. Let $g^1, g^2, ...$ be a uniformly bounded sequence in $\mathbb{R}^k$. Denote by f^n the average of the n first elements in the sequence. If for any n the inner product of f^n and g^{n+1}, denoted $\langle f^n, g^{n+1} \rangle$, is less than or equal to zero, then the sequence of the averages, f^n, converge to zero. That is, if for any n the element g^{n+1} and f^n lie on two sides of the hyperspace perpendicular to f^n, then the latter approaches zero. In a sense, by lying on the other side of the hyperspace perpendicular to f^n, the vector g^{n+1} corrects the cumulative error at time n.

We proceed next to the first extension. Suppose that at any time there is also an activeness vector $I^n \in \mathbb{R}^k$ whose coordinates are either 0 or 1. The coordinates of the vector I^n tell whether their corresponding coordinates in g^n are active or not: the i-th coordinate of g^n is active only if the i-th coordinate of I^n is 1. Let $T^n = \sum_{t=1}^n I^t$. Thus, the i-th coordinate of T^n is the number of times the i-th coordinate was active in the vectors $g^1, g^2, ..., g^n$.

Define $f^1 = g^1$ and $f^n = \frac{T^{n-1} f^{n-1} + g^n}{T^n}$, where the product (resp. quotient) of two vectors is the vector whose coordinates are the product (resp. quotient) of the corresponding coordinates. The i-th coordinate of f^n is the average of the i-th coordinates of $g^1, g^2, ..., g^n$ when these are active.

An example. Suppose that $g^1 = (1, -3)$, $g^2 = (7, 5)$, $g^3 = (-2, 9)$, $I^1 = (1, 1)$, $I^2 = (1, 0)$ and $I^3 = (1, 1)$. Then, $T^3 = (3, 2)$ and $f^3 = (2, 3)$.

Note that in case all the coordinates are always active (that is, I^n is a vector of 1's for every n), then T^n is a vector of n's and f^n is, as before, the ordinary average.

An extension of the principle of approachability to this case is the following. Suppose that all the coordinates of T^n go to infinity. If for any n the inner product of $\frac{f^n}{T^{n+1}}$ and g^{n+1}, $\langle f^n, \frac{g^{n+1}}{T^{n+1}} \rangle$, is less than or equal to zero, then f^n converges to zero. Note that when, as in the previous case, T^{n+1} is constant, then $\langle f^n, \frac{g^{n+1}}{T^{n+1}} \rangle \leq 0$ is equivalent to $\langle f^n, g^{n+1} \rangle \leq 0$. f^n is the average of $g^1, g^2, ..., g^n$ over the active times. Thus, the contribution of any coordinate of g^{n+1} (to f^{n+1}) is greater when the corresponding coordinate of T^{n+1} is smaller. More precisely, the contribution is $\frac{g^{n+1}}{T^{n+1}}$. If for any n this vector lies on the other side of the hyperspace perpendicular to f^n (i.e., $\langle f^n, \frac{g^{n+1}}{T^{n+1}} \rangle \leq 0$), meaning that the vector g^{n+1} corrects the cumulative error at time n, then f^n converges to zero.

This extension is sufficient for the case where the inspector is restricted to using only a finite number of calibration tests. In this case, each one of the k coordinates is dedicated to one test. The objective of the forecaster is, then, to obtain the average of each coordinate over the active times converging to zero. That is, to have the average vectors approaching the zero vector.

In case the inspector employs more than a finite number of tests, one needs a further extension of the principle of approachability. Let $g^1, g^2, ...$ be a sequence of random variables whose variance is finite. These random variables are measurable functions from a probability space to the set of the real numbers. Denote by λ the underlining probability. The vectors in the previous case can be thought of as functions defined over a finite set consisting of k points. Here, the coordinates of the finite vectors are replaced by the points in the probability space. Let I^n be a random variable which attains either 0 or 1 as values and let $T^n = \sum_{t=1}^{n} I^t$. As before, I^n is the activeness random variable while T^n counts the number of active times that correspond to every point in the probability space. Define $f^1 = g^1$ and $f^n = \frac{T^{n-1} f^{n-1} + g^n}{T^n}$, where the product (resp. quotient) refers to the pointwise product (resp. quotient) of two random variables. The value of f^n at a certain point is the average of the values that $g^1, g^2, ..., g^n$ take at the same point, over the active times.

Note that the expectation of the product of two random variables can serve as an inner product in the space of random variables whose variance is finite. The principle of approachability then takes the following form. If T^n goes to infinity and if, in addition, for any n the expectation of the product of f^n and $\frac{g^{n+1}}{T^{n+1}}$, $E(f^n \frac{g^{n+1}}{T^{n+1}})$, is less than or equal to zero, then f^n converges to zero. In symbols, if

$$\int f^n \tfrac{T^{n+1}}{g^{n+1}} d\lambda = \langle f^n, \tfrac{g^{n+1}}{T^{n+1}} \rangle = \int \tfrac{f^n}{T^{n+1}} g^{n+1} d\lambda = \langle \tfrac{f^n}{T^{n+1}}, g^{n+1} \rangle \leq 0,$$

then f^n converges to zero. The convergence here is in the "almost surely" sense. That is, $f^n \to 0$ with λ-probability 1.

This principle is summarized by Proposition 1 in Lehrer (1997).

Proposition 1. *Suppose that*

(a) $\{T^n\}_0^\infty$ *is a sequence of non decreasing random variables that assume integer values.* $T^{n+1} - T^n \leq 1$, $T^n \to \infty$ λ*-a.s. and* $T^0 = 0$*;*

(b) $\{g^n\}$ *is a sequence of random variables that take values in* $[-1, 1]$ *s.t.* $T^{n+1} - T^n = 0$ *implies* $g^{n+1} = 0$*;*

(c) $f^{n+1} = \frac{T^n f^n + g^{n+1}}{T^{n+1}}$*; and*

(d) $\langle \frac{f^n}{T^{n+1}}, g^{n+1} \rangle \leq 0$.

Then, f^n *converges to zero with* λ*-probability 1.*

6. An Illustration of the Manipulative Forecast

Since non-Bayesian games are hard to handle, we consider a different game where the players are the forecaster and nature. As before, the forecaster provides a prediction while nature chooses a realization. The payoffs at any round are not numbers but functions defined over the set of pure tests. For any of the players' actions the value of the function-payoff that corresponds to a test is the gap between the forecaster's prediction (of an event determined by the test) and nature's choice (1 for the case where nature's choice is in this event and 0 otherwise.) It is shown that the forecaster has a strategy that ensures that the partial average payoffs converge to zero with probability 1. In other words, the gaps between the relative frequency and the realizations that correspond to almost all tests converge to zero.

To illustrate the game and the strategy of the forecaster (his manipulative forecast), consider[3] $\Omega = \{a, b, c\}$. Suppose that the inspector chooses one of the tests $D^1, ..., D^4$ with probability $\frac{1}{4}$ each. Suppose, furthermore, that on the null history the tests are defined as follows: $C^1 = \{a\}$, $U^1 = \{a, b\}$; $C^2 = \{a\}$, $U^2 = \{a, c\}$; $C^3 = \{c\}$, $U^3 = \{a, b, c\}$; and $C^4 = \{b, c\}$, $U^4 = \{a, b, c\}$. Assume that the first realization, ω_1, is a. Thus, $I^1(D^i, \omega^1) = 1$ for $i = 1, ..., 4$. Assume that the forecast at the first stage was $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. Thus, $f^1(D^1, \omega^1, \mu) = \frac{1}{3}$; $f^1(D^2, \omega^1, \mu) = \frac{1}{3}$; $f^1(D^3, \omega^1, \mu) = -\frac{1}{3}$; and $f^1(D^4, \omega^1, \mu) = -\frac{2}{3}$.

[3] For an example with C strictly in $U \neq \Omega$, we need more than two states.

The forecast at the second stage depends on the tests employed, which depend on the history $\omega^1 = (\omega_1)$. Assume that $C^1(\omega^1) = \{a\}$, $U^1(\omega^1) = \{a, b\}$; $C^2(\omega^1) = \{c\}$, $U^2(\omega^1) = \{a, c\}$; $C^3(\omega^1) = \{c\}$, $U^3(\omega^1) = \{b, c\}$; and $C^4(\omega^1) = \emptyset$, $U^4(\omega^1) = \{a, b, c\}$.

Let[4] G_{D^i} for $i = 1, ..., 4$ be the four following 3×3 matrices.

$$\underset{G_{D^1}}{\begin{array}{c} \\ a \\ b \\ c \end{array}\begin{array}{c} \begin{array}{ccc} a & b & c \end{array} \\ \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{array}} \quad \underset{G_{D^2}}{\begin{array}{c} \\ a \\ b \\ c \end{array}\begin{array}{c} \begin{array}{ccc} a & b & c \end{array} \\ \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \end{array}} \quad \underset{G_{D^3}}{\begin{array}{c} \\ a \\ b \\ c \end{array}\begin{array}{c} \begin{array}{ccc} a & b & c \end{array} \\ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \end{array}} \quad \underset{G_{D^4}}{\begin{array}{c} \\ a \\ b \\ c \end{array}\begin{array}{c} \begin{array}{ccc} a & b & c \end{array} \\ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{array}}.$$

Note that each one of the matrices G_{D^i} is anti-symmetric[5]. These matrices are related to the tests in the following sense. Suppose that the forecast at the second stage is $\mu(\cdot|\omega^1) = (\mu_a, \mu_b, \mu_c)$. The product of the matrix G_{D^i} with the vector (μ_a, μ_b, μ_c), which is a three-dimensional vector, determines, along with the realization on the second stage, ω_2, the value of $f^2(D^i, \omega^2, \mu)$. More precisely, let the product of the matrix G_{D^i} with the vector (μ_a, μ_b, μ_c) be denoted as $(g^2(D^i, (a, a), \mu), g^2(D^i, (a, b), \mu), g^2(D^i, (a, c), \mu))$. If the realization on the second stage is ω_2, then $\omega^2 = (a, \omega_2)$ and $f^2(D^i, \omega^2, \mu) = \frac{f^1(D^i,\omega^1,\mu)+g^2(D^i,\omega^2,\mu)}{T^2(D^i,\omega^2,\mu)}$, (see (1)). For instance, if the realization on the second stage is b, then $g^2(D^1, (a, b), \mu) = -\mu_a$ and $f^2(D^1, \omega^2, \mu) = f^2(D^1, (a, b), \mu) = \frac{\frac{1}{3}+(-\mu_a)}{2}$. Since ω_2 is not in $U^2(\omega^1)$, the second test, D^2, is inactive on the second stage and, hence, $f^2(D^2, \omega^2, \mu) = f^1(D^2, \omega^1, \mu)$.

Consider now the matrix $\overline{G} = \sum_{i=1}^{4} \frac{1}{4} \frac{f^1(D^i,\omega^1,\mu)}{2} G_{D^i}$. This is a linear combination of four anti-symmetric matrices and is, therefore, anti-symmetric in itself. The matrix $\overline{G}$ is equal to

$$\underset{\overline{G}}{\begin{array}{c} \\ a \\ b \\ c \end{array}\begin{array}{c} \begin{array}{ccc} a & b & c \end{array} \\ \begin{pmatrix} 0 & \frac{1}{24} & \frac{-1}{24} \\ \frac{-1}{24} & 0 & \frac{1}{24} \\ \frac{1}{24} & \frac{-1}{24} & 0 \end{pmatrix} \end{array}}.$$

Considered as a zero-sum game, the matrix $\overline{G}$ has 0 as a value. Thus, the column player has an optimal strategy that ensures that the payoff will not exceed 0. This strategy, $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, is chosen to be the forecast at the second stage.

Refer to $f^1(D^i, \omega^1, \mu)$ and $g^2(D^i, \omega^2, \mu)$ as four-dimensional vectors (the coordinate i corresponds to the checking rule D^i, i=1,2,3,4). The fact that the column player can ensure that the payoff in $\overline{G}$ will be at most 0, ensures that $\langle f^1, g^2 \rangle \leq 0$. This is so no

[4] The notation corresponds to that used in the general construction given in Section 5.

[5] A square matrix $A = \big(a_{ij}\big)$ is *anti-symmetric* if $a_{ij} = -a_{ji}$ for any i and j.

matter what the realization at the second stage may be. Thus, for any realization, g^2 and f^1 lie on different sides of the hyperspace in $I\!R^4$ that is perpendicular to f^1. Following a similar procedure in all the subsequent stages will enable one to use the principle of approachability and to prove that f^n converges to the zero of $I\!R^4$ or, in other words, to prove that all four tests are passed.

7. The Proof of the Existence of a Manipulative Forecast

Proof of Theorem 1.

Let $\overline{\omega}^\infty$ be fixed throughout. For any $D \in \mathcal{D}$, let

$$G_{(D,\overline{\omega}^n)}(\omega',\omega'') = \mathbb{1}_{\{\omega',\omega'' \in U(\overline{\omega}^n)\}}(\mathbb{1}_{\omega' \in C(\overline{\omega}^n)} - \mathbb{1}_{\omega'' \in C(\overline{\omega}^n)}).$$

For a given D and $\overline{\omega}^n$ one may consider $G_{(D,\overline{\omega}^n)}(\omega',\omega'')$ as a zero-sum game, where player 1 chooses $\omega' \in \Omega$ and player 2 chooses $\omega'' \in \Omega$. Since $G_{(D,\overline{\omega}^n)}(\omega',\omega'') = -G_{(D,\overline{\omega}^n)}(\omega'',\omega')$, it follows that $G_{(D,\overline{\omega}^n)}$ is a zero-sum game whose matrix is anti-symmetric. As such, its value is 0.

One may think of the forecaster as the column player and of nature, which chooses the realization, as the row player. A forecaster's prediction at time $n+1$, $\mu(\cdot|\overline{\omega}^n)$, is a distribution over the columns. In other words, for a given D, the forecaster's prediction is the column player's mixed strategy in $G_{(D,\overline{\omega}^n)}$. Nature's choice at time $n+1$ is a row, ω_{n+1}. In case the forecaster predicts $\mu(\cdot|\overline{\omega}^n)$ and nature chooses $\overline{\omega}_{n+1}$, the expected payoff (corresponding to $G_{(D,\overline{\omega}^n)}(\omega',\omega'')$) is

$$(3) \quad g^{n+1}(D,\overline{\omega}^{n+1},\mu) = \mathbb{1}_{\{\overline{\omega}_{n+1} \in U(\overline{\omega}^n)\}}\Big(\mu\big(U(\overline{\omega}^n)|\overline{\omega}^n\big)\mathbb{1}_{\{\overline{\omega}_{n+1} \in C(\overline{\omega}^n)\}} - \mu\big(C(\overline{\omega}^n)|\overline{\omega}^n\big)\Big).$$

This term represents the $n+1$ summand in the numerator of (1). Note that by Remark 3, whenever $I^{n+1}(D,\overline{\omega}^{n+1}) = 0$, that is, whenever the checking rule D is inactive, $g^{n+1}(D,\overline{\omega}^{n+1},\mu) = 0$. Thus, $f^{n+1}(D,\overline{\omega}^{n+1},\mu)$ is the average payoff, over the times when D was active, of the expected payoffs in the games $G_{(D,\overline{\omega}^t)}$, $t = 1,...,n$. Alternatively,

$$(4) \qquad f^{n+1}(D,\overline{\omega}^{n+1},\mu) = \frac{\sum_{t=1}^{n+1} g^t(D,\overline{\omega}^t)}{T^{n+1}(D,\overline{\omega}^{n+1})} = \frac{f^n(D,\overline{\omega}^n,\mu)T^n(D,\overline{\omega}^n) + g^{n+1}(D,\overline{\omega}^{n+1},\mu)}{T^{n+1}(D,\overline{\omega}^{n+1})}.$$

The forecaster and nature are, then, engaged in a sequential game. At round $n+1$ they take actions (the forecaster takes a mixed action, a prediction, and nature takes a pure action, a realization) and receive payoffs from many zero-sum games ($G_{(D,\overline{\omega}^n)}$, $D \in \mathcal{D}$). The problem arising is how to manage many games played simultaneously and to arrive at $f^{n+1}(D,\overline{\omega}^{n+1},\mu)$ converging to zero for λ-almost all D. The key idea is to combine all the games played into one game and then to play in the game obtained, optimally.

Let h be a bounded random variable over $\mathcal{D}$. Consider the game $\overline{G}^h_{\overline{\omega}^n}(\omega',\omega'') = \int h(D)G_{(D,\overline{\omega}^n)}(\omega',\omega'')d\lambda(D)$. $\overline{G}^h_{\overline{\omega}^n}(\omega',\omega'')$ is, like $G_{(D,\overline{\omega}^n)}(\omega',\omega'')$, a game where the forecaster chooses a mixture over the columns, ω'', and nature chooses a realization, a row ω'. The game $\overline{G}^h_{\overline{\omega}^n}$ is a linear combination of the games $G_{(D,\overline{\omega}^n)}$. Since each one of the games $G_{(D,\overline{\omega}^n)}$ has an anti-symmetric matrix, so does the game $\overline{G}^h_{\overline{\omega}^n}$. Therefore, the column player can guarantee a payoff that does not exceed zero. In other words, there exists a distribution, call it $\mu^h(\cdot|\overline{\omega}^n)$, over the columns of the matrix $\overline{G}^h_{\overline{\omega}^n}$ such that for every row, ω', the expected payoff, $\sum_{\omega''}\mu^h(\omega''$ at time $n+1|\overline{\omega}^n)\int h(D)G_{(D,\overline{\omega}^n)}(\omega',\omega'')d\lambda(D)$, is less than or equal to zero. That is,

$$\sum_{\omega''}\mu^h(\omega'' \text{ at time } n|\overline{\omega}^n)\int h(D)G_{(D,\overline{\omega}^n)}(\omega',\omega'')d\lambda(D) =$$

$$\int h(D)\sum_{\omega''}\mu^h(\omega'' \text{ at time } n|\overline{\omega}^n)G_{(D,\overline{\omega}^n)}(\omega',\omega'')d\lambda(D) =$$

$$\int h(D)\sum_{\omega''}\mu^h(\omega'' \text{ at time } n+1|\overline{\omega}^n)(\mathbb{1}_{\{\omega',\omega''\in U(\overline{\omega}^n)\}})(\mathbb{1}_{\omega'\in C(\overline{\omega}^n)} - \mathbb{1}_{\omega''\in C(\overline{\omega}^n)}) \leq 0.$$

Rearranging the terms, we obtain that when the realization ω' is $\overline{\omega}_{n+1}$,

$$\int h(D)\mathbb{1}_{\{\overline{\omega}_{n+1}\in U(\overline{\omega}^n)\}}\Big(\mu^h\big(U(\overline{\omega}^n)|\overline{\omega}^n\big)\mathbb{1}_{\{\overline{\omega}_{n+1}\in C(\overline{\omega}^n)\}} - \mu\big(C(\overline{\omega}^n)|\overline{\omega}^n\big)\Big)d\lambda(D) =$$

$$\int h(D)g^{n+1}(D,\overline{\omega}^{n+1},\mu^h)d\lambda(D) \leq 0. \tag{5}$$

Recall that $I^{n+1}(D,\overline{\omega}^{n+1})$ indicates whether the checking rule D is active or not at time $n+1$ and that $T^{n+1}(D,\overline{\omega}^{n+1})$ counts the number of active times up to $n+1$. Both figures depend not only on $\overline{\omega}^n$ but also on ω_{n+1}, the realization at time $n+1$. However, the forecaster is unaware of this realization, and he may rely only on the information available to him, $\overline{\omega}^n$, before providing the prediction at time $n+1$.

Note that $T^n(D,\overline{\omega}^n)+1 \neq T^{n+1}(D,\overline{\omega}^{n+1})$, only if $\overline{\omega}_{n+1} \notin U(\overline{\omega}^n)$. In this case, $g^{n+1}(D,\overline{\omega}^{n+1},\mu) = 0$ for any μ (see(3)). Thus,

$$\tfrac{g^{n+1}(D,\overline{\omega}^{n+1},\mu)}{T^n(D,\overline{\omega}^n)+1} = \tfrac{g^{n+1}(D,\overline{\omega}^{n+1},\mu)}{T^{n+1}(D,\overline{\omega}^{n+1})}. \tag{6}$$

Now we are ready for the key step of the proof: defining the forecast at time $n+1$. Let $h(D) = \frac{f^n(D,\overline{\omega}^n,\mu)}{T^n(D,\overline{\omega}^n)+1}$. The function h depends only on past information, $\overline{\omega}^n$. Define $\mu(\cdot|\overline{\omega}^n) = \mu^h(\cdot|\overline{\omega}^n)$.

The term $h(D)$ is the coefficient of the game $G_{(D,\varpi^n)}$ in the linear combination $\overline{G}^h_{\varpi^n}$. The greater the cumulative error of test D, $f^n(D,\varpi^n)$, the greater the coefficient. Moreover, The greater the value of $T^n(D,\varpi^n)$ the smaller the coefficient. The intuition of the last statement is that when $T^n(D,\varpi^n)$ is relatively large, meaning that D was relatively active up to time n, the contribution of the next result, $g^{n+1}(D,\varpi^{n+1},\mu)$, and hence the coefficient of $G_{(D,\varpi^n)}$ are relatively small.

By (5) and by the definition of h, one obtains,

(7) $$\int \tfrac{f^n(D,\varpi^n,\mu)}{T^n(D,\varpi^n)+1} g^{n+1}(D,\varpi^{n+1},\mu)d\lambda(D) \leq 0.$$

Stated differently, $\frac{g^{n+1}(D,\varpi^{n+1},\mu)}{T^n(D,\varpi^n)+1}$ corrects (on average) the cumulative error, $f^n(D,\varpi^n,\mu)$ in the sense described in Section 5. (6) and (7) imply that

(8) $$\int \tfrac{f^n(D,\varpi^n,\mu)}{T^{n+1}(D,\varpi^{n+1})} g^{n+1}(D,\varpi^{n+1},\mu)d\lambda(D) \leq 0.$$

Let ϖ^∞ be the sequence of realized states, and let μ be the forecast just defined. We complete the proof by applying Proposition 1 to the random variables $f^n = f^n(D,\varpi^n,\mu)$, $g^n = g^n(D,\varpi^n,\mu))$, and $T^{n+1} = T^{n+1}(D,\varpi^{n+1})$, $n = 1,2,\ldots$. Note that ϖ^∞ is fixed; therefore, f^n, g^n and T^{n+1} are functions of D only. Note that all the random variables involved are functions of finite histories, and are, therefore, measurable.

Condition (a) of the proposition is satisfied due to the definition of $T^{n+1}(D,\varpi^{n+1})$. Condition (b) is satisfied by the definition of $g^n(D,\varpi^n,\mu)$ (see (3)) and by Remark 3. Condition (c) is implied by (4) and condition (d) is satisfied due to (8). Proposition 1 implies that whenever $T^{n+1}(D,\varpi^{n+1})$ goes to infinity, $f^{n+1}(D,\varpi^{n+1},\mu)$ converges to zero, λ-almost surely. This means that the forecast μ passes the tests induced by λ-almost all checking rules at ϖ^∞. Since ϖ^∞ was arbitrary, Theorem 1 is proven. ∎

One may also conceive of the game played by the forecaster and nature as a game with payoffs in the space of random variables over $\mathcal{D}$. At stage $n+1$, the one-shot payoff is the random variable $g^{n+1}(\cdot,\varpi^{n+1},\mu)$ and the average payoff is $f^{n+1}(\cdot,\varpi^{n+1},\mu)$.

It turns out that for any specific checking rule and at any period of time, the payoffs are obtained from an anti-symmetric game. Since any linear combination of anti-symmetric games is also anti-symmetric, the combined game, $\overline{G}^h_{\varpi^n}$, is also anti-symmetric. The value of an anti-symmetric game is zero; therefore, the column player, the forecaster, can guarantee that the stage-payoff will not exceed zero no matter what state nature may choose. This is why (7) can be guaranteed.

The geometric meaning of (7) is that the forecaster can ensure that the contribution of the one-shot payoff at time $n+1$, $\frac{g^{n+1}}{T^{n+1}}$ (which is a point in the Hilbert space of the random variables whose second moment is finite), to the average payoff is on the other side of the hyperspace perpendicular to the average payoff up to time n (i.e., the point f^n).

The intuition is as follows: $f^n(D, \overline{\omega}^n, \mu)$ can be thought of as the error related to the test D at time n. Thus, the point f^n (in the Hilbert space) represents the errors of all the tests employed. The direction of f^n in that space is the direction of the error; the opposite direction is the direction of the error correction. g^{n+1}, on the other hand, is the result of the $n+1$ stage; its contribution to the average at time $n+1$, f^{n+1}, is $\frac{g^{n+1}}{T^{n+1}}$. The fact that $\frac{g^{n+1}}{T^{n+1}}$ is on the other side of the hyperspace perpendicular to f^n, means that $\frac{g^{n+1}}{T^{n+1}}$ lies in the direction of the error correction. Proposition 1 ensures that if g^{n+1} corrects the error in this way, at any time $n+1$, then, the error diminishes to zero. That is, f^n goes to the 0 point of the Hilbert space.

8. The Span

Definition. a. Let $\mathcal{D}'$ be a set of checking rules. We say that the checking rule D is in the span of $\mathcal{D}'$ if for every ω^∞, whenever a forecast μ passes the calibration test induced by D' at ω^∞ for every $D' \in \mathcal{D}'$, μ also passes the calibration test induced by D at ω^∞. Denote by $\mathrm{sp}(\mathcal{D}')$ the set of all the checking rules in the span of $\mathcal{D}'$. This set is called the *span of* $\mathcal{D}'$.

b. A set $\mathcal{D}'$ of checking rules is called *minimal* if $D \in \mathcal{D}'$ implies $D \notin \mathrm{sp}(\mathcal{D}' \setminus \{D\})$.

As Oakes (1985) showed, any μ has a D and ω^∞ s.t. μ does not pass the calibration test induced by D at ω^∞. This fact and Corollary 1 show the following:

Proposition 2. *The set of all checking rules cannot be the span of any countable set.*

9. Sampling Rules

Definition. A *sampling rule* is a function, F, from the set of histories, $\cup_t \Omega^t$, to $\{0, 1\}$.

A sampling rule indicates whether the realization that comes after the history $h \in \cup_t \Omega^t$ will be included in the sample (in case $F(h) = 1$) or not (in case $F(h) = 0$).

For a given sampling rule F, and a sequence ω^∞, denote by $N^n(F, \omega^n)$ the number of observations that were sampled according to F up to time n. That is, $N^n(F, \omega^n) = \sum_{t=1}^{n} F(\omega^t)$. Let the empirical frequency of the sample, according to F up to time n, be denoted by $e^n(F, \omega^n)$. In other words, e^n is a distribution over Ω which assigns to $\omega \in \Omega$

a probability that is equal to the number of times ω appeared in the sample up to time n divided by $N^n(F, \omega^n)$.

Definition. A forecast μ is F-good at ω^∞ if $\lim_n N^n(F, \omega^n) = \infty$ implies

$$\lim_{n\to\infty} \|e^n(F, \omega^n) - \frac{\sum_{t=1}^n \mu(\cdot|\omega^t)\mathbf{1}_{\{F(\omega^t)=1\}}}{N^n(F,\omega^n)}\| = 0.$$

In other words, the forecast μ is F-good if, over the sample determined by F, the gap[7] between the empirical frequency and the average prediction tends to zero.

Fudenberg and Levine (1999b) referred to countably many sampling rules that generate a partition of the entire sample. That is, each observation is sampled by exactly one F. Corollary 1 implies a result which holds for any countable family of sampling rules, as follows.

Proposition 3. *For any given countable family of sampling rules, there is a forecasting scheme μ such that μ is F-good at any ω^∞ and for every F in the family.*

Proof: Apply Corollary 1 to the following countable set of checking rules. For every sampling rule F, and for every $\omega \in \Omega$, consider the checking rule defined as $U(\omega^{n-1}) = \Omega$ and $C(\omega^{n-1}) = \{\omega\}$ whenever $F(\omega^{n-1}) = 1$ and $U(\omega^{n-1}) = C(\omega^{n-1}) = \emptyset$ otherwise. Verbally, the checking rule corresponding to F and ω is the one that checks the singleton $\{\omega\}$ whenever the observation is to be sampled according to F and is inactive otherwise. ∎

10. Concluding Remarks and Open Problems

10.1 Checking Rules that Depend on Previous Predictions.

The checking rules defined above are, at any time, functions of the history of realizations. One may extend the checking rules so that the events checked may depend in addition also on all historical predictions. It turns out that Theorem 1 remains correct when applied to such extended checking rules.

The key step of the proof in Section 7 is to find the function h that depends only on past information, $\overline{\omega}^n$, and based on that function to define $\mu(\cdot|\overline{\omega}^n) = \mu^h(\cdot|\overline{\omega}^n)$. As long as the checking rules, whether deterministic or not, depend solely on past information (i.e., whether the rules depend on past realizations or on past predictions), one may define a function h that depends on the history alone. This enables one to define $\mu(\cdot|\overline{\omega}^n)$ as before.

This means that Theorem 1^* remains true, even if the inspector is allowed to condition his choice on past predictions.

[7] Note that this time, the gap is between distributions over Ω.

10.2 Checking Rules which Also Depend on Current Predictions. The main result tells us that when the inspector uses only past information, the forecaster can manipulate. The checking rules we dealt with are not forecast based and the proof of Section 7 cannot be applied to such checking rules.

Sandroni, Smorodinsky and Vohra (1999) showed that checking rules that may depend on the stage prediction are manipulable as well. The forecasting scheme provided in Section 7 is such that after every history, the forecaster chooses deterministically one prediction. In contrast, the forecasting scheme of Sandroni, Smorodinsky and Vohra (1999) chooses a prediction randomly out of a finite set of possible predictions.

10.3 Failing a Test in an Uncountable Set.

Corollary 1 states that if $\mathcal{D}'$ is countable, then there is a forecast which passes any test induced by any $D \in \mathcal{D}'$ at any ω^∞. The question then arises as to whether for any uncountable $\mathcal{D}'$, such a statement is always incorrect. Formally, let μ be given and let $\mathcal{D}'$ be an uncountable minimal set of checking rules. Are there always a checking rule $D \in \mathcal{D}'$ and a sequence ω^∞ of realizations such that $f^{n+1}(D, \omega^{n+1}, \mu)$ does not converge to zero?

10.4 An Impossibility Result.

Definition. A set of tests, $\mathcal{D}$, is said to be *Type I-error free* if for any μ, on a set of ω^∞'s whose μ-probability is 1, μ passes the calibration test induced by D at ω^∞, for every $D \in \mathcal{D}$.

Obviously, any countable $\mathcal{D}$ is Type I-error free. It was shown previously that for any countable $\mathcal{D}$ there exists a forecasting scheme which passes all tests in the set on every ω^∞. It is conjectured that any Type I-error free set is manipulable.

While the definition of Type I-error seems to be the only natural definition, there is no obvious way to define what is Type II-error (not rejecting wrong models). I believe that one can find a reasonable sense of Type II-error so that an affirmative answer to the conjecture would mean that there is no set of tests which is immunized against both Type I-error and against Type II-error.

11. References

Banõs, A. (1968) "On Pseudo-Games," *The Annals of Mathematical Statistics*, **39**, 1932-145.

Blackwell, D. (1956) "A Vector Valued Analog of the Mini-Max Theorem," *Pacific Journal of Mathematics*, **6**, 1-8.

Dawid, A. P. (1982), "The Well-Calibrated Bayesian," *Journal of the American Statistical Association*, **77**, #379, 605-613.

Foster, D., and R. Vohra (1997), "Calibrated Learning and Correlated Equilibrium," *Games and Economic Behavior*, 21, 40-55.

Foster, D., and R. Vohra (1999), "Regret in On-Line Decision Problem," *Games and Economic Behavior*, 29, 7-35.

Fudenberg, D., and D. Levine (1999a), "An Easier Way to Calibrate," *Games and Economic Behavior*, 29, 131-137.

Fudenberg, D., and D. Levine (1999b), "Conditional Universal Consistency," *Games and Economic Behavior*, 29, 104-130.

Hart, S., and A. Mas-Colell (1996), "A Simple Adaptive Procedure Leading to Correlated Equilibrium," Center for Rationality and Interactive Decision Theory, DP #126, The Hebrew University, Jerusalem, Israel.

Kalai, E., E. Lehrer and R. Smorodinsky (1999), "Calibrated Forecasting and Merging," *Games and Economic Behavior*, **29**, 151-169.

Lehrer, E. (1997), "Approachability in Infinite Dimensional Spaces and an Application: A Universal Algorithm for Generating Extended Normal Numbers," mimeo.

Megiddo, N. (1980), "On Repeated Games with Incomplete Information Played by Non-Bayesian Players," **9** , 3, 157-167.

Mertens, J.-F., S. Sorin, and S. Zamir (1994), "Repeated Games," CORE discussion paper #9420.

Oakes, D. (1985), "Self-Calibrated Priors Do Not Exist," *Journal of the American Statistical Association*, **80**, 339.

Sandroni, A., R. Smorodinsky, and R. Vohra (1999), "Calibration with Many Checking Rules," manuscript, Northwestern University.

Shiryayev, A.N. (1984), *Probability*, Springer-Verlag, New York.