Interpreted and Generated Signals

Lu Hong[*]and Scott Page[†]

July 20, 2007

Abstract

Models of incomplete information rely on probabilistic signals to capture the uncertain values of variables allowing the relationship between signals and the relevant variables to be captured by a joint probability distribution function. In this paper, we define and contrast two types of signals: *generated signals* and *interpreted signals*. Generated signals are distortions of the true values or outputs of some process. Interpreted signals are predictions based on inputs of a process or attributes of an object. Binary discrete interpreted signals are negatively correlated in their correctness, be it conditional or unconditional; moreover, the amount of negative correlation is uniquely determined in important cases. In contrast, generated signals can be independent conditional on the true value. In other words, these two types of signals produce distinct statistical signatures. Thus, our findings limit the contexts in which many well known models of information aggregation and strategic choices in auctions, markets, and voting apply.

[*]Department of Finance, Loyola University Chicago
[†]Departments of Political Science and Economics, University of Michigan

1 Introduction

When making economic or political decisions, agents rarely have the benefit of full information. They rely on partial or distorted information or on crude predictive models. Social scientists model this incomplete information in the form of signals. A joint probability distribution characterizes the statistical relationships between these signals and the variable of interest: the better the information, the tighter the link between the signals and the underlying variable of interest. A signal is an abstraction. A signal could represent a noisy measure of a variable or a prediction based on partial information. In this paper, we analyze whether these different types of real world signals constrain the class of feasible joint probability distributions. Put differently, do standard assumptions of independence of signals or independence conditional on the true value, constrain the set of environments that could produce those signals.

The statistical properties of signals influence strategic choices and therefore outcomes. Therefore, any analysis of economic and political institutions in incomplete information environments implicitly hinges on statistical assumptions about signals. In this paper, we connect those assumptions, and therefore those findings, to specific environments. In doing so, we contribute to the literature on the design and the performance of markets and political mechanisms. Our results imply that in some cases, auction markets and democracies may work better than previously thought and that in other cases, existing results may rely on assumptions that prove unrealistic.

To show how properties of the joint distribution function depend on the micro level foundations of the signals, we describe two frameworks that capture distinct environments for producing signals. The first framework assumes that signals are generated by a process. The other assumes that agents construct predictive models that act as signals. Both types of signals, of course, can be modeled by joint probability distributions. However, the statistical properties of those distributions differ depending on the process. Binary signals generated by a process are typically independent. One person's sample of a new chocolate candy is independent of another's. In contrast, binary signals that result from strategically chosen predictive models cannot be independent. They must be negatively correlated in their correctness. Moreover, the degree to which they are negatively correlated is often uniquely determined by their accuracy.

In what follows, we refer to the first type of signals, those produced by some process, as *generated signals*. Two agents' generated signals differ because the agents draw different samples or experience idiosyncratic shocks or distortions. We refer to the second type of signals in which agents filter a high dimensional reality onto dimensions or into categories (or attributes) and make predictions as *interpreted signals* (Fryer and Jackson 2003, Nisbett 2003, Page 2007).[1] Two agents' interpreted signals differ if the agents rely on different predictive models; if they differ in how they categorize or classify objects, events, or data.[2]

---

[1] Organizations also differentially filter reality. See for example, Stinchecombe 1990.

[2] Our introduction of the term interpreted signals is new to the literature. Similarities exist be-

Given our construction, generated signals have a probabilistic relationship to the relevant outcome variable. Thus, no number of generated signals, independent or non independent, pins down the outcome value with certainty. In addition, any number of generated signals can be produced satisfying the same distributional and correlational assumptions – scientists can sample air quality for days on end. Neither of these characteristics need to be true of interpreted signals. First, if all relevant attribute values can be learned, then the value of an outcome is known with certainty. If a poker player knows all of the cards in her opponents' hands, she knows whether she wins or loses. Second, given that interpreted signals depend on attributes, the number of distinct interpreted signals is bounded by the power set of the set of attributes. This latter implication implies that any model that assumes a large number of signals assumes generated signals and rules out interpreted signals, an insight we take up in more depth in the discussion at the end of this paper.

In this paper, we focus on *independent interpretations* – interpretations that consider distinct attributes of a common representation, or perspective (Hong and Page 2001, Page 2007). This captures situations in which agents look at distinct but relevant pieces of information when making predictions. Later, we discuss why, if people were to make strategic choices of interpretations, we might expect those interpretations to be independent. We also consider the special case where the outcome of interest takes on only binary values. Doubtless, the model generalizes, but the binary case provides the cleanest entry into the issues of concern.

Several common distributional assumptions are inconsistent with independent interpreted signals. For example, many models of information aggregation in common value auctions, voting, and in information cascades assume the following signaling structure that involves two outcomes G (good) and B (bad) and two signals g and b: The two outcomes are equally likely, the two signals predict the two outcomes with equal probability, the signals are independently correct (i.e. knowing that one signal is correct tells nothing about the probability that the other is correct) and the signals are informative, i.e. they are correct more than half of the time. We show that these assumptions are inconsistent with independent interpreted signals, which, as we show, must be negatively correlated conditional on at least one outcome. Moreover, these interpreted signals cannot be independently correct. They must be negatively correlated in their correctness both unconditionally and conditionally on at least one outcome. Therefore, these assumptions would not be consistent with a world in which theses signals are, for example, predictions of the values of investments.

The negative correlation result bodes well for information aggregation but has a contingent normative implication for auctions depending on whether we're aligned

---

tween our interpreted signal framework and models of causal inference (Pearl 2000), fact free learning (Aragones, Gilboa, Postlewaite, and Schmeidler 2005), and complexity (Al-Najjar, Casadesus-Masanell and Ozdenoren 2003). Interpreted signals resemble predictions from PAC learning (Valiant 1984) and classification theory (Barwise and Seligman 1997). The acronym PAC refers to Probably Approximately Correct. We differ from PAC learning and classification models in that we consider multiple agents each with his/her own classification rule.

with the buyers or the seller. Were our result only that the correlation is negative, it could be seen as moving the literature backward. However, if as often assumed, good and bad predictions are equally likely, the probability of a correct interpreted signal uniquely determines the extent of this negative correlation. Thus, we have a benchmark correlation assumption for interpreted signals that differs from zero. The uniqueness of this correlation has an unexpected corollary, namely that the *complexity* of the outcome function (the extent of nonlinearity and interaction terms) has no effect on pairwise correlation beyond that captured by signal accuracy.

We also relax the independent interpreted signals and ask whether interpreted signals can satisfy the standard assumption in the signaling literature of independence conditional on the outcome. We show that interpreted signals can satisfy the conditional independence assumption, the outcome function and the associated interpreted signals have a unique (up to isomorphism) structure that requires overlapping interpretations in a particular way: each agent must ignore a different piece of information. Intuitively, this runs counter to incentive effects, since by knowing that piece of information, the agent would have complete information. This result provides a powerful argument that even though conditional independence is a reasonable benchmark assumption for generated signals, it imposes a specific, and unlikely structure on interpreted signals making it an improper benchmark assumption for interpreted signals.

To summarize, in this paper we establish that the statistical properties of generated and interpreted signals differ. These differences call into question the generality of existing results that rely on an abstract model of signals – the auction, voting, or signaling literatures are replete with models that include assumptions based on tractability and not on descriptive realism. For example, in that part of the auction literature that relates to the derivation of optimal bidding strategies, the assumptions align with our interpreted signal framework (Klemperer 2004). However, in that part concerned with information aggregation in common value auctions, particularly those papers with large numbers of bidders, the assumptions are inconsistent with interpreted signals. Similarly, the bulk of the political science literature, including almost all jury models and election models (Ladha 1992, Feddersen and Pesendorfer 1997), makes assumptions consistent with generated signals, not interpreted signals, even though by definition, jurors make predictions based on information, i.e. use interpreted signals.

The remainder of the paper is organized as follows. In Section 2, we provide examples of generated and interpreted signals and compare them. In Section 3, we introduce a framework for interpretations and predictions and discuss three notions of independence: *independent interpretations*, *independent signals*, and *independently correct signals*. In Section 4, we prove the result that independent predictions are negatively correlated in their correctness. In Section 5, we cast our interpretations and predictions framework as signals, interpreted signals, and compare interpreted signals to generated signals. We demonstrate how to construct non independent interpreted signals that satisfy conditional independence. We show that an assumption

4

of conditionally independent signals imposes a unique (up to isomorphism) mapping from attributes to outcomes. In Section 6, we present an example in the context of voting to illustrate how our interpreted signals framework can be used to investigate information aggregation properties of the majority rule voting mechanism. We demonstrate why using interpreted signals framework requires a fundamentally different technique and lead to fundamentally different types of results compared to the standard information aggregation results. We conclude with a discussion of the implications of our results and future research questions.

2 Generated and Interpreted Signals: Examples

We begin with an example that highlights the statistical differences between generated and interpreted signals. Consider two venture capitalists who receive signals about the quality of an investment, say a restaurant. This restaurant can either be a good investment ($V = G$) or a bad one ($V = B$). Each outcome occurs with equal probability. We first describe the standard *generated signal* model.

2.1 Generated Signals

Generated signals can be thought of as noisy glimpses or distortions of an outcome value V. Imagine, for example, that two potential investors eat meals prepared at the restaurant. We can think of those meals as generated signals denoted by s_1 and s_2 that either take value g (for good) or b (for bad). Conditional on the restaurant being a good investment, i.e. conditional on outcome G, we assume that the probability of getting a good meal, i.e. of receiving the signal g, equals $2/3$. Similarly, conditional on B, the probability of getting the signal b equals $2/3$. Each of these signals is further assumed to be an independent draw from this distribution. Thus, we can write the joint probability distributions of signals conditional on the restaurant's value as follows:

<div align="center">

Generated Signals Conditional on G

s_1/s_2	b	g
b	1/9	2/9
g	2/9	4/9

</div>

<div align="center">

Generated Signals Conditional on B

s_1/s_2	b	g
b	4/9	2/9
g	2/9	1/9

</div>

Using this information, we can compute the probability that the restaurant is a good investment conditional on the signals of the two investors.

Probability of a Good Investment
Conditional on Generated Signals

s_1/s_2	b	g
b	1/5	1/2
g	1/2	4/5

The above table can be read as follows. If both investors get the signal b then the probability that the restaurant is a good investment equals $\frac{1}{5}$.

2.2   Interpreted Signals

We now turn to interpreted signals. Interpreted signals are predictions based on *interpretations*. Interpretations create partitions (or categorizations) of the set of possible restaurants. Interpretations partition the set of attributes that define a restaurant. In this example, we consider those attributes to be the restaurant's *location* and its *prices*. We assume that each investor sees only one of these attributes and bases her prediction on that attribute's value.

Interpreted signals require an *outcome function*, V, that maps the restaurant's *attributes*, into a probability that the restaurant is a good investment. Here, we assume that the location ℓ and the prices \$ can be either good 1 or bad 0, and that each combination of attribute pairs is equally likely. We assume the following functional form for the outcome function.[3]

$$V(\ell, \$) = \begin{cases} \frac{1}{3} & \text{if } (\ell + \$) \leq 1 \\ 1 & \text{otherwise} \end{cases}$$

In this example, we assume that the first investor looks only at the location and the second looks only at prices. The investors then construct *predictive models* based on their attributes' values. In brief, if the attribute's value equals 1 (resp. 0), the investor predicts the restaurant will be good (bad). The investors' *interpreted signals* equal the value of the investors' attributes. In the general framework, predictions

---

[3]In the paper, we consider deterministic outcome functions. Any probabilistic outcome function can be transformed into a deterministic outcome function by adding attributes. In this example, we need only add a third independent attribute h that takes three values, say 0, $\frac{1}{2}$, and 1, with equal probability. The outcome function can then be written as a deterministic function as follows:

$$V(\ell, \$, h) = \begin{cases} 1 & \text{if } h = 1 \text{ or } (\ell + \$) > 1 \\ 0 & \text{otherwise} \end{cases}$$

can be based on multiple attributes and therefore are not identical to the attributes themselves. Given these interpreted signals, we can now write a joint probability distribution for the signals and the value of the outcome function.

Probability of a Good Investment
Conditional on Interpreted Signals

$\ell/\$$	0	1
0	1/3	1/3
1	1/3	1

We can then calculate the joint probability distribution of the interpreted signals conditional on the restaurant's quality. As with the generated signals, the probability of an agent getting the good (bad) signal conditional on the restaurant being of good (bad) quality equals $\frac{2}{3}$.

Interpreted Signals Conditional on G

$\ell/\$$	0	1
0	1/6	1/6
1	1/6	1/2

Interpreted Signals Conditional on B

$\ell/\$$	0	1
0	1/3	1/3
1	1/3	0

Comparing the generated and interpreted signals reveals several differences. First, the interpreted signals are not conditionally independent. Conditional on the restaurant being a bad investment, having a good location *reduces* the likelihood that the prices are also good. In fact, the probability of having good prices equals zero. Second, the generated signals are independently correct, but the interpreted signals are not. Again, conditional on the restaurant being a bad investment, if one interpreted signal is incorrect, the other *must* be correct. Third, with interpreted signals, the correlation between the signals and values depends on the outcome function. With generated signals, whatever correlation exists is just assumed. Finally, in the case of interpreted signals, we are limited to two attributes: location and prices. The only constraint on the number of generated signals is the chef's time.[4]

---

[4]This example creates a clean distinction between generated and interpreted signals, whereas

3 The Interpretation and Prediction Framework

We now develop our framework for interpretations and predictions. Only when we recast them as signals later in the paper will we call them **interpreted signals**. To avoid confusion, we also distinguish among several types of independence.

3.1 Interpretations and Predictions

In constructing interpretations and predictions, we differentiate between the set of objects or events and their outcome values. We define *the environment*, Ω, to be a finite collection of objects or events with a cardinality equal to N. Each of these events or objects has associated with it an outcome. We denote the set of outcomes by S and the deterministic mapping from events to outcomes as an outcome function $\widetilde{O} : \Omega \to S$. [5] Agents are to predict outcomes.

The problems we consider are equivalent to classification problems in which an agent has to place the N objects into $|S|$ bins representing possible outcome values and are related to the problem of selecting regressors (Aragones, et al 2005). In this paper, we restrict attention to cases in which the cardinality of S equals two. To make a prediction, an agent first partitions the environment into non-overlapping sets. We denote the partition of agent i, Π^i, to be the sets $\{\pi_1^i, \pi_2^i, \ldots, \pi_{n_i}^i\}$, where n_i is the number of sets in agent i's partition. Π^i is agent i's understanding of the environment and it is incomplete as long as not all sets in the partition are singletons. When an agent sees an object or event, she associates it with the set in her partition that contains this object. We call these partitions **interpretations**.

Let $P : \Omega \to [0, 1]$ be the probability distribution over Ω where $P(\omega)$ denotes the probability that event ω arises. Given this distribution over events and an interpretation of the environment, an agent makes predictions about the outcome. For example, an agent might use a Bayesian approach to making these predictions and assume that the most probable outcome arises conditional on the set in her interpretation. We refer to these as **experience generated predictions**. At this point, though, we do not specify how predictions are generated. Agent i's **prediction** $\widetilde{\phi}_i$ is simply defined as a function from Ω to S with the restriction that $\widetilde{\phi}_i$ is measurable with respect to agent i's interpretation Π^i.[6] The following example illustrates the main components of the interpretation and prediction framework:

---

often the differences can be subtle. Consider a university whose quality q in $\{0, 1\}$ depends on its faculty's abilities: each faculty member i has an ability a_i in the set $\{0, 1\}$ drawn from some distribution F. Assume that this university produces students at regular intervals and that each graduate j receives an added value x_j in $[0, 1]$. The graduating students' added values would be generated signals and the faculty's abilities would be interpreted signals. Thus, their statistical properties would differ in the ways that we describe.

[5] As previously mentioned, the framework extends to include probabilistic mappings.

[6] Here, we do not allow predictions to be probability distributions over outcomes simply because we want to be consistent with the signal framework so that our comparison later is meaningful.

Example 1 *The environment,* $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$. *All events are equally likely,* $P(\omega_i) = \frac{1}{6}$. *The set of outcomes,* $S = \{G, B\}$. *The outcome function maps the first three events to G and the rest to B. Let $\Pi = \{\pi_1, \pi_2\}$ be an agent's interpretation of the environment where $\pi_1 = \{\omega_1, \omega_2, \omega_4\}$ and $\pi_2 = \{\omega_3, \omega_5, \omega_6\}$. If this agent makes experience generated predictions, then her predictions can be described by the following function*

$$\widetilde{\phi}(\omega_i) = \begin{cases} G & \text{for } \omega_i \in \pi_1 \\ B & \text{for } \omega_i \in \pi_2 \end{cases}$$

In the example and generally, the outcome function $\widetilde{O} : \Omega \to S$ together with a prediction $\widetilde{\phi} : \Omega \to S$ induce another random variable, $\widetilde{\delta}$, called the **correctness of predictions**, which is defined as follows: $\widetilde{\delta} : \Omega \to \{c, i\}$ such that

$$\widetilde{\delta}(\omega) = \begin{cases} c & \text{for } \omega \in \left\{\omega' \in \Omega : \widetilde{\phi}(\omega') = \widetilde{O}(\omega')\right\} \\ i & \text{for } \omega \in \left\{\omega' \in \Omega : \widetilde{\phi}(\omega') \neq \widetilde{O}(\omega')\right\} \end{cases}$$

where c means "correct" and i means "incorrect". The **accuracy** of a prediction $\widetilde{\phi}$ is defined as the probability of its associated $\widetilde{\delta}$ taking value c. In the above example, $\widetilde{\delta}(\omega) = c$ for $\omega \in \{\omega_1, \omega_2, \omega_5, \omega_6\}$ and $\widetilde{\delta}(\omega) = i$ for $\omega \in \{\omega_3, \omega_4\}$. Thus, with probability $\frac{2}{3}$, $\widetilde{\phi}$ makes correct predictions.

3.2 Definitions and Relationships of Types of Independence

We now turn to our discussion of various types of independence that relate to interpretations and predictions. Several of these definitions and results we describe are standard. We include them to place the new definitions and results in their proper context. We consider only pairwise independence. Extensions to include any finite number of interpretations can be constructed.

definition 1 *Interpretations, Π^1 and Π^2, are* **independent interpretations** *if*

$$Prob\left(\pi_i^1 \cap \pi_j^2\right) = Prob\left(\pi_i^1\right) \times Prob\left(\pi_j^2\right)$$

for all $i \in \{1, ..., n_1\}$ and $j \in \{1, ..., n_2\}$

Two interpretations are independent implies that knowing how one agent interprets an event provides no information about how the other agent interprets that same event. Note that interpretations are not predictions, they are the sets within which agents place singular events. We focus here on independent interpretations for three reasons. First, they provide a natural benchmark, and that benchmark lacks a formal definition in the literature. They formally define what might be loosely described as two people seeing the world differently. Second, in a binary classification problem, they provide the most *diverse* predictive models possible. This follows

because interpretations cannot be negatively correlated once we tack on predictions. Finally, in the case where interpretations are chosen strategically, maximizing the value of information would entail choosing an interpretation as different from those of others. That would mean an independent interpretation.[7]

We now state a surprising result: *independent interpretations imply that the environment can be written as a product space and the interpretations can be written as projections onto variables.* To make the logic that drives this result as transparent as possible, we first assume that each event in Ω is equally likely. Consider the following trivial observation: if Ω can be represented as a product of attribute spaces and if agents make partitions by looking at disjoint subsets of attributes, then the agents have independent interpretations. For example, if the environment is written as a two by two lattice and one agent considers the row and the other considers the column, then these interpretations are independent. This is not surprising. Knowledge of an event's row, tells us nothing about its column.

Surprisingly, the converse also holds. If two interpretations are independent, then the event space can be mapped into a coordinate system (a two attribute, x and y, model) where each event is represented by (x, y), and one interpretation considers only the x attribute and the other is along the y attribute. This result implies that any two independent interpretations can be rewritten as projections onto different attributes of *the same perspective* (Hong and Page 2001, Page 2007).[8] We provide the formal statement of this result below. Its proof along with all proofs of many subsequent claims is contained in the appendix.

Claim 1 *Assume that each event in Ω is equally likely. Let $\Pi^1, \cdots, \Pi^n$ ($n \geq 2$) be non-trivial interpretations of Ω. If they are independent, then Ω can be represented by an n-attribute rectangle such that Π^i is along the ith attribute. Thus $N = \prod_{h=1}^{n} a_h$ for some larger-than-1 integers a_h, $h = 1, ..., n$[9]*

---

[7]Here we take the attributes as given. If one player chooses a single factor with the highest predictive value, a second player could then choose the factor with the highest residual predictive value, which would by construction be orthogonal.

[8]Recall that a perspective is a representation of the entire space of possibilities. Two agents use different perspectives if they represent the set of the possible alternatives with different languages. These languages can be basis. For example, one agent may identify a point in the plane using Cartesian coordinates (x, y). Another agent may use polar coordinates (r, θ). The natural interpretations differ for these two perspectives. In the former, someone might partition the space into points in which $x \leq 5$ and points in which $x > 5$. In the latter, an agent might partition the space into points in which $r \leq 10$ and points in which $r > 10$.

[9]The result above is established with the assumption that all events are equally likely. This assumption is not essential. We can show that if events in the original space Ω do not have equal probability, there exists an equally probable event space Ω' that has greater cardinality (the least common denominator of probabilities of original events expressed in fractions) such that Ω' can be represented by an n-attribute rectangle and the independent interpretations of the original event space Ω correspond to interpretations of the new event space Ω' along different attributes. The original Ω consists of lumping of some events in Ω'. The key is that for independent interpretations,

Intuitively, Claim 1 implies a bound on the number of independent interpretations. It cannot exceed the number of primes in the factorization of N. As stated in the Corollary below, a finite set of events admits few independent interpretations.

Corollary 1 *Assume events are equally likely. Let $\prod_{i=1}^{k} p_i$ be the unique prime factorization of N, that is,*

$$N = \prod_{i=1}^{k} p_i$$

where each p_i is a prime. Then, the maximum number of independent non-trivial interpretations is k.

The implications of this corollary sink in when applied to a specific example such as the set of possible independent interpretations of all of the 300 million people in the United States. Such interpretations, the parsing of people into categories like soccer moms or NASCAR dads, are used to construct predictive models for economic, political, and social outcomes. The corollary implies that there exist fewer than *thirty* independent interpretations for the entire US population.[10]

Independent interpretations are distinct from independent predictions. Saying two agents's predictions are independent means that knowing one agent's prediction about the outcome of an event provides no information about the other agent's prediction.

definition 2 *Predictions, $\widetilde{\phi}_1$ and $\widetilde{\phi}_2$, are* **independent predictions** *if they are independent random variables.*

Note that if two agents have independent interpretations, then their predictions are independent as well.

Observation 1 *Independent interpretations imply independent predictions.*

Proof: Since each agent's prediction is measurable w.r.t. her interpretation, the claim follows.

In contrast, predictions can be independent without agents having independent interpretations. This result should come as no big surprise, but it drives a conceptual wedge between the two types of independence.

Observation 2 *Independent predictions may not imply independent interpretations.*

---

probabilities have the rectangle property, i.e.,

$$Prob(\pi_i^1 \cap \pi_j^2) = Prob(\pi_i^1) \times Prob(\pi_j^2)$$

[10]This result assumes that each attribute is binary such as { male, female }. To be precise, 2^{28} is slightly less than 300 million and 2^{29} exceeds it by a substantial margin.

Proof: Due to the measurability requirement, an agent's prediction defines a partition on the environment that is in general coarser than her interpretation. Recall Example 1. We can add a second agent whose interpretation is

$$\Pi^2 = \{\{\omega_1, \omega_2\}, \{\omega_4, \omega_5\}, \{\omega_3, \omega_6\}\}$$

and whose prediction is

$$\widetilde{\phi}_2(\omega) = \begin{cases} G & \text{for } \omega \in \{\omega_1, \omega_2, \omega_3, \omega_6\} \\ B & \text{otherwise} \end{cases}$$

It can be shown that this agent's prediction is independent of the prediction made by the agent in Example 1, even though the two interpretations are not independent.

Relatedly, we say that two predictions are independently correct, if knowing that one agent's prediction of an event is correct gives no information about whether the other's prediction of the same event is correct.

definition 3 *Predictions, $\widetilde{\phi}_1$ and $\widetilde{\phi}_2$, are* **independently correct predictions** *if $\widetilde{\delta}_1$ and $\widetilde{\delta}_2$ are independent random variables.*

Our next observation reveals the absence of a causal linkage between independent predictions and independently correct predictions. Even though seeing the world independently implies predicting independently, it need not imply being independently correct.

Observation 3 *Independent predictions need not be independently correct.*

Proof: Consider the following example:

Predictions	g	g	b	b
g	G	G	G	B
g	G	G	B	G
b	G	B	B	B
b	B	G	B	B

In this example, assume that each of the sixteen events is equally likely. The upper case letters represent outcomes of events. The lower case letters in the first column and in the first row are predictions of the row agent and the column agent respectively. By construction, the two predictions are independent. However, they are not independently correct. The joint probability of both agents making correct predictions is $\frac{1}{2}$ while the multiplication of the probabilities of each agent making correct predictions equals $\frac{3}{4} \times \frac{3}{4} = \frac{9}{16}$. They are not equal.

Not only the predictions of the two agents but also their interpretations are independent in this example. Therefore, the correctness of predictions need not be independent even if the interpretations are.

Observation 4 *Independent interpretations may not lead to independently correct predictions.*

4 Predictions and Correctness of Predictions

We now present results within the interpretation and prediction framework. We first show independent predictions to be inconsistent with predictions being independently correct. Since independent interpretations imply independent predictions, these inconsistency results also apply to independent interpretations. We establish these results for the case of binary outcomes in which agents' predictions have identical probability distributions and equal accuracy.[11]

As before, we let upper case letters, G and B, refer to outcomes and lower case letters, g and b refer to predictions. Let $P(G)$ and $P(B)$ denote priors, assumed to be common among all agents. $P(g)$ and $P(b)$ denote the probabilities of predicting g and b respectively. $P(g, g)$ denotes the probability of both predicting g. $P(b, b)$, $P(g, b)$ and $P(b, g)$ are similarly defined. $P(c)$ and $P(i)$ denote the probabilities of making correct and incorrect predictions, which are also assumed to be the same for both agents. Finally $P(c, c)$, $P(i, i)$, $P(c, i)$ and $P(i, c)$ denote joint probabilities of both correct, both incorrect, agent 1 correct but agent 2 incorrect and agent 1 incorrect but agent 2 correct respectively. We impose the following symmetry assumptions: $P(g, b) = P(b, g)$ and $P(c, i) = P(i, c)$.

4.1 Reasonable and Informative Predictions

Given an interpretation, an agent need not make the best possible predictions. For example, an agent who categorized agents by gender could predict that women are taller than men. To impose some degree of competence we assume that predictions are either *reasonable* or *informative*.

definition 4 *A prediction is **reasonable** if it is correct at least half of the time, i.e., $P(c) \geq \frac{1}{2}$.*

definition 5 *A prediction is **informative** if it is correct more than half of the time, i.e., $P(c) > \frac{1}{2}$.*

---

[11]We have derived a set of results that do not assume that agents's predictions have identical distributions.

In the binary outcome case, an experience generated prediction must be reasonable. Further if at least one prediction is correct more than half of the time, then an experience generated prediction is also informative. However, as we observe next, an informative prediction need not predict correctly half of the time conditional on every outcome.

Observation 5 *An informative prediction need not be reasonable conditional on every outcome.*

This observation may be obvious to some. Nevertheless, we provide an example because the underlying logic is central to our analysis.

An Informative Prediction

Prediction	Outcomes		
g	G	G	B
g	G	G	B
g	G	G	B
g	G	G	B
b	B	B	G

Assuming each outcome to be equally likely, $P(c) = \frac{2}{3}$, implying that the prediction is informative. However, conditional on outcome B, the probability of making correct prediction equals $P(c \mid B) = P(b \mid B) = \frac{1}{3}$, implying that the prediction is not reasonable conditional on the bad outcome. This example provides insight into why when predicting rare events, agents may not make reasonable predictions.

4.2   Negative Correlation in Correctness

We now state two lemmas that lead to our result about negative correlation in the correctness of predictions presented in Claim 2 below. Even though Claim 2 can be established directly, the lemmas that follow are interesting in their own right. They reveal a tension between the accuracy of predictions and the correlation of their correctness in a more general environment where the probabilities of a good prediction and a bad prediction are not required to be equal. When predictions are independent, the higher the accuracy, the less correlated their correctness. In fact, highly accurate predictions must be negatively correlated in their correctness. Without loss of generality, we assume $P(g) \geq \frac{1}{2}$. Also note that unless otherwise specified, we do not require that the prior probabilities over G and B to be equal.

Lemma 1 *The correctness of independent and reasonable predictions exhibit positive (zero, negative) correlation if and only if $P(g) > [=, <] \, P(c)$.*

The intuition that drives this result is straightforward. If the probability of predicting the good outcome is large relative to the probability of being correct, then both agents often predict good outcomes at the same time whether or not the prediction is correct. Thus, the correctness of their predictions must also be positively correlated.

Next, we reverse the assumption. We require that the predictions be independently correct. We can then show that the predictions themselves are independent or negatively or positively correlated depending again on the relationship between the probability of the more frequently picked prediction and the probability of being correct.

Lemma 2 *Independently correct and reasonable predictions exhibit positive (zero, negative) correlation if and only if $P(c) > [=, <] \, P(g)$.*

Our claim follows from Lemma 1.

Claim 2 *Independent informative predictions that predict good and bad outcomes with equal probability must be negatively correlated in their correctness. Furthermore, the negative correlation of their correctness is defined by the following expression*

$$\rho = 1 - \frac{1}{4(p - p^2)}$$

where p is the accuracy of the individual predictions.[12]

The first part of this claim follows directly from Lemma 1. The second part of this claim means that the accuracy of the individual predictions is a sufficient statistic for the correctness correlation between agents. Thus, changes in the outcome function do not lead to any change in the correctness correlation beyond the change in the accuracy. The following corollaries are immediate.

Corollary 2 *Any independent and independently correct predictions that predict good and bad outcomes with equal probability cannot be informative.*

In other words, they are as good as flipping a coin. Since independent interpretations imply independent predictions, we also have the following:

---

[12]Given that the correlation coefficient can not be less than -1, it can be easily computed based on the expression here that the accuracy of the individual predictions can not exceed $\frac{1}{2} + \frac{\sqrt{2}}{4}$ which is approximately equal to 0.8535. This means that independent predictions that predict two outcomes with equal probability can not be equally highly accurate. This observation is intuitive and appealing.

Corollary 3 *If interpretations are independent and if their associated predictions are informative and predict good and bad outcomes with equal probability, then the correctness of their predictions must be negatively correlated.*

The claim together with the corollaries reveal a fundamental conflict between an assumption that agents see the world independently (independent interpretations) and an assumption that the interpretation based predictions they make are independently correct.

5 Predictions as Signals - Interpreted Signals

In this section, we first recast our interpretation based predictions as signals and call them interpreted signals. We examine the implication on interpreted signals of the standard assumption of independent signals conditional on true outcomes. We then establish a result that shows a unique (up to isomorphism) structure of the outcome function and signals if signals are interpreted and satisfy the conditional independence assumption. This is a powerful result implying that the standard assumption of conditional independence may not be appropriate for interpreted signals in many situations.

Recall from Section 3 our framework for interpretations and predictions. Let (Ω, P) be the probability space[13] that describes the underlying uncertainty. The outcome function $\widetilde{O}$ is the random variable that is the variable of interest. Predictions $\widetilde{\phi}_i$'s are random variables on (Ω, P) that can be viewed as signals about the variable of interest. Thus, our interpreted based predictions are nothing but signals with a specific micro-foundation. We call signals with this micro-foundation **interpreted signals**. Clearly, interpreted signals differ in its foundation from signals that are outcomes generated by some common process which we called **generated signals**. However, does the difference in foundation extend to differences in their relevant statistical properties? We investigate the standard assumption of conditional independence, which is ubiquitous in the applied literature on incomplete information models and is arguably quite reasonable for generated signals, in the interpreted signal framework.

5.1 Conditional Independence and Correlated Interpreted Signals

For notational convenience, let p denote the probability that an agent predicts g conditional on the true outcome being G and q denote the probability that an agent predicts b conditional on the true outcome being B, that is,

$$p = P(g \mid G)$$

---

[13]Recall that Ω is finite. So the description of the probability space here is more than adequate.

$$q = P(b \mid B)$$

Interpreted signals can then be written as a typical binary signal model with the following conditional distributions (conditional on outcomes)

Conditional Probability Distribution of Signals

$outcome \backslash signal$	g	b
G	p	$1-p$
B	$1-q$	q

Consistent with the notation from the previous section, we have the following unconditional distribution of predictions.[14]

$$P(g) = P(G)p + P(B)(1-q)$$

and

$$P(b) = P(G)(1-p) + P(B)q$$

We now relate independent interpretations and predictions to signals that are independent conditional on outcomes. The first claim states that informative and experience generated predictions – therefore, reasonable predictions – that satisfy independence conditional on outcomes must be unconditionally positively correlated. In other words, the assumption that interpreted signals satisfy the standard assumption from signaling models (independence conditional on outcomes) implies that the predictions themselves are positively correlated and cannot come from independent interpretations.

Claim 3 *Experience generated and informative predictions that are independent conditional on outcomes must be positively correlated unconditionally.*

The intuition behind this claim can be seen in an example. Suppose that good and bad outcomes are equally likely and that conditional on outcomes each of two agents predicts correctly with probability $\frac{2}{3}$. Suppose, for example, that the environment

---

[14]Binary generated signals are often assumed to satisfy the *Strong Monotone Likelihood Ratio Property (SMLRP)*. Using the notation above, a signal satisfies the SMLRP if and only if

$$\frac{p}{1-q} > \frac{1-p}{q}$$

or equivalently

$$p + q > 1.$$

It can be shown that if predictions are experience generated and informative, then they satisfy the SMLRP.

consists of nine good outcomes and nine bad outcomes and that each is equally likely. For the predictions to be independent conditional on outcomes, two agents would have to both predict four of the good outcomes correctly and both predict one of the good outcomes incorrectly. Each would also have to predict two good outcomes correctly that the other predicted incorrectly. The same holds for the bad outcomes. We can represent this in a table.

<div align="center">

Conditionally Independent Predictions

agent 1 /agent 2	g	b
g	$4G$ and $1B$	$2G$ and $2B$
b	$2G$ and $2B$	$1G$ and $4B$

</div>

Notice that the predictions are not independent. If agent 1 predicts g then agent 2 predicts g with probability $\frac{5}{9}$. If agent 1 predicts b then agent 2 predicts g with probability $\frac{4}{9}$. The predictions are positively correlated. Independence conditional on outcomes requires that the predictions be positively correlated.

We can now state the flip side of Claim 3.

Claim 4 *Assume that predictions are experience generated and are informative. If predictions are independent, then for at least one outcome, predictions conditional on that outcome are negatively correlated.*

Either of the above two claims implies an inconsistency between conditional independence and independent interpretations.

Claim 5 *Conditional independence of predictions is inconsistent with informative and experience generated independent predictions and therefore with independent interpretations.*

This final claim parallels our previous result showing a conflict between reasonable, informative, independent predictions and independently correct predictions. These two claims together reveal a fundamental incompatibility between standard signaling assumptions and independent interpretations: *Seeing the world independently, looking at different attributes, not only does not imply, it is inconsistent with, both conditional independence of signals and independently correct signals.*

5.2 The Structure of Conditionally Independent Interpreted Signals

The above subsection establishes that conditionally independent interpreted signals are positively correlated unconditionally, a statistic property. Here we want to show

the functional property of interpreted signals that are conditionally independent. We first show how to construct a collection of interpreted signals which are independent conditional on the value of the outcome. We then prove that any collections of conditionally independent interpreted signals can be mapped into this constructed structure.

Our example considers the case where the objects have equal probability and that good and bad outcomes are equally likely. Let $p_i = \frac{r_i}{m_i}$ denote the probability that signal i is correct. We assume that $2r_i > m_i > r_i$ so that p_i lies in the open interval $(0.5, 1)$. p_i will also equal the probability that signal i is correct conditional on each outcome. To construct K ($K \geq 2$, integer) interpreted signals that are conditionally independent on the outcome, we set N, the number of objects of equal probability, equal to $2m_1 \times ... \times m_K$.

We denote an object as a vector of $K + 1$ attributes in which the first attribute takes one of two values, for convenience, 0 or 1, and each of the remaining K variables take values in the set $\{1, 2.....m_i\}$. As a matter of convention, we write an object as a vector of attributes $(\theta, x_1, x_2, ...x_K)$. We construct the payoff function so that if an even number of the last K attributes are greater than r_i, the value of the function equals θ. Otherwise, the value equals $(1 - \theta)$.

$f(\theta, \vec{x}) = \theta$ if $\mid \{i : x_i > r_i\} \mid = 2j$ for some non negative integer j

$\qquad = 1 - \theta$ else

We define the interpretations and the interpreted signals as follows. Interpretation i considers every attribute except attribute x_i, i.e. $(\theta, \vec{x}_{-i})$. The *interpreted signal*, s_i based on this interpretation equals θ if an even number of the attributes other than x_i take values greater than r_j and equals $1 - \theta$ otherwise.

$s_i(\theta, \vec{x}_{-i}) = \theta$ if $\mid \{j \neq i : x_j > r_j\} \mid = 2j$ for some non negative integer j

$\qquad = 1 - \theta$ else

Let's now consider the properties of the signals in this example keeping in mind that all attribute vectors are equally likely. First, s_i makes a mistake if and only if x_i has a value larger than r_i. This is true what ever the true outcome is. So, s_i is correct with probability $\frac{r_i}{m_i}$. Also, s_i is equal to the true outcome plus an error term (independent of the true outcome) which is determined entirely by x_i. By construction, x_i and x_j are independent. Therefore s_i and s_j are independent conditional on either outcome value.

Unlike generated signals, in which an agent gets a signal that is correct with some probability, these interpreted signals correspond to models that leave out one piece of information – the value of an attribute. Most of the time, the realization of that value will not change the value of the outcome, but sometimes it will. Thus, it is possible for conditionally independent signals as occurring provided each agent constructs a predictive model that ignores a different attribute and if the functional form has the property that each attribute has a proportionally similar effect on the outcome regardless of the outcome value.

19

As we now show, this last restriction rules out all functions over attributes except those isomorphic to the class of examples we just described.

Claim 6 *Let $\widetilde{O} : \Omega \rightarrow \{0, 1\}$ be an outcome function. For any $i \in \{1, ...K\}$, let $\widetilde{\phi}_i$ be agent i's interpreted signal associated with the outcome function $\widetilde{O}$. Assume that each signal $\widetilde{\phi}_i$ is correct with probability $p_i = \frac{r_i}{m_i}$ given either outcome. If the interpreted signals $\widetilde{\phi}_1..., \widetilde{\phi}_K$ are independent conditional on each outcome, then the original event space Ω can be represented by the space of attributes in the above example such that the outcome function $\widetilde{O}$ and the signals $\widetilde{\phi}_1..., \widetilde{\phi}_K$ are represented by the outcome function f and the interpreted signals $s_1, ..., s_K$ in the above example.*

This claim implies that while it is possible to create conditionally independent interpreted signals, doing so implies a unique outcome function and requires that each agent neglects one piece of information. Thus given a specific function mapping attributes to outcomes, paradoxically, conditionally independent signals are not consistent with agents looking at different attributes but they are only consistent with agents neglecting different pieces of information. Therefore, conditional independence is a difficult condition to defend as a benchmark assumption for interpreted signals.

6 Outcome Function Complexity and Aggregation of Interpreted Signals

The interpreted signal framework relies on an outcome function. We might expect that the complexity of that outcome function has implications for both signal accuracy and correlation. For example, we would expect that a linear outcome function allows more accurate interpreted signals than does an outcome function with lots of interaction terms. The relationship between complexity and correlation appears less obvious. In fact, our earlier result that gives correlation as a function of signal accuracy suggests that complexity only matters for correlation in so far as it affects accuracy. In this section, we discuss that relationship in greater detail by introducing a crude complexity measure. We then show how outcome function complexity, through its effect on signal accuracy, influences the accuracy of aggregated interpreted signals.

6.1 Function Complexity, Signal Accuracy, and Signal Correlation

In this section, we restrict attention to the special case of a two dimensional attribute space in which each attribute takes on one of $2K$ values. We further assume that good and bad outcomes as well as good and bad interpreted signals are equally likely. We can write each event as a vector (x, y), where x and y take values in $\{1, ..2K\}$.

Moreover, without loss of generality, we can assume that if $x \leq K$ (resp $> K$), then the interpreted signal based on the first attribute equals g (resp b) and that the same conditions hold for the interpreted signal based on the second attribute.

An abundance of complexity measures exist in the literature (Miller and Page 2007). Here, by complexity we mean nonlinearity and interaction terms in the outcome function. Given that our function only takes two values, we can appeal to a crude but simple measure that we call *attribute based value changes*, which we denote by $\Delta(V)$. For each attribute value, we can count the number of times the value changes as that attribute ranges from 1 to K. For each attribute, we denote this value as $\delta(z)$. To compute $\Delta(V)$, we compute the sum of the $\delta(z)$'s. We show two examples below for the case $K = 3$.

	y_1	y_2	y_3	y_4	y_5	y_6	$\delta()$
x_1	1	1	1	1	1	1	**0**
x_2	1	1	1	1	1	1	**0**
x_3	1	1	1	1	0	0	**1**
x_4	1	1	0	0	0	0	**1**
x_5	1	0	0	0	0	0	**1**
x_6	1	0	0	0	0	0	**1**
$\delta()$	**0**	**1**	**1**	**1**	**1**	**1**	$\Delta(V) = 9$

	y_1	y_2	y_3	y_4	y_5	y_6	$\delta()$
x_1	0	1	1	0	1	1	**3**
x_2	1	0	1	1	0	0	**3**
x_3	1	1	1	0	1	1	**2**
x_4	0	1	1	0	0	0	**2**
x_5	1	0	0	0	1	0	**3**
x_6	0	1	1	0	0	0	**2**
$\delta()$	**4**	**4**	**2**	**2**	**5**	**3**	$\Delta(V) = 35$

Note in this second example that we can reduce $\Delta(V)$ by switching attributes 2 and 3.

	y_1	y_2	y_3	y_4	y_5	y_6	$\delta()$
x_1	0	1	1	0	1	1	**3**
x_3	1	1	1	0	1	1	**2**
x_2	1	0	1	1	0	0	**3**
x_4	0	1	1	0	0	0	**2**
x_5	1	0	0	0	1	0	**3**
x_6	0	1	1	0	0	0	**2**
$\delta()$	**4**	**4**	**2**	**2**	**3**	**1**	$\Delta(V) = 31$

Given an outcome function V, we can define $\Delta^*(V)$ to be the minimum value of $\Delta(V)$ given any permutation of the attributes such that interpreted signals for the first K attribute values equal g for both attributes.

To show how this measure of complexity correlates with interpreted signal accuracy and correlation, we need only partition the set of outcomes into two sets: the *agreement set* and the *disagreement set*.

definition 6 *The* **agreement set** $A = \{(x,y) : s_1(x,y) = s_2(x,y)\}$.

definition 7 *The* **disagreement set** $D = \{(x,y) : s_1(x,y) \neq s_2(x,y)\}$.

We first argue that any change in the outcome function in the agreement set A that lowers accuracy for both interpreted signals is more likely to increase complexity than to decrease it. Consider an outcome that both interpreted signals predict correctly. Without loss of generality, assume that this outcome has value one. If we switch the value of this outcome, we decrease signal accuracy. Whether we increase or decrease complexity depends upon the values of the outcome function in neighboring columns and rows. The complexity can only change if either both neighboring columns or both neighboring rows take the value one. If the interpreted signals are informative, i.e. correct more than half of the time, then the neighboring rows or columns are more likely to be of value one than to be of value zero. Therefore, complexity is likely to increase. Thus, on average, complexity would be more likely to increase than to decrease. Counterexamples where complexity decreases are, of course, easy to construct.

Next, let's consider changes to the outcome function in D. So that both interpreted signals have the same accuracy, we choose two changes to the outcome function in D so that one change improves the accuracy of one signal and one change improves the accuracy of the other. (Note: given the changes are in D both signals have the same accuracy they had prior to the changes.) Those changes could increase complexity or they could decrease complexity but from our previous result that correlation depends only on p, they can have no effect on the correlation between the signals.

Claim 7 Changes in Δ within D that have no effect on the accuracy of the signals can have no effect on signal correlation

This claim implies that if we make a function more (resp. less) "complex" or more (less) "nonlinear" within the disagreement set D, while keeping accuracy unchanged, we have no effect on the correlation of the predictive signals. In other words, details of the complexity of the function have no effect on interpretive signal correlation other than what is captured by signal accuracy. This claim allows us to think of accuracy as a sufficient statistic for complexity when computing pairwise correlation. However, we should emphasize that these results hinge on independent interpreted signals. [15]

_____

[15]See *Prediciting Success* in Page 2007 for an example using a non interpreted signal in which the complexity matters.

6.2 An Example: Interpreted Signals and Information Aggregation

We next describe a toy model to show how to apply the interpreted signal framework in the context of a mechanism. We leave a more general analysis for future work. Through this example, we show that the key to whether private information represented by interpreted signals aggregate through a mechanism, specifically the majority rule voting mechanism, lies in the structure of the outcome function relative to agents' interpretations. In other words, in the interpreted signal framework, the information being aggregated is the functional relation between the outcome and the interpreted signals. Thus if this functional relationship is "complex", then a "simple" mechanism may not be able to aggregate information perfectly.

In our toy model, agents vote on candidates for a political office. Whether a candidate is a good or a bad choice depends on five attributes. Specifically, assume that it is represented by the following outcome function: $F : \{0,1\}^5 \to \{0,1\}$ such that $F(a_1, ..., a_5) = 1$ iff $\sum_{i=1}^5 a_i \geq 3$. Since we want to focus purely on information aggregation as opposed to preference aggregation, we assume that agents agree on what makes a good or a bad choice. Assume also that a priori, each attribute vector is equally likely. Let's consider five agents, each observing the value of one and a different attribute prior to casting their votes.

We can think of these attributes as interpreted signals. We now consider using the majority rule voting mechanism to aggregate information. Straightforward calculations show that conditional on the true outcome, (i) each attribute is correct with probability $\frac{11}{16}$ (ii) the probability of a pair of attributes being correct at the same time is $\frac{7}{16}$ and (iii) since $\frac{7}{16} < \frac{11}{16} \times \frac{11}{16}$,, the attributes are pairwise negatively correlated, which by our previous results must be true.

All agents voting their information constitutes a Nash equilibrium (NE) of this model. To see this, note that given all other agents vote their information, agent i is pivotal if and only if other agents' attributes split, two 1's and two 0's. In those situations, agent i's vote alone determines the outcome of the game. Since the true outcome in these situations always matches with agent i's attribute value, it is in the interest of agent i to vote her attribute value. Therefore all agents voting their information is a NE. It is obvious that the outcome of this majority rule voting game when agents vote their information always matches with the true outcome. Therefore, ex ante, the majority rule voting mechanism perfectly aggregates information in this example.

We now consider a slightly different outcome function and our conclusion about the performance of the majority rule voting mechanism changes. Let

$$F'(a_1, ..., a_5) = F(a_1, ..., a_5)$$

for all $(a_1, ..., a_5)$ such that $\sum_{i=1}^5 a_i \neq 0$ or 5; otherwise,

$$F'(a_1, ..., a_5) = 1 - F(a_1, ..., a_5).$$

23

In other words, the outcome function here is the same as before except that when all attributes agree, the outcome switch from either 0 to 1 or 1 to 0. Again, it is easy to compute that conditional on the true outcome, (i) each attribute is correct with probability $\frac{10}{16}$ (ii) the probability of a pair of attributes being correct at the same time is $\frac{6}{16}$ and (iii) since $\frac{6}{16} < \frac{10}{16} \times \frac{10}{16}$, the attributes are pairwise negatively correlated. The analysis to establish that informative voting by all agents still constitutes a NE is the same as that in the previous example because the true outcome switch happens when no agent is pivotal. Since informative voting combined with the majority rule voting mechanism leads to outcomes that match with the previous outcome function perfectly, they do not match with the outcome function in this example. When $\sum_{i=1}^{5} a_i = 0$ or 5, information does not aggregate.

These examples demonstrate that whether a particular mechanism aggregates information perfectly depends on the structure of the outcome function. In both examples, agents have informative signals, conditional on the true outcome, signals are pairwise negatively correlated, voting informatively is a NE, but in one case information is perfectly aggregated and in the other case, information fail to always aggregate. The structure of the two outcome functions relative to their respective interpreted signals differs. F as a function of the interpreted signals (attributes) is monotonic and no externality is present. On the other hand, F' as a function of the interpreted signals is not monotonic and there are interaction terms that include externalities. The fact that information fails to aggregate exactly when externalities are present means that the majority rule voting mechanism can not resolve externalities in private information effectively.

These examples show that a different type of analysis is needed when we consider aggregating information represented by interpreted signals as opposed to generated signals. The functional relationship as opposed to the statistical relationship between outcomes and signals should be the focus. Our companion paper (Hong &Page 2005) develops this approach a lot further for studying information aggregation properties of general voting mechanisms.

7   Discussion

In this paper, we have contrasted generated and interpreted signals and demonstrated differences between them that have implications for the generality of claims of the optimality of strategies and of institutional design. At a minimum, our results suggest that modelers should relate their assumptions to the specific context: are the signals generated, interpreted, or possibly both? In addition, the distinction between the two types of signals might enable us to better understand differences between experimental and real world results. In experiments, information is often generated using the standard conditionally independent signal model. In practice, it may not be. Therefore, testing our theory using experiments may not be testing one of the most important assumptions: the assumption of conditionally independent signals.

This insight also applies to attempts to calibrate computational models with standard models of signals. These efforts may also run up against this fundamental inconsistency. In an agent based model (Tesfatsion 1997, Holland and Miller 1991), the signals are often lower dimensional projections of a larger reality. In rich, fine detailed computer models, such as the trading agent competition (Wellman et al 2003), agents do not take into account all of the information in the environment. Instead, they monitor a lower dimensional world than the one within which they interact. In spatial models and network models, something close to dimensional reduction also occurs. Agents only see what happens in a local region, thereby creating interpreted signals.

Owing to its close connections to computer science, the interpreted signal framework can be seen as a computational approach to incomplete information. Ideally, computational and mathematical models inform and complement one another (Judd 1997, Judd and Page 2004). However, our ability to align computational and mathematical models is hindered if the assumptions about signals that we make in our mathematical models are not consistent with the information that agents realize in the computational implementations of those models.

The interpreted signal framework allows for the modeling of endogenous signals. If agents want to predict correctly individually, this could lead to correlated signals as they might all learn to look at the same attributes. If agents are concerned with collective performance, such as in the case of voting to aggregate information, agents have an incentive to look at different attributes. These insights are not surprising. What is surprising is that by looking at different attributes, the correctness of agents' predictions are negatively correlated (by a specific amount), so information aggregates better than our standard models assume. Unfortunately, the number of different attributes that can be considered depends upon the dimensionality of the problem. Therefore, large groups of agents may do worse than generated signal theory predicts because they necessarily have lots of overlap. Small groups, in contrast, may do better than the generated signal theory predicts.

The choice over which attributes to include in interpretations in competitive situations, such as auctions, is among the most interesting questions to consider. Competitive situations create both types of incentives: an incentive to be correct and an incentive to be different. These two counterbalancing effects may allow us to model the structure of incomplete information as an equilibrium phenomenon as opposed to being exogenously given.

Finally, the interpretation framework permits more fine grained analysis of the link between complexity and uncertainty. The complexity-uncertainty link is also the focus of a paper by Al-Najjar, Casadesus-Masanell, and Ozdenoren (2003). They consider the continual addition of more and more attributes. As the number of attributes considered increases, the signals should improve. In their model, a problem is complex if no matter how many attributes are considered, the uncertainty never goes away. Relatedly, in our framework, within some sets in a partition, both good and bad outcomes can exist. Our formulation differs in that it emphasizes the nonlinearity and

interaction terms in the mapping from attributes to outcomes. As this mapping becomes more complex, the inference problem becomes more difficult unless interaction terms are absorbed by an interpretation. This suggests the need for deeper investigations into the relationship between the complexity of mappings from attributes to outcomes and the uncertainty of interpreted signals based on those mappings.

8   Appendix: Proofs

Proof of Claim 1. We prove the claim for $n = 2$. The proof for more general cases follows the same procedure.

Without loss of generality, assume that $\Pi^i = \left\{ \pi_1^i, \cdots, \pi_{n_i}^i \right\}$ where for each $i = 1, 2$, $n_i \geq 2$. We write the event space, Ω, in the following form which helps to visualize the proof.

$$
\begin{array}{cccc}
\pi_1^1 \cap \pi_1^2 & \pi_1^1 \cap \pi_2^2 & \cdots & \pi_1^1 \cap \pi_{n_2}^2 \\
\pi_2^1 \cap \pi_1^2 & \pi_2^1 \cap \pi_2^2 & \cdots & \pi_2^1 \cap \pi_{n_2}^2 \\
\cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot \\
\pi_{n_1}^1 \cap \pi_1^2 & \pi_{n_1}^1 \cap \pi_2^2 & \cdots & \pi_{n_1}^1 \cap \pi_{n_2}^2
\end{array}
\tag{1}
$$

Each cell above can be represented by a 2-dimensional rectangle with the property that cells so represented in the same row have the same height and cells in the same column have the same width.

To show this, we first show that for each $j = 2, ..., n_2$, that the number of events contained in each cell in any given column is proportional to the number of events in each cell in the first column:

$$
\frac{\left| \pi_1^1 \cap \pi_j^2 \right|}{\left| \pi_1^1 \cap \pi_1^2 \right|} = \frac{\left| \pi_2^1 \cap \pi_j^2 \right|}{\left| \pi_2^1 \cap \pi_1^2 \right|} = \cdots = \frac{\left| \pi_{n_1}^1 \cap \pi_j^2 \right|}{\left| \pi_{n_1}^1 \cap \pi_1^2 \right|}
\tag{2}
$$

where $|\cdot|$ denotes the cardinality of a set. By independence (recall that each event in Ω is equally likely), for all $i = 1, ..., n_1$ and all $j = 2, ..., n_2$,

$$
\frac{\left| \pi_i^1 \cap \pi_j^2 \right|}{N} = \frac{\left| \pi_i^1 \right|}{N} \cdot \frac{\left| \pi_j^2 \right|}{N}
$$

and

$$
\frac{\left| \pi_i^1 \cap \pi_1^2 \right|}{N} = \frac{\left| \pi_i^1 \right|}{N} \cdot \frac{\left| \pi_1^2 \right|}{N}
$$

Therefore,

$$
\frac{\left| \pi_i^1 \cap \pi_j^2 \right|}{\left| \pi_i^1 \cap \pi_1^2 \right|} = \frac{\left| \pi_j^2 \right|}{\left| \pi_1^2 \right|}
$$

This proves (2) above.

For each $j = 2, ..., n_2$, let the ratio in (2) be equal to $\frac{u_j}{d_j}$ where both u_j and d_j are positive integers and $\frac{u_j}{d_j}$ can not be further simplified. That is, for each $i = 1, ..., n_1$, we can write the number of events in the ith row and jth column as $\frac{u_j}{d_j}$ times the number of events in the first column of the ith row.

$$
\left| \pi_i^1 \cap \pi_j^2 \right| = \frac{u_j}{d_j} \cdot \left| \pi_i^1 \cap \pi_1^2 \right|
$$

This implies that for each $i = 1, ..., n_1$, $\left|\pi_i^1 \cap \pi_1^2\right|$ is divisible by all d_j's, $j = 2, ..., n_2$. Let d be the smallest positive integer that is divisible by all d_j's. Then for each $i = 1, ..., n_1$, there exists a unique positive integer k_i such that

$$\left|\pi_i^1 \cap \pi_1^2\right| = k_i \cdot d.$$

Thus,

$$\left|\pi_i^1 \cap \pi_j^2\right| = k_i \cdot \left(u_j \cdot \frac{d}{d_j}\right)$$

for all $i = 1, ..., n_1$ and $j = 2, ..., n_2$. Notice that $\frac{d}{d_j}$ is a positive integer in the above expression.

The above argument proves that for any $i = 1, ..., n_1$ and $j = 1, 2, ..., n_2$, $\pi_i^1 \cap \pi_j^2$ can be represented by a 2-dimensional rectangle of k_i rows (height) and $u_j \cdot \frac{d}{d_j}$ columns (width). Here we have implicitly defined $u_1 = d_1 = 1$. Therefore, each cell in (2) can be represented by a 2-dimensional rectangle such that cells in row i all have the same height of k_i and cells in column j all have the same width of $u_j \cdot \frac{d}{d_j}$. Therefore, (2) can be represented by a 2-dimensional rectangle with $\sum_{i=1}^{n_1} k_i$ rows and $\sum_{j=1}^{n_2} u_j \cdot \frac{d}{d_j}$ columns. That means, $N = \left(\sum_{i=1}^{n_1} k_i\right) \cdot \left(\sum_{j=1}^{n_2} u_j \cdot \frac{d}{d_j}\right)$. It is obvious that the number in each parenthesis is larger than 1.

Proof of Corollary 1. By Claim 1, a necessary condition for n non-trivial interpretations to be independent is that N can be written as the multiplications of n larger-than-1 integers. Thus, the largest number of independent non-trivial interpretations is bounded by the number of prime factors which is k. Now we only need to show that there exist k many non-trivial interpretations that are independent. When $N = \prod_{i=1}^{k} p_i$, Ω can be represented by a k-dimensional rectangle where the ith dimension has a length of p_i. Let Π^i be the interpretation that can only identify events along the ith dimension. Showing that these k interpretations are independent is a straightforward exercise.

Proof of Lemma 1, Lemma 2. First, observe the following identity:

$$P(g, b) + P(b, g) = P(c, i) + P(i, c)$$

Each side of this equation expresses the probability that agents disagree. Then by symmetry,

$$P(g, b) = P(c, i)$$

Second, notice that the function $x(1 - x)$ is a decreasing function of x for $x \geq \frac{1}{2}$. Therefore,

$$P(g)(1 - P(g)) < [=, >] P(c)(1 - P(c))$$

if and only if

$$P(g) > [=, <] P(c)$$

That is,
$$P(g)P(b) < [=, >] P(c)P(i)$$
if and only if
$$P(g) > [=, <] P(c)$$
Now we prove Lemma 1. If predictions are independent, then
$$P(g, b) = P(g)P(b)$$
Also, by definition, the correctness of predictions are positively correlated (independent or negatively correlated) iff
$$P(c, i) < [=, >] P(c)P(i)$$
Combine the above two equations with the identity at the beginning of the proof, we have the correctness of predictions are positively correlated (independent or negatively correlated) iff $P(g)P(b) < [=, >] P(c)P(i)$. The result then follows. Lemma 2 can be similarly proved.

Proof of Claim 2. The first part of Claim 2 is a special case of Lemma 1. Now we establish the exact expression for the negative correlation coefficient. Define variable χ_i as the indicator of the correctness of prediction $\widetilde{\phi}_i$. In other words, $\chi_i = 1$ if $\widetilde{\phi}_i$ predicts correctly and otherwise $\chi_i = 0$. Then the correlation coefficient of the correctness of the predictions, ρ, is defined as
$$\rho = \frac{Cov\,(\chi_1, \chi_2)}{\sqrt{Var\,(\chi_1)}\sqrt{Var\,(\chi_2)}}$$
Consider the following relationship between events and their probabilities. Let p_i denote the probability that $\widetilde{\phi}_i$ is correct. Then

$$
\begin{aligned}
p_1 \quad &= \Pr\,(1 \text{ agrees with } 2 \text{ and } 1 \text{ is correct}) + \Pr\,(1 \text{ disagree with } 2 \text{ and } 1 \text{ is correct }) \\
&= \Pr\,(1 \text{ and } 2 \text{ agree and both correct}) + \Pr\,(1 \text{ and } 2 \text{ disagree and } 2 \text{ is incorrect})
\end{aligned}
$$

And, similarly,

$$p_2 \quad = \Pr\,(1 \text{ and } 2 \text{ agree and both correct}) + \Pr\,(1 \text{ and } 2 \text{ disagree and } 2 \text{ is correct})$$

Add the two equations, we have

$$p_1 + p_2 \quad = 2\Pr\,(\text{both agree and correct}) + \Pr\,(1 \text{ and } 2 \text{ disagree})$$

The above equation holds without any specific assumptions. Now we assume that $p_i = p_j = p$ and that predictions $\widetilde{\phi}_1$ and $\widetilde{\phi}_2$ are independent. Then

$$\Pr\,(1 \text{ and } 2 \text{ disagree}) \quad = 2P(g)P(b) = 2P(g)(1 - P(g))$$

So,

$$\Pr(\text{both agree and correct}) = p - P(g)(1 - P(g))$$

We can now compute the correlation coefficient of the correctness of predictions, ρ.

$$Var(\chi_i) = E\chi_i^2 - (E\chi_i)^2 = p - p^2$$
$$Cov(\chi_1, \chi_2) :=: E(\chi_1\chi_2) - (E\chi_1)(E\chi_2) :=: p - P(g)(1 - P(g)) - p^2$$

Therefore,

$$\rho = \frac{p - P(g)(1 - P(g)) - p^2}{p - p^2} = 1 - \frac{P(g)(1 - P(g))}{p - p^2}$$

To complete the proof, note that $P(g) = \frac{1}{2}$, then

$$\rho = 1 - \frac{1}{4(p - p^2)}$$

Proof of Claim 3. We need to show

$$P(g)^2 < P(g, g)$$

Here, $P(g, g)$ denote the joint probability of both agents predicting g. We know

$$P(g) = P(G)p + P(B)(1 - q)$$

Now we compute $P(g, g)$. Since predictions are conditionally independent,

$$P(g, g) = P(G)p^2 + P(B)(1 - q)^2$$

Therefore,

$$P(g, g) - P(g)^2$$
$$= P(G)p^2 + P(B)(1 - q)^2 - [P(G)p + P(B)(1 - q)]^2$$
$$= P(G)P(B)(p + q - 1)^2$$

Since the predictions are experience generated and informative,

$$p + q > 1$$

This means that

$$P(g, g) - P(g)^2 > 0$$

Therefore, predictions are unconditionally positively correlated.

Proof of Claim 4. We prove this claim by way of contradiction. Suppose both conditional distributions of predictions are not negatively correlated. Then

$$P(g, g \mid G) \geq p^2$$

and
$$P(g, g \mid B) \geq (1 - q)^2$$

Since predictions are independent,
$$P(g)^2 = P(g, g)$$

By definition,
$$P(g) = P(G)p + P(B)(1 - q)$$

and
$$P(g, g) = P(G)P(g, g \mid G) + P(B)P(g, g \mid B)$$

Thus,
$$[P(G)p + P(B)(1 - q)]^2 \geq P(G)p^2 + P(B)(1 - q)^2$$

which simplifies to
$$P(G)P(B)(p + q - 1)^2 \leq 0$$

Since
$$p + q > 1$$

a contradiction.

Proof of Claim 6. Let $\delta \in \{0, 1\}$. Let J denote a subset of $\{1, ..., K\}$. For a given δ and a given $J \subseteq \{1, ..., K\}$, consider the event

$$\left\{ \omega \in \Omega : \tilde{O}(\omega) = \delta, \ \tilde{\phi}_j(\omega) = \delta \text{ for } j \in J, \ \tilde{\phi}_j(\omega) = 1 - \delta \text{ for } j \in J^C \right\}.$$

If $\left| J^C \right|$ is even, identify the above event with the following event in the example:

$$\left\{ \begin{array}{c} (\theta; x_1, ..., x_K) \in \{0, 1\} \times \{1, ..., m_1\} \times ... \times \{1, ..., m_K\} : \\ \theta = \delta, \ x_j \leq r_j \text{ for } j \in J, \ x_j > r_j \text{ for } j \in J^C \end{array} \right\}$$

If $\left| J^C \right|$ is odd, identify it with the following event in the example:

$$\left\{ \begin{array}{c} (\theta; x_1, ..., x_K) \in \{0, 1\} \times \{1, ..., m_1\} \times ... \times \{1, ..., m_K\} : \\ \theta = 1 - \delta, \ x_j \leq r_j \text{ for } j \in J, \ x_j > r_j \text{ for } j \in J^C \end{array} \right\}$$

This identification is valid for two reasons. First, the probabilities of these events are equal. By the assumption of conditional independence and the assumption that each signal i is correct with probability $\frac{r_i}{m_i}$ conditional on each outcome,

$$P\left(\left\{ \omega \in \Omega : \tilde{O}(\omega) = \delta, \ \tilde{\phi}_j(\omega) = \delta \text{ for } j \in J, \ \tilde{\phi}_j(\omega) = 1 - \delta \text{ for } j \in J^C \right\} \right)$$
$$= \prod \lim_{j \in J} P\left(\tilde{\phi}_j(\omega) = \delta \mid \tilde{O}(\omega) = \delta \right) \prod_{j \in J^C} P\left(\tilde{\phi}_j(\omega) = 1 - \delta \mid \tilde{O}(\omega) = \delta \right) \cdot P\left(\tilde{O}(\omega) = \delta \right)$$
$$= \frac{1}{2} \prod_{j \in J} \frac{r_j}{m_j} \prod_{j \in J^C} \left(1 - \frac{r_j}{m_j} \right)$$

31

By the structure of our example,

$$P\left(\left\{\begin{array}{c}(\theta;x_1,...,x_K)\in\{0,1\}\times\{1,...,m_1\}\times...\times\{1,...,m_K\}:\\ \theta=\delta,\ x_j\le r_j\text{ for }j\in J,\ x_j>r_j\text{ for }j\in J^C\end{array}\right\}\right)$$
$$=P\left(\left\{\begin{array}{c}(\theta;x_1,...,x_K)\in\{0,1\}\times\{1,...,m_1\}\times...\times\{1,...,m_K\}:\\ \theta=1-\delta,\ x_j\le r_j\text{ for }j\in J,\ x_j>r_j\text{ for }j\in J^C\end{array}\right\}\right)$$
$$=\tfrac{1}{2}\prod_{j\in J}\tfrac{r_j}{m_j}\prod_{j\in J^C}\left(1-\tfrac{r_j}{m_j}\right)$$

Second, the events of the original environment Ω we considered are the only events relevant to our interpreted signals. So from our proof so far, these events are represented by combining events in the attribute vector space of our example. By the definition of f and $s_1,...,s_K$ in our example, if $\left|J^C\right|$ is even,

$$\left\{\begin{array}{c}(\theta;x_1,...,x_K)\in\{0,1\}\times\{1,...,m_1\}\times...\times\{1,...,m_K\}:\\ \theta=\delta,\ x_j\le r_j\text{ for }j\in J,\ x_j>r_j\text{ for }j\in J^C\end{array}\right\}$$
$$=\left\{\begin{array}{c}(\theta;x_1,...,x_K)\in\{0,1\}\times\{1,...,m_1\}\times...\times\{1,...,m_K\}:\\ f=\delta,\ s_j=\delta\text{ for }j\in J,\ s_j=1-\delta\text{ for }j\in J^C\end{array}\right\}$$

If $\left|J^C\right|$ is odd,

$$\left\{\begin{array}{c}(\theta;x_1,...,x_K)\in\{0,1\}\times\{1,...,m_1\}\times...\times\{1,...,m_K\}:\\ \theta=1-\delta,\ x_j\le r_j\text{ for }j\in J,\ x_j>r_j\text{ for }j\in J^C\end{array}\right\}$$
$$=\left\{\begin{array}{c}(\theta;x_1,...,x_K)\in\{0,1\}\times\{1,...,m_1\}\times...\times\{1,...,m_K\}:\\ f=\delta,\ s_j=\delta\text{ for }j\in J,\ s_j=1-\delta\text{ for }j\in J^C\end{array}\right\}$$

In other words, the outcome function $\widetilde{O}$ and its associated interpreted signals $\widetilde{\phi}_1,...,\widetilde{\phi}_K$ are identical to the outcome function f and its associated interpreted signals $s_1,...,s_K$ once Ω is appropriately represented by the attribute vector space of our example.

References

[1] Al-Najjar, N., R. Casadesus-Masanell and E. Ozdenoren (2003) "Probabilistic Representation of Complexity", *Journal of Economic Theory* 111 (1), 49 - 87.

[2] Aragones, E., I. Gilboa, A. Postlewaite, and D. Schmeidler (2005) "Fact-Free Learning", *The American Economic Review* 95 (5), 1355 - 1368.

[3] Barwise and Seligman, (1997) *Information Flow: The Logic of Distributed Systems* Cambridge Tracts In Theoretical Computer Science, Cambridge University Press, New York.

[4] Feddersen, T. and W. Pesendorfer (1997) "Voting Behavior and Information Aggregation in Elections with Private Information", *Econometrica* 65 (5), 1029 -1058.

[5] Fryer, R. and M. Jackson (2003), Categorical Cognition: A Psychological Model of Categories and Identification in Decision Making, Working Paper, California Institute of Technology.

[6] Holland, J. and J. Miller (1991) "Artificial Agents in Economic Theory", *The American Economic Review Papers and Proceedings* 81, 365 - 370.

[7] Hong, L. and S. Page (2001) "Problem Solving by Heterogeneous Agents", *Journal of Economic Theory* 97, 123 - 163.

[8] Hong, L. and S. Page (2005), Diversity and Complexity in Information Aggregation, Working Paper, University of Michigan.

[9] Judd, K. (1997) "Computational Economics and Economic Theory: Complements or Substitutes?" *Journal of Economic Dynamics and Control.*

[10] Judd, K. and S. Page (2004) "Computational Public Economics", *Journal of Public Economic Theory* forthcoming.

[11] Klemperer, P. (2004) *Auctions: Theory and Practice* Princeton University Press.

[12] Ladha, K. (1992) "The Condorcet Jury Theorem, Free Speech, and Correlated Votes", *American Journal of Political Science* 36 (3), 617 - 634.

[13] Miller and Page (2007) *Complex Adaptive Systems: An Introduction to Computational Models of Social Life* Princeton University Press.

[14] Nisbett, R. (2003) *The Geography of Thought: How Asians and Westerners Think Differently...and Why* Free Press, New York.

[15] Page, S. (2007) *The Difference: How the Power of Diversity Creates Better Firms, Schools, Groups, and Societies* Princeton University Press.

[16] Pearl, Judea (2000) *Causality* New York: Oxford University Press.

[17] Stinchecombe, A. (1990) *Information and Organizations* California Series on Social Choice and Political Economy I University of California Press.

[18] Tesfatsion, L. (1997) "How Economists Can Get A-Life" in *The Economy as a Complex Evolving System II* W. Brian Arthur, Steven Durlauf, and David Lane eds. pp 533–565. Addison Wesley, Reading, MA.

[19] Valiant, L.G. (1984) "A Theory of the Learnable" *Communications of the ACM,* 17(11),1134-1142.

[20] Wellman, M.P., A. Greenwald, P. Stone, and P.R. Wurman (2003) "The 2001 Trading Agent Competition" *Electronic Markets* 13(1).