Identifying Wisdom (of the Crowd): A Regression Approach

Jonathan Libgober *

June 16, 2024

Abstract

Experts in a population hold (a) beliefs over a state and (b) beliefs over the population's belief distribution. If these are generated via Bayesian updating from a common prior using signal observations from a fixed information structure, a linear regression using the experts' beliefs identifies the information structure, provided there are no more states than signals. Furthermore, an eigenvector equation derived from the experts' beliefs identifies the prior. Thus, the ex-ante informational environment (i.e., how signals are generated) can be determined using ex-post data (i.e., the experts' beliefs). I interpret these findings and also discuss identification when states outnumber signals.

*University of Southern California. Email: libgober@usc.edu. This is a revision of "Hypothetical Beliefs Identify Information," first circulated in May 2021. I thank Alex Bloedel, Juan Carrillo, Nicolas Lambert, Xiaosheng Mu, Mallesh Pai, Larry Samuelson, and the editor and referees for helpful comments, as well as the hospitality of the Cowles Foundation and Yale University which hosted me during part of this research.

1 Introduction

Consider a population of experts independently forming opinions over certain possible events—for instance, meteorologists predicting the weather, medical professionals diagnosing a patient, or consultants determining the profitability of an investment. Whenever a new prediction confronts the group, each expert's belief corresponds to a Bayesian update from a fixed common prior using signals drawn independently conditional on the state according to a fixed information structure.

This paper is interested in the problem of how an outside analyst (e.g., an econometrician) facing such a situation may be able to learn about the underlying *informational environment*—i.e., (1) the (initial) prior over the set of possible states (e.g., weather events, diseases, success outcomes); and (2) the Blackwell experiment—i.e., the function which maps each state of the world to a distribution over the set of possible signals. The main contributions of this paper are to describe (1) how simple linear regressions enable the analyst to recover (parts of) the Blackwell experiment, (2) how this procedure together with an eigenvector characterization allows the analyst to learn the entire informational environment, and (3) the mathematical restrictions implied by the aforementioned setup.

The goal of identifying (at least some part of) informational environments has precedent in economics—a non-exhaustive list of work motivated by this task includes Arieli and Mueller-Frank (2017); Lu (2016); Miyashita (2022); Lomys and Tarantino (2022); Jakobsen (2021). Relative to these papers, the defining features of my exercise are (a) the setting of interest and (b) the data available. To elaborate on the latter, I assumes the analyst has access to two types of data: First, the beliefs of each expert regarding the relevant state, which I refer to as "state beliefs"; and second, the beliefs of each expert regarding the distribution of his or her peers' beliefs, which I refer to as "signal beliefs," foreshadowing that these beliefs pertain to the population's signals determined by the Blackwell experiment. Signal beliefs are second-order beliefs in the context of the specific setting studied in this paper. Note that an expert's signal beliefs will typically differ from the true distribution of signals in the population whenever the expert does not know the true state.

The environment which this paper focuses on is also studied in the active literature on *crowd wisdom*. In this literature, the data available is often precisely the kind of data I consider, and my basic formal framework is borrowed from Prelec et al. (2017). To avoid trivialization, this literature assumes—as I do—that it would be prohibitive for an analyst to elicit the distribution over signals conditional on the state itself. This could reflect limited memory among the group, or simply that experts only form beliefs "as-if" updating from a common prior, without direct knowledge of the prior or Blackwell experiment.[1] Instead, the analyst makes inferences based on data that would emerge after a single state is realized and beliefs are subsequently formed.

[1] An example from Prelec et al. (2017) is the question of whether Philadelphia is the capital of Pennsylvania. The questions (i) "what probability do you assign to Philadelphia being the capital of Pennsylvania?" and (ii) "given a probability p, what probability would you assign to a randomly selected person believing that Philadelphia is the capital of Pennsylvania with probability p?" are fairly straightforward to formulate; it may be harder to answer the question "if Philadelphia were the capital of Pennsylvania, what probability would you assign to someone believing Philadelphia were the capital with probability p?"

The main difference between this literature and my paper is that the former typically focuses on determining the true *state of the world* given the group's ex-post beliefs, rather than the underlying informational environment.

But there are many reasons why the informational environment may itself be of interest to an analyst. First, and perhaps most immediately, identifying the Blackwell experiment nests the problem of identifying the true state. If the analyst has data on the realized distribution over beliefs in the group (rather than just the beliefs themselves), then the true state could be determined simply by matching the observed signal distribution to the one predicted by the Blackwell experiment. Second, knowledge of the Blackwell experiment generating expert information would enable the analyst to determine the state with significantly less data in other instances when the state is independently drawn[2]—the analyst would only need the population's signal distribution. Third, the prior itself may be of interest, again since it could be of use in independently drawn problems. Finally, the analyst may be independently interested in assessing how informed the population is—to determine, for instance, whether any individual expert would benefit from access to better or additional information sources.

Typically, it will not be possible to determine either the information structure or the underlying state using state beliefs alone. Prelec et al. (2017) points out that the ex-post (state) beliefs of the population may fail to reflect the truth if instead they simply reflect a concentrated prior (see also Arieli et al. (2020) for a similar point). If one state, say θ^*, is weighted heavily by the prior and the signal observed is weak, then the experts will always simply believe θ^* is more likely, even when it is not the true state. The issues related to inference of the information structure are slightly different, but again stem from the observation that the prior is confounding—a given set of beliefs over the state could emerge given *any* prior in the interior of the convex hull of these beliefs. In turn, different priors may very well correspond to different Blackwell experiments.[3] Thus, beliefs over the state are insufficient to pin down the Blackwell experiment.

I first describe the linear regression procedure that can recover the Blackwell experiment. The signal beliefs of the population play a crucial role. The simplest case is when there are at least as many signals as states. Here, the procedure is to regress a vector of signal beliefs—where each element of this vector denotes an expert's belief that a peer chosen uniformly at randomly will have observed some fixed signal, say s_i–on the matrix of state beliefs. The resulting object is the probability that the fixed signal s_i is observed in every possible state.

Turning to the prior, I show that this object has a geometric interpretation as an eigenvector

[2]There are many examples where states could be plausibly treated as independent and identically distributed. Examples of problems studied by Prelec et al. (2017) include dermatologists assessing lesions and professionals assessing art; for both of these, there are natural interpretations of the prior (the proportion of lesions that are malignant, the fraction of representative art pieces that have high market value) which may be of interest, and if each problem is drawn from a homogenous population, then this assumption would appear plausible.

[3]For instance, suppose there are two possible states, $\{1, 2\}$ and two possible signals, $\{1, 2\}$. Suppose that following signal 1, state 1 has posterior probability 3/4, and under signal 2, state 1 has posterior probability 1/4. If each state is ex-ante equally likely, then these ex-post beliefs are consistent with the signal being equal to the state with probability 3/4. But it may be that the prior probability of state 1 is higher, with these same ex-post beliefs being induced. The same information structure would not yield the same ex-post beliefs with the different prior.

of a matrix derived from the martingale condition on beliefs; in particular, the prior is the unique eigenvector corresponding to eigenvalue 1, normalized so that the entries sum to 1. Under mild assumptions, one can back out both the information structure and the prior from the possible state beliefs, together with the signal beliefs. I also discuss what it means for these conditions to fail, and what can be done when they do.

These results implicitly assume there are more signals than states. If not, then this method cannot be used. The issue is familiar from linear regression, where this takes the form of an identification problem that emerges if there are more explanatory variables than observations. It turns out that one proposal from statistics for how to address this problem can allow the analyst to learn something in my setting as well—I describe a regularization process which essentially allows us to perform the inversion step required by linear regression. The process is known as "ridge regression." The idea is to add a small perturbation to the singular matrix to avoid the invertibility issue that arises with the identification failure. I show that this procedure identifies a subspace on which the information structure must lie, up to a vector in the null space in the matrix of beliefs. In some cases this restriction may allow the analyst to determine the information structure. But perhaps more surprisingly, I show that even if the procedure does not identify the information structure, it nevertheless does induce the same eigenvector interpretation of the prior (even with the "incorrect" information structure), providing another path for it to be recovered.

My analysis suggests geometric relationships between the geometric relationship between the ex-ante determinants of the experts' environment (i.e., the prior and Blackwell experiment) and the ex-post data (i.e., the beliefs observed by the analyst). I also ask whether it is restrictive to assume that belief matrices correspond to informational environments. I show that the dimensionality of the set of signal beliefs which can emerge, given a fixed matrix of beliefs over the state, is typically smaller than the set of signal belief matrices which can emerge based only on the requirement that probabilities are non-zero and sum to one—sometimes significantly so. I interpret this as a word of caution, as "most" possible signal belief matrices will not correspond to any informational environment. I then ask whether the same dimensionality gap exists if the analyst allows for non-common priors. I relax the assumption that all subjects use a common prior by allowing priors to depend on signal realizations themselves (even holding fixed the relevant Blackwell experiment). In this case, there are no restrictions at all—any signal beliefs can be rationalized.

As the setting I study here has been studied in past work as well, my proposal is not the only one aimed at arriving at the informational environment. While the focus of Prelec et al. (2017) is on inferring the state, some of the results described also speak to the problem of inferring the relevant Blackwell experiment; a point of departure is that the latter is my paper's main focus. However, Prelec and McCoy (2022) contemporaneously describes a different eigenvector characterization which can be used to determine the informational environment; to avoid distractions, I defer an in-depth comparison to this alternative in Section 6. The extent to which regression is useful in practice may hinge on some considerations beyond the scope of the current paper—for instance, real-world data may exhibit properties which make one procedure preferred to the other.

Independently of these concerns, however, I discuss two primary reasons the regression procedure may have use in practice. First, regression allows the analyst to identify part of the information structure even with incomplete data on signal beliefs, unlike the eigenvector characterization. Sometimes the difference in elicitation requirements can be quite dramatic—suppose the "state" is the answer to a yes-no question, and each "signal" refers to a probability assessment over the answer, which may take n possible values. The regression procedure I propose would allow the analyst to learn the state by asking each expert only two questions; approaches aimed at eliciting all signal and state beliefs would require that the experts answer n questions. If n is large, regression therefore significantly reduces demands on the experts. As shorter surveys may be more reliable and easier to conduct for various reasons, such a gain could indeed be significant. Second, as each proposal should in principle lead to the same end, the extent to which they differ can be used by the analyst to assess robustness. On the other hand, there are cases where the approach described in Prelec and McCoy (2022) is more direct and simpler to implement than regression, in particular when there are more states than signals. Thus, I conclude the approaches are complementary.

Beyond these considerations, however, part of the goal for this paper is to clarify and articulate the mathematical structure underlying expert beliefs in crowd wisdom settings. These observations may prove useful in subsequent analysis of related environments. While many of the objects I study formally arise thanks to this particular setting, the relationship between ex-post beliefs to ex-ante informational environments has attracted significant attention following Kamenica and Gentzkow (2011). This line of work typically reduces information structures to the distribution over posterior beliefs induced by them. While often this distribution is ultimately the necessary object to solve certain problems (e.g., in Kamenica and Gentzkow (2011)'s case, optimal communication with Sender commitment when facing a single Receiver), the natural question raised is how the ex-ante informational environment relate to the ex-post beliefs more generally. The regression procedure and the geometric structure I identify describes the correspondence between these two objects explicitly, and thus may be of interest in other settings where this question is of interest as well.

2 Related Literature

This paper borrows the framework of Prelec et al. (2017), on how an analyst may be able to infer a true state using predictions regarding the reports of other members of a population. Their use of higher-order beliefs built on Prelec (2004), which proposed using this additional data as part of the *Bayesian Truth Serum* in order to induce *truthful reporting* (as opposed to aggregating information) in cases where the objective truth is unknown (in contrast to much of the literature on scoring rules). Prelec et al. (2017) considered instead how an analyst might infer the true state using a mechanism referred to as the *surprisingly popular algorithm*, showing theoretically how it could aggregate information in the two-state case, as well as some other special cases. While their focus is not on inferring the informational environment, some of their analysis sheds light on this problem. For instance, as mentioned Prelec et al. (2017) demonstrated that knowledge of the state-beliefs alone would not suffice to recover the true state; results illustrating the general difficulty of learning the true state *without* higher-order beliefs information also appear in Arieli et al. (2020).

A contemporaneous paper, Prelec and McCoy (2022), filled in gaps from Prelec et al. (2017) by providing an elegant argument for how to recover the informational environment without the restrictions on the number of signals and states. I defer a detailed discussion of their procedure to Section 6, after my results are presented, as the details are helpful in understanding and interpreting the precise distinction. It is worth mentioning that while the primitives of my paper are precisely as in Prelec and McCoy (2022), the surprisingly popular algorithm itself does *not* make use of state beliefs, but rather (a) signal beliefs and (b) the distribution over signals in the population (although, as mentioned, this algorithm does require restrictive assumptions not made by either this paper or Prelec and McCoy (2022)). Still, Prelec et al. (2017) discuss how an analyst may use state beliefs in order to infer the state for the general-state and general-signal case.

Among work in this literature, most related to my approach is the contemporaneous paper Chen et al. (2021). This framework takes as a primitive an infinite population and a probability distribution over states and signals, without presuming the structure of an informational environment as described in Definition 1 (whereas this paper is focused on recovering this experiment when it exists). The main results show how, in the large sample limit, inverting a matrix of average beliefs allows the analyst to asymptotically recover the true state using predictions of the average population beliefs, doing so using assumptions which roughly speaking allow law of large number arguments to be applied. An important property of this procedure is that it can recover the true state under weaker assumptions than those imposed by Prelec et al. (2017). On the one hand, Chen et al. (2021)'s matrix inversion is reminiscent of regression; but unlike this paper and Prelec and McCoy (2022), their proposal with a general number of states essentially requires as many distinct beliefs as states to be observed. My assumptions on signals and states, by contrast, implies a simple inversion need not be feasible.

I do not consider the practical elicitation of the matrices B and Q, as was the focus of Prelec (2004). See, for instance, Cvitanić et al. (2019) or Witkowski and Parkes (2012b) which suggests ways to incentive-compatibly elicit beliefs with non-verifiable states given natural limitations an analyst might face (especially finiteness of the population). Hussam et al. (2022) uses the Witkowski and Parkes (2012b) method in a field experiment in India aimed at identifying high-ability entrepreneurs. Robustness of the Bayesian Truth Serum with respect to common priors is also discussed in Witkowski and Parkes (2012a), although recall that in the context of my exercise, relaxing common priors too strongly essentially implies no restrictions at all on B and Q.

3 Preliminaries

A continuum population of mass one forms beliefs over a finite state space Θ. The belief each individual holds over Θ is summarized by a *belief type*. Let S denote the set of belief types which are possible in the population; I assume throughout that $|S| < \infty$. Write $b_{s,\theta}$ for the probability an individual with belief type $s \in S$ assigns to the state being $\theta \in \Theta$; I call such a belief a *state belief* when necessary to avoid confusion with signal beliefs (introduced below). I denote the $|S|$-by-$|\Theta|$ matrix of beliefs (over the state) by $B = (b_{s,\theta})_{s \in S, \theta \in \Theta}$. Here, rows are belief types, and columns are states $\theta \in \Theta$ over which the beliefs are formed. I refer to the matrix B as the state belief matrix. I assume throughout that B does not have a zero vector for any column, and that B is full rank.[4]

Each individual forms not only a belief over the state, but also a conjecture of the distribution of belief types in the population. I call such beliefs *signal beliefs*. Let $q_{s,\tilde{s}}$ denote the probability an individual i with belief type s assigns to an individual j being of belief type $\tilde{s}$, when j is drawn uniformly at random from the population. I let $Q = (q_{s,\tilde{s}})_{s \in S, \tilde{s} \in S}$ denote the corresponding $|S|$-by-$|S|$ matrix, where rows index belief types of individual i, and the columns index the belief types the randomly drawn individual j observes in this exercise (so that Q is row-stochastic). The row corresponding to s gives the expected distribution over belief types in the population held by an individual of belief type s. I refer to the matrix Q as the signal belief matrix.[5]

I refer to the combination of B and Q as the *belief landscape*. In this paper, I take a belief landscape as a primitive. I will call a belief landscape *plausible* if B and Q are row stochastic and non-negative. If either of these conditions are violated, then it is immediately apparent that it is not possible to interpret B and Q as I have so far, namely reflecting Bayesian probabilities.

In principle, a belief landscape could be otherwise arbitrary. I am interested in determining when the belief landscape is consistent with the population Bayesian updating based on conditionally IID signals from a fixed Blackwell experiment and common prior—and in particular, identifying these objects. A *Blackwell experiment* or *information structure* is a function $\mathcal{I} : \Theta \to \Delta(S)$. Let $\mathcal{I}(\theta)[s]$ denote the probability that the decision-maker observes s in state θ. Given a prior $p \in \Delta(\Theta)$, a decision-maker who observes a signal s from a Blackwell experiment can update beliefs over each state $\theta \in \Theta$ via Bayes rule; that is, if the decision-maker's belief type s is determined by the information structure $\mathcal{I}$, then:

$$b_{s,\theta} = \frac{p(\theta)\mathcal{I}(\theta)[s]}{\sum_{\tilde{\theta} \in \Theta} p(\tilde{\theta})\mathcal{I}(\tilde{\theta})[s]}. \tag{1}$$

In addition, $(q_{s,s_1}, \ldots, q_{s,s_n})$ is pinned down as well. If all individuals have access to the same Blackwell experiment, then $\mathcal{I}(\theta)[\tilde{s}]$ is the fraction of the population that obtains signal $\tilde{s}$ in state θ (and thus the probability that a randomly drawn individual has belief type $\tilde{s}$). Thus, if an individual of belief type s holds beliefs $(b_{s,\theta})_{\theta \in \Theta}$, then the law of total probability implies:

$$q_{s,\tilde{s}} = \sum_{\theta} \mathcal{I}(\theta)[\tilde{s}] b_{s,\theta} \tag{2}$$

The main question of this paper is whether one can take the population's beliefs to be generated by all individuals having access to the same $\mathcal{I}$ and updating from the same prior p.

[4]A matrix is full rank if its rank is equal to the lesser of the number of columns and the number of rows.

[5]Prelec and McCoy (2022) refers to the matrix Q as the meta-knowledge matrix.

Definition 1. A belief landscape (B, Q) is generated by an informational environment $(\mathcal{I}, p)$ if $b_{s,\theta}$ and $q_{s,\tilde{s}}$ can be derived using prior p and taking the belief type s to be drawn according to $\mathcal{I}$ using (1) and (2).

This paper is concerned with identifying the informational environment; however, as stated in the introduction, were the analyst to observe the fraction of the population comprising each belief type, then the state itself could be inferred as well by comparing this fraction of population with belief type s to $\mathcal{I}(\theta)[s]$. This question is often the primary focus of the crowd wisdom literature (reviewed more completely below), with the identification of information being secondary, but I return to this observation when I compare my approach to others from this literature.

Example 1. I walk through a simple example to illustrate the key definitions from the previous section. Suppose $\Theta = \{\theta_1, \theta_2, \theta_3\}$, and suppose there is an initial prior over Θ that assigns probability $p(\theta_i)$ to state θ_i. Consider the following information structure: With probability $\varepsilon \in (0, 1)$, a "null signal" is drawn. With probability $1 - \varepsilon$, the state is observed. Using the above formalism, one can write the state belief matrix as follows:

$$B = \begin{pmatrix} p(\theta_1) & p(\theta_2) & p(\theta_3) \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Notice that, consistent with the definitions above, the rows refer to different signals an individual might observe, and the columns refer to different states.

The signal belief matrix corresponding to this information structure is therefore:

$$Q = \begin{pmatrix} \varepsilon & (1-\varepsilon)p(\theta_1) & (1-\varepsilon)p(\theta_2) & (1-\varepsilon)p(\theta_3) \\ \varepsilon & 1-\varepsilon & 0 & 0 \\ \varepsilon & 0 & 1-\varepsilon & 0 \\ \varepsilon & 0 & 0 & 1-\varepsilon \end{pmatrix}.$$

To understand why the signal belief matrix takes this form, note that if someone were to observe no information, then she would still understand this event to be an ε probability occurrence; this corresponds to the entry in the upper left corner. On the other hand, since she obtains no information following the uninformative signal, the probability she would then assign to observing each of the other three signals, *conditional on observing an informative signal*, simply coincides with the prior distribution. Since the probability of observing such a signal in the first place is $1 - \varepsilon$, she therefore assigns $(1 - \varepsilon)p(\theta_i)$ to the event that she would have seen the signal revealing state θ_i. Following every other signal, while she *would* know the state, she would also understand that there was an ε chance that she would have remained uninformed. Thus, on the one hand, she assigns probability ε to observing the uninformative signal, but also assigns probability 0 to observing any signal that would reveal any other state. Importantly, this matrix is row stochastic, each row itself being a probability distribution over S.

3.1 Discussion

The framework above assumes a continuum population. This is for expositional simplicity. As long as all $|S|$ signals are observed, one can immediately write the matrix B. The matrix Q is slightly harder to conceptualize, but also an object considered in the crowd wisdom literature; it represents the probability distribution over signals that an individual would expect given a conditionally independent draw from the same experiment. Equivalently, it represents the proportion of each belief type a decision-maker would expect to observe given an infinite population.

More substantive is that the matrix B and Q are observed without noise. If the analyst had access to an infinite sample, this could indeed be reasonable, but in practice one would expect B and Q to be observed only imprecisely. On the one hand, in the method outlined below, a small perturbation in B and Q will correspondingly lead to only a small change in the solved $\mathcal{I}$; thus, a slight error in B or Q will translate to a slight error in $\mathcal{I}$ (although how slight will itself depend on the accuracy B and Q).

Note that I also assume that the matrix B does not have any column equal to 0. This assumption is not substantive. If posterior beliefs always put probability 0 on some state, then given updating from a common prior and a common Blackwell experiment, the prior probability on this state must also be 0. As my main case of interest is in when this is how beliefs are generated, the case where this assumption is violated is not relevant for my main results. The full rank assumption is more restrictive, and I comment on it in more detail in Appendix B.1. There, I argue that the assumption that B is full rank is generic, in a sense I describe, and show how the regression procedure I propose can be extended in cases where columns of B feature linear dependencies.

4 Identifying Information if $|S| \geq |\Theta|$

The previous section showed how to construct both kinds of belief matrices from an information structure given a prior belief; the state belief matrix B is computed from Bayes rule, whereas the signal belief matrix additionally can be derived using rules of conditional and total probabilities. The goal of this paper is to go in the opposite direction. A well-known result (originally due to Aumann and Maschler (1995)) states that, given a prior belief and a set of posteriors, there is an essentially unique information structure inducing these beliefs. The main question I study in this paper is how to infer the decision-maker's information structure and prior using data from the (state and signal) belief matrices.

The following result provides an answer to this question when $|S| \geq |\Theta|$, which we assume throughout this Section. To state it, we consider the Blackwell experiment $\mathcal{I}$ as a matrix, whose rows are indexed by states and columns by signals (and where each entry yields the probability of observing a given signal in a given state). Say that a parameter is *uniquely identified* if its value can be determined from the analyst's ex-post data:

Theorem 1. *Suppose B and Q are generated by an informational environment $(\mathcal{I}, p)$ with $|\Theta| \leq |S|$.*[6]

A The jth column of $\mathcal{I}$ is uniquely identified by B and the jth column of Q. In particular:

$$\mathcal{I} = (B^T B)^{-1} B^T Q. \tag{3}$$

B If, for the $\mathcal{I}$ given by (3), $\mathcal{I}B$ is irreducible,[7] *then p is uniquely identified by B and Q. Furthermore, the irreducibility condition holds for generic Blackwell experiments.*[8]

I discuss the two parts of this theorem separately. The first part shows that the information structure arises from *regressing a given column of the signal belief matrix on the columns of the state belief matrix.* As I explain, the novel parts of the first part of Theorem 1 relate to the regression interpretation of the Blackwell experiment, together with the fact that the jth column of $\mathcal{I}$ is identified by B and a single column of Q. However, Prelec et al. (2017) and Prelec and McCoy (2022) discuss distinct ways of arriving at $\mathcal{I}$ if Q is fully specified in addition to B.

I illustrate the regression procedure using the truth-or-noise information structure from the previous section. If $p(\theta_1) = p(\theta_2) = p(\theta_3) = 1/3$, then:

$$(B^T B)^{-1} B^T = \begin{pmatrix} \frac{1}{4} & \frac{11}{12} & -\frac{1}{12} & -\frac{1}{12} \\ \frac{1}{4} & -\frac{1}{12} & \frac{11}{12} & -\frac{1}{12} \\ \frac{1}{4} & -\frac{1}{12} & -\frac{1}{12} & \frac{11}{12} \end{pmatrix}$$

For an arbitrary vector v, $(B^T B)^{-1} B^T v$ identifies the coefficients β in the equation $v = B \cdot \beta$ which provide the "best-fit" (according to the least square error). Note that in this example, the rows of $(B^T B)^{-1} B^T$ sum to 1 (which turns out to be a general property). Letting $\mathbf{1}_n$ denote a vector of 1s of length n, since the first column of Q is $\varepsilon \cdot \mathbf{1}_4^T$, the coefficients corresponding to this regression are $\varepsilon \cdot \mathbf{1}_3^T$. This is exactly the vector of probabilities of observing the null signal in each of the three states. On the other hand, if one were to instead consider the second column of Q:

$$(B^T B)^{-1} B^T \begin{pmatrix} \frac{1-\varepsilon}{3} \\ 1-\varepsilon \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1-\varepsilon \\ 0 \\ 0 \end{pmatrix},$$

which is precisely the vector of probabilities that the decision-maker observes the signal saying the state is θ_1 (that is, the probability that this signal is seen in each of the three states).

[6] Recall that I assume throughout the paper that B is full rank with no zero vectors in any column; in particular, this implies $B^T B$ is invertible when $|S| \geq |\Theta|$.

[7] A matrix M is irreducible if there is no permutation matrix P such that PMP^{-1} is block upper triangular.

[8] I take generic to mean that the set of Blackwell experiments which fail this condition fall within a lower-dimensional subspace of the space of all Blackwell experiments. I will use this notion of non-genericity (i.e., belonging to a lower-dimensional subset) again in Section 4.1.

Notice that Theorem 1, and in particular the regression interpretation, also clarifies *exactly* what is decided by each column of Q. Each column of Q gives a unique column of the matrix determining the information structure $\mathcal{I}$—in particular, the column it yields is the vector of probabilities (with different coordinates corresponding to different states) that that signal emerges.

To see why the regression interpretation is natural given the mathematical structure of the environment, consider the relationship between the state beliefs, signal beliefs, and Blackwell experiment implied by the law of total probability, (2). According to this expression, $q_{s,\tilde{s}}$—the posterior belief an expert observing s would assign to a randomly drawn peer having observed $\tilde{s}$—is linear in (a) the expert's beliefs following signal s, and (b) the probability of observing $\tilde{s}$ in each state. The latter is given by the Blackwell experiment, and crucially, does not depend on s. Now, linear regression finds a solution for $y_i = \beta \cdot x_i$, for some set of vectors x_i and scalars y_i, assuming a solution exists (and of course, if none exists, it finds the "best fitting" solution). If we let $x_i = (b_{s,\theta})_{\theta \in \Theta}$ and $y_i = q_{s,\tilde{s}}$, then setting β in the linear regression to be equal to $\mathcal{I}(\theta)[\tilde{s}]$ yields (2).

I note that if B and Q were fully specified, an argument outlined in Prelec et al. (2017) (Section 1.4) provides an alternative way of arriving at the Blackwell experiment (although not the prior). Specifically, they note that Bayes rule implies the following relationship:

$$\frac{\mathcal{I}(\theta)[s_i]}{\mathcal{I}(\theta)[s_j]} = \frac{\mathbb{P}[\theta \mid s_i]}{\mathbb{P}[\theta \mid s_j]} \frac{\mathbb{P}[s_j \mid s_i]}{\mathbb{P}[s_i \mid s_j]}. \tag{4}$$

Now, the right hand side can be determined from B and Q, implying that this data yields $|S| - 1$ distinct equations for each θ; imposing the requirement that $\sum_k \mathbb{P}[s_k \mid \theta] = 1$ yields $\mathcal{I}$. Crucially, however, this approach does not work if Q is not fully specified. For instance, suppose the analyst only sees the column of Q corresponding to some signal s^*. Setting $s_i = s^*$, one might think $\mathcal{I}(\theta)[s^*]$ could be identified by setting $s_j = S \backslash \{s^*\}$; but, doing so to attempt to infer $\mathcal{I}(\theta)[s^*]$ as per (4) requires $\mathbb{P}[s_i \mid S \backslash \{s^*\}]$, which is not a primitive.

I now consider Part (B) of Theorem 1. Note that an information structure can be identified by the prior belief and the set of posteriors induced by it (which in this case is B). However, this requires the prior to be in the interior of the convex hull of the posterior beliefs. While the regression procedure finds the information structure, it does not guarantee that the resulting prior satisfies the interiority condition.

Addressing this issue yields another insight into the geometry of information structures. Note that the probability of observing any particular signal determined from $\mathcal{I}$ and p:

$$\sum_\theta \mathcal{I}(\theta)[s] \cdot p(\theta) = \mathbb{P}[s].$$

The martingale property of beliefs states that:

$$\sum_s b_{\theta,s} \mathbb{P}[s] = p(\theta).$$

Thus, substituting in for $\mathbb{P}[s]$ and rewriting in matrix form gives the following identity:

$$B^T\mathcal{I}^T p = p \Leftrightarrow (B^T\mathcal{I}^T - I)p = \mathbf{0}.$$

This equation demonstrates that the prior is therefore a *unit eigenvector with eigenvalue 1 of the matrix* $B^T\mathcal{I}^T$ (or, by taking transposes, a left eigenvector of $\mathcal{I}B$).[9] And in fact, given the previous observation that the information structure $\mathcal{I}$ can be identified from B and Q, this shows that the prior can as well. The proof verifies that indeed this eigenvalue can be guaranteed to exist, via an appeal to the Perron-Frobenius theorem. It is worth emphasizing that the Perron-Frobenius theorem ensures that the eigenvector which satisfies this equation is *unique* (up to scaling), implying that the prior is pinned down as well.

This derivation introduces the object $B^T\mathcal{I}^T$, whose ith row and jth column is:

$$\sum_{\tilde{s}} \mathcal{I}(\theta_j)[\tilde{s}] b_{\theta_i,\tilde{s}}.$$

This expression represents the expectation, in state θ_j, of the probability assigned to state θ_i by a population member selected uniformly at random. Roughly speaking, with a more informative experiment, this should be larger for $\theta_i = \theta_j$ and smaller for $\theta_i \neq \theta_j$. By comparing the diagonal of the matrix $\mathcal{B}^T\mathcal{I}^T$ to the other entries, the analyst can get a rough sense of the accuracy of the beliefs in the population. The idea that the prior can be expressed as an eigenvector of a matrix over beliefs of a state was similarly made in Samet (1998); I defer a more detailed contrast between these arguments to Section 6.1.

A dual operation would be to instead substitute in for $p(\theta)$ instead of $\mathbb{P}[s]$; this operation is dual in the sense that the matrix obtained in the analogous eigenvector equation is $B \cdot \mathcal{I}$ in place of $B^T \cdot \mathcal{I}^T$. However, note that given the expression for $\mathcal{I}$, we have that $B^T \cdot \mathcal{I}^T = B^T B(B^T B)^{-1} Q = Q$. This result, that $\mathbb{P}[s]$ is the unit eigenvector of Q with eigenvalue 1, appeared in Prelec et al. (2013), and further exploited by Prelec and McCoy (2022); notice that the derivation of $\mathbb{P}[s]$ is thus more direct than the derivation of $p(\theta)$, as Q is a primitive of the environment. I return to a discussion of this contrast in Section 6.

The condition needed for uniqueness of the prior is that the matrix of which the prior is a Perron eigenvector is irreducible. See the discussion in Appendix B.2 for a discussion of what is identified when irreducibility fails. Briefly, if $B^T\mathcal{I}^T$ is not irreducible, then one can find a Perron eigenvector for each irreducible class of the stochastic process which $B^T\mathcal{I}^T$ defines. In this case, any convex combination between these eigenvectors will be a prior inducing B and Q.[10] Irreducibility will fail if, for instance, an agent can always distinguish between two subsets of the state space.[11] In this

[9]By unit eigenvector, in this paper I mean the entries sum to 1; such an eigenvector has a norm equal to 1 when $||x|| = \sum_i |x_i|$.

[10]Note that if $B^T\mathcal{I}^T$ is not irreducible, then Q will not be either; this follows from the observation that the signal space can be partitioned using the states that can be distinguished. Similarly, the only way a decision-maker can know a certain signal would never be observed in the population is if they also know that any state where that signal is drawn is not the true state; so, if Q is not irreducible, then neither is $B^T\mathcal{I}^T$.

[11]To clarify this most transparently, Appendix B.2 focuses on the case of deterministic information structures, where signals are generated deterministically as a function of the state. In this case, no information is conveyed within each

case, the prior can only be uniquely defined within each irreducible class, but not overall.

4.1 Are Generic B and Q Generated by Informational Environments?

A natural question that emerges from the previous discussion is how restricted the set of signal belief matrices are. Can an arbitrary plausible signal belief matrix be consistent with a state belief matrix? The answer to this question turns out to be no. For instance, suppose the belief matrix induced is the following:

$$B = \begin{pmatrix} 1/4 & 3/4 \\ 3/4 & 1/4 \end{pmatrix} \Rightarrow (B^T B)^{-1} B^T = \begin{pmatrix} -\frac{1}{2} & \frac{3}{2} \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix}. \quad (5)$$

Now, given some candidate $Q = \begin{pmatrix} a & 1-a \\ 1-b & b \end{pmatrix}$, right multiplying by Q gives:

$$\begin{pmatrix} -\frac{a}{2} + \frac{3(1-b)}{2} & -\frac{1-a}{2} + \frac{3b}{2} \\ \frac{3a}{2} - \frac{1-b}{2} & \frac{3(1-a)}{2} - \frac{b}{2} \end{pmatrix}.$$

Now, Q is plausible whenever $a, b \in [1/4, 3/4]$; however, given a value for a, only a single value for b will be consistent with a belief landscape. Thus, while any such value for a and b would yield an information structure, it does not necessarily yield one that can induce B given any prior p. This point is formalized in the next Proposition. The prior, together with B, pins down the information structure; from this, the matrix Q can always be inferred.

Note that the set of n-by-k matrices naturally form an nk dimensional linear space. I use the natural notion of dimensionality implied by this embedding. I have the following:

Proposition 1. *Consider a belief matrix B with linearly independent columns. The set of all possible Q such that B and Q are generated by an informational environment is $|\Theta|-1$ dimensional.*

By Proposition 1, for binary state informational environments, the set of Q inducing B is one dimensional, verifying the claim that given a possibly valid a, there is a unique value of b which corresponds to a valid information structure. Figure 1 shows, given the belief matrix B from (5), which a and b choices correspond to a fixed feasible prior p. For instance, $a = b = 5/8$ is the solution when $p = 1/2$. The choice of $a = b = 9/16$ is therefore invalid; nevertheless, for these choices:

$$(B^T B)^{-1} B^T Q = \begin{pmatrix} 3/8 & 5/8 \\ 5/8 & 3/8 \end{pmatrix}$$

Upon inspection, one can see that this *is* indeed a perfectly valid information structure $\mathcal{I}$, and in fact one that is symmetric. But, it is also straightforward to see that it cannot induce the belief matrix B for any prior; indeed, since B is symmetric as well, symmetry would require $p = 1/2$,

irreducible class—i.e., within each partition element—even though it is evident that agents can distinguish between elements of the partition.

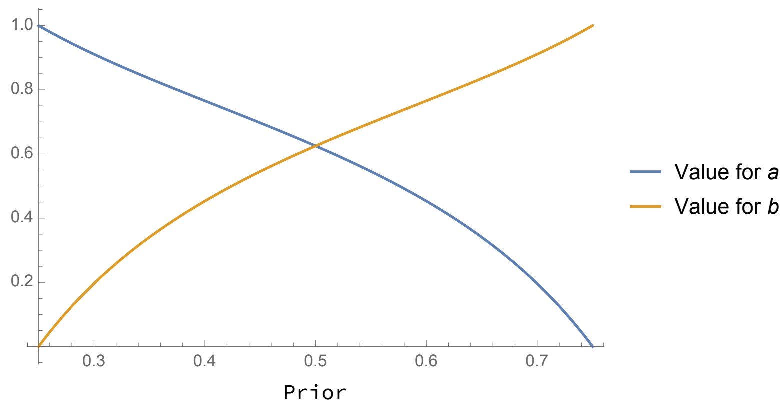

Figure 1: Value for a and b in the example which yield a valid signal belief matrix, given a prior probability of state 1 equal to $p \in [1/4, 3/4]$.

which in turn would suggest distinct beliefs given the information structure than those given by B. Thus, while the choices of $a = b = 9/16$ induce a Q matrix which is row stochastic and non-negative, and is such that $(B^T B)^{-1} B^T Q$ is an information structure, it cannot be generated by any informational environment.

One observation that is helpful for interpreting this statement is the following:

Proposition 2. *Given a row stochastic B, the space of plausible Q (as defined in Section 3) such that the regression procedure in Theorem 1 induces an information structure has dimension $|S|(|S|-1)$.*

The proof shows that if Q is a matrix rows sum to 1, then $(B^T B)^{-1} B^T Q$ will be a matrix whose rows sum to 1. The result then follows by noting that one can find an open set of plausible Q inducing some valid information structure within the set of row stochastic Q under the relative topology.[12] Note that the conditions of the theorem restrict consideration to Q such that $(B^T B)^{-1} B^T Q$ has exclusively non-negative entries; this need not hold for arbitrary plausible Q,[13] meaning that not all plausible Q necessarily yield information structures. Still, the set of Q which do is of the same dimensionality.

An implication of the previous two results is that the space of B, Q generated by informational environments is non-generic in the space of plausible B and Q. If, for instance, the Q observed by the analyst were drawn from an atomless distribution supported on the space of all plausible signal-belief matrices (say, concentrated on the true one), then the probability the observed Q would be generated by an informational environment is 0. Intuitively, this restriction is due to the prior. A matrix B and prior p determine an information structure, so the set of possible information

[12] That is, viewing a row stochastic matrix of dimension n as a subset of $\mathbb{R}^{n^2}$ and considering the relative topology on the set of row-stochastic matrices.

[13] Consider, for instance, Example 1 with $p(\theta_1) = p(\theta_2) = p(\theta_3) = 1/3$ and $\varepsilon = 1/2$. The matrix $\tilde{Q} = \begin{pmatrix} 1/2 & 1/6 & 1/6 & 1/6 \\ 1/2 & 1/2 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix}$ is plausible, but $(B^T B)^{-1} B^T \tilde{Q} = \begin{pmatrix} 25/48 & 23/48 & -1/48 & 1/48 \\ 25/48 & -1/48 & 23/48 & 1/48 \\ 13/48 & 11/48 & 11/48 & 13/48 \end{pmatrix}$.

structures inducing B has the same dimension as the set of possible priors. But the set of possible Q that could emerge (e.g., if randomly generated according to a full-support distribution) is *much* larger than the set of possible priors.

4.2 Relaxing Common Priors

Note that an assumption of this framework is that beliefs are updated from a fixed prior, p; this property, in turn, is part of what drives the dimensionality of the set of Q, since $|\Theta| - 1$ is the set of priors which induce a fixed belief matrix B (which will have linearly independent columns whenever B is full rank). This raises the question of how restrictive this assumption is; I now comment how relaxations of the common prior can rationalize any Q.

I maintain the framework that there are $|S|$ belief types in the population, with state belief matrix B and signal belief matrix Q. I emphasize that the kth row of the state belief matrix represents the kth belief type's probability assessment over Θ; the kth row of the signal belief matrix, as before, represents the kth belief type's expected proportion of belief types in the population.

However, rather than suppose that each belief type is generated by a fixed information structure $\mathcal{I}$ and prior p, consider the case where each belief type s maintains that *their* beliefs were arrived at from updating a belief-type dependent prior, say $p^s \in \Delta(\Theta)$, and belief-type-dependent information structure, say $\mathcal{I}^s : \Theta \to \Delta(S)$. Under these assumptions, the matrix B and Q could always be rationalized as follows:

$$p^s = (b_{s,\theta})_{\theta \in \Theta}, \quad \mathcal{I}^s(\theta)[\tilde{s}] = q_{s,\tilde{s}}; \tag{6}$$

that is, where every individual in the population thinks signals are uninformative, yet different belief types arrive at distinct posterior beliefs due to their differing priors. Note that this alternative not only allows for disagreement over priors, but also disagreement over the Blackwell experiment.

A seemingly more constrained formulation is to assume that while the entire population updates beliefs using the *same information structure*, each belief type also has its own prior. That is, consider the same framework above, but assume that $\mathcal{I}$ does not depend on the belief type—thus, the population agrees over the informational content of $\mathcal{I}$, even though belief type s updates using prior p^s. Despite appearances, it turns out that this alternative is no less constrained than one where individuals can disagree over the informational content of signals. The following result shows that this framework can completely rationalize any Q:

Proposition 3. *Suppose B and Q are plausible. Then there exists priors for each belief type, $p^s \in \Delta(\Theta)$, such that the state belief matrix B and signal belief matrix Q emerge via Bayesian updating using prior p^s and some information structure $\mathcal{I}$.*

While the observation in this section are suggestive that allowing for non-common priors can rationalize arbitrary B and Q, one response could be that the framework present is not satisfactory and that it may be necessary to elicit richer objects to determine whether the population shares a common prior or not. For instance: While the information environment in (6) yields identical B

and Q as the information environment in Proposition 3, one can distinguish between them by asking someone of belief type s, "what probability would you have assigned to state θ had your signal been $s' \neq s$?" If information were generated as in (6), someone observing a signal s would answer this question with $b_{s,\theta}$, for all s', since in that example the entire population views signals as drawn independently from θ. However, if $\mathcal{I} = (B^T B)^{-1} B^T Q$ is informative, then beliefs should respond to signals and not necessarily be constant, provided the prior is non-degenerate. While these kinds of questions may or may not be natural depending on the application, they are reminiscent of designs in experimental work employing the strategy method (see Brandts and Charness (2011)), whereby a subject would be asked to state their beliefs following any possible signal realization that might be potentially observed (whether or not it actually is). Addressing this point fully would take this paper too far afield and so I leave it as an open question.

However, Proposition 3 does clarify one way of interpreting $(B^T B)^{-1} B^T Q$ in cases where Q lies outside of the $|\Theta| - 1$ dimensional subspace identified in Proposition 1. In these cases, it need not be possible to find a single prior such that, when individuals adopt $\mathcal{I}$ and update from this prior, B and Q are induced. On the other hand, this is possible if one allows for each belief type to have its own prior.

I illustrate this idea using the example from Section 4.1. Recall that when $a = b = 9/16$, $(B^T B)^{-1} B^T Q$ cannot induce B and Q given a single prior. However, suppose signal s_1 (corresponding to the first row of B) starts with prior $(5/14, 9/14)$. In this case, the probability assigned to the first state, when updating according to $(B^T B)^{-1} B^T Q$ is $\frac{(5/14)(3/8)}{(5/14)(3/8)+(9/14)(5/8)} = 1/4$, which coincides with the relevant entry from B. And indeed, $(1/4)(3/8) + (3/4)(5/8) = 9/16$, coinciding with relevant entry from Q. Thus, with this prior, belief type s_1 holds the appropriate state beliefs and signal beliefs, provided the information structure is $(B^T B)^{-1} B^T Q$. Identical calculations show that, if signal s_2 (corresponding to the second row of B) were to start with prior $(9/14, 5/14)$, then state beliefs would be the second row of B and signal beliefs would be the second row of Q. Thus, while B and Q could not be induced given a common prior, they can be induced via a model with non-common priors and a common Blackwell experiment. Under these assumptions, the regression procedure will still identify the common $\mathcal{I}$.

5 The $|\Theta| > |S|$ case

The regression characterizations no longer apply when there are more states than signals, in the same way that an Ordinary Least Squares requires more observations than covariates in order to yield an identified solution (i.e., B must have more rows than columns, in addition to having no linear dependencies). The issue with my approach is that the question used to identify $\mathcal{I}$ has multiple solutions. Letting v be any vector in the null space of B, for any candidate solution $\mathcal{I}(\cdot)[\tilde{s}]$:

$$q_{\cdot,\tilde{s}} = B \cdot \mathcal{I}(\cdot)[\tilde{s}] \Rightarrow q_{\cdot,\tilde{s}} = B \cdot (\mathcal{I}(\cdot)[\tilde{s}] + v). \tag{7}$$

Thus, the set of solutions to the equation used to identify the column of $\mathcal{I}$ corresponding to $\tilde{s}$ has dimension equal to the null space of B, which is (generically) $|\Theta| - |S|$ dimensional.

Despite this multiplicity, one can still use regression ideas in order to determine the subspace of matrices which satisfy (7), even though $B^T B$ is not invertible when $|\Theta| > |S|$. Seeking to solve for a candidate information structure, as before, pre-multiply the matrix equation for the signal beliefs by the state belief matrix B:

$$B^T Q = B^T B \mathcal{I}$$

If $|S| < |\Theta|$, $B^T B$ is not invertible. The idea is to augment this equation so that an inversion can occur. Adding $M\mathcal{I}$, for any matrix $\mathcal{I}$, to both sides of this equation:

$$B^T Q + M\mathcal{I} = (B^T B + M)\mathcal{I}.$$

If $B^T B + M$ is invertible, one can obtain an expression where $\mathcal{I}$ only appears on the right hand side (multiplied by a matrix factor). A natural choice is $M = \lambda I$, where I is the identity matrix, yielding:

$$B^T Q + \lambda\mathcal{I} = (B^T B + \lambda I)\mathcal{I} \Rightarrow (B^T B + \lambda I)^{-1} B^T Q = \mathcal{I}(1 - (B^T B + \lambda I)^{-1}\lambda I) \tag{8}$$

A necessary and sufficient condition for the right hand side to converge to $\mathcal{I}$ as $\lambda \to 0$ is:

$$(B^T B + \lambda I)^{-1}\lambda I \to 0. \tag{9}$$

In statistics, the idea of adding λI in order to be able to invert $B^T B$ is used to motivate *ridge regression*; in cases where more general perturbations are considered, this process is also known as *Tikhonov regularization.*

Unfortunately, the condition (9) is not necessarily satisfied; but as Theorem 2 shows, the solution in the limit as $\lambda \to 0$ will nevertheless be a solution to $Q = B\mathcal{I}$.[14]

Theorem 2. *Suppose $|\Theta| > |S|$. The matrix:*

$$\tilde{\mathcal{I}} = \lim_{\lambda \to 0} (B^T B + \lambda I)^{-1} B^T Q,$$

exists, is well-defined, and solves the equation $B = Q\tilde{\mathcal{I}}$. Furthermore,

- *Let $v_1, v_2, \ldots, v_k$ be a basis for the null-space of B. Then $\mathcal{I}(\cdot)[s_i] = \tilde{\mathcal{I}}(\cdot)[s_i] + \sum_j \alpha_j v_j$, for some $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$.*

- *The prior p is the unique (unit) eigenvector of $B^T \tilde{\mathcal{I}}^T$ with eigenvalue 1.*

The Theorem shows that the $\lambda \to 0$ limit is equal to a solution of the equation defining signal beliefs in terms of state beliefs; it turns out that this property is sufficient in order to recover the

[14] In statistical applications where this method is used, taking λ too small is often undesirable. See van Wieringen (2015) for more on what guides the choice of λ in practice.

eigenvector interpretation of the prior, *whether or not the solution is in fact the true information structure*. While this does not imply that $\tilde{\mathcal{I}}$ is itself the information structure inducing B—or, for that matter, even an information structure—as per the discussion above, this does nevertheless provide some meaningful information which can in some cases be used to pin down the information structure.

The proof shows that the same argument illustrating that the prior is a unit eigenvector of $B^T\mathcal{I}^T$ with eigenvalue 1 can be used to show that it is an eigenvector of $B^T\tilde{\mathcal{I}}^T$, and proceeds to show that indeed one must exist. The following example illustrates:

Example 2. Suppose:

$$B = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} & 0 \\ \frac{4}{9} & \frac{1}{9} & \frac{4}{9} \end{pmatrix}, \qquad Q = \begin{pmatrix} \frac{7}{18} & \frac{11}{18} \\ \frac{11}{54} & \frac{43}{54} \end{pmatrix}.$$

In this case, I compute $\tilde{\mathcal{I}}$ to be:

$$\tilde{\mathcal{I}} = \begin{pmatrix} \frac{29}{63} & \frac{52}{63} \\ \frac{31}{126} & \frac{23}{126} \\ -\frac{4}{63} & \frac{58}{63} \end{pmatrix}$$

Obviously, $\tilde{\mathcal{I}}$ is not an information structure, as each row violates the two requirements to be probability distributions: non-negative entries and summing to 1. The nullspace of B is spanned by a single vector, $(-2, 4, 1)$. Adding a multiple of this vector to the first column of $\tilde{\mathcal{I}}$ to make sure all entries are non-negative yields:

$$\begin{pmatrix} \frac{1}{3} & \frac{52}{63} \\ \frac{1}{2} & \frac{23}{126} \\ 0 & \frac{58}{63} \end{pmatrix}$$

Adding another multiple of $(-2, 4, 1)$ to the second column allows each row to sum to 1, and yields:

$$\mathcal{I} = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{2} & \frac{1}{2} \\ 0 & 1 \end{pmatrix}$$

which indeed induces these belief matrices. I compute:

$$B^T\mathcal{I}^T = \begin{pmatrix} 14/27 & 5/9 & 4/9 \\ 5/27 & 2/9 & 1/9 \\ 8/27 & 2/9 & 4/9 \end{pmatrix}, \qquad B^T\tilde{\mathcal{I}}^T = \frac{1}{567}\begin{pmatrix} 382 & 139 & 208 \\ 139 & 58 & 46 \\ 208 & 46 & 232 \end{pmatrix}.$$

Indeed, for both of these matrices, there is a unique (up to scaling) eigenvector with eigenvalue 1, and it is $(1/2, 1/6, 1/3)$. Bayes rule verifies that this matrix, together with $\mathcal{I}$, induces B.

Note that knowledge of the prior would allow the analyst to learn the information structure itself. That is, given the prior $p(\theta)$, the analyst could determine signal weights $(\alpha_s)_{s\in S}$ such that $p(\theta) = \sum_s \alpha_s b_{\theta,s}$ (which exist since $p(\theta)$, as the prior, is in the interior of the convex hull of the columns of B). As pointed out in Proposition 1 of Kamenica and Gentzkow (2011), one then has $\mathcal{I}(\theta)[s] = b_{\theta,s}\alpha_s/p(\theta)$.

Despite this observation, as I make clear in the next Section, I view Theorem 2 as largely of theoretical interest, as it demonstrates that the martingale condition for beliefs pins down the prior given any candidate solution to the equation defining signal beliefs. While this helps clarify what drives the identification of the prior, I do not wish to overstate its practical significance; it seems more direct to simply obtain $\mathbb{P}[s]$ from Q rather than $p(\theta)$ from $B^T\mathcal{I}^T$; as I discuss in the next section, this provides a more immediate path to the information structure. The key point is that (i) regression techniques can still recover solutions for I of the equation $Q = BI$, even though (ii) with more states than signals, that may not pin down the information structure, but that (iii) the eigenvector interpretation of the prior remains even in this case.

6 The Signal Priors Approach and Eigenvector Characterizations

Prelec and McCoy (2022) provides an alternate procedure to derive the matrix $\mathcal{I}$, without restrictions on the cardinality S and Θ. The crux of this approach is to use Q directly to obtain the (ex-ante) distribution over the signals in the population. That is: Given the matrix Q, note that if v is an eigenvector of Q with eigenvalue 1 (which, again, exists by Perron-Frobenius), then v is proportional to $\mathbb{P}[s]$; indeed, $\sum_s \mathbb{P}[s^* \mid s]\mathbb{P}[s] = \mathbb{P}[s^*]$. Thus, normalizing the eigenvector of Q so that the entries sum to 1 provides a way of arriving at the (ex-ante) probability distribution over *belief types*. In this case, one has $p(\theta)\mathcal{I}(\theta)[s] = b_{s,\theta}\mathbb{P}[s]$. To obtain $p(\theta)$, simply consider $\sum_{\tilde{s}} b_{\tilde{s},\theta}\mathbb{P}[\tilde{s}]$. While this argument identifies the prior, the information structure $\mathcal{I}$ emerges via the identity $\mathcal{I}(\theta)[s] = b_{s,\theta}\mathbb{P}[s]/p(\theta)$.

This method, called the Signal Priors Approach (henceforth SP), is complementary to the regression approach, and each may have merits depending on the application. For instance, while SP clearly requires Q, it can identify the relevant signal distribution even without any information about B at all. Hence if Q is fully specified and B is not, then this approach allows for part of the information structure to be recovered nevertheless.

On the other hand, there are instances where the regression approach of this paper may be beneficial. Perhaps most significantly, it may be that B is fully specified, but not all of Q is. In that case, one can compute $(B^TB)^{-1}B^Tq$ given a column q, and still find the probability that that signal is induced in each state θ. Note that, while this would not identify the full information structure, it would still identify part of it; in particular this procedure could still be used to determine the state θ, the typical focus of the crowd wisdom problem. More precisely: if q corresponds to signal s, and if $(B^TB)^{-1}B^Tq$ yields $|\Theta|$ distinct entries, then the analyst could learn the true θ by seeing which entry matched the empirically observed proportion of the population receiving signal s.

While the extent to which this would work in practice is difficult to assess theoretically (e.g., it may depend on noise in the population, or how many beliefs are elicited), one can imagine there being cases where eliciting the full matrix Q is significantly more difficult to obtain than a single column of it. As discussed in the introduction, some informational environments may feature a large number of signals, e.g., with a binary state but a rich set of probability assessments. In that case, to form the matrix B the analyst only needs to sample a sufficiently large number of population members. But a large number of belief types implies that it might be prohibitive to ask the *experts* for their full vector of signal beliefs. Thus, if the analyst were simply interested in finding one column of $\mathcal{I}(\cdot)[\cdot]$ to determine the state, then rather than eliciting $|\Theta| + |S|$ beliefs, as SP requires, my method implies that (generically) the state can be determined by eliciting only $|\Theta| + 1$ beliefs. If $|S|$ is very large, then this simplification can be substantial.

This discussion suggests that the regression approach is perhaps more suitable when the number of signals (i.e., belief types) is larger than the number of states. However, the results I provide on the case of $|\Theta| > |S|$ are largely theoretical, in that they suggest properties of the information structure, but actual computation of it can be more cumbersome; more to the point, if the analyst were able to elicit beliefs over Θ, then the $|\Theta| > |S|$ case is precisely when eliciting beliefs over S entails fewer questions than eliciting beliefs over Θ in the first place. Thus, it appears that SP is more appropriate from a practical perspective when $|\Theta| > |S|$.[15]

A second reason that regression might be useful, even when SP is preferable, is as a way of assessing robustness. Prelec and McCoy (2022) points out that this paper's regression method may fail to yield an information environment if B and Q are misspecified (pointed out here in Footnote 13). Without any misspecification, both methods should yield the same results. And indeed, as pointed out in Section 4.1, "most" B and Q will not induce informational environments, and indeed this may be impossible if these objects are estimated imprecisely. From this perspective, the regression method could be useful as a means of assessing the extent to which the analyst should trust the estimated B and Q. If in fact the regression method estimates for the information structure are negative, then the analyst should be more cautious in interpreting the results even when using SP, since negative entries would only be consistent with an error (either with the estimation of B and Q or the framework as a whole). An interesting open question this point raises is whether empirically SP or regression (or another method) performs better in these cases, as well as whether there are theoretical reasons such advantages might emerge.

6.1 Priors as Eigenvectors and Representations of Information

The eigenvector interpretations of signal-priors from Prelec and McCoy (2022) (also mentioned in Prelec et al. (2013), the working paper version of Prelec et al. (2017)) and state-priors from this paper are related to important and influential insights from Samet (1998). I note that part of this paper's goal is to clarify which parts of the informational environment can be determined with limited data, even if the analyst cannot identify (signal or state) priors, so to some extent this connection is secondary. But since (to my knowledge) the first interpretation of priors as

[15] I credit Prelec and McCoy (2022) with originally making this point when comparing its method with this paper's.

eigenvectors is due to Samet (1998), and since some of my results are inspired by it, it is worth explicitly explaining the relationship.

Samet (1998) considered a standard general framework with multiple agents, where each agent is endowed with a partition of the state space. An agent updates beliefs after observing which element of this partition the true state belongs to. This framework, where information is generated deterministically as a function of the state, is clearly a special case of one where information is generated by Blackwell experiments (albeit one imposing $|S| \leq |\Theta|$).[16] But as pointed out by Green and Stokey (1978), the converse is true in the following sense: any framework where information is generated stochastically can instead be described using information partitions *after enriching the state space* (see also Gentzkow and Kamenica (2017) for a discussion). In crowd wisdom applications, one typically represents analyst data using *belief-based representations of information*. This type of ex-post data imposes particular restrictions on the ex-ante informational environment, and interpreting these uses the structure of the crowd wisdom setting—for instance, the distinction "beliefs over θ" and "beliefs over s_{-i}."

Samet (1998) characterized (in Propositions 3 and 5) a *common* prior as the eigenvector of stochastic transition matrices which yielded an *interim* characterization of the common prior assumption, i.e., after signals have been observed. I briefly describe this argument. Denote states by $\omega \in \Omega$. Let $M_i(\omega, \omega')$ be a matrix where the entry corresponding to ω (rows) and ω' (columns) is the probability that player i assigns to state ω' in state ω. If q is agent i's prior over Ω (i.e., beliefs before observing which partition element the state is in), then $q \cdot M_i(\omega, \omega') = q$—i.e., q will be a unit eigenvector of $M_i(\omega, \omega')$. Typically many vectors will solve this equation for a given agent i.[17] However, if q is a *common* prior, then q should also satisfy $q \cdot M_j(\omega, \omega') = q$ for $j \neq i$, so that we also have $q \cdot M_i(\omega, \omega') \cdot M_j(\omega, \omega') = q$. Identification of the (common) prior follows from the requirement that $M_1 M_2 \cdots M_n$, is irreducible, where n is the number of agents.[18]

This argument also applies to the present paper, in the sense that the ex-ante distribution over Ω can be "backed out" as an eigenvector of $M_i(\omega, \omega')$. And, given the aforementioned equivalence, any expert who understands $\mathcal{I}$ and p could alternatively describe the informational environment using information partitions. With this in mind, there are two key ways in which the exercise in this paper builds upon these insights.

First, this paper is not concerned with the problem of the *expert*, but rather the problem of the *analyst*. Here, the analyst's data, following the crowd wisdom literature, takes the form of

[16] To help clarify this connection even further, Appendix B.2 provides an illustration of the regression approach when information is deterministic.

[17] This will be the case whenever $M_i(\omega, \omega')$ is not irreducible, i.e., whenever the agent observes at least two different signals. In this case, interim beliefs do not pin down the relative probability of each information partition element.

[18] The existence of the common prior is equivalent to this eigenvector being independent of the order in which each M_i matrix is multiplied together when defining this stochastic transition matrix. See Hellman (2011) for a generalization of this result to infinite spaces, and Golub and Morris (2017) for an application of these ideas to an eigenvector characterization of the average population expectations; note that the latter paper also features an explicit invocation of the Perron-Frobenius theorem. Several other papers have used either the Markov chain interpretation of interim beliefs introduced by Samet (1998), or properties implied by the stationary distribution characterization of the common prior; see, for instance, Morris and Shin (2002), Cripps et al. (2008), or Angeletos and Lian (2018).

the matrices B and Q. But $M_i(\omega, \omega')$ is *not* a primitive. While the analyst could write down the matrix $M_i(\omega, \omega')$ using $\mathcal{I}$ and p, the question of this paper is precisely how to derive the latter objects. While p is characterized by an eigenvector equation similar to that of Samet (1998), it is one based on the primitives of Section 3; these, in turn, are inspired by the typical data considered in information aggregation problems. To use Samet (1998)'s characterization, one would need to write M_i matrices, and to do this the analyst would need to know p,[19] the precise object *this paper's eigenvector characterization* is used to determine.

Second, the resulting eigenvector of M_i, q, is actually analogous to $p(\theta)\mathcal{I}(\theta)[s_i]$ rather than $p(\theta)$, since ω relates to both θ and s_i. This leaves open the issues raised in the above discussion of Prelec and McCoy (2022), namely whether part of $\mathcal{I}$ can still be inferred without enough data to determine the relevant eigenvector. By contrast, this paper addresses this question. Relatedly, since nesting this paper's framework within Samet (1998) requires identifying ω with both θ and s realizations, it is not clear how to consider the possibility of a "non-common prior but a common information structure" without the structure of the crowd wisdom setting, as this paper has done.

These observations underlie the differences between the approaches and contributions. Belief-based frameworks typically specify $\mathcal{I}$ and p separately, and some of the results in this paper (particularly those in Section 4.1 and 4.2 below) use the distinction between these objects. My paper clarifies how data on ex-post beliefs in the form of B and Q, objects germane to the belief-based approach, restrict the ex-ante informational environment. More broadly, this discussion highlights one of this paper's goals: to clarify how important insights due to Samet (1998) are actually more general and also apply under belief-based approaches, which have become popularized in economic theory particularly following Kamenica and Gentzkow (2011) (see also Mathevet et al. (2020) for a discussion of this approach in multi-player settings).

7 Conclusion

The main lesson from this paper is that, while an informational environment can generally not be identified using only the set of ex-post beliefs over the states that generate information, the informational environment can be recovered with access to data on the beliefs over the realized signal distribution of the expert's peers. I have also highlighted one potentially disconcerting feature of the analysis: the dimensionality of the set of plausible signal belief matrices is typically larger than the dimensionality of the set of signal belief matrices which could emerge given $\mathcal{I}$. This suggests that the underlying framework may be easily falsified in practice—if individuals only report beliefs imprecisely, it may be that the observed signal beliefs are outside of this smaller dimensional space. While there are ways to accommodated this possibility (e.g., by relaxing common priors),

[19] Concretely, consider the truth-or-noise specification in Example 1. If player i were to receive a null signal, then the probability player i assigns to player j having received a signal revealing the state to be $\tilde{\theta}$ would be $(1-\varepsilon)p(\tilde{\theta})$, which clearly varies with the prior p. If defining states ω to be triples of θ and signals for a pair of experts, then specifying the corresponding $M_i(\omega, \omega')$ therefore requires knowledge of p.

it may be desirable to have a relaxation that accommodates richer signal beliefs without simply stating that "anything goes." These questions seem worthwhile but are beyond this paper's scope.

Two further avenues seem worthwhile of pursuit. First, the comparison to the Signal Priors approach suggests that there are multiple methods at an analyst's disposal to identify either the information structure or the underlying state. If there are many signals, then it will generically be possible to recover the state by eliciting a fewer set of beliefs relative to what the Signal Priors approach requires. An interesting question is which kinds of data are best suited (both theoretically and practically) to answering a relevant question an analyst might be interested in. Along these lines, it also seems worthwhile to grapple with the practical issues related to elicitation which this paper has abstracted from (but addressed by some papers discussed above).

Second, the use of regression in order to identify information seems worthy of further exploration. A Bayesian agent who computes an expected value, by definition, computes as a linear combination (i.e., a weighted average) of several parameters. Regression is simply a way of inverting this relationship, to find the parameters as a function of beliefs. This paper is not alone in suggesting that these tools can be deployed powerfully in a wide variety of settings—other examples sharing this theme include Aguiar et al. (2023); Miyashita (2022); Caradonna (2021); Ball (2021). But it may very well be that more elaborate regression techniques could allow for more realistic assumptions or speak to different applications and questions.

While this paper has been focused on the application of crowd wisdom, I view the observations here as relevant to a richer set of single-agent decision problems where agents form beliefs using Blackwell experiments. Frameworks of this form have become popular recently (see, for instance, Lu (2016, 2019); Jakobsen (2021)). An equivalent formulation of my model would be to posit that the belief matrices B and Q emerge from a single-agent problem, without imposing the structure of crowd wisdom models. The main challenge with such a formulation is that typically, a decision maker will only observe one signal. It is therefore unclear how an analyst could observe the beliefs an agent would have had following signals that were not realized. This criticism does not apply to crowd wisdom settings, as the different signal realizations are observed by different experts. On the other hand, there may be cases where this criticism also has no bite. In any event, even if the relevant data take a different form than what I have considered here, the results in this paper may speak to issues relevant to other single-agent decision problems where a decision maker updates beliefs from some prior following the outcome of a Blackwell experiment.

References

Aguiar, V., M. J. Boccardi, N. Kashaev, and J. Kim (2023): "Random Utility and Limited Consideration," *Quantitative Economics*, 14, 71–116.

Angeletos, G.-M. and C. Lian (2018): "Forward Guidance without Common Knowledge," *American Economic Review*, 108, 2477–2512.

Arieli, I., Y. Babichenko, and R. Smorodinsky (2020): "Identifiable information structures," *Games and Economic Behavior*, 120, 16–27.

Arieli, I. and M. Mueller-Frank (2017): "Inferring Beliefs from Actions," *Games and Economic Behavior*, 102, 455–461.

Aumann, R. J. and M. Maschler (1995): *Repeated Games with Incomplete Information*, MIT Press.

Ball, I. (2021): "Scoring Strategic Agents," *Working Paper.*

Brandts, J. and G. Charness (2011): "The strategy versus the direct-response method: a first survey of experimental comparisons," *Experimental Economics*, 14, 375–398.

Caradonna, P. (2021): "Preference Regression," *Working Paper.*

Chen, Y.-C., M. Mueller-Frank, and M. Pai (2021): "The Wisdom of the Crowd and Higher Order Beliefs," *Working Paper.*

Cripps, M., J. Ely, G. Mailath, and L. Samuelson (2008): "Common Learning," *Econometrica*, 76, 909–933.

Cvitanić, J., D. Prelec, B. Riley, and B. Tereick (2019): "Honesty via Choice-Matching," *American Economic Review: Insights*, 1, 179–192.

Gentzkow, M. and E. Kamenica (2017): "Bayesian persuasion with multiple senders and rich signal spaces," *Games and Economic Behavior*, 104, 411–429.

Golub, B. and S. Morris (2017): "Higher-Order Expectations," *Working Paper, Northwestern University and MIT.*

Green, J. and N. Stokey (1978): "Two Representations of Information Structures and Their Comparisons," *Unpublished.*

Hellman, Z. (2011): "Iterated Expectations, Compact Spaces and Common Priors," *Games and Economic Behavior*, 72, 163–171.

Horn, R. A. and C. R. Johnson (2013): *Matrix Analysis*, Cambridge University Press, second ed.

Hussam, R., N. Rigol, and B. N. Roth (2022): "Targeting High Ability Entrepreneurs Using Community Information: Mechanism Design in the Field," *American Economic Review*, 112, 861–898.

Jakobsen, A. M. (2021): "An Axiomatic Model of Persuasion," *Econometrica*, 89, 2081–2116.

Kamenica, E. and M. Gentzkow (2011): "Bayesian Persuasion," *American Economic Review*, 101, 2590–2615.

Lomys, N. and E. Tarantino (2022): "Identification in Search Models with Social Information," *Working Paper.*

Lu, J. (2016): "Random Choice and Private Information," *Econometrica*, 84, 1983–2027.

——— (2019): "Bayesian Identification: A Theory for State-Dependent Utilities," *American Economic Review*, 101, 2590–2615.

Mathevet, L., J. Perego, and I. Taneva (2020): "On Information Design in Games," *Journal of Political Economy*, 128, 1370–1404.

Meyer, C. D. (1995): *Matrix Analysis and Applied Linear Algebra*, SIAM.

Miyashita, M. (2022): "Identification of Information Structures in Bayesian Games," *Working paper.*

Morris, S. and H. S. Shin (2002): "Social Value of Public Information," *American Economic Review*, 92, 1521–1534.

Prelec, D. (2004): "A Bayesian Truth Serum for Subjective Data," *Science*, 306, 462–466.

Prelec, D. and J. McCoy (2022): "General identifiability of possible world models for crowd wisdom." *Working Paper.*

Prelec, D., H. S. Seung, and J. McCoy (2013): "Finding Truth Even if the Crowd is Wrong," *Working Paper.*

——— (2017): "A solution to the single-question crowd wisdom problem," *Nature*, 541, 532–535.

Samet, D. (1998): "Iterated Expectations and Common Priors," *Games and Economic Behavior*, 24, 131–141.

van Wieringen, W. N. (2015): *Lecture Notes on Ridge Regression*, ArXiv Preprint.

Witkowski, J. and D. C. Parkes (2012a): "Peer Prediction Without a Common prior," *Proceedings of the 13th ACM Conference on Electronic Commerce (EC12).*

——— (2012b): "A Robust Bayesian Truth Serum for Small Populations," *Proceedings of the 26th AAAI Conference on Artificial Intelligence (AAAI 13).*

A Proofs of the Main Result

Proof of Theorem 1. The first part of the theorem follows the arguments as laid in the main text. Writing the information structure so that rows are states and columns are signals, the definition of signal beliefs matrix tells us that:

$$Q = B \cdot \mathcal{I},$$

since the ith row and jth column of Q is the inner product of the ith row of B (since rows of B index signals) and the jth column of $\mathcal{I}$ (using the convention that columns index signals). Given this expression, and using that $(B^T B)^{-1}$ is invertible, the solution for $\mathcal{I}$ comes from left-multiplying both sides by B^T and then left multiplying by $(B^T B)^{-1}$. From this, the formula for $\mathcal{I}$ stated in (3) follows immediately.

Next, I show that the prior is identified. As shown in the main text, the prior is a unit eigenvector (with eigenvalue 1) of the matrix $B^T \mathcal{I}^T$, and therefore a unit eigenvector of $B^T Q^T B (B^T B)^{-1}$. I show that this matrix is *always* row-stochastic. Recall that $\mathbf{1}^T B^T$ and $\mathbf{1}^T Q^T$ are both $\mathbf{1}$, since both B and Q are row-stochastic (so that the transposes are column-stochastic). Therefore:

$$\mathbf{1}^T B^T Q^T B (B^T B)^{-1} = \mathbf{1}^T Q^T B (B^T B)^{-1} = \mathbf{1}^T B (B^T B)^{-1}$$

Taking the transpose of this expression gives:

$$(B^T B)^{-1} B^T \mathbf{1}.$$

Now, recall the regression interpretation of the linear mapping $(B^T B)^{-1} B^T$; when applied to a vector, it gives the coefficient of the regression of the vector onto the columns of B. However, the columns of B sum to 1. Therefore, to write the vector $\mathbf{1}$ as a linear combination of the columns of B, I need only write it using coefficients equal to 1, and hence this expression is itself a vector of 1s, demonstrate that the matrix is row-stochastic.

Hence, the Perron-Frobenius theorem holds provided the matrix $B^T \mathcal{I}^T$ is irreducible. This property holds generically; indeed, the entries of $B^T \mathcal{I}^T$ are generically positive, and all such matrices are irreducible. This theorem therefore yields the existence of a Perron eigenvector, which is positive and sums to 1. While the argument in the main text shows that being a Perron eigenvector is a *necessary* condition for the prior to be identified, the Perron-Frobenius theorem implies that this vector is unique, and therefore this condition is also sufficient. As a result, the prior is additionally identified from B and Q, in addition to the information structure, as desired. □

Proof of Proposition 1. By Proposition 1 of Kamenica and Gentzkow (2011), given any belief matrix B and prior p, there exists an information structure $\mathcal{I}$ inducing this belief matrix.[20] On the other

[20] Their proof is constructive; in my notation, one can set $\mathcal{I}(\theta)[s] = B_{s,\theta} \mathbb{P}[s]/p[\theta]$. Now, note that $\mathbb{P}[s]$ is a left unit eigenvector of the matrix Q. Note, however, that $\mathbb{P}[s]$ would be derived from Q in this paper, and not the prior p and the posterior beliefs, as in theirs.

hand, Theorem 1 shows that *any* vector v of length $|S|$ yields a vector of length Θ when considering $(B^T B)^{-1} B^T v$. Thus, the set of information structures is spanned by the set of Q that emerge in some informational environment. Putting the previous observations together, the set of Q which induce an informational environment given a belief matrix B is isomorphic to the set of priors inducing B, which has dimensionality equal to $|\Theta| - 1$, for any belief matrix satisfying the linear independence condition. □

Proof of Proposition 2. It is immediate that the set of row stochastic matrices has dimension $|S|(|S|-1)$. I show that this is also the dimensionality of the set of plausible matrices which induce valid information structures.

I first argue that the rows of $(B^T B)^{-1} B^T Q$ sum to 1 if Q has rows that sum to 1. I note that the sum of the rows of this matrix is given by right multiplying by $\mathbf{1}_{|S|}$, a vector of length $|S|$ which is all 1s. If Q has rows which sum to 1, then $(B^T B)^{-1} B^T Q \cdot \mathbf{1}_{|S|} = (B^T B)^{-1} B^T \mathbf{1}_{|S|}$. On the other hand, recall that this expression is also the coefficients $\beta_1, \ldots, \beta_n$ solving:

$$\mathbf{1}_{|S|} = \sum_{i=1}^{n} \beta_i b_i,$$

where b_i is the ith column of the belief matrix B_i. While there is a unique set of coefficients solving this equation, I also have that the columns of a belief matrix sum to 1. Hence $\beta_1 = \cdots = \beta_n = 1$ is the solution. I thus conclude that $(B^T B)^{-1} B^T Q \cdot \mathbf{1}_{|S|}$ is a vector of 1s, as claimed.

Now, consider an arbitrary information structure $\tilde{\mathcal{I}}$ with *only strictly positive entires*, and let $Q := B^T \tilde{\mathcal{I}}$. Then considering the regression coefficients as in Theorem 1, I have $\tilde{\mathcal{I}} = (B^T B)^{-1} B^T Q$, which is a valid information structure.

Furthermore, given that B is full rank, $(B^T B)^{-1} B^T$ is a continuous transformation on the set of row stochastic matrices. Since $(B^T B)^{-1} B^T Q$ is strictly positive, by the definition of continuity, I have there is an open subset (assuming the relative topology) of row-stochastic matrices $\tilde{Q}$ such that $(B^T B)^{-1} B^T Q$ is strictly positive. This implies that the set of plausible Q has the same dimensionality as the set of row-stochastic Q, since an open set within a topological space has the same dimensionality as the space. This completes the proof. □

Proof of Proposition 3. Note that, given any B satisfying the assumptions of my framework, that $B^T B$ is invertable, so that $(B^T B)^{-1} B^T Q$ is well-defined. I consider an information environment where all belief types use this information structure, but each belief type arrives at posterior beliefs by updating an belief-type specific prior p^s. I write $(b_{s,\theta})_\theta$ as the belief vector of type for belief type s, and consider some fixed s. Choose $r(\theta)$ to satisfy:

$$r(\theta)\mathcal{I}(\theta)[s] = b_{s,\theta},$$

noting that, since all other terms in this expression are non-negative, $r(\theta)$ is as well. Further note that one cannot have $\mathcal{I}(\theta)[s] = b_{s,\theta} = 0$ for all θ, since $\sum_\theta b_{s,\theta} = 1$ by assumption. Hence $r(\theta) \geq 0$ for all θ and $r(\theta) \neq 0$ for some θ. Therefore, I can set:

$$p_s(\theta) = \frac{r(\theta)}{\sum_{\tilde{\theta}} r(\tilde{\theta})},$$

which is an element of $\Delta(\Theta)$.

Suppose belief type s updates using $\mathcal{I}$ and prior p^s. In that case, following signal s, the Bayesian belief is:

$$\frac{\frac{r(\theta)}{\sum_{\tilde{\theta}} r(\tilde{\theta})}\mathcal{I}(\theta)[s]}{\sum_{\theta'} \frac{r(\theta')}{\sum_{\tilde{\theta}} r(\tilde{\theta})}\mathcal{I}(\theta')[s]} = \frac{r(\theta)\mathcal{I}(\theta)[s]}{\sum_{\theta'} r(\theta')\mathcal{I}(\theta')[s]} = \frac{b_{s,\theta}}{\sum_{\theta'} b_{s,\theta'}} = b_{s,\theta},$$

where all equations follow from either cancellations or definitions. Furthermore, the probability belief type s assigns to a randomly selected individual having belief type $\tilde{s}$ is $q_{s,\tilde{s}} = \sum_\theta \mathcal{I}(\theta)[\tilde{s}]b_{s,\theta}$, since posterior beliefs are $b_{s,\theta}$ and the assumed information structure is $\mathcal{I}(\theta)$. This completes the proof. □

Proof of Theorem 2. First, I claim that the limit $\lim_{\lambda\to 0}(B^TB + \lambda I)^{-1}B^TQ$ exists and is a finite matrix; I provide an independent proof of this result after completing the rest of the proof, though I mention that this fact appears known (though requiring some additional detours the proof below avoids).[21]

I now turn to the second bulletpoint of the Theorem. Note that the ridge estimator defined by $\tilde{\mathcal{I}}$ solves the following minimization problem:

$$\tilde{\mathcal{I}}_\lambda(\cdot)[s] = \underset{x}{\operatorname{argmin}} \quad ||q_{s,\cdot} - Bx||^2 + \lambda\,||x||^2 . \tag{10}$$

By contrast, the information structure $\mathcal{I}$ solves $q_{s,\cdot} = B\mathcal{I}(\cdot)[s]$. Now, given the claim that $\lim_{\lambda\to 0}\tilde{\mathcal{I}}_\lambda(\cdot)[s]$ exists, it follows from this expression that the resulting limit must also be a solution to the equation $q_{s,\cdot} = Bx$; if it weren't, then one would have the objective in (10) would converge to some strictly positive amount as $\lambda \to 0$; by contrast, *any* solution to this equation makes this objective equal to 0 in the limit. Hence any vector x which does not satisfy $q_{s,\cdot} = Bx$ cannot be the limit of $\tilde{\mathcal{I}}_\lambda(\cdot)[s]$ as $\lambda \to 0$.

On the other hand, for *any* vector x satisfying $q_{s,\cdot} = Bx$, subtracting the equation for $\tilde{\mathcal{I}}$ from this equation yields $0 = B(x - \tilde{\mathcal{I}}(\cdot)[s])$, so that $x - \tilde{\mathcal{I}}(\cdot)[s]$ is in the nullspace of B. But the information structure generating the decision-maker's information is one possible choice of x; therefore, $\mathcal{I}(\cdot)[s] - \tilde{\mathcal{I}}(\cdot)[s] = \sum_i \alpha_i v_i$, where $\{v_1, \ldots, v_k\}$ is a basis for the nullspace of B (assuming the dimension of this space is k); adding $\tilde{\mathcal{I}}(\cdot)[s]$ to both sides of this expression proves the second bulletpoint.[22]

[21] As per van Wieringen (2015), the limit of the ridge estimator as $\lambda \to 0$ is precisely the least square estimate of smallest norm; showing this, however, requires a significant detour in defining ridge estimators. Note that van Wieringen (2015) also shows that multiplying by a matrix M as described in the main text amounts to a rescaling of the design matrix (in this case, B).

[22] An identical argument shows the claim that *every* solution to the equation $Q = BX$ differs from $\mathcal{I}$ in this way.

Next, I show that p is a unit eigenvector with eigenvalue 1. Suppose $Q = B\tilde{\mathcal{I}}$, and let q denote left unit eigenvector with eigenvalue 1. Recall that (see Section 6) that $q = (\mathbb{P}[s])_{s\in S}$, the vector of probabilities that each signal in s is observed given the informational environment. However, note that the martingale properties of beliefs states that $\mathbb{P}[s] \cdot B = p(\theta)$; putting this together gives us that, for every s, I have $\sum_\theta \tilde{\mathcal{I}}(\theta)[s] \cdot p(\theta) = q_s$ (i.e., the entry of q corresponding to s. On the other hand, the martingale property of beliefs, written $qB = p$ (where p is the prior), does not depend on the information structure. Thus, I can apply the same argument to say that the prior is a left unit eigenvector (with eigenvalue 1) of $\tilde{\mathcal{I}}B$. Note that this does not complete the proof since I still have to show this eigenvector is unique.

Letting $p_X(\lambda)$ be the characteristic polynomial for a matrix X, I note that for matrices A and B where A is m-by-n and B is n-by-m, with $n \geq m$ one has $p_{BA}(\lambda) = \lambda^{n-m} p_{AB}(\lambda)$. I apply this result to $B^T\mathcal{I}^T$ and $B^T\tilde{\mathcal{I}}^T$ (see Theorem 1.3.22 in Horn and Johnson (2013)). In particular, using the previous result, write $\tilde{\mathcal{I}}$ as $\mathcal{I} + W$, where each row of W is in the null space of B. In particular, since each W is in the null space of B, I have:

$$(\mathcal{I} + W)^T B^T = \mathcal{I}^T B^T.$$

Putting this together with the previous results, using that $|\Theta| > |S|$ (so that B^T has more rows than columns), I have:

$$p_{B^T\tilde{\mathcal{I}}^T}(\lambda) = \lambda^{|\Theta|-|S|} p_{\tilde{\mathcal{I}}^T B^T}(\lambda) = \lambda^{|\Theta|-|S|} p_{\tilde{\mathcal{I}}^T B^T}(\lambda) = p_{B^T\mathcal{I}^T}(\lambda)$$

As argued in Theorem 1, $B^T\mathcal{I}^T$ has a unique eigenvector with eignvalue 1 (i.e., there is no multiplicity in the eigenspace). Thus, since $B^T\mathcal{I}^T$ has a unique eigenvector with eigenvalue 1, so does $B^T\tilde{\mathcal{I}}$; as argued above, this must be the prior, completing the proof.

I conclude by showing that $\lim_{\lambda\to 0}(B^TB + \lambda I)^{-1}B^TQ$ exists and is a finite matrix, as claimed. I first determine the rate at which the determinant of $B^TB + \lambda I$ tends to 0 as $\lambda \to 0$. Note that, by the Matrix Determinant Lemma (see (6.2.3) of Meyer (1995) for a version of this result), I have:

$$\det(\frac{1}{\lambda}B^TB + I) = \det(I_{|S|} + \frac{1}{\lambda}BB^T).$$

Note that the matrix involved in the left-hand side of this equation is $|\Theta| - by - |\Theta|$ and the matrix involved in the right hand side of this equation is $|S| - by - |S|$. I therefore have, multiplying through by $\lambda^{|\Theta|}$ and using that $\det(cA) = c^n \det(A)$ for $c \in \mathbb{R}$ and A an n-by-n matrix,

$$\det(B^TB + \lambda I) = \det(\lambda^{|\Theta|/|S|} I_{|S|} + \lambda^{(|\Theta|-|S|)/|S|} BB^T).$$

Note that this determinant is a polynomial in λ which evaluates to 0 at $\lambda = 0$, and hence this approaches 0 at a rate equal to the rate of the smallest term in this polynomial. I claim the degree is strictly less than $|\Theta|$. This is clear from examining the right hand side of the equation above. While every term on the diagonal in this matrix is of the order $\lambda^{|\Theta|/|S|}$, every term *off* the diagonal is of the order $\lambda^{|\Theta|/|S|-1}$. To show that there is a term in the polynomial defined by the determinant that is of order less than $|\Theta|$, it suffices to show that some term in this expression reflects off diagonal terms. Note that the determinant is a sum over permutations $\sigma : |S| \to |S|$:

$$\sum_{\sigma} \text{sgn}(\sigma) \prod_{i=1}^{n} b_{i,\cdot} b_{\cdot,\sigma(i)},$$

where $b_{i,\cdot}$ is the ith row of B. Then the permutations which simply swap two elements (of which there are $|S| \cdot (|S|-1)/2$ of) contribute to the determinant; for any such permutation, $\text{sgn}(\sigma) = -1$. Since any coefficient in this sum where λ is of degree $|\Theta| - 2$ must correspond to one of these permutations, the exponent on λ reflecting these permutations is at most $|\Theta| - 2$; and thus, the smallest non-zero degree of the characterisic polynomial must be less than $|\Theta|$.[23]

Therefore, as $\lambda \to 0$, $\det\left((B^T B + \lambda I)\frac{1}{\lambda} I\right) = \det(B^T B + \lambda I)\frac{1}{\lambda^{|\Theta|}} \not\to 0$. Taking inverses, $(B^T B + \lambda I)^{-1} \lambda I$ must have a limit; indeed, in the definition of the matrix inverse, each term is scaled by the inverse of the determinant, and otherwise comes from multiplying and adding matrix elements together—so, since each term is scaled by a term that does not approach infinity, each term converges to a finite limit. Using Equation 8, I conclude that the limit defining $\tilde{\mathcal{I}}$ exists. □

B Additional Results and Discussion

B.1 Removing Linear Dependencies

This section comments on the full rank assumptions. First, note that this assumption is generic: Indeed, the set of all possible belief matrices is of dimension $|S| \times (|\Theta| - 1)$ (with one degree of freedom lost for every row, since every row is restricted to sum to 1); in general, the space spanned by k belief vectors, restricted to sum to 1, is $(k-1) \times |S|$ dimensional. Hence a generic belief matrices is full rank, in that the set of belief matrices which fail to satisfy this belong to a lower dimension subspace. Given this observation, B is generically of rank equal to $|\Theta|$. On the other hand, the rank of $B^T B$ is equal to the rank of B, and therefore $B^T B$ has full rank. Since a square matrix is invertible if and only if it has full rank, I have that $B^T B$ is invertible.

I now comment on the implicit assumption behind the full rank assumption that the columns of B are linearly independent. While the argument in the previous paragraph shows this assumption is generic, one may be interested in cases in where it is violated or understanding the substance of this assumption. Note that if this condition fails, then the matrix $B^T B$ is not invertible.[24]

I show how the case of linearly dependent columns can be interpreted as reflecting the case where a state is "split apart." Rather than discussing this in full generality, I present an illustrative

[23]More generally, the lowest degree of the polynomial should be $|\Theta| - |S|$; showing this, however, requires that some permutations which influence the determinant do not fix any elements on the diagonal. Determining that not all terms cancel out, while certainly intuitive, appears less direct than this argument. However, provided this is the case, then any entry corresponding to exclusively off-diagonal term will be a polynomial of order $(\lambda^{(|\Theta|-|S|)/|S|})^{|S|}$, since the matrix is $|S|$-by-$|S|$.

[24]Indeed, as discussed in Section 5, this condition always *fails* when $|S| > |\Theta|$.

example.

Suppose that $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$, where the third column of B is a linear combination of the first two:

$$B = \begin{pmatrix} 2/3 & 0 & 1/3 & 0 \\ 1/3 & 1/3 & 1/3 & 0 \\ 0 & 2/5 & 1/5 & 2/5 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad Q = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/2 & 1/4 & 0 \\ 0 & 3/10 & 1/2 & 1/5 \\ 0 & 0 & 1/2 & 1/2 \end{pmatrix}$$

Even though B is 4-by-4, $B^T B$ is not invertible, as the linear independence condition is not satisfied; specifically, the third column is $1/2$ times the first column and $1/2$ times the second column. Ideally, one could "remove" the third state responsible for the linear dependencies. Importantly, since the signal belief matrix makes no reference to the underlying states, it would not change if states were removed, provided the distributions over the signals were to not change.

I now show how to remove the state θ_3, and subsequently interpret the original state space as an auxiliary one where θ_3 is induced with equal probabilities following θ_1 and θ_2 (*after* these are already drawn). After doing this, it will be possible to recover the information structure and prior. Renormalize B so that it does not include θ_3; that is, consider the belief matrix that would emerge conditional on $\{\theta_1, \theta_2, \theta_4\}$. Considering the belief matrix that emerges when I remove θ_3 in this way, I have:

$$\tilde{B} = \begin{pmatrix} 1 & 0 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 3/5 & 2/5 \\ 0 & 0 & 1 \end{pmatrix}.$$

If one were to have started with $\tilde{B}$, then B could be obtained by considering the case where the state is "flipped" to θ_3 following θ_1, with probability $1/3$, and the state is "flipped" to θ_3 following θ_2, again with probability $1/3$ (and never following θ_4). Indeed, the third column of B is the sum of the first two columns, times $1/2$; and the first two columns of B are the same as the first two columns of $\tilde{B}$, divided by $2/3$ (and $2/3$ is the probability that the state is "unflipped"). This is the sense in which θ_3 is a linear combination of θ_1 and θ_2. In particular: whenever one column is a convex combination of other columns, one can simply eliminate it from the belief matrix, and then "regenerate" it in this way.

Now, the matrix $\tilde{B}^T \tilde{B}$ *is* invertible, and regressing each column of Q on $\tilde{B}$ gives an information structure. In this case:

$$(\tilde{B}^T \tilde{B})^{-1} \tilde{B} Q = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 1/2 & 1/2 \end{pmatrix}.$$

One can check that this information structure generates $\tilde{B}$ and Q, as Theorem 1 suggests it should, using the prior $\mathbb{P}[\theta_1] = \mathbb{P}[\theta_2] = 3/8$ and $\mathbb{P}[\theta_4] = 1/4$.

Now, notice that in the above interpretation, θ_3 is induced with equal probabilities following θ_1 and θ_2. So consider the following information structure on the original state space $\{\theta_1, \theta_2, \theta_3, \theta_4\}$:

$$\mathcal{I} = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 1/4 & 1/2 & 1/4 & 0 \\ 0 & 0 & 1/2 & 1/2 \end{pmatrix}.$$

Where does the signal distribution following state θ_3 come from? The third row of this vector is one half the first row plus one half the second row. In other words, the signal distribution is exactly what it would be if "the state is θ_3" is equivalent to "the state is θ_1 with probability 1/2 and θ_2 with probability 1/2." And indeed, one can check that this information structure, under a uniform prior (which, again, is what would the prior would be under the specification of how θ_3 is determined from θ_1 and θ_2), generates B.

B.2 Deterministically Generated Signals

In this section I discuss a special case of the framework, where signals are generated deterministically as a function of the state. There are three reasons why this is of interest. One is simply practical—there are cases where a decision-maker may observe a partition of the state space, and this alternative describes this model. Second, it is likely the simplest case where $B^T \mathcal{I}^T$ fails to be irreducible, allowing me to illustrate that the prior is not identified (and in particular why one should not expect the prior to be identified, and what is identified instead). Third, this case imposes that $|\Theta| > |S|$, since a partition of a state space by definition cannot have more elements than a state space. We will see that the ridge regression procedure can, in this case, produce an information structure, albeit one using a different limiting equation.

Suppose $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$, and consider an information structure where the decision-maker observes which element of $\mathcal{P} = \{\{\theta_1\}, \{\theta_2, \theta_3\}, \{\theta_4\}\}$ the state belongs to. If $p(\theta_i)$ is the prior probability over state θ_i, then this corresponds to the following state belief matrix and signal belief matrix:

$$B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{p(\theta_2)}{p(\theta_2)+p(\theta_3)} & \frac{p(\theta_3)}{p(\theta_2)+p(\theta_3)} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad Q = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Of course, in this example my main exercise is fairly straightforward, but this directness will be helpful in understanding the implications of irreducibility of $B^T \mathcal{I}^T$ as well as why the ridge estimator does not yield the correct information structure. I compute:

$$\lim_{\lambda\to 0} (B^TB+\lambda I)^{-1}\lambda I = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{p(\theta_3)(p(\theta_3)-p(\theta_2))}{p(\theta_2)^2+p(\theta_3)^2} & 0 \\ 0 & \frac{p(\theta_2)(p(\theta_2)-p(\theta_3))}{p(\theta_2)^2+p(\theta_3)^2} & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

This matrix converges to 0 only when $p(\theta_2) = p(\theta_3)$. And indeed, the procedure fails to produce an information structure:

$$\lim_{\lambda\to 0} (B^TB+\lambda I)^{-1}B^TQ = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{p(\theta_2)(p(\theta_2)+p(\theta_3))}{p(\theta_2)^2+p(\theta_3)^2} & 0 \\ 0 & \frac{p(\theta_3)(p(\theta_2)+p(\theta_3))}{p(\theta_2)^2+p(\theta_3)^2} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

What went wrong? Notice that in the above derivation, adding λI to B^TB was only one way to ensure the inversion step would be possible. In this case, the procedure arrives at a solution for $Q = B\tilde{\mathcal{I}}$ which does not correspond to the true information structure. As discussed, this equation *should* have multiple solutions in the case of $|\Theta| > |S|$, while the limit only considers one of them.

Note that in this case, a different regularization would deliver the information structure. For instance, one can compute that:

$$\lim_{\lambda\to 0} \left(B^TB + \lambda \overbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{p(\theta_3)}{p(\theta_2)} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}^{:=M} \right)^{-1} B^TQ = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

which is indeed the deterministic information structure in this example. One can derive this expression by following the same steps as outlined in Equation (8), but considering a different perturbation; namely, using λM instead of λI.

And indeed, there is no unique prior identified by B and Q in this case—while one can determine the prior beliefs "within a signal," in the deterministic information structure case, it is not possible to determine the relative probability across different partition elements. More generally, as illustrated by this example, if $B^T\mathcal{I}^T$ is not irreducible, then while one can use the Perron-Frobenius theorem to determine a unique prior within each irreducible class, one cannot determine the *relative* prior probabilities across classes. These relative probabilities cannot possibly be determined using *any* ex-post data, reflecting the substantive implications of the failure of irreducibility.

To conclude this section, I note that an information structure is deterministic if and only if Q is the identity. Note that in this case, $|\Theta| > |S|$ whenever the information structure does not reveal the state.

A deterministic information structure involves the decision-maker observing an element of the partition of Θ. Suppose an information structure is partitional. This implies that the probability of observing any signal given any state is either 0 or 1. On the other hand, $b_{s,\theta}$ is positive if and only if $\mathcal{I}(\theta)[s] = 1$, meaning that $q_{s,\tilde{s}}$ is equal to 1 if $s = \tilde{s}$ and 0 otherwise. Therefore, Q is the identity.

Now suppose Q is the identity. Notice that each entry of Q is a convex combination of $\mathcal{I}(\theta)[\cdot]$, weighted according to a row of B. So, if $b_{s,\theta} > 0$, then I must have $\mathcal{I}(\theta)[s] = 1$. Notice that this immediately implies that this partitions the state space, since one cannot have two signals s, s' for which $b_{s,\theta} > 0$, since this would imply the rows of $\mathcal{I}$ sum to a number greater than 1. Therefore, I obtain a partition of a subset of the state space Θ; for any $\theta \in \Theta$ that is not in this subset, I have $b_{s,\theta} = 0$ for all $s \in S$. In this case, B and Q are generated according to a partitional information structure, where each element of the partition is the support of $b_{s,\theta}$ for some s, and where the prior assigns probability 0 to any state where $b_{s,\theta} = 0$ for all s. Hence, the information is generated by a partition.