Supplement to:

Distributional Uncertainty and Persuasion

Evan Piermont*

July 27, 2018

Abstract

Piermont (2018) investigates a Sender who tries to persuade a Receiver about many realizations of the state simultaneously. This supplement extends the analysis the static analysis to a model of long-run beliefs: signals arrive sequentially and the players maximize long-run average payoffs.

1 Long Run Persuasion

In many economic situations, information revelation is a slow process; the Sender might send signals in a sequential manner. This supplement shows that under suitable continuity assumptions, and players who care about long-run average payoffs, assuming a dynamic structure does not change the equilibrium analysis provided above.

Towards this, let $\Omega = X \times S$ denote the space of period-by-period realizations and set $\boldsymbol{X} = \prod_{t\geqslant 0} X$, $\boldsymbol{S} = \prod_{t\geqslant 0} S$, and $\boldsymbol{\Omega} = \prod_{t\geqslant 0} \Omega$ denote the corresponding sets of all infinite sequences of elements. Equip $\boldsymbol{\Omega}$ with the σ-algebra generated by all finite sequences of state/signal realizations, $\mathcal{F}^{\Omega}$. Moreover, for any $t \in \mathbb{N}$, let $\mathcal{F}^{\Omega}_t$ denote the sigma algebra generated by all sequences on length less than t, so that $\{\mathcal{F}^{\Omega}_t\}_{t\geqslant 0}$ is a filtration of $\mathcal{F}^{\Omega}$. In analogy, let $\big(\boldsymbol{X}, \mathcal{F}^{X}, \{\mathcal{F}^{X}_t\}_{t\geqslant 0}\big)$ and $\big(\boldsymbol{\mathcal{S}}, \mathcal{F}^{\mathcal{S}}, \{\mathcal{F}^{\mathcal{S}}_t\}_{t\geqslant 0}\big)$ denote the filtered spaces of infinite sequences of state realizations and signal realizations, respectively. For some product space $M \times M'$, and $E \subset M$, let $\mathrm{cyl}_{M\times M'}(E) = \mathrm{proj}^{-1}_M E$ denote the cylinder set generated by E (where the subscript delineating the ambient space will be suppressed when it is not confusing to do so). Let $\mathrm{sig} \equiv \mathrm{proj}_S$ denote the map that takes an event and returns the signal component. Notice, $\mathrm{sig}^{-1} : \mathcal{F}^S \to \mathcal{F}^{\Omega}$ as $E \mapsto \mathrm{cyl}_{\boldsymbol{\Omega}}(E)$.

For any $\boldsymbol{\omega} \in \boldsymbol{\Omega}$, denote the t^{th} realization by $\boldsymbol{\omega}^t = \mathrm{proj}_t \boldsymbol{\omega}$, and the first t realizations by $\boldsymbol{\omega}_t = (\boldsymbol{\omega}^1, \ldots, \boldsymbol{\omega}^t) = ((x_i, s_i))_{i\leqslant t}$. Define $\boldsymbol{s}$, $\boldsymbol{s}^t$, and $\boldsymbol{s}_t$ in analogy. Associate any finite sequence with the cylinder it creates in the corresponding product space. For example $\boldsymbol{s}_t \cong \{\boldsymbol{s} \in \boldsymbol{S} | \mathrm{proj}_{\mathcal{F}^S_t} \boldsymbol{s} = \boldsymbol{s}_t\}$.

In each period the state, $x \in X$, is drawn i.i.d. from some underlying, invariant distribution $\mu^\star \in \Delta(X)$. The decision maker (henceforth abbreviated DM, and a stand in for the Receiver or Sender) does not observe the realization of X but rather a signal of it; the DM observes the outcome of the chosen experiment $e^\star \in \mathcal{E}$. Thus, the periodic state is a pair (x, s), where x is drawn according to $\mu^\star$ and s is drawn according to $e^\star(\cdot|x)$. A *generator*, $(\mu, e) \in \Delta(X) \times \mathcal{E}$, determines the period-by-period realization of the state as well as the signal structure of the observed experiment. Let the true situation be denoted with stars: $(\mu^\star, e^\star)$.

*Contact Information: Department of Economics, Royal Holloway, University of London; Email: evan.piermont@rhul.ac.uk

It is natural to think of a pair (μ, e) as a distribution over Ω. Indeed, (μ, e) induces $\sigma^{(\mu,e)} \in \Delta(\Omega)$ defined by the following:

$$\sigma^{(\mu,e)}(x, s) = \mu(x)e(s|x). \tag{1.1}$$

This association will prove helpful in analyzing the DM's belief evolution conditional on the history, since elements of the history live in the same space.

Of course, the true generation process will in general not be known. The DM has a *second order* belief regarding the generators, a distribution in $\Delta(\Delta(X) \times \mathcal{E})$, which corresponds to a second order belief $\psi \in \Delta(\Delta(\Omega))$. As the DM observes a sequence signals (the outcome of the experiment that is carried out each period), she updates her belief in a Bayesian manner.

From the perspective of the Persuasion game, we are interested in the period-by-period evolution of the Receiver's beliefs regarding X, which he will use as the motivation for choosing his action. Given the exchangeable process, it is easy to understand the evolution these beliefs. Define $\mu_{\boldsymbol{s}_t,n}$, for any sequence of signals, $\boldsymbol{s}_t$ and time period $n \geqslant t$, as the DM's beliefs about the realization of X, in period n, given the observation of $\boldsymbol{s}_t$. This is defined in the obvious way

$$\mu_{\boldsymbol{s}_t,n}(x) = \zeta(x_n = x|\boldsymbol{s}_t) = \frac{\zeta((x_n = x) \cap \boldsymbol{s}_t)}{\zeta(\boldsymbol{s}_t)}. \tag{1.2}$$

Long-Run Beliefs. In the Persuasion game, where the equilibrium notion concerns long-run beliefs, we are interested how the Receiver's beliefs and behavior change in response to the observation of more and more information. In particular, when is the Receiver is able to properly forecast the future (when does $\mu_{\boldsymbol{s}_t,t} \to \mu^\star$)? And, when this convergence fails, what is are the limiting beliefs?

If at each t, the full history $\boldsymbol{\omega}_t$ is observed, the question of convergence was answered by Blackwell and Dubins (1962).

Remark 1.1 (Blackwell and Dubins)**.** *Let $\sigma^\star$ denote the joint distribution induced by true generator. Then if the product measure, $\boldsymbol{\sigma}^\star$, is absolutely continuous with respect to the DM's second order belief, ψ, then $\zeta(\cdot|\boldsymbol{\omega}_t)$ converges in norm to $\boldsymbol{\sigma}^\star(\cdot|\boldsymbol{\omega}_t)$ for $\boldsymbol{\sigma}^\star$-almost all $\boldsymbol{\omega} \in \boldsymbol{\Omega}$.*

In the language of Blackwell-Dubbins, $\zeta(\cdot|\boldsymbol{\omega}_t)$ *strongly merges* to $\boldsymbol{\sigma}^\star(\cdot|\boldsymbol{\omega}_t)$.[1] The following assumption is a simple way of ensuring that merging will occur.

Assumption 1. *The DM's ex-ante belief ψ has finite support and $\psi(\sigma^\star) > 0$.*

Of course, the exercise at hand considers a DM who does not observe the full realization each period, but only the signal. As such, the standard results above cannot be directly applied. What the DM can actually observe are sequences of signal realizations, $\boldsymbol{S}$. So, if the first t realizations are $\boldsymbol{\omega}_t = \{(x_i, s_i)\}_{i \leqslant t}$ she observes $\text{sig}(\boldsymbol{\omega}_t) = \{(s_i)\}_{i \leqslant t}$; importantly, she cannot distinguish between events that coincide on their signal components.

Since each period the state/signal pair is realized in an conditionally i.i.d. manner, it follows that, conditional on the true generation process, the marginal distribution over S is also going to be an i.i.d. process. Therefore, the DMs beliefs regarding sequences of signal realizations, *without observing the realized state*, will be invariant to permutations. This ensures that the techniques used to analyze exchangeable processes are still valid when inferences must be made on less than full information –when the DM does not observe the full evolution of the state. In particular:

Remark 1.2. *Under Assumption 1, the DM's beliefs regarding the distribution of signals will tend to the true distribution over S; i.e., $\text{marg}_S\zeta$ merges with $\text{marg}_S\boldsymbol{\sigma}^\star$.*

[1] In the literature on asymptotic learning, there is a distinction between learning to forecast the infinite horizon (strong merging) and learning to forecast the near future, for example, the next period's realization (weak merging). Kalai and Lehrer (1994) show that the later requires a strictly weaker condition than absolute continuity. Because the focus of this paper is on Receiver's period-by-period forecast of the payoff-state, it might seem constructive to consider the conditions for weak merging instead. However, because we are also working in an exchangeable model, where conditional on $\boldsymbol{\sigma}^\star$ realizations are i.i.d. across periods, the two notions of merging coincide.

Proof. In appendix B. ■

So, after observing sufficiently many signals the DM will learn the true distribution thereover. When the distribution over signals fully reveals the state, full learning takes place, and the DM learns the true distribution over X.

On the other hand, when the distribution of signals is not a sufficient statistic, learning is limited. The characterization of learning with state-dependent signal structures is therefore captured by the following two conditions: (1) if the distribution over signals fully characterizes the state/signal generation process then the DM must also learn the distribution over the payoff state, and, (2) if two states induce the same distribution over signals then these states cannot be separated. Towards making these observations formal, take the following definition:

Definition. *For any $\sigma \in \Delta(\Omega)$ let*

$$[\sigma] = \{\rho \in \Delta(\Omega) | \mathrm{marg}_S \rho = \mathrm{marg}_S \sigma\}.$$

*Call two generators **$\boldsymbol{s}$-equivalent**, denoted $(\mu, e) \overset{\boldsymbol{s}}{\sim} (\mu', e')$, whenever $\sigma^{(\mu',e')} \in [\sigma^{(\mu,e)}]$.*

Two generating processes are $\boldsymbol{s}$-equivalent if they induce the same periodic distribution over S. If the DM entertains the possibility of two distinct $\boldsymbol{s}$-equivalent states, then she will never be able to distinguish between them. The intuition is clear, if the DM can only observe signals, and the distribution of signals is identical across different generators, then the DM has no information that separates the generators from one another. The following result codifies this intuition, showing that the limit of asymptotic learning under limited observability is fully characterized by $\boldsymbol{s}$-equivalence.

Theorem 1.3. *Let the DMs ex-ante belief $\psi \in \Delta(\Delta(X \times S))$ satisfy Assumption 1. Then $\zeta(\cdot|\boldsymbol{s}_t)$ converges in norm to*

$$\int_{\Delta(\Omega)} \boldsymbol{\rho}(\cdot|\boldsymbol{s}_t) \, \mathrm{d}\psi(\rho|[\sigma^\star]) \tag{1.3}$$

for $\boldsymbol{\sigma}^\star$-almost all $\boldsymbol{s} \in \boldsymbol{S}$.

Proof. In appendix B. ■

As time progresses, the DM becomes very confident the empirical frequency of signals is close to the true distribution. Thus, she *learns* which element of $\Delta(\Omega)/\overset{\boldsymbol{s}}{\sim}$ is true, i.e., she believes with high probably the true state is contained in $[\sigma^\star]$. Of course, since any two $\sigma, \sigma' \in [\sigma^\star]$ generate the same distribution over S, the relative likelihood between σ and σ' will remain the same after any sequence of signals. This last point implies two things: first, the DM will never be able to separate the states in $[\sigma^\star]$, and second, the limiting distribution is exactly the ex-ante distribution conditional on the event $[\sigma^\star] \subset \Delta(\Omega)$.

A Auxiliary Results

Proposition A.1. *Let P be a measure space, with $\boldsymbol{P}$ the corresponding infinite product space. $\zeta \in \Delta(\boldsymbol{P})$ be exchangeable process with representation $\psi \in \Delta(\Delta(P))$. The for any $E, F \in \mathcal{F}^\infty$*

$$\zeta(F|E) = \int_{\Delta(P)} \boldsymbol{\sigma}(F|E) \, \mathrm{d}\psi(\sigma|E),$$

where

$$\psi(\sigma|E) = \frac{\psi(\sigma)\boldsymbol{\sigma}(E)}{\int_{\Delta(P)} \boldsymbol{\sigma}(E) \, \mathrm{d}\psi(\sigma)}.$$

Proof of Proposition A.1. Let E and F, be as given in the theorem. Then,

$$\begin{aligned}\zeta(F|E) &= \frac{\int_{\Delta(P)} \boldsymbol{\sigma}(F\cap E)\ \mathrm{d}\psi(\sigma)}{\int_{\Delta(P)} \boldsymbol{\sigma}(E)\ \mathrm{d}\psi(\sigma)} \\ &= \int_{\Delta(P)} \frac{\boldsymbol{\sigma}(F\cap E)}{\int_{\Delta(P)} \boldsymbol{\sigma}(E)\ \mathrm{d}\psi(\sigma)}\ \mathrm{d}\psi(\sigma) \\ &= \int_{\Delta(P)} \boldsymbol{\sigma}(F|E)\frac{\boldsymbol{\sigma}(E)}{\int_{\Delta(P)} \boldsymbol{\sigma}(E)\ \mathrm{d}\psi(\sigma)}\ \mathrm{d}\psi(\sigma) \\ &= \int_{\Delta(P)} \boldsymbol{\sigma}(F|E)\ \mathrm{d}\psi(\sigma|E).\end{aligned}$$

■

The map $(\mu, e) \mapsto \sigma^{(\mu,e)}$, defined by (1.1), is not invertible, and so, dealing only with $\Delta(\Omega)$ would collapse information regarding the DM's beliefs about the true type. For example, if for some $x \in X$, $\mu^\star(x) = 0$, then $e^\star(\cdot|x)$ cannot be identified from the joint distribution of state/signal pairs –since (x, s) is never observed for any s. However, under a full support assumption it is without loss of generality to consider only the joint distributions induced by (μ, e):

B Proofs

Proof of Remark 1.2. Let $\zeta \in \Delta(\boldsymbol{\Omega})$ be exchangeable with state $\psi \in \Delta(\Delta(\Omega))$. Then let $\zeta_S \in \Delta(\boldsymbol{S})$ be defined by

$$\zeta_S(\boldsymbol{s}_t) = \int_{\Delta(\Omega)} \boldsymbol{\sigma}(\mathrm{sig}^{-1}(\boldsymbol{s}_t))\ \mathrm{d}\psi(\sigma),$$

for all $t \geqslant 0$ and $\boldsymbol{s}_t \in S_t$. ζ_S is an exchangeable random variable. Indeed, let $E = \prod_{n\in\mathbb{N}} E_n \subseteq \boldsymbol{S}$. Then,

$$\zeta_S(E) = \zeta(\mathrm{sig}^{-1}(E)) \tag{B.1}$$

$$= \zeta(\mathrm{sig}^{-1}(\prod_{n\in\mathbb{N}} E_n)) \tag{B.2}$$

$$= \zeta(\prod_{n\in\mathbb{N}} \mathrm{sig}^{-1}(E_n)) \tag{B.3}$$

$$= \zeta(\prod_{n\in\mathbb{N}} \mathrm{sig}^{-1}(E_{\pi(n)})) \tag{B.4}$$

$$= \zeta(\mathrm{sig}^{-1}(\prod_{n\in\mathbb{N}} E_{\pi(n)})) = \zeta_S(\prod_{n\in\mathbb{N}} E_{\pi(n)}). \tag{B.5}$$

The equality of (B.2) and (B.3) and of (B.4) and (B.5) both stem from fact that the projection operator, sig, acts independently on each coordinate, and therefore commutes with the cartesian-product; the equality of (B.3) and (B.4) follows from exchangeability; all other equalities are definitional.

The result, therefore follows from Blackwell and Dubins (1962) applied to ζ_S and $\mathrm{marg}_S\boldsymbol{\sigma}^\star$. ■

Proof of Theorem 1.3. First, notice, for any measurable $E, F \subset [\sigma]$ such that $E \cap F = \varnothing$ and $\psi(E), \psi(F) \neq 0$, we have

$$\frac{\psi(E|\boldsymbol{s}_t)}{\psi(F|\boldsymbol{s}_t)} = \frac{\frac{\psi(E)\boldsymbol{\sigma}(E)}{\int_{\Delta(P)} \boldsymbol{\rho}(E)\ \mathrm{d}\psi(\rho)}}{\frac{\psi(F)\boldsymbol{\sigma}(F)}{\int_{\Delta(P)} \boldsymbol{\rho}(F)\ \mathrm{d}\psi(\rho)}} = \frac{\psi(E)}{\psi(F)}.$$

Hence, $\psi(\cdot|\boldsymbol{s}_t, [\sigma]) = \psi(\cdot|[\sigma])$. Now, consider ζ_S. By remark 1.2, $\mathrm{marg}_S\zeta_S$ converges to $\mathrm{marg}_S\boldsymbol{\sigma}^\star$.

Moreover, it is immediate that $\psi_S \in \Delta(\Delta(P))$ defined by
$$\psi_S(E) = \psi(\mathrm{marg}_S^{-1}(E)),$$
is well defined (by the measurability of the marginal operator) and represents ζ_S. Moreover, $\psi([\sigma]) = \psi_S(\mathrm{marg}_S\sigma)$. Appealing to Remark A.1, delivers: $\int_{\Delta(S)} \boldsymbol{\rho}(\cdot)\, \mathrm{d}\psi_S(\rho|\boldsymbol{s}_t)$ converges in norm to $\mathrm{marg}_S\boldsymbol{\sigma}^\star$ for almost all $\boldsymbol{s} \in \boldsymbol{S}$. Cleary, this implies, $\lim_{t\to\infty} \psi_S(\mathrm{marg}_S\boldsymbol{\sigma}^\star|\boldsymbol{s}_t) = 1$, for almost all $\boldsymbol{s} \in \boldsymbol{S}$. Now notice, for any $\boldsymbol{s} \in \boldsymbol{S}$ and $t \in \mathbb{N}$, we have
$$\psi([\sigma]|\boldsymbol{s}_t) = \frac{\psi([\sigma])\boldsymbol{\sigma}(\boldsymbol{s}_t)}{\int_{\Delta(\Omega)} \boldsymbol{\rho}(\boldsymbol{s}_t)\ \mathrm{d}\psi(\rho)} = \frac{\psi_S(\mathrm{marg}_S\sigma)\mathrm{marg}_{\boldsymbol{S}}\boldsymbol{\sigma}(\boldsymbol{s}_t)}{\int_{\Delta(S)} \mathrm{marg}_{\boldsymbol{S}}\boldsymbol{\rho}(\boldsymbol{s}_t)\ \mathrm{d}\psi_S(\mathrm{marg}_S\rho)} = \psi_S(\mathrm{marg}_S\sigma|\boldsymbol{s}_t),$$
and therefore, $\lim_{t\to\infty} \psi([\boldsymbol{\sigma}^\star]|\boldsymbol{s}_t) = 1$. As such, for any measurable $E \subseteq \Delta(\Omega)$,
$$\lim_{t\to\infty} \psi(E|\boldsymbol{s}_t) = \lim_{t\to\infty} \psi(E \cap [\sigma]|\boldsymbol{s}_t) = \lim_{t\to\infty} \frac{\psi(E \cap [\sigma]|\boldsymbol{s}_t)}{\psi([\sigma]|\boldsymbol{s}_t)} = \psi(E|\boldsymbol{s}_t, [\sigma]) = \psi(E|[\sigma]).$$
The result follows from another application of Remark A.1. ■

References

David Blackwell and Lester Dubins. Merging of opinions with increasing information. *The Annals of Mathematical Statistics*, 33(3):882–886, 1962.

Ehud Kalai and Ehud Lehrer. Weak and strong merging of opinions. *Journal of Mathematical Economics*, 23(1):73–86, 1994.

Evan Piermont. Distributional Uncertainty and Persuasion. 2018. *Working paper.*