Supplemental Material to

Distributional Uncertainty and Persuasion: An Application to Polarization and Speculative Trade*

Evan Piermont†

October 11, 2016

Abstract

In this note, I show the same informational conditions as in Piermont (2016), but in non-strategic environments, can create polarization of beliefs in the short-run, but leads to a consensus opinion in the long-run. Initial disagreement regarding the distribution of the state induces disagreement about the interpretation of new information. As signals accumulate and Bayesian decision makers learn the true distribution of signals, they also learn the true distribution regarding the state. I show that the optimal persuasion information structure can never lead to polarization.

* "Distributional Uncertainty and Persuasion" is available at http://www.pitt.edu/~ehp5/DUaP.pdf

† Contact Information: Email: ehp5@pitt.edu. Telephone: (412) 648 7039. University of Pittsburgh, Department of Economics.

1 Distributional Uncertainty as a Source of Polarization

The phenomenon where two decision makers (DMs), after observing the same evidence, update their beliefs so as to become *more* opposed to one another is referred to as *belief polarization*(Lord et al., 1979; Darley and Gross, 1983; Dixit and Weibull, 2007). Belief polarization is found in numerous economic environments. Voters of opposing parties both become more convinced their candidate will win after a debate (Sigelman and Sigelman, 1984). In the wake of the 2008 crisis, the initial failure of stimulus to stymie economic decline led proponents to demand more government intervention and opponents to demand less (Krugman, 2009).

Under the assumption that DMs update their beliefs in a Bayesian manner, it is well known that polarization is not possible whenever the observed evidence has a commonly known interpretation (Baliga et al., 2013). Thus, an obvious (and intuitive) explanation of polarization is that DMs disagree on the interpretation of signals.[1] This indicates that the statistical model, presented in Section 2, may provide a foundation for polarization.

In this section, I show this model can explain the divergence and polarization of beliefs, without relaxing the Baysian paradigm. Specifically, two decision makers, after seeing the same signals, rationally update their beliefs in opposite directions. This can occur even when the observed signals have agreed upon qualitative interpretations.[2] It is well known that Bayesianism precludes such a divergence in beliefs when the signal structure is commonly known. I show this impossibility result can be strengthened: polarization cannot occur whenever both DMs believe the signal structures are independent of the underlying distributions (even if they have different beliefs regarding the relative likelihood of different signal structures).

Although short-run polarization is possible, this model predicts long-run agreement among decision makers with probability 1. Moreover, $\boldsymbol{s}$-equivalence fully characterizes the knife-edge cases where divergence is persistent. Putting these two results together, the main prediction of this section is that *belief divergence is a temporary phenomenon*; while rational decision makers might initially disagree about the interpretation of information, they will eventually find consensus.

The intuition is thus. A signal is indicative of two things. First, it is informative about the realized state. For example, a media report claiming government abuse might increase the DM's belief that abuse has taken place that period. Because this mechanism affects the DM's belief regarding the current period realization, it is referred to as the *contemporaneous channel*. Second, a signal is informative about the signal distribution itself –about the type of government. For example, if the report claiming government abuse is much more likely when the media is free rather than when it is manipulated, seeing such a report may increase the DM's belief that the media is unobstructed. Because a change in the belief about the type of government affects the DM's belief regarding all future realizations, it is referred to as the *intertemporal channel*.

1.1 Corrupt Governments and Biased Reporting

Consider the following simple example. An international investor, deciding to invest in a particular county's economy, cares each period about the country's political stability –represented by

[1] There are other potential explanations: ambiguity attitude (Baliga et al., 2013; Zimper and Ludwig, 2009); a mixture of public and private signals (Kondor, 2012; Andreoni and Mylovanov, 2012); psychological biases **?**

[2] I.e., both DM's agree that for every interpretation of signals, s is more likely in state a than state b.

whether or not the government has done something unconstitutional;

$$X = \{x_{uphold} - \text{ constitution is upheld}, x_{abuse} - \text{constitution is abused}\}.$$

This abuse of power, however, is not directly observable; the investor sees only the media's report regarding the governments actions. Assume further, that the media can make three types of reports: (s_e) a report exculpates the government of any legal abuses, (s_g) a report on celebrity gossip, unrelated to the governments workings, or (s_s) a report there is a scandal involving an abuse of power.

The investor believes there are two possible *types* of government: honest and corrupt; with distributions over $\{x_u, x_a\}$, $\mu_H = [\frac{9}{10}, \frac{1}{10}]$ and $\mu_C = [\frac{1}{10}, \frac{9}{10}]$, respectively. Moreover, the media of the two governments are dictated by the signal structures, conditional distributions over $S = \{s_e, s_g, s_s\}$), e_C and e_H,

$$e_H(\cdot|x_u) = [\tfrac{2}{10}, \tfrac{6}{10}, \tfrac{2}{10}] \qquad e_C(\cdot|x_u) = [\tfrac{6}{10}, \tfrac{4}{10}, \tfrac{0}{10}]$$
$$e_H(\cdot|x_a) = [\tfrac{1}{10}, \tfrac{5}{10}, \tfrac{4}{10}] \qquad e_C(\cdot|x_a) = [\tfrac{4}{10}, \tfrac{5}{10}, \tfrac{1}{10}].$$

The joint distributions corresponding to these beliefs are as follows.

σ_H

	s_e	s_g	s_s
x_u	$\frac{18}{100}$	$\frac{54}{100}$	$\frac{18}{100}$
x_a	$\frac{1}{100}$	$\frac{5}{100}$	$\frac{4}{100}$

σ_H

	s_e	s_g	s_s
x_u	$\frac{6}{100}$	$\frac{4}{100}$	$\frac{0}{100}$
x_a	$\frac{36}{100}$	$\frac{45}{100}$	$\frac{9}{100}$

Consider the investor who ex-ante has the belief, $\theta(H) = \alpha$ and $\theta(C) = 1 - \alpha$, so her prior subjective distribution regarding the state is

$$\mu^{prior}(x_u) = \alpha\frac{9}{10} + (1-\alpha)\frac{1}{10},$$

and her updated belief after observing s_e, as given by (2.1),

$$\mu^{\alpha}_{s_e}(x_u) = \frac{\alpha 18 + (1-\alpha)6}{\alpha 19 + (1-\alpha)42}.$$

In particular, notice that for all α which are sufficiently central (i.e., approximately between .1 and .94), $\mu^{\alpha}_{s_e}(x_u) < \mu^{\alpha}_0(x_u)$. That is, after seeing signal s the DM's updated belief places *less* probability on the state being x_u. For example, if the DM initially believes there is an equal chance between corruption and honesty, then $\mu^{\alpha}_0(x_u) = \frac{1}{2}$. After seeing signal s_e, her updated belief is $\mu^{\alpha}_{s_e}(x_u) \approx \frac{39}{100}$. See Figure 1.

So two investors (i and j) who initially disagree on the likelihood that the government is corrupt, may exhibit polarizing beliefs after the observance of signal s_e. The media of an honest government very rarely reports s_e, and does so with considerable accuracy, whereas the corrupt government reports s_e rather liberally. If investor i is very confident the government is honest ($\alpha^i \approx 0$), she will interpret the signal s_e as a very strong signal of no constitutional violation, thus lowering her belief is an abuse. On the other hand, if j is uncertain about the type of government ($\alpha \approx \frac{1}{2}$), she might interpret s_e as a signal not of political stability but of media manipulation! As such, she increases her belief the government at hand is corrupt, and thus that a constitutional violation has taken place.

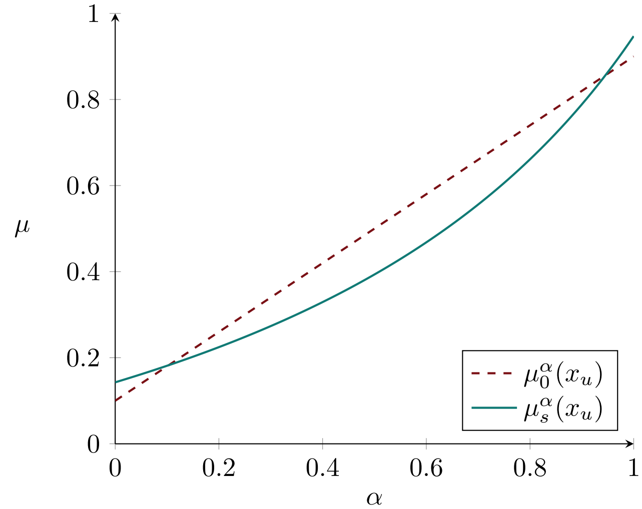

Figure 1: A plot of the Investor's prior, $\mu_0^\alpha(x_u)$, and posterior, $\mu_{s_e}^\alpha(x_u)$, after seeing signal s, plotted as a function of α. Near $\alpha = 0$ and $\alpha = 1$, beliefs are revised upwards, whereas less extreme beliefs are biased downwards.

Notice, if the type of government was known to be either type, s_e would increase the belief in political stability: the *qualitative* interpretation of the signal is known. Nonetheless, i and j's beliefs polarize. For investor i, who is very confident the government is honest, the intertemporal effect of a signal is negligible –a signal will not move her belief regarding the type of government. Thus her change in belief is entirely dominated by the contemporaneous effect, interpreting the signal as if it came from a known signal structure. On the other hand, because investor j is very uncertain about the governments type, a single signal might induce a non-trivial change in her belief about the type. Thus the intertemporal channel is much larger for j, and, since it has the opposite effect as the contemporaneous channel, can cause her beliefs regarding the likelihood of abuse to move in the opposite direction as well.

2 The Contemporaneous and Intertemporal Channels

Recall from Section 3 of Piermont (2016), the DM's beliefs is given by exchangeable process, from which we can define the period-by-period beliefs: $\mu_{\boldsymbol{s}_t,n}$, for any sequence of signals, $\boldsymbol{s}_t$ and time period $n \geqslant t$, is 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

$$\mu_{\boldsymbol{s}_t,n}(x) = \zeta(x_n = x|\boldsymbol{s}_t) = \frac{\zeta((x_n = x) \cap \boldsymbol{s}_t)}{\zeta(\boldsymbol{s}_t)}. \tag{2.1}$$

The observation of each signal has two effects on the Receivers actions. The first, an intertemporal effect, is the effect on the Receivers beliefs regarding the likelihood of different underly distributions in D. The second, contemporaneous effect, is the persuasion component; the effect on the Receivers belief regarding the likelihood of different states in X. This idea can be made formal.

Definition. *Let $\boldsymbol{s} \in \boldsymbol{S}$, and $\{\mu_{\boldsymbol{s}_t,n}\}_{t,n\in\mathbb{N}}$ be generated from some exchangeable model $\zeta \in \Delta(\Delta(\Omega))$. Then the* ***intertemporal effect*** *of observing $\boldsymbol{s}_{t+1}$ conditional on having observed $\boldsymbol{s}_t$ is the change* *in beliefs:* $\mu^{int}_{\boldsymbol{s}_{t+1}} = \mu_{\boldsymbol{s}_{t+1},t+2} - \mu_{\boldsymbol{s}_t,t+2}$. *The* ***contemporaneous effect*** *is the change in beliefs:* $\mu^{ctmp}_{\boldsymbol{s}_{t+1}} = \mu_{\boldsymbol{s}_{t+1},t+1} - \mu_{\boldsymbol{s}_{t+1},t+2}$.

To see how μ^{int} captures the change in second order beliefs, note, conditional on the true generator, the state/signal is drawn in an i.i.d. manner. This implies, of course, that the DM's belief about the state in period $t+2$ is conditionally independent of the t^{th} and the $(t+1)^{th}$ signal (where *conditionally* refers to the true generator). Hence, any change between $\mu_{\boldsymbol{s}_t,t+2}$ and $\mu_{\boldsymbol{s}_{t+1},t+2}$ must have stemmed from a change in second order belief. Likewise, $\mu_{\boldsymbol{s}_{t+1},t+1}$ and $\mu_{\boldsymbol{s}_{t+1},t+2}$ are predicated on the same sequence of signals, and so, any difference in belief arises from the correlation between the $(t+1)^{th}$ state and the $(t+1)^{th}$ signal. Hence, μ^{ctm} captures the remaining difference in belief after any change second order order beliefs has been taken into account. Of course, for these definitions to make sense, it must be that they collectively describe the full updating process.

Remark 2.1. *Let* $\boldsymbol{s} \in \boldsymbol{S}$*, and* $\{\mu_{\boldsymbol{s}_t,n}\}_{t,n\in\mathbb{N}}$ *be generated from some exchangeable model* $\zeta \in \Delta(\Delta(\Omega))$*. Then,*

$$\mu_{\boldsymbol{s}_{t+1},t+1} - \mu_{\boldsymbol{s}_t,t+1} = \mu^{int}_{\boldsymbol{s}_{t+1}} + \mu^{ctmp}_{\boldsymbol{s}_{t+1}}.$$

Remark 2.1 reduces to an identity under the observation that $\mu_{\boldsymbol{s}_t,t+1} = \mu_{\boldsymbol{s}_t,t+2}$, itself an obvious consequence of exchangeability. We can also decompose the evolution of belief into these channels by examining the direct updating effect the on second order belief ψ^θ, and the distributions in its support, $\boldsymbol{\sigma} \in \operatorname{supp}(\psi^\theta)$. Towards this:

Definition. *Let* $\psi \in \Delta(\Delta(\Omega))$*, then for any* $\boldsymbol{s} \in \mathcal{S}$ *and* $t \in \mathbb{N}$*, define* $\psi(\cdot|\boldsymbol{s}_t) \in \Delta(\Delta(\Omega))$ *according to*

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

Then the effect of an additional signal, $\boldsymbol{s}_t \to \boldsymbol{s}_{t+1}$, which changes beliefs (over all events) according to $\zeta^\theta(\cdot|\boldsymbol{s}_t) \to \zeta^\theta(\cdot|\boldsymbol{s}_{t+1})$, can be decomposed into the following two transitions:

$$\begin{aligned}
\zeta^\theta(\cdot|\boldsymbol{s}_t) &= \int_{\Delta(\Omega)} \boldsymbol{\sigma}(\cdot|\boldsymbol{s}_t)\ \mathrm{d}\psi^\theta(\sigma|\boldsymbol{s}_t) \\
&\to \int_{\Delta(\Omega)} \boldsymbol{\sigma}(\cdot|\boldsymbol{s}_t)\ \mathrm{d}\psi^\theta(\sigma|\boldsymbol{s}_{t+1}) \\
&\to \int_{\Delta(\Omega)} \boldsymbol{\sigma}(\cdot|\boldsymbol{s}_{t+1})\ \mathrm{d}\psi^\theta(\sigma|\boldsymbol{s}_{t+1}) = \zeta^\theta(\cdot|\boldsymbol{s}_{t+1}),
\end{aligned}$$

where the first arrow represents the change to the second order belief, and the second arrow change to the periodic distribution. At promised, these two transitions are exactly the intertemporal and contemporaneous channels.

Proposition 2.2. *Let* $\boldsymbol{s} \in \boldsymbol{S}$*, and* $\{\mu_{\boldsymbol{s}_t,n}\}_{t,n\in\mathbb{N}}$ *be generated from some exchangeable model* $\zeta \in \Delta(\Delta(\Omega))$*. Then*

$$\begin{aligned}
\mu^{ctmp}_{\boldsymbol{s}_{t+1}} &= \operatorname{marg}_{X_{t+1}}\left(\zeta^\theta(\cdot|\boldsymbol{s}_{t+1}) - \int_{\Delta(\Omega)} \boldsymbol{\sigma}(\cdot|\boldsymbol{s}_t)\ \mathrm{d}\psi^\theta(\sigma|\boldsymbol{s}_{t+1})\right), \text{ and,} \\
\mu^{int}_{\boldsymbol{s}_{t+1}} &= \operatorname{marg}_{X_{t+1}}\left(\int_{\Delta(\Omega)} \boldsymbol{\sigma}(\cdot|\boldsymbol{s}_t)\ \mathrm{d}\psi^\theta(\sigma|\boldsymbol{s}_{t+1}) - \zeta^\theta(\cdot|\boldsymbol{s}_t)\right).
\end{aligned}$$

Proof. In appendix A. ■

2.1 When Can Polarization Occur?

Because the general literature on polarization has been focused on the case of two states, and because this greatly simples the notation required to state results, throughout this section I make the simplifying assumption that X contains 2 elements (as in threshold environments).

Definition. *Let θ_i and θ_j denote the second order beliefs of two DMs. Then we say that the beliefs regarding states **polarize** given $\boldsymbol{s}_t$ if*

$$\mu^j_{\boldsymbol{s}_{t+1},t+1}(x) < \mu^j_{\boldsymbol{s}_t,t+1}(x) < \mu^i_{\boldsymbol{s}_t,t+1}(x) < \mu^i_{\boldsymbol{s}_{t+1},t+1}(x).$$

for some $x \in X$.

Definition. *Let θ_i denote a second order belief of DM i. Then, the **qualitative interpretation of signals is known** if for all $(\mu, e), (\mu', e') \in \operatorname{supp}(\theta_i)$, we have $\frac{e(s|x)}{e(s|y)} > 1 \iff \frac{e'(s|x)}{e'(s|y)} > 1$, for $\{x, y\} = X$ and $s \in S$. Moreover, if $\operatorname{supp}(\theta_j) = \operatorname{supp}(\theta_i)$, then we say the qualitative interpretation of signals is **commonly** known.*

The above example shows commonly knowledge of the qualitative interpretation of signals is not sufficient to bar polarization. This is because the restriction that signals have known qualitative interpretations concerns only the contemporaneous channel—this is most easily seen in light of Proposition 2.2. Still, the following is true.

Proposition 2.3. *Let θ_i be such that the qualitative interpretation of signals is known. Then, if $\frac{e(s|x)}{e(s|y)} > 1$, for some $(\mu, e) \in \operatorname{supp}(\theta_i)$, then $\mu^{i,ctmp}_{\boldsymbol{s}_{t+1}}(x) > 0$. In particular, if the qualitative interpretation of signals is commonly known, $\mu^{i,ctmp}_{\boldsymbol{s}_{t+1}}(x) > 0 \iff \mu^{j,ctmp}_{\boldsymbol{s}_{t+1}}(x) > 0$.*

Proof. In appendix A. ■

Proposition 2.3 states that whenever the qualitative interpretation of signals are known, then direction of the contemporaneous channel does not depend on the relative likelihood of different signal structures. So, polarization can only happen when the contemporaneous and intertemporal channels are working in opposite directions.

In the example, s_e, was more likely to occur when the government is corrupt. As such, observation of s_e, qualitatively known to be indicative of x_u, nonetheless shifts the *second order* beliefs to a distribution more likely to realize x_a. The contemporaneous and intertemporal channels have conflicting effects on the direction of updating; which effect dominates depends on the DM's prior belief.

Of course, if the channels work in the same direction, polarization cannot occur. The condition that guarantees this is an adaptation of the *monotone likelihood ratio property* (MLRP) to the environment at hand. A signal, s, will unambiguously increase the the belief that state x is realized if s is more likely to occur whenever the underlying distribution over X is more biased towards x. In other words, the signal s and the completely uninformative signal jointly satisfy the MLRP. Formally,

Proposition 2.4. *For any $\mu \in \Delta(X)$ let $[\mu]$ denote the equivalence class of all (μ, e) with $e \in \mathcal{E}$. Then, if $\int_{[\mu]} \sigma(\boldsymbol{s}^{t+1})\, \mathrm{d}\psi^\theta(\sigma|\boldsymbol{s}_t) > \int_{[\mu']} \sigma(\boldsymbol{s}^{t+1})\, \mathrm{d}\psi^\theta(\sigma|\boldsymbol{s}_t)$ whenever $\mu(x) > \mu'(x)$, then $\mu^{int}_{\boldsymbol{s}_{t+1}}(x) > 0$.*

Proof. In appendix A. ■

These restrictions should not come as any surprise, since they correspond to the natural association of s with state x in the contemporaneous and intertemporal domains, respectively. That the qualitative interpretation is known is to say s is more likely when the state is x *today*; that the adapted MLRP condition holds is to say s is more likely when the state is more likely to be x *on average*. These two propositions are important, however, because they apply in a variety of common and relevant specifications.

For example, assume that two decision makers both believe that the realized experiment is independent of the realized distribution over X and that the qualitative interpretation of signals is commonly known. Then, no matter how much the two DMs disagree on the likelihood of different types, their beliefs will not polarize. Again, this is a corollary of Proposition 2.4.

Corollary 2.5. *Let θ_i and θ_j be as in Proposition 2.3, and in addition assume that $\theta_i(e|\mu) = \theta_i(e)$ and $\theta_j(e|\mu) = \theta_j(e)$ for all $\mu \in \Delta(X)$ and $e \in \mathcal{E}$. Then polarization cannot occur.*

Proof. In appendix A. ■

2.2 Generic Signals

Of course, when $[\sigma^\star] \cap \text{supp}(\varphi^\theta) = \sigma^\star$, then the DM will learn the true fundamental distribution. So, when the DM believes that each possible type induce a unique distribution of signals, then she will learn the true distribution over *the state space*.

Assumption 3. *If $(\mu, e), (\mu', e') \in \text{supp}(\theta)$ then* $\text{marg}_S \sigma^{(\mu,e)} \neq \text{marg}_S \sigma^{(\mu',e')}$.

Corollary 2.6. *Let θ satisfy Assumptions 1 and 3, then then $\zeta(\cdot|\boldsymbol{s}_t)$ converges in norm to $\boldsymbol{\sigma}^\star(\cdot|\boldsymbol{s}_t)$, for $\boldsymbol{\sigma}^\star$-almost all $\boldsymbol{s} \in \boldsymbol{S}$.*

The question remains, however, of whether or not Assumptions 3 is reasonable. I now argue that it is, but only *when signal structures are chosen generically*. In other words, when there is not an a priori reason to believe the signals across different types will align themselves (which might happen, for example, in strategic environments), then Assumptions 3 will hold (with high probability).

To make this precise, assume that there $N \in \mathbb{N}$ possible *types* of government $\{\tau_i\}_{i=1}^N$. Each type is associated with particular level of corruption and media reporting style, and so, corresponds to a type: σ^τ. Assume this is commonly known, so that $\text{supp}(\varphi^\theta) = \{\sigma^{\tau_i}\}_{i=1}^N$. The DM's beliefs about a particular government, therefore, embodies her subjective assessment of likelihood of each of these N possible types.

Now, further assume the N types arose from an unbiased process –so that for each τ the corresponding type was determined in a uniform manner. While this may not hold when the type is the outcome of an endogenous process –if, for example, the government choose the media's bias strategically– it seems reasonable when the signal structure is the result of environmental constraints (technology, demographics, legal restrictions, etc.) that are not chosen strategically.

Theorem 2.7. *Assume that $\{\sigma^{\tau_i}\}_{i=1}^N$ is drawn uniformly from $\prod_{i=1}^N \Delta(\Omega)$. Then the probability that for any $i \neq j$, $\sigma^{\tau_i} \overset{\boldsymbol{s}}{\sim} \sigma^{\tau_j}$ is 0.*

Proof. In appendix A. ■

It is important to note, Theorem 2.7 regards the mechanism by which type are themselves generated, and *not* the beliefs of the DM. What Theorem 2.7 does imply is that in any environment where the set of possible types is known, and this set was generated in a uniform way, Assumption 3 will be met with probability 1. This is important as it indicates that, in generic cases, the DM will fully learn the type, and therefore properly forecast future realizations, even under limited observability.

3 An Application to Speculative Trade.

When investors hold the same beliefs, both about the payoff relevant variables and the interpretation of new evidence there regarding, new information should not incentivize trade. When the new information is public, it should change the value of the asset in same way for buyers and sellers alike. When the information is private, the knowledge that the other party *wants* to trade on the basis of this new information is enough to inform the investor the proposition is a losing one. Nonetheless, *speculative trade* abounds (Harris and Raviv, 1993). Perhaps the most obvious explanation, and one which is empirically supported, is that different investor's beliefs diverge in response to the same evidence (Kandel and Pearson, 1995). Such behavior serves as a clear explanation for speculative trade: for example, after Apple posted a disappointing earnings report, pessimistic analysts revised there estimates downward, as expected, while optimistic analysts raised their exceptions, going so far as to issue new buy recommendations (Fisher, 1993).

This section contains a very simple application to speculative trading. The environment, while simple, conveys the following basic intuition. Disagreement about the interpretation of signals creates an incentive to engage in speculative trading, but in the absence of other informational issues, this incentive will diminish over time. Under regularity conditions (in particular, genericity of signals), the fact that trade is proposed is informationally sufficient to guarantee learning about the signal structure, and therefore, the attenuation of any incentive to engage in speculative trading.

3.1 Set Up

Each period two decision makers, a buyer, i, and a seller, j, are randomly paired. The seller, j, owns an asset which has a commonly known payoff structure that depends on on the state in $X = \{x_a, x_b\}$. The value of the asset is given by $a(x_a) = 1$ and $a(x_b) = 0$. Both the buyer and seller are risk/uncertainty neutral. The seller can propose to sell the asset at any price p, and if the buyer accepts, the asset is traded for the price less some outside transaction cost c levied (without loss of generality) on both parties. That is, if trade occurs, payoffs in state x are given by $\pi_i(x, trade) = a(x) - p - c$ for the the buyer and $\pi_j(x, trade) = p - c$ for the seller. If no trade occurs, payoffs are $\pi_i(x, no\ trade) = 0$ for the buyer and $\pi_j(x, no\ trade) = a(x)$ for the seller.

We will assume the distribution over X is *commonly known.* That is the realization of the state occurs with known probabilities. However, before each period, the seller observes a private signals, s, drawn according to some experiment from $S = \{s_a, s_b\}$. The identity of this experiment will not necessarily be known nor the particular uncertainty there regarding agreed upon. If there is no disagreement about the signal structure, then no trade theorems apply, and trade will never occur (Milgrom and Stokey, 1982). Of course, when there exists (initial) disagreement about the likelihood of different signal structures, the common prior assumption is not satisfied (i.e., the prior over $X \times S$), and speculative trade can reign freely.

Example 1. *It is common knowledge that each period the economy is in state x_a with probability $\frac{2}{3}$ and state x_b with probability $\frac{1}{3}$. Let $\mu = [\frac{2}{3}, \frac{1}{3}]$, denote this distribution. There exists a single asset, a, owned by decision maker j. The asset is such that $a(x_a) = 1$ and $a(x_b) = 0$. Let c denote the external cost of making a trade. Without seeing any signal, both i and j perceive the value of a to be $\frac{2}{3}$.*

Each period, an economic reporting agency can make two possible announcements $S = \{s_a, s_b\}$. It is commonly known that s_a is an informative signal of x_a and s_b of x_b but the accuracy of the reporting agency is unknown. In particular, there are two possibilities, high accuracy: e_h defined by $e_h(s_a|x_a) = e_h(s_b|x_b) = \frac{8}{10}$; and, low accuracy: e_l defined by $e_l(s_a|x_a) = e_l(s_b|x_b) = \frac{7}{10}$. Thus, there are two possible types of the economy: $h = (\mu, e_h)$ and $l = (\mu, e_l)$. The two decision makers disagree on their ex-ante assessment of the accuracy of the ratings agency.

Therefore, given an initial belief α that the economy is h, the updated belief that the state is x_a, after seeing signal s_a and s_b are given by

$$\mu^{\alpha}_{s_a} = \frac{\alpha 16 + (1-\alpha)14}{\alpha 18 + (1-\alpha)17} \qquad \mu^{\alpha}_{s_b} = \frac{\alpha 4 + (1-\alpha)6}{\alpha 12 + (1-\alpha)13}.$$

Notice that the former is strictly increasing in α while the latter is strictly decreasing.

The two decision makers disagree on their ex-ante assessment of the accuracy of the ratings agency. Let $\alpha_i(\boldsymbol{s}_t)$ denote the probability i assigns to economy h conditional on the sequence of signals $\boldsymbol{s}_t$, and let $\alpha_j(\boldsymbol{s}_t)$ be defined analogously. Further assume that $\alpha_i(\varnothing) < \alpha(\varnothing)$.

It is straightforward to check, when c is sufficiently small (i.e., when $\mu^{\alpha_j(\varnothing)}_{s_b} + c < \mu^{\alpha_i(\varnothing)}_{s_b} - c$, then for any $p(\varnothing) \in (\mu^{\alpha_j(\varnothing)}_{s_b} + c, \mu^{\alpha_i(\varnothing)}_{s_b} - c)$ the following strategies constitute an equilibrium for the first period. After observing s_a the seller does not propose trade, and retains the asset. After observing s_b the seller proposes a trade for $p(\varnothing)$. The seller will accept $p(\varnothing)$ and no other offer.

In the equilibrium in Example 1, which signal was observed becomes common knowledge. Because the buyer will only accept trade at a price of $p(\varnothing) < \frac{2}{3}$, the seller will propose trade only after seeing the bad signal. Therefore, in such an equilibrium, the buyer knows the signal that the seller observed. Although this echoes the intuition for the impossibility of trade in standard environments, here, common knowledge of the signal realization does not impede costly trade, since the buyer and seller interpret the single idiosyncratically. In fact, it is common knowledge that their posteriors are different. As witnessed in the example, polarization is not necessary for speculative trading to occur –only divergence. Trade is a possible equilibrium outcome because after the observation of s_b, there is a wedge between the buyer's and the seller's beliefs (and the cost of trade is sufficiently small).

Example 1 (continued)**.** *Consider the setup introduced in Example 1. Assume that the true economy is type h. Let $\boldsymbol{s}$ denote the infinite sequence of economic reports that can be assessed by sellers each period. Because, in equilibrium, it is common knowledge which signals were observed, both the buyer's and seller's beliefs will evolve as if signals were public. Then, the classic results on merging imply that, with probability 1 over the space of signal sequences,* $\lim_{t\to\infty} \alpha_j(\boldsymbol{s}_t) = \lim_{t\to\infty} \alpha_j(\boldsymbol{s}_t) = 1$.

Thus, with probability 1, both $\alpha_j(\boldsymbol{s}_t)$ and $\alpha_j(\boldsymbol{s}_t)$ will eventually be, and remain, larger than $\frac{39x-10}{3c-10}$. By construction, this implies with some algebraic manipulation, that $\mu_{s_b}^{\alpha_i(\boldsymbol{s}_t)}$ and $\mu_{s_b}^{\alpha_j(\boldsymbol{s}_t)}$ will be, and remain, within $2c$ of $\frac{1}{3} = \lim_{\alpha\to 1}\mu_{s_b}^{\alpha}$. But, within the region, trade cannot occur.

As the buyer and seller learn the true parameterization of the signal structure, the amount by which their beliefs diverge after the observation of a signal will tend to zero. As evidenced by Theorem 3.4 of Piermont (2016), asymptotically, the buyer and seller will find consensus. Furthermore, the difference in the decision makers' valuation of the asset is a continuous function of their posteriors, and so, as these converge, eventually, the difference in subjective value will be less than the cost of making a trade. This intuition generalizes.

Theorem 3.1. *Fix some X, S, $a : X \to \mathbb{R}$, and $c > 0$. Then if all buyers and sellers satisfy Assumption 1, trade will occur in at most a finite number of periods.*

Proof. In appendix A. ■

The knowledge that other parties want to engage in trade reveals some aspect of private information. Celebrated results show, in the canonical case, where signals have common interpretations, that this is sufficient to ensure all parties know trade cannot be mutually beneficial, impeding trade. Here, when the interpretation of signals is held in contention, trade is still possible. However, some of the private information is still revealed by the knowledge that other parties want to trade. When this knowledge is sufficient to *eventually* ensure parties learn the parameters governing the market, speculative trade cannot continue forever.

A Proofs

Proof of Proposition 2.2. By the definition of the contemporaneous channel and the conditional beliefs, we have $\mu^{ctmp}_{\boldsymbol{s}_{t+1}}(x) = \mu_{\boldsymbol{s}_{t+1},t+1}(x) - \mu_{\boldsymbol{s}_{t+1},t+2}(x) = \zeta^\theta(x_{t+1} = x|\boldsymbol{s}_{t+1}) - \zeta^\theta(x_{t+2} = x|\boldsymbol{s}_{t+1})$. From Remark C.1 of Piermont (2016) we know this last expression is equivalent to

$$\int_{\Delta(\Omega)} \boldsymbol{\sigma}(x_{t+1} = x|\boldsymbol{s}_{t+1})\ \mathrm{d}\psi^\theta(\sigma|\boldsymbol{s}_{t+1}) - \int_{\Delta(\Omega)} \boldsymbol{\sigma}(x_{t+2} = x|\boldsymbol{s}_{t+1})\ \mathrm{d}\psi^\theta(\sigma|\boldsymbol{s}_{t+1}). \tag{A.1}$$

By the product structure of $\boldsymbol{\sigma}$, $\boldsymbol{\sigma}(x_{t+2} = x|\boldsymbol{s}_{t+1})$ can be replaced with $\boldsymbol{\sigma}(x_{t+2} = x|\boldsymbol{s}_{t+1})$ is equal to $\boldsymbol{\sigma}(x_{t+1} = x|\boldsymbol{s}_t)$ and (A.1) reduces to the desired expression. Now, it is a consequence of remark 2.1 and the previous claim that both $\mu^{int}_{\boldsymbol{s}_{t+1}}(x)$ and $\mathrm{marg}_{X_{t+1}}\left(\int_{\Delta(\Omega)} \boldsymbol{\sigma}(\cdot|\boldsymbol{s}_t)\ \mathrm{d}\psi^\theta(\sigma|\boldsymbol{s}_{t+1}) - \zeta^\theta(\cdot|\boldsymbol{s}_t)\right)$ are fully characterized as the difference between the aggregate change in beliefs and the contemporaneous effect, and therefore coincide. ■

Proof of Proposition 2.3. Denote $\boldsymbol{s}_{t+1}$ by s. Assume $e'(s|x) > e'(s|y)$ for some $(\mu', e') \in \mathrm{supp}(\theta_i)$. Thus, by the hypothesis of the theorem, $e(s|x) > e(s|y)$ for all $(\mu, e) \in \mathrm{supp}(\theta_i) = \mathrm{supp}(\theta_j)$. Hence,

$$\sigma(x|s) = \frac{\mu(x)e(s|x)}{\mu(x)e(s|x) + \mu(y)e(s|y)} > \mu(x) = \sigma(x), \tag{A.2}$$

and therefore, $\boldsymbol{\sigma}(x_{t+1} = x|\boldsymbol{s}_{t+1}) - \boldsymbol{\sigma}(x_{t+1} = x|\boldsymbol{s}_t) > 0$, for all $\sigma \in \mathrm{supp}(\psi^{\theta^i})$. Therefore,

$$\int_{\Delta(\Omega)} [\boldsymbol{\sigma}(x_{t+1} = x|\boldsymbol{s}_{t+1}) - \boldsymbol{\sigma}(x_{t+1} = x|\boldsymbol{s}_t)]\ \mathrm{d}\psi^{\theta^i}(\sigma|\boldsymbol{s}_{t+1}) > 0. \tag{A.3}$$

Appealing to Proposition 2.2 delivers the result for a single DM. Of course, since (A.2) holds for DM j be virtue of the fact that $\text{supp}(\theta_j) = \text{supp}(\theta_i)$, (A.3) holds as well; the second result follows in turn. ■

Proof of Proposition 2.4. First, consider the augmented environment in which an addition signal, denoted s', is completely uninformative for all $e \in \mathcal{E}$, but the relative likelihood of all other signals remains unchanged. Then, we can describe the intertemporal effect as

$$\mu^{int}_{\boldsymbol{s}_{t+1}} = \int_{\Delta(X)} \mu(x)\ \mathrm{d}\psi^\theta([\mu]|\boldsymbol{s}_t, s) - \int_{\Delta(X)} \mu(x)\ \mathrm{d}\psi^\theta([\mu]|\boldsymbol{s}_t, s'). \tag{A.4}$$

Under the interpretation of $\mu(x)$ as real valued random variable, and the signals s and s' as signals regarding this random variable –the hypothesis of the lemma implies the signals s and s' satisfy the strict monotone likelihood ratio property. By Proposition 1 of Milgrom (1981), $\psi^\theta([\mu]|\boldsymbol{s}_t, s')$ first order stochastically dominates $\psi^\theta([\mu]|\boldsymbol{s}_t, s)$. It follows from the definition of first order stochastic dominance that (A.4) is positive. ■

Proof of Corollary 2.5. Let s be a signal with $\frac{e(s|x)}{e(s|y)} > 1$ for all e. Further let $\pi(e)$ denote the probability of signal structure e according to the μ-independent beliefs over $\mathcal{E}$. Now, define $e^*(s|x) = \int_{\mathcal{E}} e(s|x)d\pi(e)$; $e^*(s|x)$ is the probability of seeing signal s when the state is x, which is clearly independent on the underlying distribution on the state space. Define $e^*(s|y)$ analogously; since $e(s|x)$ point wise dominates $e(s|y)$, $e^*(s|x) > e^*(s|y)$. So, the probability of seeing s unconditionally is $\mu(x)e^*(s|x) + (1 - \mu(x))e^*(s|y)$, which is obviously monotone in $\mu(x)$. ■

Proof of Theorem 2.7. Let $|X| = n_X$ and $|S| = n_S$, so that $|\Omega| = n = n_X n_S$. Identify $\Delta(\Omega)$ with the n-dimensional simplex, so that $\Delta(\Omega) \times \Delta(\Omega)$ is associated with a full-dimensional subset, R, of $\mathbb{R}^{2(n-1)}$. Let a point in $r \in R$ be described as (r^1, r^2) where each $r^i = (r^i_{1,1} \dots r^i_{|X|,|S|})$. Let

$$\bar{R} = \left\{ (r \in R| \sum_{x \leqslant |X|} r^1_{x,s} = \sum_{k \leqslant |X|} r^2_{x,s} \text{ for all } s \leqslant |S| \right\}. \tag{A.5}$$

Then $\sigma^{\tau_1} \overset{\boldsymbol{s}}{\sim} \sigma^{\tau_2}$ if any only if $r \in \bar{R}$. But $\bar{R}$ clearly has dimension $\mathbb{R}^{2(n-1)-n_S}$, and therefore has $2(n-1)$-Lebesgue measure 0. ■

Proof of Theorem 3.1. The result follows directly from the fact that, whenever trade is possible, then stage game outcome perfectly reveals the signal. Assume that trade is possible in period k (that is there is some signal that induces trade in equilibrium). Since this (statically) improves the seller's payoff, and since rematching ensures there is effect on continuation value, trade *will* occur. Further, since both players beliefs are martingales, it cannot be that trade occurs after both signals. Moreover, because μ is commonly known, no two signal structures can $\boldsymbol{s}$-equivalent. Therefore, if trade occurs in an infinite number of periods, by Theorem 3.4 of Piermont (2016), so does both players posteriors converge to the true experiment. By the continuity of evaluations with respect to beliefs, this implies eventually the valuation of the asset after every possible signal will remain within $2c$ of one another. But, within this region, trade cannot occur; this is a contradiction to the assumption trade occurs for an infinite number of periods. ■

References

James Andreoni and Tymofiy Mylovanov. Diverging opinions. *American Economic Journal: Microeconomics*, 4(1):209–232, 2012.

Sandeep Baliga, Eran Hanany, and Peter Klibanoff. Polarization and ambiguity. *The American Economic Review*, 103(7):3071–3083, 2013.

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

John M Darley and Paget H Gross. A hypothesis-confirming bias in labeling effects. *Journal of Personality and Social Psychology*, 44(1):20, 1983.

Bruno de Finetti. Funzione caratteristica di un fenomeno aleatorio. 1931.

Bruno de Finetti. La prévision: ses lois logiques, ses sources subjectives. In *Annales de l'institut Henri Poincaré*, volume 7, pages 1–68, 1937.

Avinash K Dixit and Jörgen W Weibull. Political polarization. *Proceedings of the National Academy of Sciences*, 104(18):7351–7356, 2007.

Lawrence M. Fisher. Market place. *The New York Times*, Apr 19, 1993, `http://www.nytimes.com/1993/04/19/business/market-place.html`, 1993. Accessed: 2016-07-28.

Milton Harris and Artur Raviv. Differences of opinion make a horse race. *Review of Financial studies*, 6(3):473–506, 1993.

Edwin Hewitt and Leonard J. Savage. Symmetric measures on Cartesian products. *Transactions of the American Mathematical Society*, 80(2):470, 1955.

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

Eugene Kandel and Neil D Pearson. Differential interpretation of public signals and trade in speculative markets. *Journal of Political Economy*, pages 831–872, 1995.

Péter Kondor. The more we know about the fundamental, the less we agree on the price. *The Review of Economic Studies*, 79(3):1175–1207, 2012.

Paul Krugman. Failure to rise. *The New York Times*, Feb 13, 2009, `http://www.nytimes.com/2009/02/13/opinion/13krugman.html?ref=paulkrugman`, 2009. Accessed: 2016-06-01.

Charles G Lord, Lee Ross, and Mark R Lepper. Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of personality and social psychology*, 37(11):2098, 1979.

Paul Milgrom and Nancy Stokey. Information, trade and common knowledge. *Journal of Economic Theory*, 26(1):17–27, 1982.

Paul R Milgrom. Good news and bad news: Representation theorems and applications. *The Bell Journal of Economics*, pages 380–391, 1981.

Evan Piermont. Distributional Uncertainty and Persuasion. 2016. *Working paper.*

Lee Sigelman and Carol K Sigelman. Judgments of the carter-reagan debate: The eyes of the beholders. *Public Opinion Quarterly*, 48(3):624–628, 1984.

Alexander Zimper and Alexander Ludwig. On attitude polarization under bayesian learning with non-additive beliefs. *Journal of risk and uncertainty*, 39(2):181–212, 2009.