## VALUE OF PERSISTENT INFORMATION

MARCIN PĘSKI
*Department of Economics, University of Toronto*

JUUSO TOIKKA
*Department of Economics, Massachusetts Institute of Technology*

# VALUE OF PERSISTENT INFORMATION

MARCIN PĘSKI
Department of Economics, University of Toronto

JUUSO TOIKKA
Department of Economics, Massachusetts Institute of Technology

We develop a theory of how the value of an agent's information advantage depends on the persistence of information. We focus on strategic situations with strict conflict of interest, formalized as stochastic zero-sum games where only one of the players observes the state that evolves according to a Markov operator. Operator $Q$ is said to be better for the informed player than operator $P$ if the value of the game under $Q$ is higher than under $P$ regardless of the stage game. We show that this defines a convex partial order on the space of ergodic Markov operators. Our main result is a full characterization of this partial order, intepretable as an ordinal notion of persistence relevant for games. The analysis relies on a novel characterization of the value of a stochastic game with incomplete information.

KEYWORDS: Dynamic games, incomplete information, Markov chains.

## 1. INTRODUCTION

IN MANY STRATEGIC SETTINGS, an agent has to decide how to make use of an information advantage, trading off short-term benefits from the use of the information against the cost of revealing some of it through his actions. Examples of such situations range from insider trading to international conflicts. In more formal terms, the problem is embedded in dynamic game-theoretic models where players have private information about the state of the world. The optimal use of information then depends on how likely, and for how long, the information is expected to remain relevant in the future. For instance, if the state is distributed independently across the periods of interaction, then revealing information about its current value is costless, which contrasts with the decidedly more careful calculus by Aumann and Maschler (1995) for the case where the state remains fixed forever.

Motivated by this observation, we develop a theory of how the value of an agent's information advantage depends on the persistence of information. Following Aumann and Maschler (1995), we focus on strategic situations with strict conflict of interest, formalized as zero-sum games, as for such games the value provides an unambiguous, single-valued solution concept that facilitates comparison of information structures. However, as we discuss below, the notion of persistence that emerges from our analysis also characterizes the comparative statics of the (limit) equilibrium-payoff set in a class of repeated Bayesian non-zero-sum games studied by Athey and Bagwell (2008), Escobar and Toikka (2013), and Hörner, Takahashi, and Vieille (2015), among others.

Specifically, we consider discounted stochastic zero-sum games where only the maximizer observes the current state of a Markov chain that is governed by a transition matrix, or operator, $P$. In every period, the players are engaged in the same zero-sum stage game

$g$, where payoffs depend on the state observed by the maximizer. The minimizer's only information about the state is given by the maximizer's past actions and his understanding of the data generating process.

As a discounted zero-sum game, the above game has a value; we focus on the limit value as the discount factor converges to 1. This is crucial for tractability. It also makes our results directly comparable to those in the literature. Importantly, the trade-off between short-term gains and future losses from the use of information—absent in Aumann and Maschler's analysis of undiscounted games with a permanent state—remains relevant even in the limit if the state is ergodic.

Our first result, Theorem 1, shows that the (limit) value of the game can be computed from an auxiliary one-shot problem where the maximizer can choose the information structure subject to it being stable under operator $P$ in a certain sense. Theorem 1 yields Aumann and Maschler's formula for the permanent state as a special case and complements the results of Renault (2006) and Neyman (2008), who showed the existence of the value and optimal strategies for Markov games with one-sided incomplete information. Our auxiliary problem can be viewed as a dynamic persuasion or information design problem where a principal commits to an information revelation policy, though in contrast to Kamenica and Gentzkow (2011), or Ely (2017), here also the principal takes an action.

In order to study the effects of persistence—our main object of interest—we restrict attention to ergodic operators. Given two such operators $P$ and $Q$, we say that $Q$ is better for the maximizer than $P$ if, for every game $g$, the (limit) value of the stochastic game under $Q$ is higher than under $P$. We show that this relation is a partial order on the space of ergodic operators. Our main result, Theorem 2, characterizes this partial order. A series of corollaries describes it further.

Theorem 2 suggests interpreting our partial order as an ordinal notion of persistence. To explain this, we adopt the standard view of an information structure (also known as an experiment) as a probability distribution over posterior beliefs about the state. In the game, the minimizer's information changes from one period to another due to the state transition described by operator $P$. If the distribution of his beliefs is $\nu$ at the end of the period, then it is $P\nu$ at the beginning of the next one. Intuitively, because the transition is stochastic, we would expect it to lead to loss of information. If this is true in the precise sense that $\nu$ is more informative than $P\nu$ according to the Blackwell ordering, then we say that $\nu$ is stabilizable under $P$. (Any stationary $\nu$ has this property since then the information revealed by the maximizer during the period leads from $P\nu$ back to $\nu$, and hence $\nu$ is more informative than $P\nu$.)

Now consider two operators, $P$ and $Q$. If $P\nu$ is (Blackwell) more informative than $Q\nu$, then this means that information is more persistent under $P$ than under $Q$. Theorem 2 shows that $Q$ is better for the maximizer than $P$ if and only if, for all distributions $\nu$ stabilizable under $P$, $P\nu$ is more informative than $Q\nu$.

To illustrate, consider the following simple example due to Renault (2006).

EXAMPLE 1: There are two states. The maximizer chooses $U$ or $D$, the minimizer chooses $L$ or $R$. Stage-game payoffs are given in Figure 1, where in each cell the first entry is the payoff in state 1 and the second the payoff in state 2. The state remains the same with probability $\rho$ and changes with probability $1 - \rho$. The invariant distribution assigns equal probability to both states (for any $0 < \rho < 1$).

In Example 1, the maximizer wants to match his action to the state, but doing so reveals the state to the minimizer. This is harmless if $\rho = \frac{1}{2}$ (i.e., if the state is i.i.d.), but comes at

| | L | R |
|---|---|---|
| U | 1, 0 | 0, 0 |
| D | 0, 0 | 0, 1 |

FIGURE 1.—Stage-game payoffs for Example 1.

a cost if $\rho > \frac{1}{2}$ as then the minimizer knows that the state is likely to remain the same. In particular, Hörner, Rosenberg, Solan, and Vieille (2010) showed that the (limit) value of the game is $\frac{\rho}{4\rho-1}$ for $\rho \in [\frac{1}{2}, \frac{2}{3}]$, and Aumann and Maschler (1995) showed that it equals $\frac{1}{4}$ for $\rho = 1$ (given a symmetric prior).[1]

It follows from our results that the value of the game in Example 1, albeit still elusive, is nevertheless monotone decreasing in $\rho$ on $[\frac{1}{2}, 1)$. And this is true for any stage game, not just for the one depicted in Figure 1. More generally, in the special case of two states, whether one ergodic operator is better for the maximizer than another depends on a simple comparison of the eigenvalues of the two operators.

The best case for the maximizer in Example 1 is $\rho = \frac{1}{2}$. This finding is general: the maximal elements of our partial order are operators under which the state is i.i.d. across periods. The worst case in Example 1 is a permanent state (i.e., $\rho = 1$). Surprisingly, this is not a general result. For each ergodic operator, there does exist a game in which the maximizer would be worse off if the state was permanent. But for a large class of operators, there also exist games for which the ranking is reversed. Example 2 in Section 4 illustrates the idea. As a result, for any ergodic operator $P$, the permanent state case is either worse than or incomparable to $P$.

One can interpret our analysis as pertaining to the (limit) minmax value in repeated Bayesian non-zero-sum games with ergodic types. Hörner, Takahashi, and Vieille (2015) showed that under independent private values and perfect monitoring, the limit equilibrium-payoff set consists of all incentive-feasible payoffs above the minmax value. Building on their result, we show in the Supplemental Material (Pęski and Toikka (2017)) that the limit equilibrium-payoff set is decreasing (in the sense of set inclusion) in our partial order. Heuristically, this follows because players can be punished more effectively when their types are more predictable. Requiring the comparative static to hold for all stage games yields the converse that identifies our relation as the appropriate one.

Long-run values of Markov decision problems and repeated zero-sum games were studied in independent prior work by Renault and Venel (2012, 2017). We comment on the connection to their results in Section 3.

Anderson and Smith (2013) studied deception using a continuous-time game with a permanent state. The trade-off between short-term gains and future losses from the use of information arises in their model because of discounting.

Finally, our analysis is partly inspired by Blackwell's (1953) comparison of experiments and its extension to zero-sum games by Gossner and Mertens (2001), Pęski (2008), and Shmaya (2006). (See Gossner (2000), and Bergemann and Morris (2016) for non-zero-sum games.) To see the connection, note that the maximizer's action provides a signal about the state to the minimizer. Noise is then added to this signal by the state transition. That is, different operators "garble" the signal to varying degrees. The analogy is imperfect as the signal here is endogenous, and the operator determines not only the in-

---

[1]Bressaud and Quas (2014) extended the range for which the value is known to about $\rho \in [\frac{1}{2}, 0.719]$.

formation, but also the payoffs in the next period. Nevertheless, we develop a formal link between the problems in the ergodic case.

## 2. MODEL

Throughout this paper, $S$ is the finite state space with at least two elements, and $\Delta S$ denotes the set of all probability distributions on $S$. A Markov transition matrix on $S$ is a non-negative $|S| \times |S|$ matrix whose columns sum to 1. As usual, we can view any such matrix as a (linear) operator $P : \Delta S \to \Delta S$. Operator $P$ is ergodic if it is irreducible and aperiodic. We say that $\pi \in \Delta S$ is an invariant distribution of $P$ if $\pi = P\pi$. If $P$ is ergodic, its unique invariant distribution is denoted $\pi_P$.

A zero-sum game on the state space $S$ is a tuple $(A, B, g)$, where $A$ and $B$ are finite action sets and $g : A \times B \times S \to \mathbb{R}$ is a payoff function extended to $\Delta(A \times B \times S)$ by expected utility, as usual. The interpretation is that $g(a, b, s)$ is the payoff of the maximizer if he plays $a$, the minimizer plays $b$, and the state is $s$. In what follows, we vary the game $(A, B, g)$, keeping the state space $S$ fixed. To simplify notation, we typically suppress the action sets and denote a zero-sum game simply by $g$.

A stochastic zero-sum game with incomplete information is a tuple $(\delta, \pi, g, P)$ consisting of a discount factor $\delta \in [0, 1)$, an initial distribution $\pi \in \Delta S$, a zero-sum game $g$, and an operator $P$. The game proceeds in discrete time, indexed by $t \in \mathbb{N}$, as follows. The initial state $s_0$ is drawn according to $\pi$ and it transitions between time periods according to $P$. In each period $t$, the maximizer is informed of the state $s_t$ before choosing an action; the minimizer never observes the state. The players choose actions simultaneously, with both observing the realized action profile $(a_t, b_t)$. The maximizer's payoff is given by the discounted average $(1 - \delta) \sum_t \delta^t g(a_t, b_t, s_t)$. Payoffs are not observable.[2]

Because only the maximizer observes the state, we refer to him as the informed player and to the minimizer as the uninformed player. It turns out that the latter has a myopic best-response, so an alternative interpretation is that there is an informed long-run player facing a sequence of uninformed short-run players who observe their predecessors' actions, but not their payoffs.

Let $v^\delta(\pi; g, P)$ be the value of the stochastic zero-sum game $(\delta, \pi, g, P)$. (Existence for $\delta < 1$ is standard; see, e.g., Mertens, Sorin, and Zamir (2015).) Renault (2006) showed that, as discounting vanishes, the value converges to a well-defined limit that in the ergodic case is independent of the initial distribution.[3]

LEMMA 0: *For every initial distribution $\pi$, every game $g$, and every operator $P$, the limit $v(\pi; g, P) = \lim_{\delta \to 1} v^\delta(\pi; g, P)$ exists. Furthermore, if $P$ is ergodic, then the limit is independent of $\pi$, and we denote it by $v(g, P)$.*

In what follows, we focus on the limit value $v(\pi; g, P)$, henceforth referred to simply as the value. As is clear from Example 1 in the Introduction, the value depends on the operator $P$. Our goal is to describe this relationship. Below, we introduce the notation and recall some standard definitions needed for this.

We work with the space $\Delta^2 S = \Delta(\Delta S)$ of Borel probability measures on $\Delta S$, interpreted as the space of distributions of beliefs about the state, endowed with the weak topology.

---

[2]This assumption has bite only for the minimizer as the maximizer observes the triplet $(a_t, b_t, s_t)$.

[3]Renault (2006) considered undiscounted games, but it is well known that $\lim_{\delta \to 1} v^\delta(\pi; g, P)$ equals the limit of values of undiscounted $T$-period games as $T \to \infty$ (see, e.g., Sorin (2002), Lemma 3.1).

Given any $\mu \in \Delta^2 S$ and operator $P$, let $\mathcal{E}\mu = \int p \, d\mu(p) \in \Delta S$ denote the expected value of $\mu$, and let $P\mu \in \Delta^2 S$ denote the distribution for which $\int f \, dP\mu = \int f \circ P \, d\mu$ for each continuous function $f : \Delta S \to \mathbb{R}$.

A *mean-preserving spread* (or a *splitting*) is a measurable mapping $m : \Delta S \to \Delta^2 S$ such that $\mathcal{E}m(p) = p$ for each $p$. Let $\mathcal{M}$ be the space of mean-preserving spreads. For any $\mu, \mu' \in \Delta^2 S$, we say that $\mu'$ is *(Blackwell) more informative* than $\mu$, or $\mu \leq^B \mu'$, if there exists $m \in \mathcal{M}$ such that for each continuous function $f : \Delta S \to \mathbb{R}$,

$$\int f(p) \, d\mu'(p) = \int \left( \int f(q) \, dm(q|p) \right) d\mu(p).$$

We also write $\mu' = \mu * m$. That is, if $\mu'$ and $\mu$ are distributions of the minimizer's posterior beliefs, then his information about the state under $\mu'$ is better than under $\mu$ in the sense of Blackwell (1953). Finally, we recall the following familiar characterization: $\mu \leq^B \mu'$ if and only if for each concave function $f : \Delta S \to \mathbb{R}$, $\int f \, d\mu' \leq \int f \, d\mu$.

## 3. CHARACTERIZATION OF VALUE

We start by characterizing the value. To sketch the idea, note that the relevant state variable in the game is the minimizer's belief about the state. This well-known observation generalizes the fact that, in a repeated zero-sum game of complete information, it is impossible to provide intertemporal incentives and the value is just the value of the stage game. Here, since there is a payoff-relevant state unobservable to the minimizer, the analogous result is that play is Markov in the minimizer's belief.

We refer to the minimizer's belief at the beginning of the period as the prior, reserving the term posterior for the belief the minimizer forms at the end of the period after observing the maximizer's action, but before the state transitions.

Any strategy for the maximizer that is Markov with respect to the minimizer's prior induces via Bayes's rule a mean-preserving spread $m : \Delta S \to \Delta^2 S$ that associates to each prior $p$ a distribution $m(p)$ over posteriors. The mapping $m$ encapsulates how much information the maximizer's actions reveal about the state as a function of the prior. The key to our characterization is to note that, rather than having the maximizer choose a strategy, we can think of him first choosing the mapping $m$ determining the information revelation policy, and then choosing a stage-game strategy subject to it not revealing more information than $m$.[4]

For any information revelation policy $m$, there are many strategies that reveal no more information than $m$. For each prior $p$, we denote by $\hat{g}(m(p))$ the maximizer's optimal stage-game payoff among such strategies. Formally, given a game $g$ and a distribution $\nu \in \Delta^2 S$ over posteriors, let

$$\hat{g}(\nu) = \min_{\beta \in \Delta B} \int \max_{a \in A} g(a, \beta, q) \, d\nu(q). \tag{3.1}$$

As the minimum over terms linear in $\nu$, this defines a concave function $\hat{g} : \Delta^2 S \to \mathbb{R}$.

---

[4] The idea of decomposing a strategy into revelation of information and an action choice measurable with respect to this information is central to Aumann and Maschler's (1995) celebrated cav $u$ theorem. As we explain below, our characterization is a generalization of that result. Similar ideas have been used elsewhere in the literature on dynamic games with incomplete information; see, for example, De Meyer (2010) or Gensbittel (2015). (We are grateful to a referee for these examples.)

We can interpret $\hat{g}(m(p))$ as the value of the following auxiliary game: The common prior on the state is $p$. First, the minimizer chooses a mixed strategy $\beta \in \Delta B$. Then an exogenous device sends a signal about the state such that the induced distribution of posteriors is $m(p)$. Finally, the maximizer chooses an action $a \in A$ conditional on the realized posterior $q$. (The action being measurable with respect to the posterior ensures that it does not reveal more information than the signal.)

Together, the operator $P$ and an information revelation policy $m$ induce a Markov chain on the minimizer's prior beliefs: given a distribution of priors $\mu \in \Delta^2 S$, the distribution of posteriors is $\mu * m$, and the distribution of the next period's priors is $P(\mu * m)$. The maximizer's long-run payoff in the stochastic game is the expectation of $\hat{g}(m(p))$ under an invariant distribution of priors, that is, under some $\mu$ with $P(\mu * m) = \mu$. The (limit) value is obtained by letting the maximizer choose $m$ and $\mu$ subject to $\mu$ being invariant and containing at least as much information about the state as the minimizer would have if the maximizer did not reveal any information.

The relevant measure of the minimizer's information in the absence of information revelation is given by the long-run distribution of his priors, defined as

$$\psi_{\pi,P} = \lim_{\delta \to 1} (1 - \delta) \sum_t \delta^t P^t e_\pi \in \Delta^2 S,$$

where $e_\pi \in \Delta^2 S$ is a distribution concentrated on $\pi$ (i.e., a Dirac measure at $\pi$). Note that $\psi_{\pi,P}$ is the long-run average distribution of beliefs, not the long-run average belief; the latter (viewed as a Dirac measure) can be less informative than the former.[5]

We then have the following novel characterization of value.

THEOREM 1: *For every initial distribution $\pi$, every game $g$, and every operator $P$,*

$$v(\pi; g, P) = \max_{(\mu, m) \in \Delta^2 S \times \mathcal{M} : \psi_{\pi,P} \leq^B \mu \text{ and } P(\mu * m) \leq^B \mu} \int \hat{g}(m(p)) \, d\mu(p). \tag{3.2}$$

Consistent with the heuristic derivation preceding the theorem, the proof in Section 6 shows that the maximum is in fact achieved by some $(\mu, m)$ such that $P(\mu * m) = \mu$. However, allowing for inequality in the sense of Blackwell ordering in the second constraint is convenient for the purposes of establishing Theorem 2 below.

It is instructive to consider the following special cases. First, let $\pi$ be an invariant distribution of $P$. If the maximizer never reveals any information, then the minimizer's belief stays constant at $\pi$. Thus, in this case, we have $\psi_{\pi,P} = e_\pi$ and the constraint $e_\pi \leq^B \mu$ is equivalent to the familiar martingale condition $\mathcal{E}\mu = \pi$.

Second, suppose the state is permanent. Then $P$ is the identity operator $I$ and the second constraint becomes $\mu * m \leq^B \mu$. This can hold only if $m$ maps each $p$ to a point mass at $p$ and hence reveals no information. But the constraint places no restrictions on $\mu$. As any $\pi$ is invariant for $I$, we thus recover Aumann and Maschler's (1995) formula

$$v(\pi; g, I) = \max_{\mu \in \Delta^2 S : \mathcal{E}\mu = \pi} \int \hat{g}(p) \, d\mu(p) = \operatorname{cav} \hat{g}(\pi),$$

---

[5]For example, if $\pi = (1, 0)$ and $P = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ so that the state alternates, then the minimizer knows the state in every period. In this case, $\psi_{\pi,P} \in \Delta^2 S$ assigns equal weight to beliefs $(1, 0)$ and $(0, 1)$, whereas the average belief is $(\frac{1}{2}, \frac{1}{2}) \in \Delta S$. Clearly, $\psi_{\pi,P}$ is strictly more informative than $e_{(\frac{1}{2}, \frac{1}{2})}$.

where, with abuse of notation, $\hat{g}(p) = \hat{g}(e_p)$ is the value of our auxiliary game given prior $p$ when no information is revealed (also known as the non-revealing game), and $\mathrm{cav}\,\hat{g}$ is the smallest concave function $\Delta S \to \mathbb{R}$ greater than $\hat{g}(p)$ at every $p$.

Third, let $P$ be ergodic. Then absent information revelation, the minimizer's beliefs converge to the invariant distribution $\pi_P$. Thus, $\psi_{\pi,P} = e_{\pi_P}$ and the first constraint becomes $\mathcal{E}\mu = \pi_P$. This shows that the value is independent of $\pi$. Moreover, $\mathcal{E}\mu = \pi_P$ is implied by the second constraint.[6] Therefore, we have

$$v(g, P) = \max_{(\mu,m)\in\Delta^2 S\times\mathcal{M}:P(\mu*m)\leq^B\mu} \int \hat{g}\big(m(p)\big)\,d\mu(p). \qquad (3.3)$$

It is in this form that Theorem 1 is used in the next section.

Finally, in the i.i.d. case, the ergodic formula (3.3) simplifies further. We then have $P(\mu*m) = e_{\pi_P}$ for every $\mu$ and $m$, so the constraint becomes $e_{\pi_P} \leq^B \mu$, or $\mathcal{E}\mu = \pi_P$. As it no longer involves $m$, it is clearly optimal to have $m$ fully reveal the state by definition of $\hat{g}$ in (3.1). But then, since the probability of each state $s$ is just $\pi_P(s)$ under any $\mu$ such that $\mathcal{E}\mu = \pi_P$, we have

$$v(g, P) = \min_{\beta\in\Delta B} \int \max_{a\in A} g(a, \beta, s)\,d\pi_P(s).$$

Of course, the expression on the right is just the value of the stage game $g$.

Although Theorem 1 plays a key role in our analysis, the characterization is in general not tractable enough for the purposes of computing the value. For example, we have not been able to calculate the value of the game in Example 1. Given $(\mu, m)$, the integral in (3.2) is relatively easy to compute. However, the constraint $P(\mu*m) \leq^B \mu$ has fixed point flavor and is not easy to characterize.

Independent prior work by Renault and Venel (2017) characterized the long-run value for a class of Markov decision problems and repeated games, which includes our setting as a special case. (The results were circulated several years earlier; see Renault and Venel (2012).) Because of its generality, Renault and Venel's formula is much more complicated than that in Theorem 1. Importantly, as they allow actions to affect state transitions, it is impossible to separate the choice of actions from the choice of stationary belief distributions in a meaningful way as is done in (3.2) via the function $\hat{g}$; this separation is instrumental for the comparison of operators. Our proof of Theorem 1 and Renault and Venel's proof of their result are unrelated. In fact, because the results are stated in very different terms, the connection between them is not apparent at first sight. We elaborate on this in the Supplemental Material, where we sketch a direct proof of their equivalence for ergodic operators.

## 4. COMPARISON OF OPERATORS

We are now ready to describe how the value of the game to the informed player depends on the operator governing the evolution of the state. We focus on ergodic operators as the value is then independent of the initial distribution, allowing for cleaner comparisons. (See Section 5 for comparison of all operators.)

Formally, we seek to characterize the following relation.

---

[6]If $\mu = P(\mu*m)*m'$ for some $m' \in \mathcal{M}$, then $\mathcal{E}\mu = \mathcal{E}[P(\mu*m)*m'] = P\mathcal{E}\mu$. Thus, $\mathcal{E}\mu = \pi_P$.

DEFINITION 1: For any ergodic operators $P$ and $Q$, we say that $Q$ *is better for the informed player than* $P$, or $P \preceq Q$, if for every game $g$, $v(g, P) \leq v(g, Q)$.

We are interested in representing the relation $\preceq$ in terms of conditions that depend only on the operators $P$ and $Q$, making no reference to strategic behavior or games. The following result provides two solutions to this problem.

THEOREM 2: *For any ergodic operators $P$ and $Q$, the following are equivalent*:
(a) $P \preceq Q$.
(b) *For every* $(\mu, m) \in \Delta^2 S \times \mathcal{M}$ *such that* $P(\mu * m) \leq^B \mu$, *we have* $Q(\mu * m) \leq^B \mu$.
(c) *For every* $\nu \in \Delta^2 S$ *such that* $P\nu \leq^B \nu$, *we have* $Q\nu \leq^B P\nu$.

Conditions (b) and (c) give alternative characterizations of $\preceq$. Condition (b) shows the connection to Theorem 1: it simply says that every information structure $(\mu, m)$ that is feasible in the maximization problem in (3.3) under $P$ is also feasible under $Q$. It is then immediate that (b) implies (a). There is a simple intuition for this.

Recall that the maximum is achieved by some $(\mu, m)$ with $P(\mu * m) = \mu$. Suppose that operator $P$ is replaced by $Q$ such that $Q(\mu * m) \leq^B \mu$. Consider an augmented game where at the beginning of each period, the maximizer can reveal information about the state using a costless signalling device. If the maximizer plays the Markov strategy corresponding to $m$, with $\mu$ the distribution of priors, then the distribution of the next period's priors is $Q(\mu * m)$. Thus, if the maximizer uses the device at the beginning of each period to reveal information from $Q(\mu * m)$ to $\mu$, then his payoff in the augmented game under $Q$ is equal to the value of the original game under $P$. Because the device reveals information without any benefit, not using it should only increase the maximizer's payoff.

For the converse, we assume in negation of (b) that the feasible sets are not nested for operators $P$ and $Q$, and then construct a game $g$ such that $v(g, P) > v(g, Q)$ in negation of (a). We turn to the proof in Section 7 where we also prove the equivalence of (b) and (c), the latter a simple restatement of the former in terms of posteriors.

Condition (c) suggests interpreting $\preceq$ as a notion of persistence. To see this, let $\nu$ be the distribution of posteriors. Then $Q\nu \leq^B P\nu$ is exactly the condition for the distribution of the next period's priors to be more informative under $P$ than under $Q$. In this sense the state—or, more precisely, the information about it—is more persistent under $P$ than under $Q$. Condition (c) requires this to hold only for any $\nu$ that is stabilizable under $P$ in the sense that $P\nu \leq^B \nu$. Note that such $\nu$ are exactly the distributions that arise as invariant distributions of posteriors under some information revelation policy $m$, since invariance requires $P\nu * m = \nu$.

We now present several corollaries to Theorem 2 that describe the relation $\preceq$ further. The first one shows that $\preceq$ is a convex partial order under which operators can be ranked only if their invariant distributions coincide.

COROLLARY 1: *Let $P$, $Q$, and $Q'$ be ergodic*.
(a) *If $P \preceq Q$ or $Q \preceq P$, then $\pi_P = \pi_Q$*.
(b) *If $P \preceq Q$ and $Q \preceq P$, then $P = Q$*.
(c) *If $P \preceq Q$ and $P \preceq Q'$, then for every $\lambda \in (0, 1)$, $P \preceq \lambda Q + (1 - \lambda) Q'$*.

Part (a) can be deduced from Theorem 2, but the intuition is best seen from a direct argument. Namely, for any $P$ and $Q$ such that $\pi_P \neq \pi_Q$, we can find trivial games $g'$ and $g''$ such that $v(g', P) > v(g', Q)$ and $v(g'', P) < v(g'', Q)$. Say, pick any states $s'$

and $s''$ such that $\pi_P(s') > \pi_Q(s')$ and $\pi_P(s'') < \pi_Q(s'')$. Let $g'(s, a, b) = \mathbf{1}\{s = s'\}$ and $g''(s, a, b) = \mathbf{1}\{s = s''\}$. This gives $v(g', P) = \pi_P(s') > \pi_Q(s') = v(g', Q)$ and $v(g'', P) = \pi_P(s'') < \pi_Q(s'') = v(g'', Q)$ as desired.

Part (b) shows that $\preceq$ is antisymmetric. Since $\preceq$ is obviously reflexive and transitive, it thus partially orders the space of ergodic operators. Part (c) shows that $\preceq$ is convex in the sense that, for any operator $P$, the set of operators that are better for the informed player than $P$ (i.e., the upper-contour set of $P$) is convex.

We then give a partial characterization of $\preceq$, which refines the interpretation of $\preceq$ as a form of persistence suggested by Theorem 2. To this end, let $\Lambda$ be the set of non-negative infinite sequences $\lambda_1, \lambda_2, \ldots, \lambda_\infty$ such that $\sum \lambda_n = 1$. (Note that the sequence is indexed by $n \in \mathbb{N} \cup \{\infty\}$.) Given an ergodic operator $P$ and $\lambda \in \Lambda$, define $P(\lambda) = \lambda_1 P + \lambda_2 P^2 + \cdots + \lambda_\infty P^\infty$. Here, $P^\infty$ is the operator that maps any belief to the unique invariant distribution of the operator $P$.

COROLLARY 2: *Let $P$ and $Q$ be ergodic.*
(a) *If $Q = P(\lambda)$ for some $\lambda \in \Lambda$, then $P \preceq Q$.*
(b) *If $P \preceq Q$, then for every $p \in \Delta S$, there exists $\lambda_p \in \Lambda$ such that $Qp = P(\lambda_p)p$.*
(c) *If $P \preceq Q$ and $P$ has only real eigenvalues, then $Q = P(\lambda)$ for some $\lambda \in \Lambda$.*

The first part shows that $Q$ being a convex combination of the powers of $P$ is a sufficient condition for $Q$ to be better for the informed player than $P$. Heuristically, the operator $P(\lambda)$ is less persistent than $P$ in that it amounts, roughly, to applying $P$ multiple times between each period. This means that more noise is added to the uninformed player's information between periods under $P(\lambda)$ than under $P$. Thus, revealing information under $P(\lambda)$ is less costly than under $P$, explaining part (a) of Corollary 2. (This intuition is only partial as the operator does not only determine the uninformed player's information but also the next period's payoffs. But since $\pi_P = \pi_{P(\lambda)}$, the payoffs are "the same on average.")

Part (b) of Corollary 2 shows that $Q$ being a convex combination of the powers of $P$ is also necessary for $P \preceq Q$ in the sense that for every $p \in \Delta S$, there exist some weights $\lambda_p$ such that $Qp = P(\lambda_p)p$. These weights generally depend on the distribution $p$. However, if $P$ has only real eigenvalues (say, because $P$ is symmetric, or because there are only two states), then $\lambda$ can be chosen independently of $p$ by part (c) of Corollary 2. Together, parts (a) and (c) imply that for such $P$, having $Q = P(\lambda)$ for some $\lambda$ fully characterizes $P \preceq Q$.[7]

Corollary 2 allows us to establish a connection to Blackwell's garblings.[8] To see this, reinterpret each operator $P$ as describing the conditional distribution of a signal observed by a decision maker given the value of an underlying unobservable state of the world. Recall that operator $P$ is better for the decision maker than operator $Q$ according to Blackwell's ordering if and only if there exists an operator $M$, called a *garbling*, such that $Q = MP$. Now observe that $P(\lambda) = \sum_n \lambda_n P^n = (\sum_n \lambda_n P^{n-1})P$. Thus, part (a) of Corollary 2 can be rephrased as saying that a sufficient condition for $P \preceq Q$ is that $Q = MP$ for a garbling of the form $M = \sum_n \lambda_n P^{n-1}$, $\lambda \in \Lambda$. Similarly, part (b) shows that the point-wise existence of such a garbling is necessary in general, and parts (a) and (c) show that the existence of a garbling of this form characterizes the upper-contour set of $P$ under $\preceq$ if $P$

---

[7]We show in the Supplemental Material that this characterization does not extend to operators with complex eigenvalues by constructing $P$ and $Q$ such that $P \preceq Q$, but $Q \neq P(\lambda)$ for all $\lambda \in \Lambda$.
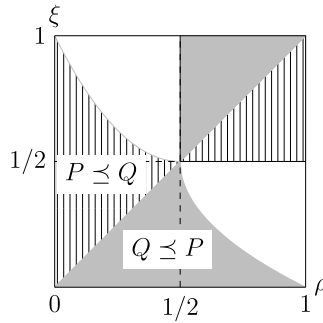
[8]We thank Lones Smith for pointing out the connection.

FIGURE 2.—Comparison of operators $P = M(\rho)$ and $Q = M(\xi)$.

has only real eigenvalues. Heuristically, garbling the minimizer's information is good for the maximizer. But in contrast to Blackwell's decision maker, not all garblings work here, because the transitions described by the operator are payoff-relevant in the game.

Corollary 2 yields an eigenvalue characterization when there are only two states:

COROLLARY 3: *If* $|S| = 2$, *then every ergodic operator P has exactly two eigenvalues*, 1 *and* $\phi_P \in (-1, 1)$. *For any two such operators P and Q, we have* $P \preceq Q$ *if and only if* (a) $\pi_P = \pi_Q$, *and* (b) *if* $\phi_P \geq 0$, *then* $\phi_Q \in [0, \phi_P]$, *and if* $\phi_P \leq 0$, *then* $\phi_Q \in [\phi_P, \phi_P^2]$.

PROOF: Every stochastic operator has the eigenvalue 1. Ergodicity implies the existence of the second eigenvalue $\phi_P \in (-1, 1)$. Since $P$ has only real eigenvalues, parts (a) and (c) of Corollary 2 show that $P \preceq Q$ is equivalent to having $Q = P(\lambda)$ for some $\lambda \in \Lambda$, which in turn is equivalent to $\pi_P = \pi_Q$ and $\phi_Q = \sum_{k \geq 1} \lambda_k \phi_P^k$, implying the result.  *Q.E.D.*

Corollary 3 covers our introductory example, thus providing an answer to the question about monotonicity left open by Hörner et al. (2010):

EXAMPLE 1—Continued: The state remains the same with probability $\rho$ and changes with probability $1 - \rho$, so the operator is of the form

$$M(\rho) = \begin{bmatrix} \rho & 1 - \rho \\ 1 - \rho & \rho \end{bmatrix}, \quad 0 < \rho < 1.$$

The invariant distribution assigns equal probability to both states for any $\rho$, and the smallest eigenvalue is equal to $\phi = 2\rho - 1$. Hence, by Corollary 3, the value to the maximizer is monotone decreasing in $\rho$ on $[\frac{1}{2}, 1)$. We also see that it is increasing in $\rho$ on $(0, \frac{1}{2}]$. Finally, it is straightforward to verify Corollary 2 for this case. Say, let $Q = M(\xi)$ and $P = M(\rho)$ for $\frac{1}{2} \leq \xi < \rho < 1$. Putting $\lambda_1 = \frac{\xi - \frac{1}{2}}{\rho - \frac{1}{2}}$ and $\lambda_\infty = 1 - \lambda_1$, we then have $Q = \lambda_1 P + \lambda_\infty P^\infty$.

The operators $M(\rho)$ considered in Example 1 are indexed by the probability $\rho$ of the state not changing. The one-dimensional parameterization allows us to visualize the restriction of the partial order $\preceq$ to this class of symmetric operators—see Figure 2.

In Example 1, the best case for the maximizer is to have the state be i.i.d. across periods. This is a general result. In order to state it, for any $\pi \in \Delta S$, write $D_\pi$ for the operator that maps each $p$ to $\pi$ and thus generates a sequence of i.i.d. draws from $\pi$. Then $D_{\pi_P} = P^\infty$ for each $P$, and parts (a) of Corollaries 1 and 2 show that the i.i.d. operators $D_\pi$, $\pi \in \Delta S$, are the maximal elements of the partial order $\preceq$.

|   | L | R |
|---|---|---|
| U | $-2, 0, 3$ | $0, -2, 3$ |
| D | $-1, 1, 0$ | $1, -1, 0$ |

FIGURE 3.—Stage-game payoffs for Example 2.

COROLLARY 4: *Let $\pi \in \Delta S$. Then $P \preceq D_\pi$ for every ergodic operator $P$ with $\pi_P = \pi$.*

Corollary 4 is not surprising: with i.i.d. states, the maximizer has all the benefits and no cost of revealing information. In the other direction, one might conjecture that the value of any game under an ergodic operator $P$ is larger than the value of the same game with a permanent state. Surprisingly, this is not true for every $P$.

COROLLARY 5: *Let $P$ be ergodic.*
(a) $v(g, P) > v(\pi_P; g, I)$ *for some game $g$.*
(b) $v(g, P) \geq v(\pi_P; g, I)$ *for every game $g$ if and only if there exists $\lambda \in [0, 1)$ such that* $P = \lambda I + (1 - \lambda) D_{\pi_P}$.

Corollary 5 shows that no ergodic operator $P$ is worse than the permanent state, but not every ergodic operator $P$ is comparable to $I$ given initial distribution $\pi_P$.

Part (b) implies that if $P$ is not a convex combination of the identity operator $I$ and the i.i.d. operator $D_{\pi_P}$, then there exists a game $g'$ such that $v(\pi_P; g', P) < v(\pi_P; g', I)$. To see why such a game can be found, recall from Theorem 1 that with a permanent state, the value is computed as if the maximizer induced some distribution of priors for the minimizer through an initial revelation, and then stayed at the realized prior by playing a non-revealing strategy for the rest of the game. But when the state evolves over time, so do the minimizer's beliefs. It is only under the condition in part (b) that it is possible to replicate any distribution of priors from the permanent case. This issue arises already when there are two states, but it is best illustrated with the following three-state example. (See also the discussion following Theorem 2′ below.)

EXAMPLE 2: There are three states, $S = \{s_1, s_2, s_3\}$. The maximizer chooses $U$ or $D$, the minimizer chooses $L$ or $R$. Stage-game payoffs are given in Figure 3 where each cell contains the payoff from a given action profile in all three states. The initial distribution is $\pi = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. It is easy to see that it is optimal for the minimizer to play $L$ if $p_1 \geq p_2$ and to play $R$ if $p_2 \geq p_1$, where $p = (p_1, p_2, p_3)$ is his current prior.

If the state is permanent, it is optimal for the maximizer to play $U$ if $s = s_3$ and to play $D$ otherwise. The value of the game is $\frac{2}{3} \cdot 0 + \frac{1}{3} \cdot 3 = 1$. Note that the minimizer only ever learns whether the state is $s_3$ or not. That is, after the first period his belief is either $(\frac{1}{2}, \frac{1}{2}, 0)$ or $(0, 0, 1)$, and it remains constant for the rest of the game.

Suppose then that after the initial draw, transitions are given by the operator

$$
P = \begin{bmatrix}
1 - 2\varepsilon & \varepsilon & \varepsilon \\
\varepsilon & \dfrac{1}{2} - \dfrac{\varepsilon}{2} & \dfrac{1}{2} - \dfrac{\varepsilon}{2} \\
\varepsilon & \dfrac{1}{2} - \dfrac{\varepsilon}{2} & \dfrac{1}{2} - \dfrac{\varepsilon}{2}
\end{bmatrix},
$$

where $\varepsilon \in (0, \frac{1}{3})$. Note that $\pi_P = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. If the maximizer now plays the optimal strategy from the permanent state case, then upon observing $U$, the minimizer's next period prior is $(\varepsilon, \frac{1}{2} - \frac{\varepsilon}{2}, \frac{1}{2} - \frac{\varepsilon}{2})$. That is, the minimizer correctly anticipates $s_2$ to be more likely than $s_1$, and he plays $R$. Similarly, playing $D$ would be followed by the minimizer playing $L$. Thus the expected payoff from this strategy is now strictly less than 1, which was the value under a permanent state. On the other hand, if the maximizer does not always play $U$ when $s = s_3$, then his payoff must again be less than 1, since the minimizer's strategy ensures an expected average payoff of at most 0 from the other two states.

## 5. EXTENSIONS

1. *Comparison of all operators.* Since $S$ is finite, almost all operators are ergodic. Moreover, the proof of Corollary 5(a) can be easily adapted to show that for any ergodic operator $P$ and any non-ergodic operator $Q$ for which $\pi_P$ is invariant, there exists a game $g$ such that $v(g, P) > v(\pi_P; g, Q)$. Therefore, no ergodic operator is worse than any non-ergodic operator. Nevertheless, as Corollary 5(b) shows, unexpected behavior emerges at the boundary of the set of operators. To explain it—and to prove Corollary 5(b)—we extend Theorem 2 to all operators. As the proof is the same, it is this general version we prove in Section 7.

The value now depends on the initial distribution, motivating the next definition.

DEFINITION 2: For any initial distribution $\pi$ and operators $P$ and $Q$, we say that $Q$ is *better for the informed player than $P$ given $\pi$*, or $P \preceq_\pi Q$, if for every game $g$, $v(\pi; g, P) \leq v(\pi; g, Q)$.

Note that if $P$ and $Q$ are ergodic, then $P \preceq Q$ if and only if $P \preceq_\pi Q$ for every $\pi$.

THEOREM 2′: *For any $\pi \in \Delta S$ and operators $P$ and $Q$, the following are equivalent*:
  (a) $P \preceq_\pi Q$.
  (b) $\psi_{\pi,Q} \leq^B \psi_{\pi,P}$ and for every pair $(\mu, m) \in \Delta^2 S \times \mathcal{M}$ such that $\psi_{\pi,P} \leq^B \mu$ and $P(\mu * m) \leq^B \mu$, we have $Q(\mu * m) \leq^B \mu$.
  (c) $\psi_{\pi,Q} \leq^B \psi_{\pi,P}$ and for every $\nu \in \Delta^2 S$ such that $\psi_{\pi,P} \leq^B \nu$ and $P\nu \leq^B \nu$, we have $Q\nu \leq^B P\nu$.

The new common part in conditions (b) and (c), $\psi_{\pi,Q} \leq^B \psi_{\pi,P}$, requires the distribution of the minimizer's long-run beliefs to be less informative under $Q$ than under $P$ absent any information revelation. This is clearly necessary for $P \preceq_\pi Q$; otherwise we can find a game $g$ where only the minimizer has an action and for which $v(\pi; g, P) > v(\pi; g, Q)$ by the virtue of his better information under $Q$.

There are two important cases where the condition $\psi_{\pi,Q} \leq^B \psi_{\pi,P}$ can be omitted. First, if $\pi$ is an invariant distribution of both $P$ and $Q$, then the condition is satisfied a priori as $\psi_{\pi,P} = \psi_{\pi,Q} = e_\pi$. Second, if $P$ and $Q$ are ergodic, then $\psi_{\pi,P} \leq^B \psi_{\pi,Q}$ reduces to $\pi_P = \pi_Q$, which for ergodic operators is implied by the second part of conditions (b) or (c).

The only other change to conditions (b) and (c) compared to Theorem 2 is the need to explicitly require $\mu$ and $\nu$ to respect the minimizer's information absent revelations. These are best seen as feasibility constraints. For example, $\psi_{\pi,P} \leq^B \mu$ ensures that $\mu$ can be obtained with some information revelation policy under operator $P$ starting from $\pi$. Otherwise, the interpretation of conditions (b) and (c) is the same as in Theorem 2. Note that in the ergodic case, these extra conditions are implied by $P(\mu * m) \leq^B \mu$ and $P\nu \leq^B \nu$, respectively. Thus, in this case, the statement of Theorem 2′ reduces to that of Theorem 2.

Armed with Theorem $2'$, we can understand the finding in Corollary 5(b) as follows. Observe that (c) implies $Q\nu \leq^B \nu$, so we can view the second part of (c) as having two requirements: (c1) any (feasible) $\nu$ stabilizable under $P$ is stabilizable under $Q$ (i.e., $\psi_{\pi,P} \leq^B \nu$ and $P\nu \leq^B \nu$ imply $Q\nu \leq^B \nu$), and (c2) for such $\nu$, $P\nu$ is more informative than $Q\nu$. It is because of (c1) that there are cases where operator $P$ is in some intuitive sense more persistent than operator $Q$, but $P \preceq_\pi Q$ does not hold as some distributions stabilizable under $P$ are not stabilizable under $Q$.

To further illustrate this and the use of Theorem $2'$ for comparison of non-ergodic operators, let $|S| = 2$ and $\pi = (\frac{1}{2}, \frac{1}{2})$. Let $A$ be the operator that deterministically alternates between the states. Then condition (c) implies that $A \preceq_\pi I$, but the converse does not hold. Indeed, $\pi$ is invariant for both $I$ and $A$, and thus we have $\psi_{\pi,I} = \psi_{\pi,A} = e_\pi$. However, whereas any $\nu$ with $\mathcal{E}\nu = \pi$ is stabilizable under $I$, only a symmetric $\nu$ is stabilizable under $A$. (For an example where $I$ is strictly better, consider an asymmetric game where it is optimal to reveal one of the states when the state is permanent. Replicating this strategy under $A$ inadvertently reveals the wrong state in every other period.)

We then briefly comment on the extensions developed in the Supplemental Material.

2. *Imperfect monitoring.* The fact that the maximizer observes the minimizer's action plays no role in the analysis. As for the minimizer, the results readily extend to imperfect monitoring games where he only observes a noisy public signal whose distribution $F_a$ depends on the maximizer's action $a$. If we replace "every game $g$" with "every imperfect monitoring game $(g, F)$" in Definition 2, Theorem $2'$ holds verbatim: one direction follows as perfect monitoring is a special case, the other because our characterization of value still holds if $\hat{g}$ is suitably redefined.

3. *Public signals.* It is also possible to accommodate public information about the current state, modeled as a public signal observed at the beginning of each period. Such a signal puts a lower bound on how much information can be revealed. Formally, this appears as an exogenous mean-preserving spread that acts on the minimizer's priors at the start of each period, before the application of the information revelation policy $m$. Let $v(g, P, F)$ be the value given ergodic operator $P$ and public signal distribution $F$. Rank operator-signal pairs by writing $(P, F) \preceq (Q, G)$ if $v(g, P, F) \leq v(g, Q, G)$ for every game $g$. With these definitions, results similar to the ergodic version of Theorem 1 and Theorem 2 continue to hold. In particular, $(P, F) \preceq (Q, G)$ if and only if for every $\nu \in \Delta^2 S$ such that $P\nu * n^F \leq^B \nu$, we have $Q\nu * n^G \leq^B P\nu * n^F$. Here, $n^F$ and $n^G$ are the mean-preserving spreads induced by $F$ and $G$. Note that taking $P = Q$ gives a comparison of public signals.

4. *Discounting.* Comparison of operators for a fixed discount factor is difficult in general, but we note here a simple sufficient condition. Fix any operator $P$. If $Q = \lambda P + (1 - \lambda)D_\pi$ for some $\lambda \in [0, 1]$ and some invariant distribution $\pi$ of $P$, then for every game $g$, we have $v^\delta(\pi; g, P) \leq v^\delta(\pi; g, Q)$ for all $\delta < 1$. In particular, this implies that in Example 1, the value is decreasing in $\rho$ on $[\frac{1}{2}, 1]$ for any fixed $\delta$.

## 6. PROOF OF THEOREM 1

Fix an initial distribution $\pi$, game $g$, and operator $P$. The proof has four steps.

### 6.1. *Recursive Formula for the Value*

Given a prior $p \in \Delta S$ and a stage-game strategy $\alpha \in (\Delta A)^S$, denote by $\overline{\alpha}(a) = \sum_s \alpha(a|s)p(s)$ the expected probability of playing action $a$. For any $a$ with $\overline{\alpha}(a) > 0$, de-

fine the posterior belief $q^{\alpha,p}(a) \in \Delta S$ by

$$q^{\alpha,p}(s|a) = \frac{\alpha(a|s)p(s)}{\overline{\alpha}(a)}.$$

Let $\nu^{\alpha,p} \in \Delta^2 S$ denote the induced distribution over posterior beliefs. That is,

$$\nu^{\alpha,p}(q) = \sum_a \mathbf{1}\{q^{\alpha,p}(a) = q\}\overline{\alpha}(a). \tag{6.1}$$

Notice that $\mathcal{E}\nu^{\alpha,p} = p$, that is, the average posterior is equal to the prior.

The following recursion is well known (cf. Proposition 5.1 in Renault (2006); see Mertens, Sorin, and Zamir (2015), for the discounted case).

LEMMA 1: *For every $\delta < 1$ and every $p \in \Delta S$,*

$$v^\delta(p; g, P) = \min_{\beta \in \Delta B} \max_{\alpha \in (\Delta A)^S} \sum_a \overline{\alpha}(a)\big((1-\delta)g\big(a, \beta, q^{\alpha,p}(a)\big) + \delta v^\delta\big(Pq^{\alpha,p}(a); g, P\big)\big). \tag{6.2}$$

*Moreover, $v^\delta(p; g, P)$ is concave in $p$.*

We observe first that we can rewrite this recursion with the maximizer choosing a distribution $\nu$ over posteriors and an action $a$ for each posterior $q$.

LEMMA 2: *For every $\delta < 1$ and every $p \in \Delta S$,*

$$\begin{aligned}
v^\delta(p; g, P) &= \min_{\beta \in \Delta B} \max_{\nu \in \Delta^2 S : \mathcal{E}\nu = p} \int \Big((1-\delta) \max_{a \in A} g(a, \beta, q) + \delta v^\delta(Pq; g, P)\Big) d\nu(q) \\
&= \max_{\nu \in \Delta^2 S : \mathcal{E}\nu = p} (1-\delta)\hat{g}(\nu) + \delta \int v^\delta(Pq; g, P) \, d\nu(q).
\end{aligned} \tag{6.3}$$

Note that the second line in (6.3) shows that $v^\delta(p; g, P)$ is the value of a Markov decision problem where the state is $p$, the action is $\nu$, and the current payoff is $\hat{g}(\nu)$.

PROOF: We show first that $v^\delta(p; g, P)$ is not smaller than the right-hand side of the first line in (6.3). For each $\beta \in \Delta B$, let $\nu_\beta$ be a solution to the maximization problem on the first line of (6.3), and let $\gamma_\beta : \Delta S \to A$ be a measurable selection $\gamma_\beta(q) \in \arg\max_{a \in A} g(a, \beta, q)$.[9] Define $\alpha_\beta \in (\Delta A)^S$ by setting, for any $s$ with $p(s) > 0$,

$$\alpha_\beta(a|s) = \frac{1}{p(s)} \int q(s)\mathbf{1}\{\gamma_\beta(q) = a\} \, d\nu_\beta(q). \tag{6.4}$$

(Actions for states $s$ such that $p(s) = 0$ can be taken to be arbitrary.) We then have

$$\sum_a \overline{\alpha}_\beta(a)g\big(a, \beta, q^{\alpha_\beta,p}(a)\big) = \int \max_{a \in A} g(a, \beta, q) \, d\nu_\beta(q).$$

---

[9]The existence of $\nu_\beta$ follows from the continuity of the integral with respect to the measure $\nu$ and the compactness of the choice set. The existence of $\gamma_\beta$ follows by standard results about the existence of measurable solutions to continuous and compact maximization problems e.g., Aliprantis and Border (2007), Theorem 18.19.

Thus, the objective in the minimization problem on the first line in (6.3) equals

$$(1-\delta)\sum_a \overline{\alpha}_\beta(a)g\big(a,\beta,q^{\alpha_\beta,p}(a)\big) + \delta \int v^\delta(Pq; g, P)\, d\nu_\beta(q)$$

$$\leq (1-\delta)\sum_a \overline{\alpha}_\beta(a)g\big(a,\beta,q^{\alpha_\beta,p}(a)\big) + \delta \int v^\delta(Pq; g, P)\, d\nu^{\alpha_\beta,p}(q)$$

$$= \sum_a \overline{\alpha}_\beta(a)\big((1-\delta)g\big(a,\beta,q^{\alpha_\beta,p}(a)\big) + \delta v^\delta\big(Pq^{\alpha_\beta,p}(a); g, P\big)\big)$$

$$\leq \max_{\alpha\in(\Delta A)^S}\sum_a \overline{\alpha}(a)\big((1-\delta)g\big(a,\beta,q^{\alpha,p}(a)\big) + \delta v^\delta\big(Pq^{\alpha,p}(a); g, P\big)\big),$$

where $\nu^{\alpha_\beta,p}$ on the second line is the distribution over posteriors defined by (6.1); the first inequality follows because $\nu^{\alpha_\beta,p} \leq^B \nu_\beta$ by construction,[10] and the value is concave in the initial distribution by Lemma 1; the third line follows from the second by unraveling the definitions. By inspection, the last line is the objective function in the minimization problem in (6.2), so the claim follows by Lemma 1.

We show then that $v^\delta(p; g, P)$ is not larger than the right-hand side of the first line of (6.3). For each $\beta \in \Delta B$, let $\alpha_\beta \in (\Delta A)^S$ be a solution to the maximization problem in (6.2). Let $\nu^{\alpha_\beta,p}$ be the induced distribution over posteriors defined by (6.1), and let $\tau^{\alpha_\beta,p}(q) \in \Delta A$ denote a measurable version of the conditional distribution over actions given posterior $q$. Then,

$$\sum_a \overline{\alpha}_\beta(a)\big((1-\delta)g\big(a,\beta,q^{\alpha_\beta,p}(a)\big) + \delta v^\delta\big(Pq^{\alpha_\beta,p}(a); g, P\big)\big)$$

$$= \int \big((1-\delta)g\big(\tau^{\alpha_\beta,p}(q),\beta,q\big) + \delta v^\delta(Pq; g, P)\big)\, d\nu^{\alpha_\beta,p}(q)$$

$$\leq \max_{\nu\in\Delta^2S:\mathcal{E}\nu=p}\int \Big((1-\delta)\max_{a\in A}g(a,\beta,q) + \delta v^\delta(Pq; g, P)\Big)\, d\nu(q),$$

where the last line is the objective function in the minimization problem on the first line of (6.3). Thus, the claim follows by Lemma 1.

The second equality in (6.3) follows from Fan's (1953) minmax theorem (notice that the expression is linear in $\nu$ and convex in $\beta$) and the definition of $\hat{g}$ in (3.1).    *Q.E.D.*

### 6.2. *Discounted Value*

The recursive formula (6.3) implies the following "approximately stationary" characterization of the discounted value. For any $\delta < 1$, define

$$\psi^\delta_{\pi,P} = (1-\delta)\sum_t \delta^t P^t e_\pi \in \Delta^2 S.$$

---

[10]To see this, note that we can generate $\nu_\beta$ in two steps by first drawing an action $a$ according to $\overline{\alpha}_\beta$ (with $p$ as the prior), and then drawing a posterior belief $q$ according to the conditional distribution over posteriors given $a$. Then, by construction, $\nu^{\alpha_\beta,p}$ is the distribution of posteriors after the first step, and the second step amounts to a mean-preserving spread.

Then $\psi^\delta_{\pi,P}$ is the discounted average distribution of the minimizer's beliefs if the maximizer reveals no information.

LEMMA 3: *For every $\delta < 1$, there exist a distribution $\mu^\delta \in \Delta^2 S$ with $\psi^\delta_{\pi,P} \leq^B \mu^\delta$ and a mean-preserving spread $m^\delta \in \mathcal{M}$ such that*

$$v^\delta(\pi; g, P) = \int \hat{g}(m^\delta(p)) \, d\mu^\delta(p), \qquad (6.5)$$

*and for each continuous function $f : \Delta S \to \mathbb{R}$ with $\|f\|_\infty \leq 1$,*

$$\left| \int f \, dP(\mu^\delta * m^\delta) - \int f \, d\mu^\delta \right| \leq 2(1 - \delta).$$

The second display above implies that for $\delta$ close to 1, $(\mu^\delta, m^\delta)$ is approximately stationary in the sense that $P(\mu^\delta * m^\delta)$ is close to $\mu^\delta$ in the weak topology.

PROOF: The objective function on the second line of (6.3) is continuous in $\nu$, since $\hat{g}$ is continuous, and it is constant and thus measurable in $p$. The feasible set $\{\nu : \mathcal{E}\nu = p\}$ is compact and continuous as a correspondence of $p$. By the Measurable Maximum Theorem (Aliprantis and Border (2007), Theorem 18.19), the maximizer correspondence is nonempty valued and admits a measurable selection $m^\delta : \Delta S \to \Delta^2 S$.

Let $\mu_0 \in \Delta^2 S$ be the Dirac measure at $\pi$ (i.e., $e_\pi$). Define $\mu_t = P(\mu_{t-1} * m^\delta)$ for $t \geq 1$ by induction on $t$, and let $\mu^\delta = \sum_{t=0}^\infty (1 - \delta)\delta^t \mu_t$. By induction on $t$, we verify that $\mathcal{E}\mu_t = P^t \pi$, which implies $P^t e_\pi = e_{P^t \pi} \leq^B \mu_t$ for each $t$. This in turn implies $\psi^\delta_{\pi,P} \leq^B \mu^\delta$, since the Blackwell ordering is preserved under convex combinations.[11] Equation (6.5) then follows from repeated application of (6.3).

Finally, for each continuous function $f : \Delta S \to \mathbb{R}$ such that $|f| \leq 1$, we have

$$\int f \, dP(\mu^\delta * m^\delta) = \int f \, d\left( \sum_{t=0}^\infty (1 - \delta)\delta^t P(\mu_t * m^\delta) \right)$$

$$= \int f \, d\left( \sum_{t=0}^\infty (1 - \delta)\delta^t \mu_{t+1} \right)$$

$$= \int f \, d\left( (1 - \delta)\mu_0 + \sum_{t=0}^\infty (1 - \delta)\delta^{t+1} \mu_{t+1} \right)$$

$$+ \sum_{t=0}^\infty (1 - \delta)(\delta^t - \delta^{t+1}) \int f \, d\mu_{t+1} - (1 - \delta) \int f \, d\mu_0$$

$$= \int f \, d\mu^\delta + (1 - \delta) \sum_{t=0}^\infty \delta^t (1 - \delta) \int f \, d\mu_{t+1} - (1 - \delta) \int f \, d\mu_0.$$

The last two terms are each bounded in absolute value by $(1 - \delta)$ as desired.        *Q.E.D.*

---

[11]That is, if $\nu_n \leq^B \psi_n$, $n \in \mathbb{N}$, then for any $\lambda_1, \lambda_2, \ldots \geq 0$ with $\sum \lambda_n = 1$, $\sum_n \lambda_n \nu_n \leq^B \sum_n \lambda_n \psi_n$. Indeed, for any concave $f$, $\int f \, d(\sum_n \lambda_n \psi_n) = \sum_n \lambda_n \int f \, d\psi_n \leq \sum_n \lambda_n \int f \, d\nu_n = \int f \, d(\sum \lambda_n \nu_n)$.

### 6.3. *Maximum*

Let $V = \{(\mu, m) \in \Delta^2 S \times \mathcal{M} : \psi_{\pi,P} \leq^B \mu \text{ and } P(\mu * m) \leq^B \mu\}$, where $\psi_{\pi,P} = \lim_{\delta \to 1} \psi_{\pi,P}^{\delta}$. We show that $\int \hat{g}(m(p)) \, d\mu(p)$ attains its supremum on $V$, and this supremum is not smaller than the limit of the right-hand side of (6.5). Both claims follow from the same argument, so we combine them into a single lemma.

LEMMA 4: *For every $\delta < 1$, let $(\mu^{\delta}, m^{\delta})$ be as in Lemma 3. There exists $(\mu^*, m^*) \in V$ such that*

$$\limsup_{\delta \to 1} \int \hat{g}(m^{\delta}(p)) \, d\mu^{\delta}(p) \leq \sup_{(\mu,m) \in V} \int \hat{g}(m(p)) \, d\mu(p)$$

$$= \int \hat{g}(m^*(p)) \, d\mu^*(p).$$

PROOF: It will be useful to view $\Delta^2 S \times \mathcal{M}$ as a subset of $\Delta^3 S$, the set of probability distributions over $\Delta^2 S$, endowed with the weak topology. The idea is that each pair $(\mu, m)$ generates a distribution over $\Delta^2 S$ as follows: $m$ assigns a point $m(p) \in \Delta^2 S$ to each $p \in \Delta S$, so drawing $p$ according to $\mu$ yields an element of $\Delta^3 S$. Formally, take a mapping $\iota : \Delta^2 S \times \mathcal{M} \to \Delta^3 S$ such that for each continuous function $\phi : \Delta^2 S \to \mathbb{R}$,

$$\int \phi(m(p)) \, d\mu(p) = \int \phi(\nu) \, d\iota(\mu, m)(\nu).$$

The above equation uniquely defines the mapping $\iota$.

Take a sequence $(\mu^n, m^n) \in \{(\mu^{\delta}, m^{\delta}) : \delta \geq 1 - \frac{1}{n}\} \cup V$ such that

$$\lim_{n \to \infty} \int \hat{g}(m^n(p)) \, d\mu^n(p) = \max \left\{ \limsup_{\delta \to 1} \int \hat{g}(m^{\delta}(p)) \, d\mu^{\delta}(p), \sup_{(\mu,m) \in V} \int \hat{g}(m(p)) \, d\mu(p) \right\}.$$

Because $\Delta^3 S$ is compact, we can assume (by taking a subsequence if necessary) that the sequence $(\mu^n, m^n)$ converges. That is, there exists a distribution $\omega \in \Delta^3 S$ such that $\lim_n \iota(\mu^n, m^n) = \omega$. Furthermore, since $\hat{g}$ is continuous, we have

$$\lim_{n \to \infty} \int \hat{g}(m^n(p)) \, d\mu^n(p) = \lim_{n \to \infty} \int \hat{g}(\nu) \, d\iota(\mu^n, m^n)(\nu) = \int \hat{g}(\nu) \, d\omega(\nu).$$

The limit $\omega$ is a distribution over $\nu \in \Delta^2 S$. We let $\mu^* \in \Delta^2 S$ denote the induced distribution over $\mathcal{E}\nu$. Formally, $\mu^*$ is defined so that for each continuous $f : \Delta S \to \mathbb{R}$,

$$\int f(p) \, d\mu^*(p) = \int f(\mathcal{E}\nu) \, d\omega(\nu),$$

where the latter integral is over $\Delta^2 S$. Let $\omega(\cdot | \mathcal{E}\nu = p)$ be a measurable version of the conditional distribution, and let $m^*(p) = \int \nu \, d\omega(\nu | \mathcal{E}\nu = p) \in \Delta^2 S$ be the expected value of the conditional distribution. Jensen's inequality and the concavity of $\hat{g}$ imply

$$\int \hat{g}(\nu) \, d\omega(\nu | \mathcal{E}\nu = p) \leq \hat{g}(m^*(p)).$$

Thus, collecting from above, we have

$$\int \hat{g}(m^n(p)) \, d\mu^n(p) \to \int \hat{g}(v) \, d\omega(v)$$

$$= \int \left( \int \hat{g}(v) \, d\omega(v | \mathcal{E}v = p) \right) d\mu^*(p) \qquad (6.6)$$

$$\leq \int \hat{g}(m^*(p)) \, d\mu^*(p).$$

Given the choice of the sequence $(\mu^n, m^n)$, the lemma follows from the above display if we show that the pair $(\mu^*, m^*)$ is in fact an element of $V$.

To this end, we note first that $\mu^n \to \mu^*$, since for each continuous $f : \Delta S \to \mathbb{R}$,

$$\int f(p) \, d\mu^*(p) = \int f(\mathcal{E}v) \, d\omega(v) = \lim_{n \to \infty} \int f(\mathcal{E}v) \, d\iota(\mu^n, m^n)(v)$$

$$= \lim_{n \to \infty} \int f(\mathcal{E}m^n(p)) \, d\mu^n(p)$$

$$= \lim_{n \to \infty} \int f(p) \, d\mu^n(p),$$

where the last equality follows since $\mathcal{E}m(p) = p$ for any mean-preserving spread $m$. Similarly, we see that $\mu^n * m^n \to \mu^* * m^*$ by observing that

$$\int \left( \int f(q) \, dm^*(q|p) \right) d\mu^*(p) = \int \left( \int \left( \int f(q) \, dv(q) \right) d\omega(v | \mathcal{E}v = p) \right) d\mu^*(p)$$

$$= \int \left( \int f(q) \, dv(q) \right) d\omega(v)$$

$$= \lim_{n \to \infty} \int \left( \int f(q) \, dv(q) \right) d\iota(\mu^n, m^n)(v)$$

$$= \lim_{n \to \infty} \int \left( \int f(q) \, dm^n(q|p) \right) d\mu^n(p).$$

It follows that $P(\mu^n * m^n) \to P(\mu^* * m^*)$.

The choice of sequence $(\mu^n, m^n)$ and Lemma 3 imply, for each $f : \Delta S \to \mathbb{R}$ concave,

$$\int f \, dP(\mu^n * m^n) - \int f \, d\mu^n \geq -\frac{2}{n} \|f\|_\infty.$$

Letting $n \to \infty$ gives $\int f \, dP(\mu^* * m^*) - \int f \, d\mu^* \geq 0$, and thus $P(\mu^* * m^*) \leq^B \mu^*$.

Finally, notice that for each $n$, either there exists $\delta_n \geq 1 - \frac{1}{n}$ such that $\psi_{\pi,P}^{\delta_n} \leq^B \mu^n$, or $(\mu^n, m^n) \in V$ and $\psi_{\pi,P} \leq^B \mu^n$. Therefore, for each $f : \Delta S \to \mathbb{R}$ concave, we have

$$\int f \, d\psi_{\pi,P} = \lim_{n \to \infty} \int f \, d\psi_{\pi,P}^{\delta_n} \geq \lim_{n \to \infty} \int f \, d\mu^n = \int f \, d\mu^*,$$

where the first equality is by definition of $\psi_{\pi,P}$ as the (weak) limit of $\psi_{\pi,P}^\delta$, and the last equality follows by $\mu^n \to \mu^*$. Thus, $\psi_{\pi,P} \leq^B \mu^*$, and hence $(\mu^*, m^*) \in V$. 　　　Q.E.D.

### 6.4. *Proof of Theorem 1*

We observe first that $v(\pi; g, P)$ is not larger than the right-hand side of (3.2). Indeed, by Lemmas 0, 3, and 4, we have

$$v(\pi; g, P) = \lim_{\delta \to 1} v^\delta(\pi; g, P) \leq \max_{(\mu,m) \in V} \int \hat{g}\big(m(p)\big) \, d\mu(p).$$

We then show that $v(\pi; g, P)$ is not smaller than the right-hand side of (3.2). By Lemma 4, the maximum on the right-hand side is attained by some $(\mu^*, m^*) \in V$.[12] By the recursive formula (6.3), for each $\delta < 1$ and $p \in \Delta S$, we have

$$v^\delta(p; g, P) \geq (1 - \delta)\hat{g}\big(m^*(p)\big) + \delta \int v^\delta(Pq; g, P) \, dm^*(q|p).$$

Taking expectations with respect to $\mu^*$ gives

$$\int v^\delta(p; g, P) \, d\mu^*(p) \geq (1 - \delta) \int \hat{g}\big(m^*(p)\big) \, d\mu^*(p) + \delta \int v^\delta(p; g, P) \, dP\big(\mu^* * m^*\big)(p)$$

$$\geq (1 - \delta) \int \hat{g}\big(m^*(p)\big) \, d\mu^*(p) + \delta \int v^\delta(p; g, P) \, d\mu^*(p),$$

where the second line follows because the discounted value is concave in $p$ by Lemma 1, and we have $P(\mu^* * m^*) \leq^B \mu^*$ by $(\mu^*, m^*) \in V$. Thus,

$$\int \hat{g}\big(m^*(p)\big) \, d\mu^*(p) \leq \int v^\delta(p; g, P) \, d\mu^*(p).$$

Letting $\delta \to 1$, the Dominated Convergence Theorem and Lemma 0 imply

$$\int \hat{g}\big(m^*(p)\big) \, d\mu^*(p) \leq \int v(p; g, P) \, d\mu^*(p) \leq \int v(p; g, P) \, d\psi_{\pi,P}(p),$$

where the second inequality is by $\psi_{\pi,P} \leq^B \mu^*$ (since $(\mu^*, m^*) \in V$) and concavity. Thus, it suffices to show that $v(\pi; g, P) \geq \int v(p; g, P) \, d\psi_{\pi,P}(p)$.

For any $\delta < 1$, a feasible strategy for the maximizer is to not reveal any information before some period $s$, and then switch to the optimal strategy in period $s$. Hence,

$$v^\delta(\pi; g, P) \geq (1 - \delta) \sum_{t=0}^{s-1} \delta^t \hat{g}(e_{P^t \pi}) + \delta^s v^\delta\big(P^s \pi; g, P\big).$$

Letting $\delta \to 1$ gives $v(\pi; g, P) \geq v(P^s \pi; g, P)$. Since $s$ was arbitrary, for any $\gamma \in (0, 1)$,

$$v(\pi; g, P) \geq (1 - \gamma) \sum_s \gamma^s v\big(P^s \pi; g, P\big) = \int v(p; g, P) \, d\psi_{\pi,P}^\gamma \overset{\gamma \to 1}{\to} \int v(p; g, P) \, d\psi_{\pi,P}(p),$$

where we have used the definitions of $\psi_{\pi,P}^\gamma$ and $\psi_{\pi,P}$. This establishes Theorem 1.

---

[12]Replacing $P(\mu * m) \leq^B \mu$ with $P(\mu * m) = \mu$ in the definition of $V$, the proof of Lemma 4 shows that some $(\mu, m)$ with $P(\mu * m) = \mu$ achieves the maximum. The only change in the argument is that at the end, $|\int f \, dP(\mu^n * m^n) - \int f \, d\mu^n| \leq \frac{2}{n} \|f\|_\infty \to 0$ for all $f$ continuous, implying $P(\mu^* * m^*) = \mu^*$.

## 7. PROOF OF THEOREM 2′

We first verify that Theorem 2 follows from Theorem 2′. Let $P$ and $Q$ be ergodic. To see that 2′(c) implies 2(c), let $P\nu \leq^B \nu$. Then $P\mathcal{E}\nu = \mathcal{E}\nu$, which implies $\mathcal{E}\nu = \pi_P$. Thus, $\psi_{\pi,P} = e_{\pi_P} \leq^B \nu$, so 2′(c) implies $Q\nu \leq^B \nu$. Conversely, let 2(c) hold. Then the second part of 2′(c) holds. For the first, note that $Pe_{\pi_P} = e_{\pi_P}$. Thus, $Qe_{\pi_P} \leq^B e_{\pi_P}$ by 2(c), implying $Q\pi_P = \pi_P$. But then $\pi_Q = \pi_P$ and $\psi_{\pi,P} = \psi_{\pi,Q} = e_{\pi_P}$ as desired. Theorem 2′ now implies the equivalence of 2(a) and 2(c). The equivalence of 2(b) and 2(c) follows by the same argument as the equivalence of (b) and (c) below.

We then turn to the proof of Theorem 2′. To see that (b) implies (c), suppose $P\nu \leq^B \nu$ and $\psi_{\pi,P} \leq^B \nu$. Then $P\nu * m = \nu$ for some $m \in \mathcal{M}$, and $P(P\nu * m) = P\nu$. Moreover, $\psi_{\pi,P} = P\psi_{\pi,P} \leq^B P\nu$, because the Blackwell order is preserved under linear transformations.[13] Thus, setting $\mu = P\nu$ in (b) gives $Q(P\nu * m) \leq^B P\nu$, or $Q\nu \leq^B P\nu$. For the converse, let $P(\mu * m) \leq^B \mu$ and $\psi_{\pi,P} \leq^B \mu$. Then $P(\mu * m) \leq^B \mu * m$, so (c) implies $Q(\mu * m) \leq^B P(\mu * m) \leq^B \mu$. Hence, (b) and (c) are equivalent.

That (b) implies (a) follows immediately from Theorem 1, since (b) implies that any $(\mu, m)$ feasible in (3.2) under $P$ is feasible under $Q$. Thus, it only remains to show that (a) implies (b). We establish the contrapositive.

First, suppose $\psi_{\pi,Q} \not\leq^B \psi_{\pi,P}$. Then for some concave $f : \Delta S \to \mathbb{R}$, we have

$$\int f \, d\psi_{\pi,Q} < \int f \, d\psi_{\pi,P}. \tag{7.1}$$

Because $f$ is concave, there exists a set $L$ of affine functions $l : \Delta S \to \mathbb{R}$ such that $f(p) = \inf_{l \in L} l(p)$, $p \in \Delta S$. Because $\Delta S$ is compact, we can assume that $L$ is finite. (Otherwise, replace $f$ by an approximation obtained as the minimum over finitely many affine functions. A sufficiently good approximation satisfies (7.1).) Construct game $g$ by letting $A = \{a_0\}$, $B = L$, and $g(a_0, l, s) = l(e_s)$. Then $g(a, l, p) = l(p)$, and

$$v(\pi; g, P) = \lim_{\delta \to 1} \sum_t (1 - \delta)\delta^t f(P^t \pi) = \lim_{\delta \to 1} \int f \, d\psi_{\pi,P}^\delta = \int f \, d\psi_{\pi,P}.$$

Similarly, $v(\pi; g, Q) = \int f \, d\psi_{\pi,Q} < \int f \, d\psi_{\pi,P} = v(\pi; g, P)$ by (7.1), contradicting (a).

Second, suppose $\psi_{\pi,Q} \leq^B \psi_{\pi,P}$, but there exists $(\mu_0, m_0) \in \Delta^2 S \times \mathcal{M}$ such that $\psi_{\pi,P} \leq^B \mu_0$, $P(\mu_0 * m_0) \leq^B \mu_0$, and $Q(\mu_0 * m_0) \not\leq^B \mu_0$. Then there exists a concave function $f : \Delta S \to \mathbb{R}$ such that

$$\int f \, d\mu_0 - \int f \, dQ(\mu_0 * m_0) > 0 \geq \sup_{(\mu, m) : Q(\mu * m) \leq^B \mu} \int f \, d\mu - \int f \, dQ(\mu * m). \tag{7.2}$$

Because $f$ is concave, there exists a set $L$ of affine functions $l : \Delta S \to \mathbb{R}$ such that $f(p) = \inf_{l \in L} l(p)$. Because $\Delta S$ is compact, we can again assume that $L$ is finite.

We now construct a game $g$ such that $v(\pi; g, P) > v(\pi; g, Q)$. Let $A = B = L$. Recall that for each affine function $l$, there exist constants $g_s^l$ for each $s \in S$ such that $l(p) = \sum_s g_s^l p_s$. (Since $\sum_s p_s = 1$, this allows for an additive constant term in $l(p)$.) For each $a \in A$, $b \in B$, and $s \in S$, define

$$g(a, b, s) = g_s^b - a(Qe_s),$$

---

[13] Indeed, for all $f$ concave, $\int f \circ P \, d\nu \leq \int f \circ P \, d\psi_{\pi,P}$, since $f \circ P$ is concave and $\psi_{\pi,P} \leq^B \nu$.

where $e_s$ is the Dirac measure at state $s$. For each $\beta \in \Delta B$ and $q \in \Delta S$, observe that

$$\max_a g(a, \beta, q) = \sum_{b,s} \beta(b)q(s)g_s^b - \min_a a(Qq) = \sum_{b,s} \beta(b)q(s)g_s^b - f(Qq).$$

Thus, for each $\mu \in \Delta^2 S$ and $m \in \mathcal{M}$, we have

$$\int \hat{g}(m(p)) \, d\mu(p) = \int \left( \min_{\beta \in \Delta B} \int \sum_{b,s} \beta(b)q(s)g_s^b \, dm(q|p) - \int f(Qq) \, dm(q|p) \right) d\mu(p)$$

$$= \int \min_{b \in B} \sum_s g_s^b p(s) \, d\mu(p) - \int f(Qp) \, d(\mu * m)(p)$$

$$= \int f(p) \, d\mu(p) - \int f(p) \, dQ(\mu * m)(p).$$

Since $\psi_{\pi,P} \leq^B \mu_0$ and $P(\mu_0 * m_0) \leq^B \mu_0$, it then follows from Theorem 1 and inequality (7.2) that $v(\pi; g, P) > 0 \geq v(\pi; g, Q)$ in negation of (a), as desired.

# APPENDICES

The following appendices collect the proofs omitted in the main text. Appendix A contains the proof of Corollary 2. Appendix B contains the proof of Corollary 1 (which relies partially on Corollary 2). Appendix C contains the proof of Corollary 5.

## APPENDIX A: Proof of Corollary 2

### A.1. *Proof of Condition (a)*

Suppose that $P\nu \leq^B \nu$. Then, for any $n \geq 2$,

$$\nu \geq^B P\nu \geq P^2\nu \geq^B \cdots \geq P^n\nu,$$

where each subsequent inequality follows from the previous one as $\leq^B$ is preserved under $P$ (see footnote 13). Therefore, if $Q = P(\lambda)$ for some $\lambda \in \Lambda$, then we have

$$Q\nu = \left( \sum \lambda_n P^n \right)\nu \leq^B \sum \lambda_n (P^n \nu) \leq^B P\nu. \tag{A.1}$$

The second inequality in (A.1) follows, because $\leq^B$ is preserved under convex combinations (see footnote 11). To see the first, note that for any $\psi \in \Delta^2 S$ and $\lambda \in (0, 1)$,

$$\left( \lambda P + (1 - \lambda)Q \right)\psi \leq^B \lambda P\psi + (1 - \lambda)Q\psi. \tag{A.2}$$

Indeed, for any concave $f$, we can use Jensen's inequality to obtain

$$\int f\, d\big(\lambda P\psi + (1-\lambda)Q\psi\big) = \lambda \int f(Pp)\, d\psi(p) + (1-\lambda) \int f(Qp)\, d\psi(p)$$

$$= \int \big(\lambda f(Pp) + (1-\lambda)f(Qp)\big)\, d\psi(p)$$

$$\leq \int f\big(\lambda Pp + (1-\lambda)Qp\big)\, d\psi(p)$$

$$= \int f\big((\lambda P + (1-\lambda)Q)p\big)\, d\psi(p)$$

$$= \int f\, d\big(\lambda P + (1-\lambda)Q\big)\psi.$$

The argument clearly extends to countable convex combinations.

Condition (a) now follows by inequality (A.1) and Theorem 2(c).

### A.2. *Proof of Condition (b)*

Suppose that there exists a distribution $p \in \Delta S$ such that $Qp \notin \mathrm{con}\{Pp, P^2 p, \ldots, \pi_P\}$, where $\pi_P$ is the invariant distribution of $P$. Fix $p$ for the rest of the proof. We will show that there exists $\mu$ such that $P\mu \leq^B \mu$ and not $Q\mu \leq^B P\mu$, which contradicts $P \preceq Q$ by conditions (a) and (c) of Theorem 2.

Let $A_\varepsilon^0 = \{(1-\varepsilon)\pi_P + \varepsilon e_s : s \in S\}$, with $e_s$ a Dirac measure at $s$, be a finite set of beliefs and let $A_\varepsilon = \mathrm{con}\, A_\varepsilon^0$. Then, $PA_\varepsilon \subseteq A_\varepsilon$, $\pi_P \in \mathrm{int}\, A_\varepsilon$, and $\lim_{\varepsilon \to 0} A_\varepsilon = \{\pi_P\}$. For $n = 1, 2, \ldots, \infty$, let $D_n^0 = \{Pp, P^2 p, \ldots, P^n p\}$ and $D_n = \mathrm{con}\, D_n^0$. Then, $Qp \notin D_\infty$. Because $D_\infty$ is closed and the sets $A_\varepsilon$ are arbitrarily small neighborhoods of $\pi_P \in D_\infty$, there exists $\varepsilon > 0$ small enough so that $Qp \notin \mathrm{con}\{A_\varepsilon^0 \cup D_\infty^0\}$. Because $P^n p \to \pi_P$, and $A_\varepsilon$ is a neighborhood of $\pi_P$, there exists $n$ large enough so that $P^m p \in \mathrm{int}\, A_\varepsilon$ for all $m > n$. It follows that $P(\mathrm{con}(A_\varepsilon^0 \cup D_n^0)) \subseteq \mathrm{con}(A_\varepsilon^0 \cup D_n^0)$ and $Qp \notin \mathrm{con}\{A_\varepsilon^0 \cup D_n^0\}$.

Let $K = \{p\} \cup D_n^0 \cup A_\varepsilon^0$ and let $K' = PK$. For each $q \in K'$, we have $q \in \mathrm{con}\, K$, and there exists a probability distribution over beliefs $m_q \in \Delta K$ such that $\mathcal{E}m_q = q$ and $m_q(K) = 1$. We assume that $m_q(p) > 0$ for each $q \in P(\mathrm{int}\, A_\varepsilon^0) \subseteq K'$ and that $m_q$ is continuous in $q$. We treat $m$ as a mean-preserving spread (despite the fact that it is not defined on $q \notin K'$).

Consider a Markov process with state space $K$ and transition function $T : K \to \Delta K$ given by $T(q) = m_{Pq}$. Let $\nu \in \Delta K$ be an invariant distribution of such a process (it exists due to the continuity of $m_q$). Thus, $P\nu * m = \nu$, or $P\nu \leq^B \nu$. The choice of probability distributions $m_q$ implies $\nu(p) > 0$, which in turn implies $Qp \in \mathrm{supp}\, Q\nu$. Because $Qp \notin \mathrm{con}\, K' = \mathrm{con}\, \mathrm{supp}\, P\nu$, we see that $\mathrm{supp}\, Q\nu \nsubseteq \mathrm{con}\, \mathrm{supp}\, P\nu$, which implies $Q\nu \nleq^B P\nu$, as desired.

### A.3. *Proof of Condition (c)*

From now on, assume that $P$ and $Q$ are two ergodic operators such that $P$ has purely real eigenvalues, and for each $p \in \Delta S$, we have $Qp \in \mathrm{con}\{\pi_P, Pp, P^2 p, \ldots, \pi_P\}$.

We recall some concepts from linear algebra. Let $E = \{v \in R^S : \sum v_s = 0\}$ be an $(S-1)$-dimensional linear space. The operator $P$ can be uniquely represented as a linear mapping $\hat{P} : E \to E$ such that for each $\pi \in \Delta S$,

$$\hat{P}(\pi - \pi_P) = P\pi - \pi_P.$$

From now on, we identify $\hat{P}$ with $P$ and we drop the hat from the notation. Similarly, we identify $Q$ with the linear mapping $\hat{Q}$ it induces on $E$. For each $v \in E$, let

$$\Lambda(v) = \left\{ \lambda \in \Lambda : \sum_n \lambda_n P^n v = Qv \right\}.$$

The sets $\Lambda(v)$ are compact and, by the hypothesis of the lemma, nonempty. Our claim will be proven if we show that $\bigcap_v \Lambda(v) \neq \emptyset$.

A linear subspace $F \subseteq E$ is $P$-invariant, if $PF \subseteq F$ (i.e., $v \in F$ implies $Pv \in F$). A $P$-invariant subspace $F \subseteq E$ is $P$-cyclic, if there is a vector $v$, a (real) eigenvalue $\phi$, and a multiplicity $r$ such that $v, (P - \phi I)v, \ldots, (P - \phi I)^{r-1}v$ is a basis for $F$ and $(P - \phi I)^r v = 0$. The Jordan decomposition theorem says that $E$ can be represented as a direct sum of $P$-invariant $P$-cyclic subspaces

$$E = \bigoplus F_m.$$

Let $v_m$, $\phi_m$, and $r_m$ denote the vector, the eigenvalue, and the multiplicity associated with $F_m$. For each $l = 1, \ldots, r_m$, let $v_{m,l} = (P - \phi_m I)^{l-1} v_m$ be the basis for space $F_m$.

The next result describes a representation of operators $P(\lambda)$ for $\lambda \in \Lambda$.

LEMMA A.1: *For every $\lambda \in \Lambda$, there exist $p_l^\lambda$ for $l = 1, \ldots, r_{m-1}$ such that the linear operator $P(\lambda)$ has a representation*

$$P(\lambda) = \begin{bmatrix} p_1^\lambda & p_2^\lambda & \cdots & p_{r_m}^\lambda \\ 0 & p_1^\lambda & \cdots & p_{r_m-1}^\lambda \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & p_1^\lambda \end{bmatrix}$$

*in the basis $\{v_{m,l}\}_{l=1}^{r_m}$. In other words, for every $v = \sum v^l v_{m,l} \in F_m$,*

$$P(\lambda)v = \sum_{k=1}^{r_m} \left( \sum_{l=1}^{r_m-k} p_l^\lambda v^{k+l} \right) v_k.$$

PROOF: Because the above representation is preserved under convex combinations, it is enough to demonstrate that for each $n$, the linear operator $P^n$ has the above representation. The claim holds for $n = 1$ with $p_0^{(1)} = \phi$, $p_1^{(1)} = 1$, and $p_l^{(1)} = 0$ for $l > 1$. Suppose that the claim holds for $n$. Then, algebra shows that

$$p_l^{(n+1)} = \sum_{k=1}^l p_k^{(n)} p_{l+1-k}^{(1)},$$

which yields the desired representation.                                    *Q.E.D.*

LEMMA A.2: *If a linear subspace $F \subseteq E$ is $P$-invariant, then it is $Q$-invariant.*

PROOF: Take any $x \in F$. Because $F$ is an invariant subspace of $P$, $Px, P^2 x, \ldots \in F$. This implies that $Qx \in \text{con}\{0, Px, P^2 x, \ldots\} \subseteq F$.                    *Q.E.D.*

Lemma A.2 implies that each $F_m$ is $Q$-invariant and the restriction $Q|_{F_m}$ is a linear operator from $F_m$ to $F_m$.

LEMMA A.3: *There exist $q_l$ for $l = 0, \ldots, r_{m-1}$ such that $Q|_{F_m}$ has a representation*

$$Q = \begin{bmatrix} q_1 & q_2 & \cdots & q_{r_m} \\ 0 & q_1 & \cdots & q_{r_m-1} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & q_1 \end{bmatrix}$$

*in the basis $\{v_{m,l}\}_{l=1}^{r_m}$. Moreover, there exists $f_m^* \in F_m$ such that for every $\lambda \in \Lambda(f_m^*)$, $Q|_{F_m} = P(\lambda)|_{F_m}$.*

PROOF: Let $[q_{kl}^m]_{k,l=0,\ldots,r-1}$ be the representation of $Q$ on the $Q$-invariant subspace $F_m$ in the basis $v_{m,0}, \ldots, v_{m,r-1}$. Let $q_l = q_{r-l+1,r_m}$ for each $l$.

Define $v^\varepsilon = \sum_l v_{m,l} \varepsilon^{l-1}$. For each $\varepsilon > 0$, let $\lambda^\varepsilon \in \Lambda(v^\varepsilon)$. Also, let

$$I_\varepsilon = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & \varepsilon^{-1} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \varepsilon^{-(r_m-1)} \end{bmatrix}.$$

Suppose that there exists $l^* \leq r_m$ that is the smallest index such that there exists $\varepsilon > 0$ and $k = 1, \ldots, r_m - l + 1$ such that $q_{k,k+l^*-1} \neq p^{\lambda^\varepsilon}$.

By the choice of $\lambda^\varepsilon$, $(P(\lambda^\varepsilon) - Q)v^\varepsilon = 0$. Moreover, simple calculations show that for each $\varepsilon > 0$,

$$0 = \varepsilon^{-(l^*-1)} I_\varepsilon \left(P(\lambda^\varepsilon) - Q\right) v^\varepsilon$$

$$= \varepsilon^{-(l^*-1)} I_\varepsilon \begin{bmatrix} \left(p_{l^*}^{\lambda^\varepsilon} - q_{1,l^*}\right)\varepsilon^{l^*-1} + O(\varepsilon^{l^*}) \\ \vdots \\ \left(p_{l^*}^{\lambda^\varepsilon} - q_{r_m-l^*-2,r_m-1}\right)\varepsilon^{r_m-1} + O(\varepsilon^{r_m}) \\ \left(p_{l^*}^{\lambda^\varepsilon} - q_{r_m-l^*-1,r_m}\right)\varepsilon^{r_m} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} p_{l^*}^{\lambda^\varepsilon} - q_{1,l^*} + O(\varepsilon) \\ \vdots \\ p_{l^*}^{\lambda^\varepsilon} - q_{r_m-l^*-2,r_m-1} + O(\varepsilon) \\ p_{l^*}^{\lambda^\varepsilon} - q_{r_m-l^*-1,r_m} \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

The equality implies that (i) $p_{l^*}^{a^\varepsilon} \to q_{k,k+l^*-1}$ for each $k = 1, \ldots, r_m - l^*$, and that (ii) $p_{l^*}^{a^\varepsilon} = q_{r_m-l^*-1,r_m}$ for each $\varepsilon > 0$. Further, (i) and (ii) imply that, for each $k$,

$$p_{l^*}^{\lambda^\varepsilon} = q_{k,k+l^*-1} = q_{l^*}.$$

For the last claim, take any $\varepsilon > 0$ and let $f_m^* = v^\varepsilon$ for any $\varepsilon > 0$. Q.E.D.

LEMMA A.4: *For each $m$, take any $f_m \in F_m$, and let $f = \sum f_m$. Then, $\Lambda(f) = \bigcap_m \Lambda(f_m) \neq \emptyset$.*

PROOF: It is easy to see that $\bigcap_m \Lambda(f_m) \subseteq \Lambda(f)$. We show the other inclusion. Take $\lambda \in \Lambda(f)$. Then, $P(\lambda)f = \sum P(\lambda)f_m = \sum Qf_m$. Because the spaces $F_m$ are both $P(\lambda)$ and $Q$-invariant, $P(\lambda)f_m \in F_m$ and $Qf_m \in F_m$. Because $E$ is a direct sum of the spaces $F_m$, we have $P(\lambda)f_m = Qf_m$ for each $m$. This implies that $\lambda \in \Lambda(f_m)$ for each $m$. Q.E.D.

We can now finish the proof of our claim. Let $f_m^* \in F_m$ be as in Lemma A.3. By Lemma A.4, there exists $\lambda^* \in \bigcap_m \Lambda(f_m^*)$. Moreover, by Lemma A.3, $Q|_{F_m} = P(\lambda^*)|_{F_m}$ for each $m$. Because $E$ is a direct sum of spaces $F_m$, it must be that $Q = P(\lambda^*)$.

## APPENDIX B: PROOF OF COROLLARY 1

Part (a) is proven in the main text, so it remains to establish parts (b) and (c).

### B.1. *Proof of Part (b)*

We start with some preliminary observations. For each belief $p \in \Delta S$ and ergodic operator $P$, let $A^P(p) = \text{con}\{Pp, P^2 p, \ldots, \pi_P\}$. We show that for each $p \neq \pi_P$, $p \notin A^P(p)$. Suppose not. Then, $p$ is a convex combination of $P^n p$ for $n > 0$. Because convex combinations are preserved under $P$, this implies that each $P^n p$ is a convex combination of $P^m p$ for $m > n$, which, in turn, implies that $p$ is a convex combination of $P^n p$ for $n > 1$. By repeating the argument $m$ times, we see that $p$ is a convex combination of $P^n p$ for $n > m$ for any $m$. But $P^n p \to \pi_P$, which leads to a contradiction with $p \neq \pi_P$.

The above argument implies that $Pp \notin A^P(Pp) = \text{con}\{P^2 p, \ldots, \pi_P\}$. It follows that $Pp$ is an extreme point of the set $A^P(p)$.

Next, take any $p \neq \pi_P$ and $q \in A^P(p) \setminus \{Pp\}$. We show that $A^P(q) \subseteq A^P(p) \setminus \{Pp\}$. Indeed, suppose that $q = \sum_n \lambda_n P^n p$ for some $\lambda \in \Lambda$. Take $x \in A^P(q)$. Then, there exists $\lambda' \in \Lambda$ such that $x = \sum_k \lambda_k' P^k q = \sum_{k,n} \lambda_k' \lambda_n P^{k+n} p \in \text{con}\{P^2 p, \ldots, \pi_P\} \subseteq A^P(p) \setminus \{Pp\}$.

We move to the proof of part (b) of Corollary 1. Suppose towards a contradiction that $P \neq Q$. By part (a) of Corollary 1, there exists a belief $p_0 \neq \pi_P$ such that $Pp_0 \neq Qp_0$. We show by induction on $n$ that $Q^n p_0 \in A^P(p_0) \setminus \{Pp_0\}$. (Note that we can assume without loss of generality that $Pp_0 \neq \pi_P$; otherwise, $A^P(p_0) = \{\pi_P\}$ and the set on the right-hand side of the previous inclusion is empty.) Indeed, the claim for $n = 1$ is implied by condition (b) of Corollary 2. Moreover, if the claim holds for $n \geq 1$, then condition (b) of Corollary 2 and the above observation imply that $Q^{n+1} p = Q(Q^n p) \in A^P(Q^n p) \subseteq A^P(p) \setminus \{Pp\}$.

Because $Q^n p_0 \to \pi_Q = \pi_P$, it follows that if $P \preceq Q$, then

$$A^Q(p_0) \subseteq A^P(p_0) \setminus \{Pp_0\}.$$

The same argument implies that if $Q \preceq P$, then

$$A^P(p_0) \subseteq A^Q(p_0) \setminus \{Qp_0\}.$$

Thus, if $P \preceq Q$ and $Q \preceq P$, we get a contradiction: $A^P(p_0) \subsetneq A^Q(p_0) \subsetneq A^P(p_0)$.

## B.2. *Proof of Part (c)*

Take $\nu$ such that $P\nu \leq^B \nu$. If $P \preceq Q$ and $P \preceq Q'$, then by Theorem 2(c), $Q\nu \leq^B P\nu$ and $Q'\nu \leq^B P\nu$. By inequality (A.2), for any $\lambda \in (0, 1)$,

$$\left(\lambda Q + (1 - \lambda)Q'\right)\nu \leq^B \lambda Q\nu + (1 - \lambda)Q'\nu \leq^B P\nu,$$

where the second inequality follows because $\leq^B$ is preserved under convex combinations (see footnote 11). The claim now follows by Theorem 2(c).

## APPENDIX C: PROOF OF COROLLARY 5

### C.1. *Proof of Part (a)*

Constructing the game $g$ is easy. Say, let $A = B = \{0, 1\}$. Fix $s_0 \in S$ such that $\pi_P(s_0) \leq 1/2$ and define $u : \{0, 1\} \times S \to \mathbb{R}$ by

$$u(x, s) = \begin{cases} 0 & \text{if } x = 1, \\ 1 & \text{if } x = 0 \text{ and } s = s_0, \\ -2 & \text{if } x = 0 \text{ and } s \neq s_0. \end{cases}$$

Define $g : A \times B \times S \to \mathbb{R}$ by $g(a, b, s) = u(a, s) - u(b, s)$.

We note first that $v(\pi_P; g, I) = 0$. This follows, since for any $p \in \Delta S$,

$$\hat{g}(p) = \hat{g}(e_p) = \max_a \sum_s p(s)u(a, s) - \max_b \sum_s p(s)u(b, s) = 0,$$

which implies $v(\pi_P; g, I) = \text{cav}\,\hat{g}(\pi_P) = 0$ by Theorem 1.

We then argue that $v(g, P) > 0$. Note that given prior $p$, the minimizer's best-response is to play $b = 0$ if and only if $p(s_0) \geq 2/3 > \pi_P(s_0)$. Suppose the maximizer plays $a = 0$ if $s = s_0$ and $p(s_0) < 2/3$; otherwise he plays the same (non-revealing) action as the minimizer. Whenever both players play the same action, the payoff is zero by construction. But the maximizer earns 1 whenever the actions differ. It is easy to verify that this happens in a positive fraction of periods given the process for the minimizer's prior induced by the above strategy profile, because when the maximizer plays a non-revealing action, the belief drifts towards $\pi_P$ by ergodicity of $P$. Thus, the payoff is bounded away from zero for any $\delta > 0$, and hence $v(g, P) > 0$.

### C.2. *Proof of Part (b)*

Suppose $I \preceq_{\pi_P} P$. Because $\pi_P$ is an invariant distribution of both $P$ and $I$, condition (c) of Theorem 2′ implies that this is equivalent to condition (c′): for each $\nu \in \Delta^2 S$ such that $\mathcal{E}\nu = \pi_P$, we have $P\nu \leq^B \nu$. We show that (c′) is equivalent to condition ($\ell$): $P = \lambda I + (1 - \lambda)D_{\pi_P}$ for some $\lambda \in [0, 1]$.

To see that ($\ell$) implies (c′), fix $\nu$ with $\mathcal{E}\nu = \pi_P$ and note that $P\nu = \lambda\nu + (1 - \lambda)\pi_P$. Then $P\nu * m = \nu$ for $m$ such that $m(\pi_P) = \nu$ and $m(p) = e_p$ for $p \neq \pi_P$.

Suppose then that (c′) holds. We show first that for each $p \in \Delta S$, there exists $\lambda_p \in \mathbb{R}$ such that $Pp = \lambda_p p + (1 - \lambda_p)\pi_P$. Indeed, choose any belief $p \in \Delta S \setminus \{\pi_P\}$. Take $p' \in \Delta S \setminus \{\pi_P\}$ such that $\pi_P = \alpha p + (1 - \alpha)p'$ for some $\alpha \in (0, 1)$. Such $p'$ exists because $\pi_P$ has full support by ergodicity of $P$. Then $p$ and $p'$ lie on a line passing through $\pi_P$, and they lie on

the opposite sides of $\pi_P$. Abusing terminology, we say that such $p$ and $p'$ are *co-linear*. Let $\nu = \alpha e_p + (1 - \alpha)e_{p'}$. Then, $P\nu = \alpha e_{Pp} + (1 - \alpha)e_{Pp'}$. Because $\mathcal{E}\nu = \pi_P$, we have $P\nu \leq^B \nu$, and thus $Pp$ and $Pp'$ lie on the same line as $p$ and $p'$. But then there exists $\lambda_p \in \mathbb{R}$ such that $Pp = \lambda_p p + (1 - \lambda_p)\pi_P$.

We then show that $\lambda_p = \lambda_{p'} = \lambda$ for all $p$, $p' \in \Delta S \setminus \{\pi_P\}$ for some $\lambda \in \mathbb{R}$. The linearity of $P$ shows that $\lambda_p = \lambda_{p'}$ for all co-linear $p$ and $p'$. Suppose that $p$ and $p'$ are not co-linear. Take $q = \frac{1}{2}p + \frac{1}{2}p'$. Then $Pq = \frac{1}{2}Pp + \frac{1}{2}Pp'$, or equivalently,

$$\frac{1}{2}\left(\lambda_q p + (1 - \lambda_q)\pi_P\right) + \frac{1}{2}\left(\lambda_q p' + (1 - \lambda_q)\pi_P\right)$$
$$= \frac{1}{2}\left(\lambda_p p + (1 - \lambda_p)\pi_P\right) + \frac{1}{2}\left(\lambda_{p'} p' + (1 - \lambda_{p'})\pi_P\right).$$

It follows that

$$\frac{1}{2}(\lambda_q - \lambda_p)(p - \pi_P) = -\frac{1}{2}(\lambda_q - \lambda_{p'})\left(p' - \pi_P\right).$$

Because $p$ and $p'$ are not co-linear, it must be that $\lambda_p = \lambda_{p'} = \lambda_q$.

Finally, we show that $\lambda \in [0, 1]$. If $\lambda > 1$, then $Pe_s \notin \Delta S$ for all $s$. If $\lambda < 0$, take co-linear $p$ and $p'$ such that $\pi_P = \alpha p + (1 - \alpha)p'$ and $\lambda(1 - \alpha) < -\alpha$. Then,

$$Pp = \pi_P + \lambda(p - \pi_P) = \left(\alpha + \lambda(1 - \alpha)\right)p + \left(1 - \alpha - \lambda(1 - \alpha)\right)p' \notin \mathrm{con}\{p, p'\},$$

contradicting $P\nu \leq^B \nu$ for $\nu = \alpha e_p + (1 - \alpha)e_{p'}$. Hence, $\lambda \in [0, 1]$, establishing $(\ell)$.

## REFERENCES

ALIPRANTIS, C. D., AND K. BORDER (2007): *Infinite Dimensional Analysis: A Hitchhiker's Guide* (third Ed.). Berlin; Heidelberg; New York: Springer. [1934,1936]

ANDERSON, A., AND L. SMITH (2013): "Dynamic Deception," *American Economic Review*, 103 (7), 2811–2847. [1923]

ATHEY, S., AND K. BAGWELL (2008): "Collusion With Persistent Cost Shocks," *Econometrica*, 76 (3), 493–540. [1921]

AUMANN, R. J., AND M. MASCHLER (1995): *Repeated Games With Incomplete Information*. Cambridge, Mass: The MIT Press. [1921-1923,1925,1926]

BERGEMANN, D., AND S. MORRIS (2016): "Bayes Correlated Equilibrium and the Comparison of Information Structures in Games," *Theoretical Economics*, 11, 487–522. [1923]

BLACKWELL, D. (1953): "Equivalent Comparisons of Experiments," *The Annals of Mathematical Statistics*, 24 (2), 265–272. [1923,1925]

BRESSAUD, X., AND A. QUAS (2014): "Dynamical Analysis of a Repeated Game With Incomplete Information". arXiv:1403.1385. e-print. [1923]

DE MEYER, B. (2010): "Price Dynamics on a Stock Market With Asymmetric Information," *Games and Economic Behavior*, 69 (1), 42–71. [1925]

ELY, J. (2017): "Beeps," *American Economic Review*, 107 (1), 31–53. [1922]

ESCOBAR, J. F., AND J. TOIKKA (2013): "Efficiency in Games With Markovian Private Information," *Econometrica*, 81 (5), 1887–1934. [1921]

FAN, K. (1953): "Minimax Theorems," *Proceedings of the National Academy of Sciences*, 39, 42–47. [1935]

GENSBITTEL, F. (2015): "Extensions of the Cav(u) Theorem for Repeated Games With Incomplete Information on one Side," *Mathematics of Operations Research*, 40 (1), 80–104. [1925]

GOSSNER, O. (2000): "Comparison of Information Structures," *Games and Economic Behavior*, 30, 44–63. [1923]

GOSSNER, O., AND J.-F. MERTENS (2001): "The Value of Information in Zero-sum Games," Working paper, Université Paris-Nanterre and CORE, Université Catholique de Louvain. [1923]

HÖRNER, J., D. ROSENBERG, E. SOLAN, AND N. VIEILLE (2010): "On a Markov Game With one-Sided Information," *Operations Research*, 58 (4), 1107–1115. [1923,1930]

HÖRNER, J., S. TAKAHASHI, AND N. VIEILLE (2015): "Truthful Equilibria in Dynamic Bayesian Games," *Econometrica*, 83 (5), 1795–1848. [1921,1923]

KAMENICA, E., AND M. GENTZKOW (2011): "Bayesian Persuasion," *American Economic Review*, 101, 2590–2615. [1922]

MERTENS, J., S. SORIN, AND S. ZAMIR (2015): *Repeated Games*. Cambridge University Press. [1924,1934]

NEYMAN, A. (2008): "Existence of Optimal Strategies in Markov Games With Incomplete Information," *International Journal of Game Theory*, 37, 581–596. [1922]

PĘSKI, M. (2008): "Comparison of Information Structures in Zero-sum Games," *Games and Economic Behavior*, 62 (2), 732–735. [1923]

PĘSKI, M., AND J. TOIKKA (2017): "Supplement to 'Value of Persistent Information'," *Econometrica Supplemental Material*, 85, DOI:10.3982/ECTA14330. [1923]

RENAULT, J. (2006): "The Value of Markov Chain Games With Lack of Information on one Side," *Mathematics of Operations Research*, 31 (3), 490–512. [1922,1924,1934]

RENAULT, J., AND X. VENEL (2012): "A Distance for Probability Spaces, and Long-Term Values in Markov Decision Processes and Repeated Games Games". arXiv:1202.6259. e-print. [1923,1927]

——— (2017): "A Distance for Probability Spaces, and Long-Term Values in Markov Decision Processes and Repeated Games," *Mathematics of Operations Research*, 42 (2), 349–376. [1923,1927]

SHMAYA, E. (2006): "The Value of Information Structures in Zero-sum Games With Lack of Information on one Side," *International Journal of Game Theory*, 34 (2), 155–165. [1923]

SORIN, S. (2002): *A First Course on Zero-sum Repeated Games*. Springer. [1924]