The Annals of Statistics
2014, Vol. 42, No. 3, 1003–1028
DOI: 10.1214/14-AOS1212

MERGING AND TESTING OPINIONS

BY LUCIANO POMATTO, NABIL AL-NAJJAR AND ALVARO SANDRONI[1]

Northwestern University

We study the merging and the testing of opinions in the context of a prediction model. In the absence of incentive problems, opinions can be tested and rejected, regardless of whether or not data produces consensus among Bayesian agents. In contrast, in the presence of incentive problems, opinions can only be tested and rejected when data produces consensus among Bayesian agents. These results show a strong connection between the testing and the merging of opinions. They also relate the literature on Bayesian learning and the literature on testing strategic experts.

1. Introduction. Data can produce consensus among Bayesian agents who initially disagree. It can also test and reject opinions. We relate these two critical uses of data in a model where agents may strategically misrepresent what they know.

In each period, either 0 or 1 is observed. Let P and Q be two probability measures on $\{0, 1\}^\infty$ such that Q is absolutely continuous with respect to P. If P and Q are σ-additive then, as shown by Blackwell and Dubins (1962), the conditional probabilities of P and Q merge, in the sense that the two posteriors become uniformly close as the amount of observations increases (Q-almost surely). So, repeated applications of Bayes' rule lead to consensus among Bayesian agents, provided that their opinions were initially compatible.

Now consider Savage's axiomatization of subjective probability. He proposed postulates that characterize a preference relation over bets in terms of a nonatomic finitely additive probability P. Call such P, for short, an *opinion*. Savage's framework allows for finitely additive probabilities that are not σ-additive. In particular, the conclusions of the Blackwell and Dubins theorem hold for some, but not all, opinions. This flexibility makes Savage's framework an ideal candidate to study the connection between the merging and the testing of opinions.

We say that an opinion P satisfies the *Blackwell–Dubins property* if whenever Q is an opinion absolutely continuous with respect to P, the two conditional probabilities merge. By definition, in this subframework, sufficient data produces agreement among Bayesian agents who have compatible initial opinions. Outside this subframework, Bayesian agents may satisfy Savage's axioms, have compatible initial opinions and yet persistently disagree. See the Appendix for an example.

Received August 2013; revised December 2013.
[1]Supported by a grant from the NSF.
MSC2010 subject classifications. Primary 62A01; secondary 91A40.
Key words and phrases. Test manipulation, Bayesian learning.

Any opinion, whether or not it satisfies the Blackwell and Dubins property, can be tested and rejected. To reject an opinion P, it suffices to find an event that has low probability according to P and then reject it if this event is observed. Thus, if opinions are honestly reported then the connection between merging and testing opinions is weak. In the absence of incentive problems, subjective probabilities can be tested and rejected whether or not data produces consensus.

Now consider the case in which a self-proclaimed expert, named Bob, may strategically misrepresent what he knows. Let Alice be a tester who wants to determine whether Bob is an informed expert who honestly reports what he believes or he is an uninformed, but strategic, expert who has reputational concerns and wants to pass Alice's test. Alice faces an adverse selection problem and uses data to screen the two types of experts.

A test is likely to *control for type I error* if an informed expert expects to pass the test by truthfully reporting what he believes. A test can be *manipulated* if even completely uninformed experts are likely to pass the test, no matter how the data unfolds in the future. The word "likely" refers to a possible randomization by the strategic expert to manipulate the test. Only nonmanipulable tests that control for type I error pass informed experts and may fail uninformed ones.

Our main results are: In the presence of incentive problems, if opinions must satisfy the Blackwell–Dubins property then there exists a test that controls for type I error and cannot be manipulated. If, instead, any opinion is allowed then every test that controls for type I error can be manipulated. Thus, in Savage's framework strategic experts cannot be discredited. However, strategic experts can be discredited if opinions are restricted to a subframework where data produces consensus among Bayesian agents with initially compatible views. These results show a strong connection between the merging and the testing of opinions but only under incentive problems.

The Blackwell–Dubins theorem has an additional interpretation. In this interpretation, Q is referred to as the data generating process and P is an agent's belief initially compatible with Q. When the conclusions of the Blackwell–Dubins theorem hold, then P and Q merge and so, the agent's predictions are eventually accurate. Thus, multiple repetitions of Bayes' rule transforms the available evidence into a near perfect guide to the future. It follows that our main results also have an additional interpretation. Under incentive problems, strategic experts can only be discredited if they are restricted to a subframework where opinions that are compatible with the data generating process are eventually accurate.

Finally, our results relate the literatures on Bayesian learning and the literature on testing strategic experts (see the next section for references). They show a strong connection between the framework under which Bayesian learning leads to accurate opinions and the framework under which strategic experts can be discredited.

The paper is organized as follows. Section 2 describes the model. Section 3 reviews the Blackwell–Dubins theorem and defines the Blackwell–Dubins property. Section 4 contains our main results. Section 5 relates our results and category tests.

Section 6 considers the case where the set of per-period outcome may be infinite. The Appendix contains all proofs and a formal example of a probability that does not satisfy the Blackwell–Dubins property.

1.1. *Related literature.* Blackwell and Dubins' idea of merging of opinions is central in the theory of Bayesian learning and Bayesian statistics. In Bayesian nonparametric statistics, see the seminal work of Diaconis and Freedman (1986), D'Aristotile, Diaconis and Freedman (1988) and the more recent work by Walker, Lijoi and Pruenster (2005). In the theory of Bayesian learning, see Schervish and Seidenfeld (1990). We refer to Dawid (1985) for a connection with the theory of calibration.

In game theory, the Blackwell–Dubins theorem is central in the study of convergence to Nash equilibrium in repeated games. The main objective is to understand the conditions under which Bayesian learning leads to a Nash equilibrium [see, among many contributions, Foster and Young (2001, 2003), Fudenberg and Kreps (1993), Fudenberg and Levine (1998, 2009), Hart and Mas-Colell (2013), Jackson, Kalai and Smorodinsky (1999), Kalai and Lehrer (1993a, 1993b), Lehrer and Smorodinsky (1996a, 1996b), Monderer, Samet and Sela (1997), Nachbar (1997, 2001, 2005), Sandroni (1998) and Young (2002, 2004)].

A series of papers investigate whether empirical tests can be manipulated. In statistics, see Foster and Vohra (1998), Cesa-Bianchi and Lugosi (2006), Vovk and Shafer (2005) and Olszewski and Sandroni (2009a). In economics, see among several contributions, Al-Najjar and Weinstein (2008), Al-Najjar et al. (2010), Babaioff et al. (2011), Dekel and Feinberg (2006), Feinberg and Lambert (2011), Feinberg and Stewart (2008), Fortnow and Vohra (2009), Fudenberg and Levine (1999), Gradwohl and Salant (2011), Gradwohl and Shmaya (2013), Hu and Shmaya (2013), Lehrer (2001), Olszewski and Peski (2011), Olszewski and Sandroni (2007, 2008, 2009a, 2009b, 2011), Sandroni (2003), Sandroni, Smorodinsky and Vohra (2003), Shmaya (2008), Stewart (2011). For a review, see Foster and Vohra (2011) and Olszewski and Peski (2011). See also Al-Najjar, Pomatto and Sandroni (2013) for a companion paper.

2. Setup. In every period an outcome, 0 or 1, is observed (all results generalize to the case of finitely many outcomes). A *path* is an infinite sequence of outcomes and $\Omega = \{0, 1\}^\infty$ is the set of all paths. Given a path ω and a period t, let $\omega^t \subseteq \Omega$ be the cylinder of length t with base ω. That is, ω^t is the set of all paths which coincide with ω in the first t periods. The set of all paths Ω is endowed with a σ-algebra of *events* Σ containing all cylinders.

The set Ω is endowed with the product topology. In this topology, a set is open if and only if it is a countable union of cylinders. We denote by Σ_1 the set of all open subsets of Ω and by $\mathcal{B}$ the Borel σ-algebra generated by the topology. Note that $\Sigma_1 \subset \mathcal{B} \subseteq \Sigma$.

Let $\mathbb{P}$ be the set of all finitely additive probabilities on (Ω, Σ). A probability $P \in \mathbb{P}$ is strongly nonatomic, or *Savagean*, if for every event E and every $\alpha \in [0, 1]$ there is an event $F \subseteq E$ such that $P(F) = \alpha P(E)$. The term "Savagean" emphasizes the relation between strongly nonatomic probabilities and the Savage (1954) representation theorem: a finitely additive probability corresponds to a preference relation satisfying Savage's axioms if and only if it is strongly nonatomic. To simplify the language, we also refer to a Savagean probability as an *opinion*. Let Δ denotes the set of all opinions.

At time 0, a self-proclaimed expert, named Bob, announces an opinion P. A tester, named Alice, evaluates his opinion empirically. Alice announces her test at period 0, before Bob announces his opinion.

DEFINITION 1. A *test* is a function $T : \Delta \to \Sigma_1$.

A test specifies an open set $T(P)$ considered inconsistent with an opinion P. An expert who announces opinion P is rejected on every path ω belonging to $T(P)$. For the next definition, fix $\varepsilon \in [0, 1)$ and a subset Λ of Δ.

DEFINITION 2. A test Λ*-controls for type I error* with probability $1 - \varepsilon$ if for any $P \in \Lambda$,

$$P\big(T(P)\big) \leq \varepsilon.$$

If a test Λ-controls for type I error, then an expert (with an opinion in Λ) expects to pass the test by honestly reporting what he believes.

2.1. *Strategic forecasting*. We now consider the case where Bob is uninformed about the odds of future events, but may produce an opinion strategically in order to pass the test. We allow strategic experts to select opinions at random. Let $\Delta_f \Delta$ be the set of probability measures on Δ with finite support. We call each $\zeta \in \Delta_f \Delta$ a *strategy*.

DEFINITION 3. A test can be *manipulated* with probability $q \in [0, 1]$ if there is a strategy ζ such that for every $\omega \in \Omega$,

$$\zeta\big(\{P \in \Delta : \omega \notin T(P)\}\big) \geq q.$$

If a test is manipulable with high probability, then a uninformed, but strategic expert is likely to pass the test regardless of how the data unfolds and how much data is available.

DEFINITION 4. A test is *nonmanipulable* if for every strategy ζ there is a cylinder C_ζ such that for every path $\omega \in C_\zeta$,

$$\zeta\big(\{P \in \Delta : \omega \notin T(P)\}\big) = 0.$$

Nonmanipulable tests can reject uninformed experts. No matter which strategy Bob employs, there is a finite history that, if observed, discredits him. These are the only tests that are likely to pass informed experts and may reject uninformed ones.

3. Merging of opinions. We now review the main concepts behind the Blackwell–Dubins theorem.

DEFINITION 5. Let $P, Q \in \mathbb{P}$. The probability P *merges* with Q if for every $\varepsilon > 0$

$$\lim_{t\to\infty} Q\Big(\Big\{\omega : \sup_{E\in\Sigma} |P(E|\omega^t) - Q(E|\omega^t)| > \varepsilon\Big\}\Big) = 0.$$

The expression $\sup_{E\in\Sigma} |P(E|\omega^t) - Q(E|\omega^t)|$ is the distance between the forecasts of P and Q, conditional on the evidence available at time t and along the path ω. The probability P merges with Q if, under Q, this distance goes to 0 in probability. In particular, if Q accurately describes the data generating process then the predictions of P are eventually accurate with high probability.

In this paper, merging is formulated in terms of convergence in probability rather than almost sure convergence [as in Blackwell and Dubins (1962)]. As is well known, convergence in probability is particularly convenient in the context of finitely additive probabilities. See, for instance, the discussion in Berti and Rigo (2006).

It is clear that for merging to occur, P and Q must be compatible *ex-ante*. The notion of absolute continuity formalizes this intuition.

DEFINITION 6. Let $P, Q \in \mathbb{P}$. The probability Q is *absolutely continuous* with respect to P, that is, $Q \ll P$, if for every sequence of events $(E_n)_{n=1}^{\infty}$,

$$\text{if } P(E_n) \to 0 \text{ then } Q(E_n) \to 0.$$

If P is σ-additive, then the definition is equivalent to requiring that every event null under P is also null under Q. Moreover, if P is a Savagean probability and Q is a probability satisfying $Q \ll P$, then Q is Savagean as well.

Absolute continuity is (essentially) necessary for merging.

PROPOSITION 1. *Let $P, Q \in \mathbb{P}$ and $P(\omega^t) > 0$ for every cylinder ω^t. If P merges with Q then $Q \ll P$.*

In their seminal paper, Blackwell and Dubins show that when P and Q are σ-additive then absolute continuity suffices for merging.

THEOREM 1 (Blackwell and Dubins). *Let P and Q be σ-additive probability measures on $(\Omega, \mathcal{B})$. If $Q \ll P$, then P merges with Q.*

One interpretation of the Blackwell–Dubins theorem is that multiple repetitions of Bayes' rule lead to an agreement among agents who initially hold compatible opinions. Another interpretation is that the predictions of Bayesian learners will eventually be accurate (provided that absolute continuity holds). However, the Blackwell–Dubins theorem does not extend to all opinions. This motivates the next definition.

DEFINITION 7. A probability $P \in \mathbb{P}$ satisfies the *Blackwell–Dubins property* if for every $Q \in \mathbb{P}$,

$$\text{if } Q \ll P \text{ then } P \text{ merges with } Q. \tag{3.1}$$

Let Δ_{BD} be the set of all opinions that satisfy the Blackwell–Dubins property.

So, an opinion P satisfies the Blackwell–Dubins property if it merges to any compatible opinion Q. We show in the Appendix that Δ_{BD} is *strictly* contained in the set of all opinions. That is, some opinions satisfy the Blackwell–Dubins property, while others do not. We also show that the set of probabilities satisfying the Blackwell–Dubins property strictly contains the set of σ-additive probabilities. We refer the reader to Example 1 and Theorem 10, respectively.

Any exogenously given (or honestly reported) opinion can be tested and rejected, whether or not the Blackwell–Dubins property holds. Thus, in the absence of strategic considerations, the connection between the merging and the testing of opinions is weak. We now show that this connection is much stronger when there are incentive problems.

4. Main results.

THEOREM 2. *Consider the case where any opinion is allowed. Let T be a test that Δ-controls for type I errors with probability $1 - \varepsilon$. The test T can be manipulated with probability $1 - \varepsilon - \delta$, for every $\delta \in (0, 1 - \varepsilon]$.*

THEOREM 3. *Consider the case where opinions must satisfy the Blackwell–Dubins property. Fix $\varepsilon \in (0, 1]$. There exists a test T that Δ_{BD}-controls for type I error with probability $1 - \varepsilon$ and is nonmanipulable.*

If Bob is free to announce any opinion, then he cannot be meaningfully tested and discredited. Given any test that controls for type I error, Bob can design a strategy which prevents rejection. However, if Bob is required to announce opinions satisfying the Blackwell–Dubins property, then it is possible to test and discredit him. These results show a strong connection between the merging and the testing of opinions, but only when there are incentive problems and agents may misrepresent what they know.

We now illustrate the basic ideas behind the proof of the two results. The proof of Theorem 3 relies on a characterization of the set of probabilities that satisfy the Blackwell–Dubins property. This characterization is also crucial for the proof of Theorem 4 below. We show that $P \in \mathbb{P}$ satisfies the Blackwell–Dubins property if and only if it is an extreme point of the set of probabilities $E(P) \subseteq \mathbb{P}$ which agree with P on every cylinder.

The proof of necessity in this characterization is simple. Suppose, by contradiction, that P can be written as the convex combination $P = \alpha Q + (1-\alpha)R$, where Q and R belong to $E(P)$. Clearly, both Q and R are absolutely continuous with respect to P. However, P does not merge to Q or R. The intuition is that Q and R agree on every finite history and so, the available data delivers equal support to them. The converse requires a deeper argument and relies on Plachky's (1976) theorem, which states that P is an extreme point of $E(P)$ if and only if the probability of every event can be approximated by the probabilities of cylinders.

Given our characterization, the proof of Theorem 3 can be sketched as follows: Let P be an opinion satisfying the Blackwell–Dubins property. Given that P is strongly nonatomic, we can divide Ω into a partition $\{A_1, \dots, A_n\}$ of events such that each of them has probability less than ε. This property is a direct implication of Savage's postulate P6 and plays an important role in our result. For general opinions, the events $\{A_1, \dots, A_n\}$ may have no useful structure and may not even be Borel sets. However, since P is an extreme point of $E(P)$, we can invoke Plachky's theorem a second time and show that each A_i can be chosen to be a cylinder. Now fix a path ω. Let us say it belongs to A_1. By definition, there is a time t such that $\omega^t = A_1$. We now define a test T such that $T(P) = \omega^t$ (note that the period t depends on the opinion P because the partition depends on it). By definition, the test Δ_{BD}-controls for type I errors with probability $1-\varepsilon$. Furthermore, it is a nonmanipulable test. Given any strategy ζ we can find a period m large enough such that ω^m rejects all opinions in the (finite) support of ζ. Therefore, in ω^m, the probability of passing the test under ζ is 0.

We now sketch the proof of Theorem 2. Consider a zero-sum game between Nature and the expert. Nature chooses an opinion P and the expert chooses a strategy ζ (a random device producing opinions). The payoff of the expert is the probability of passing the test. For each opinion P chosen by Nature there exists a strategy for the expert (to report P) that gives him a payoff of at least $1-\varepsilon$. If Fan's (1953) Minmax theorem applies then there exists a strategy ζ that guarantees the expert a payoff of at least $1-\varepsilon$ for *every* opinion chosen by Nature. In this case, the test is manipulable.

Fan's Minmax theorem requires Nature's action space to be compact and her payoffs to be (lower semi) continuous. The main difficulty is that the set of opinions is not compact in the natural topology, the weak* topology. Hence, Fan's Minmax theorem cannot be directly applied. We consider a new game, defined as above except that Nature can choose any probability in $\mathbb{P}$ (not necessarily Savagean). By the Riesz representation and the Banach–Alaoglu theorems, the set of all finitely additive probabilities satisfy the necessary continuity and compactness conditions for Fan's Minmax theorem. However, if now Nature chooses a *non*-Savagean probability M, the expert cannot replicate her choice because he is restricted to opinions.

Based on the celebrated Hammer–Sobczyk decomposition theorem, we show the following approximation result: For every $M \in \mathbb{P}$, there is an opinion P such that $M(U) \leq P(U)$ for every union U of cylinders. Thus, $M(T(P)) \leq P(T(P) \leq \varepsilon$. It follows that if Natures chooses M and the expert chooses P then he passes the test with probability at least $1 - \varepsilon$. The proof is now concluded invoking Fan's Minmax theorem.

5. Category tests. Theorem 3 provides conditions under which it is feasible to discredit strategic experts. However, even a nonmanipulable test can be strategically passed on some paths. Under σ-additivity, Dekel and Feinberg (2006) and Olszewski and Sandroni (2009a) construct nonmanipulable *category* tests, where uninformed experts fail in all, but a topologically small (i.e., *meager*) set of paths. We now show a difficulty in following this approach in the general case of opinions that satisfy the Blackwell–Dubins property.

DEFINITION 8. A collection $\mathcal{I}$ of subsets of Ω is a *strictly proper ideal* if it satisfies the following properties:

(1) If $S \in \mathcal{I}$ and $R \subseteq S$ then $R \in \mathcal{I}$;
(2) If $R, S \in \mathcal{I}$ then $R \cup S \in \mathcal{I}$; and
(3) No cylinder belongs to $\mathcal{I}$.

A strictly proper ideal is a collection of sets which can be regarded as "small." Property (1) is the natural requirement that if a set S is considered small then a set R contained in S must also be considered small. Properties (2) and (3) are satisfied by most commonly used notions of "small" sets, such as countable, nowhere dense, meager, sets of Lebesgue measure zero and shy sets. To clarify our terminology, recall that an ideal is a collection of subsets satisfying properties (1) and (2). An ideal is proper if Ω does not belong to it. We refer to the elements of a strictly proper ideal as *small* sets and to their complements as *large* sets.

Strictly proper ideals can be defined in terms of probabilities. Given $P \in \mathbb{P}$, define a set N to be *P-null* if there exists an event E that satisfies $N \subseteq E$ and $P(E) = 0$. The collection of P-null sets is a strictly proper ideal whenever P satisfies $P(\omega^t) > 0$ for every cylinder ω^t.

THEOREM 4. *Let $\mathcal{I}$ be a strictly proper ideal. There exists an opinion $P \in \Delta_{\mathrm{BD}}$ such that $P(E) = 0$ for every event E in $\mathcal{I}$.*

There exists an opinion that satisfies the Blackwell–Dubins property and finds all small events to be negligible. The proof of this result relies on the characterization of the set Δ_{BD} of opinions satisfying the Blackwell–Dubins property that we discussed in the previous section. Theorem 4 shows a basic tension between the control of type I errors and the use of genericity arguments. Suppose Alice intends to design a test that discredits Bob on a large set of paths, irrespectively of his strategy. Then the set of paths $(T(P))^c$ that do not reject opinion P must be small [otherwise Bob could simply announce opinion P and pass the test on $(T(P))^c$, a nonsmall set of realizations]. But if P is the opinion obtained from Theorem 4, we must have $P((T(P))^c) = 0$. So, the test cannot control for type I errors. We have just proved the following corollary.

COROLLARY 1. *Let $\mathcal{I}$ be a strictly proper ideal. For every test T which Δ_{BD}-controls type I errors with positive probability there exists a strategy ζ such that the set*

$$\{\omega : \zeta(\{P : \omega \notin T(P)\}) = 1\}$$

is not small.

Thus, the stronger nonmanipulable tests in Dekel and Feinberg (2006) and Olszewski and Sandroni (2009a) cannot be obtained in the general case of opinions that satisfy the Blackwell–Dubins property.

6. Extensions. In this section, we extend our analysis to the case where the set of per-period outcomes may be infinite.

6.1. *Setup.* Let $\mathcal{X}$ be a separable metric space of outcomes and denote by Ω the set of paths $\mathcal{X}^\infty$. As before, ω^t is the cylinder of length $t \geq 0$ with base $\omega \in \Omega$ (in particular $\omega^0 = \Omega$). The set Ω is endowed with the product topology and a σ-algebra Σ containing all open sets. We denote by $\mathbb{P}$ the set of finitely additive probabilities on (Ω, Σ) and by Δ the subset of opinions (i.e., strongly nonatomic probabilities).

6.2. *Conditional probabilities.* Let $\mathcal{H}$ be the set of all cylinders. We say that a function

$$P : \Sigma \times \mathcal{H} \to [0, 1]$$

is a *conditional probability* if for every $t \geq 0$ and $\omega \in \Omega$:

(1) $P(\cdot|\omega^t) \in \mathbb{P}$;
(2) $P(\omega^t|\omega^t) = 1$; and
(3) $P(E \cap \omega^{t+n}|\omega^t) = P(E|\omega^{t+n})P(\omega^{t+n}|\omega^t)$ for any event E and $n \geq 0$.

The definition of conditional probability follows Berti, Regazzini and Rigo (1998), where properties (1)–(3) are justified on the basis of de Finetti's *coherence* principle: a real function P defined on $\Sigma \times \mathcal{H}$ satisfies properties (1)–(3) if and only if a bookie, who sets $P(E|\omega^t)$ as the price of a conditional bet on event E, cannot incur in a Dutch book. We refer the reader to Regazzini (1985, 1987), de Finetti (1990) and Berti and Rigo (2002) for a precise statement and a formal discussion.

A conditional probability is a *conditional opinion* if $P(\cdot|\Omega)$ is strongly nonatomic. We denote by $\mathbb{P}^*$ and by Δ^* the sets of conditional probabilities and conditional opinions, respectively. To simplify the exposition, given an event E and a conditional probability P, we use the notation $P(E)$ instead of the more precise $P(E|\Omega)$.

At time 0, Bob is required to announce a conditional opinion P. So, a test T is now a function mapping each conditional opinion P to an open subset $T(P)$ of Ω. The definitions of type I errors, manipulable and nonmanipulable tests are analogous to the definitions of Section 2 and can be obtained by replacing Δ with Δ^*.

6.3. *Merging*. We now extend the definition of merging of opinions. We say that $\mathcal{X}$ is a *discrete space* if it is countable and endowed with the discrete topology.

DEFINITION 9. Let $\mathcal{X}$ be a discrete space. If $P, Q \in \mathbb{P}^*$, the conditional probability P *merges* with Q if for every $\varepsilon > 0$,

$$\lim_{t\to\infty} Q\Big(\Big\{\omega : \sup_{E\in\Sigma} \big|P(E|\omega^t) - Q(E|\omega^t)\big| > \varepsilon\Big\}\Big) = 0.$$

The next definition is based on Blackwell and Dubins (1962).

DEFINITION 10. Let $\mathcal{X}$ be a discrete space. A conditional probability P satisfies the *Blackwell–Dubins property* if for every probability $Q \in \mathbb{P}$ such that $Q \ll P(\cdot|\Omega)$ there exists a conditional probability $\widetilde{Q}$ such that

$$\widetilde{Q}(\cdot|\Omega) = Q \text{ and } P \text{ merges with } \widetilde{Q}.$$

Let Δ^*_{BD} be the set of all conditional opinions that satisfy the Blackwell–Dubins property.

We now show that the connection between testability and merging of opinions extends to this setup.

6.4. *Results*.

THEOREM 5. *Let $\mathcal{X}$ be a separable metric space. Consider the case where any conditional opinion is allowed. Let T be a test that Δ^*-controls for type I errors with probability $1-\varepsilon$. The test T can be manipulated with probability $1-\varepsilon-\delta$, for every $\delta \in (0, 1-\varepsilon]$.*

THEOREM 6. *Let $\mathcal{X}$ be a discrete space. Consider the case where conditional opinions must satisfy the Blackwell–Dubins property. Fix* $\varepsilon \in (0, 1]$. *There exists a test T that Δ^*_{BD}-controls for type I error with probability* $1 - \varepsilon$ *and is nonmanipulable.*

If it is possible for Bob to announce any conditional opinion, then he cannot be meaningfully tested and discredited. If Bob is restricted to conditional opinions satisfying the Blackwell–Dubins property, then it is possible to test and discredit him.

The proof of Theorem 5 follows the proof of Theorem 2. The proof of Theorem 6 is based on the following result: a conditional opinion $P \in \Delta^*_{\mathrm{BD}}$ satisfies $\lim_{t\to\infty} P(\omega^t) = 0$ for every path ω. This step requires a new argument, because the characterization of Blackwell–Dubins property used in the proof of Theorem 3 does not readily extend to the case where $\mathcal{X}$ is infinite. Once this continuity property is shown to hold, the proof continues as in Theorem 3. We fix a path ω and for each $P \in \Delta^*_{\mathrm{BD}}$ we choose a large enough period t_P such that $P(\omega^{t_P}) < \varepsilon$. Because $\mathcal{X}$ is assumed to be a discrete space, each cylinder ω^{t_P} is open. Therefore, we can define test T such that $T(P) = \omega^{t_P}$ for every $P \in \Delta^*_{\mathrm{BD}}$. Following the proof of Theorem 3, we show that T is nonmanipulable.

APPENDIX A

We now provide an example of an opinion that violates the Blackwell–Dubins property.

EXAMPLE 1. Let $\Omega = \{0, 1\}^\infty$ and $\Sigma = \mathcal{B}$. Denote by $X_1, X_2, \ldots$ the coordinate projections on Ω. For every $n \geq 1$, let P_n be the σ-additive probability defined as

$$P_n(X_k = 0) = 2^{-k} \qquad \text{for } k \leq n \quad \text{and} \quad P_n(X_k = 0) = 1 \qquad \text{for } k > n$$

and let P_∞ be the σ-additive probability defined as $P_\infty(X_k = 0) = 2^{-k}$ for all k.

Consider the opinion $P = \frac{1}{2}P_\infty + \frac{1}{2}\int P_n \, d\lambda(n)$, where λ is a finitely additive probability on $(\mathbb{N}, 2^{\mathbb{N}})$ such that $\lambda(\{n\}) = 0$ for every n. The finite additivity of the mixture $\int P_n \, d\lambda(n)$ may reflect the difficulty of predicting *when* the per-period probability of observing the outcome 0 will change from 0.5 to 1.

Clearly, $P_\infty \ll P$. However, P does not merge with P_∞. To this end, let A be the set of all paths where the outcome 1 appears infinitely often. Then $P_n(A) = 0$ for every n and $P_\infty(A) = 1$. For every cylinder ω^t, we have

$$P(\omega^t) = \frac{1}{2}P_\infty(\omega^t) + \frac{1}{2}\int_{\{n:n>t\}} P_n(\omega^t) \, d\lambda(n) = P_\infty(\omega^t)$$

and moreover,

$$
\begin{aligned}
P_\infty(A|\omega^t) - P(A|\omega^t) &= 1 - \frac{(1/2)P_\infty(A \cap \omega^t) + (1/2)\int_n P_n(A \cap \omega^t)\,d\lambda(n)}{P_\infty(\omega^t)} \\
&= 1 - \frac{1}{2}P_\infty(A|\omega^t) \\
&= \frac{1}{2}
\end{aligned}
$$

for every ω and every t. Thus, P does not merge with P_∞.

A.1. Preliminaries. To minimize repetitions, throughout the Appendix Ω stands for either $\{0, 1\}^\infty$ or $\mathcal{X}^\infty$. For every algebra $\mathcal{A}$ of subsets of Ω denote by $\mathbb{P}(\mathcal{A})$ the space of finitely additive probabilities defined on $(\Omega, \mathcal{A})$. When $\mathcal{A} = \Sigma$, we write $\mathbb{P}$ instead of $\mathbb{P}(\Sigma)$. We denote by $\Delta \subseteq \mathbb{P}$ the set of opinions (strongly nonatomic probabilities). The space $\mathbb{P}(\mathcal{A})$ is endowed with the weak* topology. It is the coarsest topology for which the functional $P \mapsto \int \varphi\, dP$ is continuous for every function $\varphi : \Omega \to \mathbb{R}$ that has finite range and is measurable with respect to $\mathcal{A}$. This should not be confused with the more common weak* topology generated by bounded *continuous* functions.

APPENDIX B: MERGING OF OPINIONS

In this subsection, we describe of the set of opinions on $\Omega = \{0, 1\}^\infty$ that satisfy the Blackwell–Dubins property. We first show that absolute continuity is essentially a necessary condition for merging.

PROOF OF PROPOSITION 1. Assume Q is not absolutely continuous with respect to P. Then there exists a sequence of events $(E_n)_{n=1}^\infty$ and some $\alpha > 0$ such that $P(E_n) \to 0$ but $Q(E_n) > \alpha$ for every n. Suppose P merges with Q. For every t, let $\mathcal{C}_t$ be a collection of pairwise disjoint cylinders of length t such that $Q(\omega^t) > 0$ for every $\omega^t \in \mathcal{C}_t$ and $Q(\cup\mathcal{C}_t) = 1$. Fix $\delta \in (0, \frac{\alpha}{4})$. There exists a time T large enough such that for every $t \geq T$ there is a subset $\mathcal{D}_t \subseteq \mathcal{C}_t$ such that $Q(\cup\mathcal{D}_t) \geq 1 - \delta$ and $\sup_{E\in\Sigma} |Q(E|\omega^t) - P(E|\omega^t)| \leq \delta$ for every $\omega^t \in \mathcal{D}_t$. Because $P(\omega^t) > 0$ for every ω^t, the expression $P(E|\omega^t)$ is well defined. For every event E_n,

$$Q(E_n) = \sum_{\omega^t \in \mathcal{D}_t} Q(E_n|\omega^t)Q(\omega^t) + \sum_{\omega^t \in \mathcal{C}_t - \mathcal{D}_t} P(E|\omega^t)P(\omega^t)$$

hence, $\sum_{\omega^t \in \mathcal{D}_t} Q(E_n|\omega^t)Q(\omega^t) \geq \alpha - \delta$. Define

$$\mathcal{E}_n = \left\{\omega^t \in \mathcal{D}_t : Q(E_n|\omega^t) \geq \frac{\alpha}{2}\right\}.$$

We have

$$
\begin{aligned}
\alpha - \delta &\leq \sum_{\omega^t \in \mathcal{D}_t} Q(E_n|\omega^t) Q(\omega^t) \\
&= \sum_{\omega^t \in \mathcal{E}_n} Q(E_n|\omega^t) Q(\omega^t) + \sum_{\omega^t \in \mathcal{D}_t - \mathcal{E}_n} Q(E_n|\omega^t) Q(\omega^t) \\
&\leq Q(\cup \mathcal{E}_n) + \frac{\alpha}{2} Q(\cup \mathcal{D}_t - \cup \mathcal{E}_n) \\
&\leq Q(\cup \mathcal{E}_n) + \frac{\alpha}{2}
\end{aligned}
$$

hence, $Q(\cup \mathcal{E}_n) \geq \frac{\alpha}{2} - \delta$. Now let n^* be large enough such that $\frac{\alpha}{2} - P(E_{n^*}|\omega^t) > \delta$ for every ω. Then, for every $\omega^t \in \mathcal{E}_{n^*}$,

$$
\begin{aligned}
\sup_{E \in \Sigma} \left| Q(E|\omega^t) - P(E|\omega^t) \right| &\geq Q(E_{n^*}|\omega^t) - P(E_{n^*}|\omega^t) \\
&\geq \frac{\alpha}{2} - P(E_{n^*}|\omega^t) \\
&> \delta.
\end{aligned}
$$

To summarize, $Q(\{\omega : \sup_{E \in \Sigma} |Q(E|\omega^t) - P(E|\omega^t)| > \delta\}) > \frac{\alpha}{2} - \delta$ for every $t \geq T$. Therefore, $Q(\cup \mathcal{D}_t) \leq 1 - (\frac{\alpha}{2} - \delta)$. By definition, $Q(\cup \mathcal{D}_t) \geq 1 - \delta$; hence, $\delta \geq \frac{\alpha}{4}$. A contradiction. Hence, P does not merge with Q. $\square$

Our main result is a characterization of the set of opinions that satisfy the Blackwell–Dubins property. We first recall some results on extensions of finitely additive probabilities. Let $\mathcal{A}_1$ and $\mathcal{A}_2$ be two algebras of subsets of Ω such that $\mathcal{A}_1 \subseteq \mathcal{A}_2$. Given $P \in \mathbb{P}(\mathcal{A}_1)$ and $Q \in \mathbb{P}(\mathcal{A}_2)$, call Q an *extension of P from $\mathcal{A}_1$ to $\mathcal{A}_2$* if $P(A) = Q(A)$ for every $A \in \mathcal{A}_1$. Let $E(P, \mathcal{A}_1, \mathcal{A}_2)$ be the set of extensions of P from $\mathcal{A}_1$ to $\mathcal{A}_2$. As is well known, the set $E(P, \mathcal{A}_1, \mathcal{A}_2)$ is nonempty. Moreover, it is a convex and compact subset of $\mathbb{P}(\mathcal{A}_2)$. The set of extreme points of $E(P, \mathcal{A}_1, \mathcal{A}_2)$ has been studied in great generality. We refer the reader to Lipecki (2007) and Plachky (1976) for further results and references.

THEOREM 7 [Plachky (1976)]. *Fix two algebras $\mathcal{A}_1 \subseteq \mathcal{A}_2$ and $P \in \mathbb{P}(\mathcal{A}_1)$. A probability $Q \in E(P, \mathcal{A}_1, \mathcal{A}_2)$ is an extreme point of $E(P, \mathcal{A}_1, \mathcal{A}_2)$ if and only if for every $\varepsilon > 0$ and $A_2 \in \mathcal{A}_2$ there exists $A_1 \in \mathcal{A}_1$ such that $Q(A_2 \triangle A_1) < \varepsilon$.*

Let $\mathcal{F}$ be the algebra generated by all cylinders of $\{0, 1\}^\infty$. An event belongs to $\mathcal{F}$ if and only if it is a finite union of (pairwise disjoint) cylinders. Recall that $\mathcal{B}$ is the Borel σ-algebra induced on Ω by the product topology. Then $\mathcal{F} \subseteq \mathcal{B} \subseteq \Sigma$. For every $P \in \mathbb{P}$ let $P_{\mathcal{F}}$ be the restriction of P on $\mathcal{F}$. It is easy to see that $P_{\mathcal{F}}$ is σ-additive. By Carathéodory theorem, it admits a σ-additive extension from $\mathcal{F}$ to $\mathcal{B}$, denoted by P_σ.

The next result is well known.

LEMMA 1. *Let $\Omega = \{0, 1\}^\infty$. For all Q and P in $\mathbb{P}$, if $Q_\mathcal{F} \ll P_\mathcal{F}$ then $Q_\sigma \ll P_\sigma$.*

We can now state our main result on merging.

THEOREM 8. *Let $\Omega = \{0, 1\}^\infty$. For every $P \in \mathbb{P}$, the following are equivalent*:

(1) *P is an extreme point of $E(P_\mathcal{F}, \mathcal{F}, \Sigma)$.*
(2) *P satisfies the Blackwell–Dubins property.*
(3) *For all $Q, R \in \mathbb{P}$, if $P = \alpha Q + (1-\alpha)R$ for some $\alpha \in (0, 1)$ then P merges with Q and R.*

PROOF. (1) $\Rightarrow$ (2). Let P be an extreme point of $E(P_\mathcal{F}, \mathcal{F}, \Sigma)$. If $Q \ll P$, then $Q_\sigma \ll P_\sigma$ by Lemma 1. By the Blackwell–Dubins theorem,

$$Q_\sigma\Big(\Big\{\omega : \lim_{t\to\infty}\Big(\sup_{B\in\mathcal{B}}\big|Q_\sigma(B|\omega^t) - P_\sigma(B|\omega^t)\big|\Big) = 0\Big\}\Big) = 1.$$

In particular, the sequence of random variables

$$\Big(\omega \mapsto \sup_{F\in\mathcal{F}}\big|Q_\sigma(F|\omega^t) - P_\sigma(F|\omega^t)\big|\Big)_{t=1}^\infty$$

converges to 0, Q_σ-almost surely. Therefore, the sequence converges in probability. For each $\varepsilon > 0$,

$$\lim_{t\to\infty} Q_\sigma\Big(\Big\{\omega : \sup_{F\in\mathcal{F}}\big|Q_\sigma(F|\omega^t) - P_\sigma(F|\omega^t)\big| > \varepsilon\Big\}\Big) = 0.$$

Since the last expression only involves events belonging to $\mathcal{F}$,

$$\lim_{t\to\infty} Q\Big(\Big\{\omega : \sup_{F\in\mathcal{F}}\big|Q(F|\omega^t) - P(F|\omega^t)\big| > \varepsilon\Big\}\Big) = 0.$$

The proof is complete by showing that

$$\sup_{E\in\Sigma}\big|Q(E|\omega^t) - P(E|\omega^t)\big| = \sup_{F\in\mathcal{F}}\big|Q(F|\omega^t) - P(F|\omega^t)\big|$$

for every ω^t such that $Q(\omega^t) > 0$.

To this end, fix an event $E \in \Sigma$ and a cylinder ω^t such that $Q(\omega^t) > 0$. By Plachky's theorem, there exists a sequence of events $(F_n)_{n=1}^\infty$ in $\mathcal{F}$ such that $P(E \triangle F_n) \to 0$ as $n \to \infty$. For each n, the inequality

$$\begin{aligned}\big|Q(E|\omega^t) - P(E|\omega^t)\big| \le{}& \big|Q(E|\omega^t) - Q(F_n|\omega^t)\big| \\ &+ \big|Q(F_n|\omega^t) - P(F_n|\omega^t)\big| \\ &+ \big|P(F_n|\omega^t) - P(E|\omega^t)\big|\end{aligned}$$

implies

$$\begin{aligned}&\big|Q(E|\omega^t)-P(E|\omega^t)\big|\\&\quad\le Q(E\bigtriangleup F_n|\omega^t)+\sup_{F\in\mathcal{F}}\big|Q(F|\omega^t)-P(F|\omega^t)\big|+P(E\bigtriangleup F_n|\omega^t).\end{aligned}$$

Because $P(E \bigtriangleup F_n) \to 0$ and $Q(\omega^t) > 0$, it follows that $P(\omega^t) > 0$ and $P(E \bigtriangleup F_n|\omega^t) \to 0$. Absolute continuity implies $Q(E \bigtriangleup F_n|\omega^t) \to 0$. Therefore,

$$\big|Q(E|\omega^t)-P(E|\omega^t)\big|\le\sup_{F\in\mathcal{F}}\big|Q(F|\omega^t)-P(F|\omega^t)\big|$$

thus

$$\sup_{E\in\Sigma}\big|Q(E|\omega^t)-P(E|\omega^t)\big|\le\sup_{F\in\mathcal{F}}\big|Q(F|\omega^t)-P(F|\omega^t)\big|$$

as claimed.

(2) $\Rightarrow$ (3). If $P = \alpha Q + (1-\alpha)R$ then $Q \ll P$, hence P merges with Q.

(3) $\Rightarrow$ (1). Assume by way of contradiction that P is not an extreme point of $E(P_{\mathcal{F}}, \mathcal{F}, \Sigma)$. Then there exist Q, R in $E(P_{\mathcal{F}}, \mathcal{F}, \Sigma)$ such that $P = \alpha Q + (1-\alpha)R$, $\alpha \in (0, 1)$ and $Q \neq R$. By assumption, P merges with Q. Let $\mathcal{C}_t$ be a collection of pairwise disjoint cylinders of length t such that $Q(\omega^t) > 0$ for every $\omega^t \in \mathcal{C}_t$ and $Q(\cup\mathcal{C}_t) = 1$. For every t large enough, there exists a subset $\mathcal{D}_t \subseteq \mathcal{C}_t$ such that $Q(\cup\mathcal{D}_t) \ge 1-\varepsilon$ and $\sup_{E\in\Sigma}|Q(E|\omega^t) - P(E|\omega^t)| \le \varepsilon$ for every $\omega^t \in \mathcal{D}_t$. For every event E,

$$\begin{aligned}\big|Q(E)-P(E)\big|&=\Big|\sum_{\omega^t\in\mathcal{C}_t}Q(E|\omega^t)Q(\omega^t)-\sum_{\omega^t\in\mathcal{C}_t}P(E|\omega^t)P(\omega^t)\Big|\\&=\Big|\sum_{\omega^t\in\mathcal{C}_t}\big(P(E|\omega^t)-Q(E|\omega^t)\big)Q(\omega^t)\Big|\\&\le\sum_{\omega\in\mathcal{D}_t}\big|P(E|\omega^t)-Q(E|\omega^t)\big|Q(\omega^t)\\&\quad+\sum_{\omega\in\mathcal{C}_t-\mathcal{D}_t}\big|P(E|\omega^t)-Q(E|\omega^t)\big|Q(\omega^t)\\&\le\varepsilon Q(\cup\mathcal{D}_t)+\big(1-Q(\cup\mathcal{D}_t)\big)\\&\le2\varepsilon,\end{aligned}$$

where the first two equalities follow from the fact that $Q(\omega^t) = P(\omega^t)$ for all ω^t. Since E and ε are arbitrary, we have $P = Q$. A contradiction. Therefore, P must be an extreme point of $E(P_{\mathcal{F}}, \mathcal{F}, \Sigma)$. $\square$

The next result shows a useful property of opinions that satisfy the Blackwell–Dubins property.

THEOREM 9. *Let $\Omega = \{0,1\}^\infty$. For every $P \in \Delta_{\mathrm{BD}}$ and every $\varepsilon > 0$, there exists a partition $\{C_1, \dots, C_n\}$ of Ω such that for each $i = 1, \dots, n$, C_i is a cylinder and $P(C_i) \le \varepsilon$.*

PROOF. Let $P \in \Delta_{\mathrm{BD}}$ and fix $\varepsilon > 0$. Because P is strongly nonatomic, then there exists a partition $\{E_1, \dots, E_m\}$ of events such that $P(E_i) < \frac{\varepsilon}{2}$ for every $i = 1, \dots, m$. By Theorem 8, P is an extreme point of $E(P_{\mathcal{F}}, \mathcal{F}, \Sigma)$. By Plachky's theorem, for each i we can find a sequence $(F_{i,k})_{k=1}^\infty$ in $\mathcal{F}$ such that $P(E_i \triangle F_{i,k}) \to 0$ as $k \to \infty$. Choose K large enough such that $P(E_i \triangle F_{i,K}) < \frac{\varepsilon}{2m}$ for each i.

Let $F_1 = F_{1,K}$ and define $F_i = F_{i,K} - \bigcup_{j=1}^{i-1} F_{j,K}$ for each $i = 2, \dots, m$. Let $F_{m+1} = \Omega - \bigcup_{i=1}^m F_{i,K}$ and consider the partition $\{F_1, \dots, F_{m+1}\}$. It satisfies $P(F_i) \le P(F_{i,K}) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2m}$ for each $i = 2, \dots, m$. Moreover,

$$P(F_{m+1}) = P\left(\left(\bigcup_{i=1}^m E_i\right) - \left(\bigcup_{i=1}^m F_{i,K}\right)\right) \le P\left(\bigcup_{i=1}^m (E_i - F_{i,K})\right) \le \frac{\varepsilon}{2}.$$

Therefore, $P(F_i) \le \varepsilon$ for each $F_i \in \{F_1, \dots, F_{m+1}\}$. Because each F_i is a finite union of pairwise disjoint cylinders, the proof is complete. □

The next theorem shows that for every sigma additive probability P we can find a continuum of probabilities that agree with P on every cylinder, fail σ-additivity but satisfy the Blackwell–Dubins property. A related result appears in Lipecki (2001).

THEOREM 10. *Let $\Omega = \{0,1\}^\infty$. For every σ-additive probability $P \in \mathbb{P}$, the set*

$$\{Q \in \Delta_{\mathrm{BD}} : Q_{\mathcal{F}} = P_{\mathcal{F}}, Q \text{ is not } \sigma\text{-additive}\}$$

has cardinality at least $\mathfrak{c}$.

PROOF. Fix a collection $\{D_\xi : \xi \in [0,1]\}$ of pairwise disjoint, Borel and dense subsets of Ω [see, e.g., Ceder (1966)]. For every $\xi \in [0,1]$, let $\mathcal{A}_\xi$ be the algebra generated by $\mathcal{F} \cup \{D_\xi\}$. That is,

$$A_\xi = \{(F_1 \cap D_\xi) \cup (F_2 \cap D_\xi^c) : F_1, F_2 \in \mathcal{F}\}.$$

Let $\rho_\xi \in \mathbb{P}(\mathcal{A}_\xi)$ be defined as

$$\rho_\xi(F \cap D_\xi) = P(F)$$

for every $F \in \mathcal{F}$. Because D_ξ is dense, then $F \cap D_\xi \neq \varnothing$ for every $F \in \mathcal{F}$. Hence, ρ_ξ is well defined. It satisfies $\rho_\xi(D_\xi) = 1$.

For each $\xi \in [0,1]$ fix an extreme point P_ξ of $E(\rho_\xi, \mathcal{A}_\xi, \Sigma)$. We claim it is also an extreme point of $E(P_{\mathcal{F}}, \mathcal{F}, \Sigma)$. By construction, $P_\xi \in E(P_{\mathcal{F}}, \mathcal{F}, \Sigma)$. Now suppose $P_\xi = \alpha Q + (1-\alpha)R$, with $\alpha \in [0,1]$ and $Q, R \in E(P_{\mathcal{F}}, \mathcal{F}, \Sigma)$. Because

$P_\xi(D_\xi) = 1$, then $Q(D_\xi) = R(D_\xi) = 1$. Hence, $P(F) = Q(F \cap D_\xi) = R(F \cap D_\xi)$ for every $F \in \mathcal{F}$. Therefore, $Q, R \in E(\rho_\xi, \mathcal{A}_\xi, \Sigma)$. By assumption, P_ξ is an extreme point of $E(\rho_\xi, \mathcal{A}_\xi, \Sigma)$. Hence, $P = Q = R$. This concludes the proof of the claim.

By Theorem 8, each P_ξ satisfies the Blackwell–Dubins property. Moreover, each P_ξ agrees with P on every cylinder. Hence, each σ-additive P_ξ must agree with P on every Borel sets. Because the sets $\{D_\xi : \xi \in [0, 1]\}$ are Borel and pairwise disjoint, at most one probability in $\{P_\xi : \xi \in [0, 1]\}$ agrees with P on every Borel set. Thus, there exists at most one σ-additive probability in $\{P_\xi : \xi \in [0, 1]\}$. Therefore, the set $\{P_\xi : \xi \in [0, 1], P_\xi \neq P\}$, which is included in $\{Q \in \Delta_{\text{BD}} : Q_{\mathcal{F}} = P_{\mathcal{F}}, Q \text{ is not } \sigma\text{-additive}\}$, has cardinality $\mathfrak{c}$. This completes the proof. □

APPENDIX C: STRONGLY NONATOMIC PROBABILITIES

We now provide a technical result important for the proofs of Theorems 2 and 5. Throughout this subsection, $\Omega = \mathcal{X}^\infty$.

A $\{0, 1\}$-*probability* is a probability $Z \in \mathbb{P}$ that satisfies $Z(E) \in \{0, 1\}$ for every $E \in \Sigma$.

THEOREM 11. *Let $\mathcal{E} \subseteq \Sigma - \{\varnothing\}$ be closed under finite intersection. There exists a $\{0, 1\}$-probability Z such that $Z(E) = 1$ for every $E \in \mathcal{E}$.*

PROOF. This is a corollary of the ultrafilter theorem. See, for instance, Aliprantis and Border (2006), Theorem 2.19. □

Every $P \in \mathbb{P}$ can be decomposed into a strongly nonatomic part and a mixture of countably many $\{0, 1\}$-probabilities.

THEOREM 12 [Sobczyk and Hammer (1944)]. *For every $P \in \mathbb{P}$, there exists an opinion $P_s \in \Delta$ and a sequence $(Z_i)_{i=1}^\infty$ of $\{0, 1\}$-probabilities such that*

$$P = \alpha P_s + (1 - \alpha) \sum_{i=1}^{\infty} \beta_i Z_i,$$

where $\alpha, \beta_i \in [0, 1]$ for every i and $\sum_{i=1}^\infty \beta_i = 1$.

Given an algebra $\mathcal{A}$, $P \in \mathbb{P}(\mathcal{A})$ is *strongly continuous* if for every $\varepsilon > 0$ there exists a partition $\{A_1, \dots, A_n\}$ of Ω such that $A_i \in \mathcal{A}$ and $P(A_i) < \varepsilon$ for every i.

THEOREM 13. *Let $\mathcal{A}$ be a σ-algebra. A probability $P \in \mathbb{P}(\mathcal{A})$ is strongly continuous if and only if it is strongly nonatomic.*

PROOF. See Bhaskara Rao and Bhaskara Rao (1983), Theorem 11.4.5. □

THEOREM 14. *For every* $P \in \mathbb{P}$, *there exists an opinion* $\widetilde{P} \in \Delta$ *such that*

$$P(U) \leq \widetilde{P}(U)$$

for every open set U.

PROOF. We first prove the result for the case where P is a $\{0,1\}$-probability. To this end, fix a $\{0,1\}$-probability Z. Let $\mathcal{C} = \{U : U \text{ open}, Z(U) = 1\}$. The collection $\mathcal{C}$ is closed under finite intersection.

We now construct a sequence $(D_n)_{n=1}^{\infty}$ of countable, dense and pairwise disjoint subsets of Ω. The proof of this claim proceeds by induction. The space $\Omega = \mathcal{X}^{\infty}$ is separable, so it has a countable dense subset D_1. Assume that for some N the sets $D_1, \ldots, D_N$ have been defined and satisfy the desired properties. Let $(V_k)_{k=1}^{\infty}$ be a countable base of Ω and for each k pick a path $\omega_k \in V_k - \bigcup_{n=1}^{N} D_n$. Let $D_{N+1} = \{\omega_1, \omega_2, \ldots\}$. This completes the induction step and the proof of the claim.

For every n, the collection

$$\mathcal{C}_n = \{U \cap D_n : U \in \mathcal{C}\}$$

is closed under finite intersection and does not contain the empty set. By Theorem 11, for every n there exists a $\{0,1\}$-probability Z_n such that $Z_n(E) = 1$ for every $E \in \mathcal{C}_n$. For each n and for every open set U, if $Z(U) = 1$ then $U \cap D_n \in \mathcal{C}_n$ and $Z_n(U) = 1$. Hence, $Z(U) \leq Z_n(U)$. Let λ be a strongly continuous finitely additive probability on $(\mathbb{N}, 2^{\mathbb{N}})$ and define the function $\widetilde{Z} : \Sigma \to [0,1]$ as

$$\widetilde{Z}(E) = \int_{\mathbb{N}} Z_n(E)\, d\lambda(n)$$

for every $E \in \Sigma$.

It follows from the additivity of the integral that $\widetilde{Z} \in \mathbb{P}$. For every open set U, if $Z(U) = 1$ then $\widetilde{Z}(U) = \int_{\mathbb{N}} Z_n(U)\, d\lambda(n) = \int_{\mathbb{N}} 1\, d\lambda(n) = 1$. Hence, $Z(U) \leq \widetilde{Z}(U)$. It remains to prove that $\widetilde{Z}$ is strongly nonatomic. By Theorem 13, it is enough to prove it is strongly continuous. Fix $\varepsilon > 0$. Since λ is strongly continuous, we can find a partition $\{\Pi_1, \ldots, \Pi_k\}$ of $\mathbb{N}$ such that $\lambda(\Pi_i) \leq \varepsilon$ for every $i = 1, \ldots, k$. Consider now the partition $\{\bigcup_{m \in \Pi_1} D_m, \ldots, \bigcup_{m \in \Pi_k} D_m, (\bigcup_{m \in \mathbb{N}} D_m)^c\}$ of Ω. For every $i = 1, \ldots, k$ and n, we have $Z_n(\bigcup_{m \in \Pi_i} D_m) = 1_{\Pi_i}(n)$ where 1_{Π_i} is the indicator function of Π_i. Therefore,

$$\begin{aligned}
\widetilde{Z}\biggl(\bigcup_{m \in \Pi_i} D_m\biggr) &= \int_{\mathbb{N}} Z_n\biggl(\bigcup_{m \in \Pi_i} D_m\biggr)\, d\lambda(n) \\
&= \int_{\mathbb{N}} 1_{\Pi_i}(n)\, d\lambda(n) \\
&= \lambda(\Pi_i) \leq \varepsilon
\end{aligned}$$

and $\widetilde{Z}((\bigcup_{m \in \mathbb{N}} D_m)^c) = 0$. This proves that $\widetilde{Z}$ is strongly continuous.

Now let P be any finitely additive probability. By the Hammer–Sobczyk decomposition, we can write P as the convex combination

$$P = \alpha P_s + (1-\alpha)\sum_{i=1}^{\infty} \beta_i Z_i,$$

where P_s is strongly nonatomic and each Z_i is a $\{0, 1\}$-probability. For each i, let $\widetilde{Z}_i$ be an opinion such that $Z_i(U) \le \widetilde{Z}_i(U)$ for every open set U. Define $\widetilde{P} = \alpha P_s + (1-\alpha)\sum_{i=1}^{\infty} \beta \widetilde{Z}_i$. It is easy to see that $\widetilde{P}$ is strongly continuous. By Theorem 13, it is an opinion. For every open set U, we have

$$\begin{aligned} P(U) &= \alpha P_s(U) + (1-\alpha)\sum_{i=1}^{\infty} \beta_i Z_i(U) \\ &\le \alpha P_s(U) + (1-\alpha)\sum_{i=1}^{\infty} \beta_i \widetilde{Z}_i(U) = \widetilde{P}(U) \end{aligned}$$

as desired. □

APPENDIX D: PROOFS OF THEOREMS 2–6

THEOREM 15 [Fan (1953)]. *Let X and Y be convex subsets of two vector spaces. Let $f : X \times Y \to \mathbb{R}$. If X is compact Hausdorff and f is concave with respect to Y and convex and lower semicontinuous with respect to X, then*

$$\sup_{y\in Y} \min_{x\in X} f(x, y) = \min_{x\in X} \sup_{y\in Y} f(x, y).$$

See Fan (1953) for a more general version of this theorem.

PROOF OF THEOREM 2. Define the function $V : \mathbb{P} \times \Delta_f \Delta \longrightarrow \mathbb{R}$ as

$$V(P, \zeta) = \int \zeta\big(\{Q \in \Delta : \omega \notin T(Q)\}\big)\, dP(\omega)$$

for every $(P, \zeta) \in \mathbb{P} \times \Delta_f \Delta$. The function V is affine in each variable and continuous with respect to $\mathbb{P}$. The weak* topology is Hausdorff. Moreover, it follows from the Riesz representation and Banach–Alaoglu theorems that $\mathbb{P}$ is compact [see Aliprantis and Border (2006), Theorems 14.4 and 6.21]. All the conditions of Fan's Minmax theorem are verified, therefore,

$$\sup_{\zeta\in\Delta_f\Delta} \min_{P\in\mathbb{P}} V(P, \zeta) = \min_{P\in\mathbb{P}} \sup_{\zeta\in\Delta_f\Delta} V(P, \zeta). \tag{D.1}$$

By Theorem 14, for every $P \in \mathbb{P}$ there exists $\widetilde{P} \in \Delta$ such that $P(U) \le \widetilde{P}(U)$ for every open set U. Because $T(\widetilde{P})$ is an open set, then $P(T(\widetilde{P})) \le \widetilde{P}(T(\widetilde{P})) \le \varepsilon$. That is, $V(P, \delta_{\widetilde{P}}) = 1 - P(T(\widetilde{P})) \ge 1 - \varepsilon$. Thus,

$$\sup_{\zeta\in\Delta_f\Delta} \min_{P\in\mathbb{P}} V(P, \zeta) = \min_{P\in\mathbb{P}} \sup_{\zeta\in\Delta_f\Delta} V(P, \zeta) \ge \min_{P\in\mathbb{P}} V(P, \delta_{\widetilde{P}}) \ge 1 - \varepsilon.$$

For every $\delta \in (0, 1-\varepsilon]$, there exists a strategy ζ such that $V(P, \zeta) > 1-\varepsilon-\delta$ for every $P \in \mathbb{P}$. In particular,

$$V(\delta_\omega, \zeta) = \zeta\big(\{Q \in \Delta : \omega \notin T(Q)\}\big) \geq 1-\varepsilon-\delta$$

for every path ω. $\square$

PROOF OF THEOREM 3. Fix $\varepsilon > 0$ and a path ω. By Theorem 9, for every $P \in \Delta_{\mathrm{BD}}$ we can choose a partition $\{C_1, \dots, C_n\}$ of cylinders such that $P(C_i) < \varepsilon$ for every $i = 1, \dots, n$. Let $\omega \in C_i$. There exists a time t_P such that $\omega^{t_P} = C_i$. Hence, $P(\omega^{t_P}) < \varepsilon$. Define a test T as $T(P) = \omega^{t_P}$ for every opinion $P \in \Delta_{\mathrm{BD}}$. The test Δ_{BD}-controls for type I errors with probability $1-\varepsilon$.

Now, let $(P_1, \dots, P_n)$ be the support of a strategy ζ. Choose a time t such that $t \geq t_{P_i}$ for each $i = 1, \dots, n$. Then

$$\omega^t \subseteq \bigcap_{i=1}^{n} T(P_i)$$

hence, $\zeta(\{P \in \Delta : \tilde{\omega} \notin T(P)\}) = 0$ for every path $\tilde{\omega}$ in ω^t. $\square$

PROOF OF THEOREM 4. Recall that $\mathcal{F}$ is the algebra generated by cylinders. Fix a probability $\pi \in \mathbb{P}(\mathcal{F})$ such that $\pi(\omega^t) \to 0$ as $t \to \infty$ for every path ω. Now let $\mathcal{I}$ be a strictly proper ideal and consider the collection of events

$$\mathcal{A} = \big\{(F \cap L) \cup S : F \in \mathcal{F}, S \in \mathcal{I} \cap \Sigma, L^c \in \mathcal{I} \cap \Sigma\big\}. \tag{D.2}$$

We prove it is an algebra. If $(F \cap L) \cup S \in \mathcal{A}$, then its complement is equal to

$$(F \cap L)^c \cap S^c = \big(F^c \cup L^c\big) \cap S^c = \big(F^c \cap S^c\big) \cup \big(L^c \cap S^c\big)$$

since S^c is large and $L^c \cap S^c$ is small we have that $(F^c \cap S^c) \cup (L^c \cap S^c) \in \mathcal{A}$. Using the notation in (D.2), let $(F_1 \cap L_1) \cup S_1$ and $(F_2 \cap L_2) \cup S_2$ belong to $\mathcal{A}$. Observe that $L = L_1 \cap L_2$ is large and, therefore $L_1 - L$ and $L_2 - L$ are small. We can write

$$\begin{aligned}(F_1 \cap L_1) \cup S_1 \cup (F_2 \cap L_2) \cup S_2 &= (F_1 \cap L) \cup (F_2 \cap L) \cup S \\ &= \big((F_1 \cup F_2) \cap L\big) \cup S,\end{aligned} \tag{D.3}$$

where $S = (F_1 \cap (L_1 - L)) \cup S_1 \cup (F_2 \cap (L_2 - L)) \cup S_2$ is a union of small sets. This proves that $(F_1 \cap L_1) \cup S_1 \cup (F_2 \cap L_2) \cup S_2 \in \mathcal{A}$. We conclude that $\mathcal{A}$ is an algebra. By construction, $\mathcal{F} \subseteq \mathcal{A} \subseteq \Sigma$.

Define a set function $\widetilde{\pi} : \mathcal{A} \to [0, 1]$ as

$$\widetilde{\pi}\big((F \cap L) \cup S\big) = \pi(F)$$

for each $(F \cap L) \cup S \in \mathcal{A}$.

We verify that $\widetilde{\pi}$ is well defined. Using the notation in (D.2), let $(F_1 \cap L_1) \cup S_1 = (F_2 \cap L_2) \cup S_2$. Equivalently,

$$(F_1 \cap F_2 \cap L_1) \cup \big(F_1 \cap F_2^c \cap L_1\big) \cup S_1 = (F_2 \cap L_2) \cup S_2.$$

Therefore, $F_1 \cap F_2^c \cap L_1 \subseteq S_2$. Hence, $F_1 \cap F_2^c \cap L_1$ is small. But also $F_1 \cap F_2^c \cap L_1^c \subseteq L_1^c$ is small, hence $F_1 \cap F_2^c$ is small. The set $F_1 \cap F_2^c$ is either empty or a union of cylinders. By the definition of strictly proper ideal $F_1 \cap F_2^c$ must be empty. Similarly, $F_2 \cap F_1^c = \varnothing$. Hence, $F_1 = F_2$, and $\widetilde{\pi}((F_1 \cap L_1) \cup S_1) = \widetilde{\pi}((F_2 \cap L_2) \cup S_2)$.

We prove $\widetilde{\pi}$ is additive. Let $(F_1 \cap L_1) \cup S_1$ and $(F_2 \cap L_2) \cup S_2$ be two disjoint sets belonging to $\mathcal{A}$. The sets F_1 and F_2 are disjoint. To see this, notice that $F_1 \cap F_2 \cap L_1 \cap L_2 = \varnothing$ implies $F_1 \cap F_2 \subseteq (L_1 \cup L_2)^c$. The set $F_1 \cap F_2$ is either empty or a union of cylinders. Since $(L_1 \cup L_2)^c$ is small, it must be empty. Let $L = L_1 \cap L_2$ and $S = (F_1 \cap (L_1 - L)) \cup S_1 \cup (F_2 \cap (L_2 - L)) \cup S_2$. Similar to (D.3), we have

$$\begin{aligned}
\widetilde{\pi}\big((F_1 \cap L_1) \cup S_1 \cup (F_2 \cap L_2) \cup S_2\big) &= \widetilde{\pi}\big(\big((F_1 \cup F_2) \cap L\big) \cup S\big) \\
&= \pi(F_1 \cup F_2) \\
&= \pi(F_1) + \pi(F_2) \\
&= \widetilde{\pi}\big((F_1 \cap L_1) \cup S_1\big) + \widetilde{\pi}\big((F_2 \cap L_2) \cup S_2\big).
\end{aligned}$$

Therefore, $\widetilde{\pi}$ is a finitely additive probability defined on $(\Omega, \mathcal{A})$. By construction, it satisfies $\widetilde{\pi}(S) = 0$ for every $S \in \mathcal{I} \cap \Sigma$.

Consider the set of extensions $E(\widetilde{\pi}, \mathcal{A}, \Sigma)$ and let P be one of its extreme points. We prove that P is an extreme point of $E(\pi, \mathcal{F}, \Sigma)$. Write P as $P = \alpha Q + (1-\alpha)R$ with $Q, R \in E(\pi, \mathcal{F}, \Sigma)$. Let $\widetilde{\pi}_Q$ and $\widetilde{\pi}_R$ be the restriction of Q and R on $\mathcal{A}$. Since P is an extension of $\widetilde{\pi}$, we have $\widetilde{\pi} = \alpha\widetilde{\pi}_Q + (1-\alpha)\widetilde{\pi}_R$. We claim that $\widetilde{\pi} = \widetilde{\pi}_Q = \widetilde{\pi}_R$. For every $S \in \mathcal{I} \cap \Sigma$, since $\widetilde{\pi}(S) = 0$, then $\widetilde{\pi}_Q(S) = \widetilde{\pi}_R(S) = 0$. Therefore,

$$\widetilde{\pi}_Q\big((F \cap L) \cup S\big) = \widetilde{\pi}_Q(F \cap L) = \widetilde{\pi}_Q(F) = \pi(F) = \widetilde{\pi}\big((F \cap L) \cup S\big)$$

for every event $(F \cap L) \cup S \in \mathcal{A}$. The same is true for $\widetilde{\pi}_R$. Therefore, $\widetilde{\pi}_Q = \widetilde{\pi}_R = \widetilde{\pi}$. This proves that $Q, R \in E(\widetilde{\pi}, \mathcal{A}, \Sigma)$. Because P is an extreme point of $E(\widetilde{\pi}, \mathcal{A}, \Sigma)$, then $P = Q = R$. This concludes the proof that P is an extreme point of $E(\pi, \mathcal{F}, \Sigma)$. By Theorem 8, P satisfies the Blackwell–Dubins property.

It remains to prove that P is strongly nonatomic. Since $\pi(\omega^t) \to 0$ for every ω, π is strongly continuous. A fortiori, P is strongly continuous and also strongly nonatomic by Theorem 13. □

PROOF OF THEOREM 5. Define the function $V : \mathbb{P} \times \Delta_f \Delta^* \to \mathbb{R}$ as

$$V(P, \zeta) = \int \zeta\big(\big\{Q \in \Delta^* : \omega \notin T(Q)\big\}\big)\, dP(\omega)$$

for all $(P,\zeta) \in \mathbb{P} \times \Delta_f \Delta^*$. Given $P \in \mathbb{P}$, by Theorem 14 there exists an opinion Q such that $P(U) \le Q(U)$ for every open set U. By Theorem 4 in Regazzini (1985), we can find a conditional opinion $Q^* \in \Delta^*$ such that $Q = Q^*(\cdot|\Omega)$. Then $P(T(Q^*)) \le Q(T(Q^*)) = Q^*(T(Q^*)) \le \varepsilon$. The proof is complete by replicating the argument used in the proof of Theorem 2. $\square$

PROOF OF THEOREM 6. We first prove that for every $P \in \Delta^*_{\mathrm{BD}}$ and every path ω, $\lim_t P(\omega^t) = 0$. We argue by contradiction. Let ω_o be a path such that $\inf_t P(\omega_o^t) = \delta > 0$. Fix a sequence of positive real numbers (ξ_t) such that

$$P(\omega_o^t) = \delta + \xi_t$$

for every t.

Fix $\varepsilon \in (0, \frac{1}{2})$. Because P is strongly nonatomic, for every time t we can find an event $F^t \subseteq \omega_0^t$ such that $P(F^t) = \frac{1}{2} P(\omega_0^t)$. For every n, we have

$$\begin{aligned} P(F^t|\omega_0^{t+n}) &= \frac{P(F^t \cap \omega_o^{t+n})}{P(\omega_o^{t+n})} \\ &= \frac{P(F^t) - P(\omega_o^t - \omega_o^{t+n})}{\delta + \xi_{t+n}} \\ &= \frac{(1/2)(\delta + \xi_t) - (\xi_t - \xi_{t+n})}{\delta + \xi_{t+n}} = \frac{1}{2}\frac{\delta + \xi_t}{\delta + \xi_{t+n}} - \frac{\xi_t - \xi_{t+n}}{\delta + \xi_{t+n}}. \end{aligned}$$

We can therefore fix $\bar{t}$ large enough such that $F = F^{\bar{t}}$ satisfies $P(F|\omega_0^t) \in (\frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon)$ for every $t > \bar{t}$.

Let Q be the opinion defined as

$$Q(E) = \frac{P(E \cap F)}{P(F)}$$

for every event E. Then $Q \ll P(\cdot|\Omega)$. By Theorem 4 in Regazzini (1985), we can find a conditional opinion $\widetilde{Q}$ satisfying $\widetilde{Q}(\cdot|\Omega) = Q$. The proof of the claim will be concluded by showing that P does not merge with $\widetilde{Q}$. Note that for every $t > \bar{t}$

$$\widetilde{Q}(\omega_0^t) = \frac{P(F \cap \omega_0^t)}{P(\omega_0^t)} \frac{P(\omega_0^t)}{P(F)} = P(F|\omega_0^t) \frac{P(\omega_0^t)}{(1/2) P(\omega_0^{\bar{t}})} = P(F|\omega_0^t) 2 \frac{\delta + \xi_t}{\delta + \xi_{\bar{t}}}$$

hence, for all $t > \bar{t}$

$$\widetilde{Q}(\omega_0^t) \ge (1 - 2\varepsilon) \frac{\delta}{\delta + \xi_{\bar{t}}}.$$

Moreover, for every $t > \bar{t}$,

$$\begin{aligned} \widetilde{Q}\Big(\Big\{\omega : \sup_E |Q(E|\omega^t) - P(E|\omega^t)| > \frac{1}{2} - \varepsilon\Big\}\Big) &> \widetilde{Q}\Big(\Big\{\omega : P(F|\omega^t) < \frac{1}{2} + \varepsilon\Big\}\Big) \\ &\ge \widetilde{Q}(\omega_0^t), \end{aligned}$$

where the first equality follows from $\widetilde{Q}(F|\omega_0^t) = 1$ and the second equality follows from $P(F|\omega_0^t) < \frac{1}{2} + \varepsilon$. Because the sequence $(\widetilde{Q}(\omega_0^{\bar{t}}), \widetilde{Q}(\omega_0^{\bar{t}+1}), \dots)$ is bounded away from 0, P does not merge to $\widetilde{Q}$. Therefore, we can conclude that for every $P \in \Delta^*_{\mathrm{BD}}$ and every path ω, $\lim_t P(\omega^t) = 0$.

Now fix a path ω and $\varepsilon > 0$. We can find for every $P \in \Delta^*_{\mathrm{BD}}$ a time t_P such that $P(\omega^{t_P}) < \varepsilon$. Because $\mathcal{X}$ is endowed with the discrete topology, ω^{t_P} is an open set. Let $T(P) = \omega^{t_P}$ for every $P \in \Delta^*_{\mathrm{BD}}$. The test Δ^*_{BD}-controls for type I error with probability $1 - \varepsilon$. The same argument in the proof of Theorem 3 shows that T is nonmanipulable. □

Acknowledgements. We thank Ehud Kalai, Wojciech Olszewski, Eran Shmaya, Marciano Siniscalchi and Rakesh Vohra for useful discussions. We are grateful to the Editor and the referees for their thoughtful comments, for simplifying Example 1 and for stimulating the results in Section 6. We also thank the seminar audiences at the Fifth Transatlantic Theory Workshop, the Summer meeting of the Econometric Society 2012, XIII Latin American Workshop in Economic Theory, the 4th Workshop on Stochastic Methods in Game Theory, the Washington University seminar series and the Paris Game Theory Seminar. All errors are ours.

REFERENCES

AL-NAJJAR, N., POMATTO, L. and SANDRONI, A. (2013). An economic model of induction. Mimeo.

AL-NAJJAR, N. I. and WEINSTEIN, J. (2008). Comparative testing of experts. *Econometrica* **76** 541–559. MR2406865

AL-NAJJAR, N. I., SANDRONI, A., SMORODINSKY, R. and WEINSTEIN, J. (2010). Testing theories with learnable and predictive representations. *J. Econom. Theory* **145** 2203–2217. MR2888863

ALIPRANTIS, C. D. and BORDER, K. C. (2006). *Infinite Dimensional Analysis*: *A Hitchhiker's Guide*, 3rd ed. Springer, Berlin. MR2378491

BABAIOFF, M., BLUMROSEN, L., LAMBERT, N. and REINGOLD, O. (2011). Only valuable experts can be valued. In *Proceedings of the* 12*th ACM Conference on Electronic Commerce* 221–222. ACM.

BERTI, P., REGAZZINI, E. and RIGO, P. (1998). Well-calibrated, coherent forecasting systems. *Theory Probab. Appl.* **42** 82–102.

BERTI, P. and RIGO, P. (2002). On coherent conditional probabilities and disintegrations. *Ann. Math. Artif. Intell.* **35** 71–82. MR1899945

BERTI, P. and RIGO, P. (2006). Finitely additive uniform limit theorems. *Sankhyā* **68** 24–44. MR2301563

BHASKARA RAO, K. P. S. and BHASKARA RAO, M. (1983). *Theory of Charges*: *A Study of Finitely Additive Measures*. *Pure and Applied Mathematics* **109**. Academic Press, New York. MR0751777

BLACKWELL, D. and DUBINS, L. (1962). Merging of opinions with increasing information. *Ann. Math. Statist.* **33** 882–886. MR0149577

CEDER, J. (1966). On maximally Borel resolvable spaces. *Rev. Roumaine Math. Pures Appl.* **11** 89–94. MR0206900

CESA-BIANCHI, N. and LUGOSI, G. (2006). *Prediction*, *Learning*, *and Games*. Cambridge Univ. Press, Cambridge. MR2409394

D'ARISTOTILE, A., DIACONIS, P. and FREEDMAN, D. (1988). On merging of probabilities. *Sankhyā Ser. A* **50** 363–380. MR1065549

DAWID, A. P. (1985). Calibration-based empirical probability. *Ann*. *Statist*. **13** 1251–1285. MR0811493

DE FINETTI, B. (1990). *Theory of Probability*. Wiley, Chicester.

DEKEL, E. and FEINBERG, Y. (2006). Non-Bayesian testing of a stochastic prediction. *Rev. Econom. Stud*. **73** 893–906. MR2260750

DIACONIS, P. and FREEDMAN, D. (1986). On the consistency of Bayes estimates. *Ann*. *Statist*. **14** 1–26. MR0829555

FAN, K. (1953). Minimax theorems. *Proc*. *Natl*. *Acad*. *Sci*. *USA* **39** 42–47. MR0055678

FEINBERG, Y. and LAMBERT, N. (2011). Mostly calibrated. Mimeo.

FEINBERG, Y. and STEWART, C. (2008). Testing multiple forecasters. *Econometrica* **76** 561–582. MR2406866

FORTNOW, L. and VOHRA, R. V. (2009). The complexity of forecast testing. *Econometrica* **77** 93–105. MR2477844

FOSTER, D. P. and VOHRA, R. (2011). Calibration: Respice, adspice, prospice. Mimeo.

FOSTER, D. P. and VOHRA, R. V. (1998). Asymptotic calibration. *Biometrika* **85** 379–390. MR1649119

FOSTER, D. and YOUNG, P. (2001). On the impossibility of predicting the behavior of rational agents. *Proc*. *Natl*. *Acad*. *Sci*. *USA* **98** 12848–12853.

FOSTER, D. P. and YOUNG, H. P. (2003). Learning, hypothesis testing, and Nash equilibrium. *Games Econom*. *Behav.* **45** 73–96. MR2022861

FUDENBERG, D. and KREPS, D. M. (1993). Learning mixed equilibria. *Games Econom*. *Behav.* **5** 320–367. MR1227915

FUDENBERG, D. and LEVINE, D. K. (1998). *The Theory of Learning in Games*. *MIT Press Series on Economic Learning and Social Evolution* **2**. MIT Press, Cambridge, MA. MR1629477

FUDENBERG, D. and LEVINE, D. K. (1999). An easier way to calibrate. *Games Econom*. *Behav.* **29** 131–137. MR1729313

FUDENBERG, D. and LEVINE, D. K. (2009). Learning and equilibrium. *Annu*. *Rev. Econ*. **1** 385–420.

GRADWOHL, R. and SALANT, Y. (2011). How to buy advice. Mimeo.

GRADWOHL, R. and SHMAYA, E. (2013). Tractable falsification. Mimeo.

HART, S. and MAS-COLELL, A. (2013). *Simple Adaptive Strategies*: *From Regret-Matching to Uncoupled Dynamics*. World Scientific, Singapore.

HU, T. W. and SHMAYA, E. (2013). Expressible inspections. *Theor. Econ*. **8** 263–280. MR3062153

JACKSON, M. O., KALAI, E. and SMORODINSKY, R. (1999). Bayesian representation of stochastic processes under learning: de Finetti revisited. *Econometrica* **67** 875–893. MR1693917

KALAI, E. and LEHRER, E. (1993a). Rational learning leads to Nash equilibrium. *Econometrica* **61** 1019–1045. MR1234792

KALAI, E. and LEHRER, E. (1993b). Subjective equilibrium in repeated games. *Econometrica* **61** 1231–1240. MR1234796

LEHRER, E. (2001). Any inspection is manipulable. *Econometrica* **69** 1333–1347. MR1848782

LEHRER, E. and SMORODINSKY, R. (1996a). Compatible measures and merging. *Math*. *Oper. Res*. **21** 697–706. MR1403312

LEHRER, E. and SMORODINSKY, R. (1996b). Merging and learning. In *Statistics*, *Probability and Game Theory*. *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **30** 147–168. IMS, Hayward, CA. MR1481779

LIPECKI, Z. (2001). Cardinality of the set of extreme extensions of a quasi-measure. *Manuscripta Math*. **104** 333–341. MR1828879

LIPECKI, Z. (2007). On compactness and extreme points of some sets of quasi-measures and measures. IV. *Manuscripta Math*. **123** 133–146. MR2306629

MONDERER, D., SAMET, D. and SELA, A. (1997). Belief affirming in learning processes. *J. Econom. Theory* **73** 438–452.

NACHBAR, J. H. (1997). Prediction, optimization, and learning in repeated games. *Econometrica* **65** 275–309. MR1433485

NACHBAR, J. H. (2001). Bayesian learning in repeated games of incomplete information. *Soc. Choice Welf.* **18** 303–326. MR1838240

NACHBAR, J. H. (2005). Beliefs in repeated games. *Econometrica* **73** 459–480. MR2117239

OLSZEWSKI, W. (2011). Calibration and expert testing. Mimeo.

OLSZEWSKI, W. and PESKI, M. (2011). The principal agent approach to testing-experts. *American Economic Journal*: *Microeconomics* **3** 89–113.

OLSZEWSKI, W. and SANDRONI, A. (2007). Contracts and uncertainty. *Theor. Econ*. **2** 1–13.

OLSZEWSKI, W. and SANDRONI, A. (2008). Manipulability of future-independent tests. *Econometrica* **76** 1437–1466. MR2468602

OLSZEWSKI, W. and SANDRONI, A. (2009a). A nonmanipulable test. *Ann. Statist*. **37** 1013–1039. MR2502659

OLSZEWSKI, W. and SANDRONI, A. (2009b). Strategic manipulation of empirical tests. *Math. Oper. Res*. **34** 57–70. MR2542989

OLSZEWSKI, W. and SANDRONI, A. (2011). Falsifiability. *American Economic Review* **101** 788–818.

PLACHKY, D. (1976). Extremal and monogenic additive set functions. *Proc. Amer. Math. Soc*. **54** 193–196. MR0419711

REGAZZINI, E. (1985). Finitely additive conditional probabilities. *Rend. Sem. Mat. Fis. Milano* **55** 69–89. MR0933711

REGAZZINI, E. (1987). de Finetti's coherence and statistical inference. *Ann. Statist*. **15** 845–864. MR0888444

SANDRONI, A. (1998). Necessary and sufficient conditions for convergence to Nash equilibrium: The almost absolute continuity hypothesis. *Games Econom. Behav*. **22** 121–147. MR1604792

SANDRONI, A. (2003). The reproducible properties of correct forecasts. *Internat. J. Game Theory* **32** 151–159. MR2046685

SANDRONI, A., SMORODINSKY, R. and VOHRA, R. V. (2003). Calibration with many checking rules. *Math. Oper. Res*. **28** 141–153. MR1961271

SAVAGE, L. J. (1954). *The Foundations of Statistics*. Wiley, New York. MR0063582

SCHERVISH, M. and SEIDENFELD, T. (1990). An approach to consensus and certainty with increasing evidence. *J. Statist. Plann. Inference* **25** 401–414. MR1064433

SHMAYA, E. (2008). Many inspections are manipulable. *Theor. Econ*. **3** 367–382.

SOBCZYK, A. and HAMMER, P. C. (1944). A decomposition of additive set functions. *Duke Math. J*. **11** 839–846. MR0011164

STEWART, C. (2011). Nonmanipulable Bayesian testing. *J. Econom. Theory* **146** 2029–2041. MR2887773

VOVK, V. and SHAFER, G. (2005). Good randomized sequential probability forecasting is always possible. *J. R. Stat. Soc. Ser. B Stat. Methodol*. **67** 747–763. MR2210691

WALKER, S. G., LIJOI, A. and PRÜNSTER, I. (2005). Data tracking and the understanding of Bayesian consistency. *Biometrika* **92** 765–778. MR2234184

YOUNG, H. P. (2002). On the limits to rational learning. *European Economic Review* **46** 791–799.

YOUNG, P. (2004). *Strategic Learning and Its Limits*. Oxford Univ. Press, Oxford, UK.

MANAGERIAL ECONOMICS & DECISION SCIENCES DEPARTMENT
KELLOGG SCHOOL OF MANAGEMENT
2001 SHERIDAN ROAD
EVANSON, ILLINOIS 60208
USA
E-MAIL: l-pomatto@kellogg.northwestern.edu
al-najjar@kellogg.northwestern.edu
sandroni@kellogg.northwestern.edu
URL: http://www.kellogg.northwestern.edu/faculty/pomatto/index.htm
http://www.kellogg.northwestern.edu/faculty/directory/al-najjar_nabil.aspx
http://www.kellogg.northwestern.edu/faculty/directory/sandroni_alvaro.aspx