How Better Information Can Garble Experts' Advice*

Matthew Elliott† Benjamin Golub‡ Andrei Kirilenko§

June 11, 2012

Abstract

We model two experts who must make predictions about whether an event will occur or not. The experts receive private signals about the likelihood of the event occurring, and simultaneously make one of a finite set of possible predictions, corresponding to varying degrees of alarm. The information structure is commonly known among the experts and the recipients of the advice. Each expert's payoff depends on whether the event occurs, her prediction, and possibly the prediction of the other expert. Our main result shows that when either or both experts receive uniformly more informative signals, their predictions can become unambiguously less informative. We call such information improvements perverse. Suppose a third party wishes to use the experts' recommendations to decide whether to take some costly preemptive action to mitigate a possible bad event. The third party would then trade off the costs of two kinds of mistakes: (i) failing to take action when the event will occur; and (ii) needlessly taking the action when the event will not occur. Regardless of how this third party trades off the associated costs, he will be worse off after a perverse information improvement. These perverse information improvements can occur when each expert's payoff is independent of the other expert's predictions and when the information improvement is due to a transfer of technology between the experts.

Keywords: experts, committee, coarse information, delegation, advice.
JEL Classification Numbers: D71, D82, C72

*We thank Jeremy Bulow, Matthew O. Jackson, Paul Milgrom, and Michael Ostrovsky for helpful guidance. We are especially grateful to Alex Frankel and Ben Olken for suggestions that led to substantial improvements. The views presented here are our own and do not represent the views of the Commodity Futures Trading Commission, its Commissioners or staff. All errors are our own.

†Division of Humanities and Social Sciences, Caltech; and Microsoft Research. Email: melliott@hss.caltech.edu, website: http://www.hss.caltech.edu/∼melliott/

‡Department of Economics, MIT. Email: ben.golub@gmail.com, website: http://www.stanford.edu/∼bgolub/

§Commodity Futures Trading Commission. Email: akirilenko@cftc.gov

1 Introduction

Executive decisions are made under uncertainty and typically informed by multiple predictions from advisors. Whether a country goes to war depends on the predicted military outcome of the conflict and likely international reaction. Whether a vaccine should be stockpiled depends on predictions about the spread of the disease. Whether a government implements policies to avert a potential financial crisis depends on experts' judgment of its likelihood. Given constraints on an executive's time, experts and advisors only coarsely communicate the complex information they have that is relevant to such decisions. How useful is the information ultimately communicated to the executive?

The department of defense, epidemiologists, and economists seem to have ever better technologies for collecting and analyzing data, but do their coarse predictions become more useful to executives as a result? We examine the possibility that improving the information available to those who advise a decision-maker garbles the information ultimately available to that decision-maker, making that information strictly less useful regardless of the principal's relative costs of various sorts of errors.

Dramatic errors in decisions taken based on expert advice are not uncommon. Predictions that Iraq possessed weapons of mass destruction influenced the decision to invade it, but no such weapons were found.[1] Following epidemiologists' predictions of a possible 30,000 to 90,000 deaths due to an influenza pandemic[2] very large amounts of vaccine and other medication were stockpiled, but the actual number of deaths was about 10,000 and most of the medication went unused. Lastly, and perhaps most dramatically, the economics profession collectively failed to detect and warn about the impending financial crisis of 2008. While these are only a few specific examples, they do demonstrate that experts often make serious errors. And they raise the question of whether better information and more sophisticated analysis lead to better recommendations.

There are many similar examples in other settings despite apparently substantial improvements in technologies for analyzing information. Satellite surveillance and other technological advancements improved the abilities of the intelligence agencies in the U.S. in the

[1] A select committee report "Report of the Select Committee on Intelligence on the U.S. Intelligence Community's Prewar Intelligence Assessments on Iraq" (2004) concluded that: "Most of the major key judgments in the Intelligence Community's October 2002 National Intelligence Estimate (NIE), *Iraq's Continuing Programs for Weapons of Mass Destruction*, either overstated, or were not supported by, the underlying intelligence reporting. A series of failures, particularly in analytic trade craft, led to the mischaracterization of the intelligence."

[2] This prediction was made by the President's Council of Advisors on Science and Technology in the "Report to the president on U.S. preparations for 2009 – H1N1 influenza" (2009).

1990s. The same is true of the intelligence agencies in the U.K. Yet since then large terrorist attacks have occurred in both of these countries. In the medical profession, technological and scientific advances have improved the ability of doctors to detect different illnesses, yet it is not clear how well these improvements have translated into better diagnoses (Raftery and Chorozoglou, 2011).

While there are clearly many possible explanations specific to each of the above examples, several of which are likely to play important roles, we identify a possible contributing factor to all of the above outcomes. Our partial explanation is that the optimizing behavior of experts, be they economists, different intelligence agencies, doctors, or financial traders, can react in counterintuitive ways to strict improvements in the quality of their information. We assume that the experts cannot communicate their information fully, and must instead make coarse predictions. In a considerable set of circumstances, information improvements can then lead to the combined predictions of the experts becoming less informative. The effects of better information can be strictly bad in equilibrium, simultaneously for all potential consumers of the advice, no matter what their trade-offs between different kinds of errors are. These perverse outcomes do not require the experts' payoffs to depend on each other's predictions. We also show that these perverse information improvements can come from experts learning each other's technologies.

We now sketch a simplified version of the framework in which we derive these conclusions. Consider two experts, say Alice and Bob. They are called experts because they are in the business of analyzing large amounts of data using sophisticated models. Each must decide whether the results of the analysis become worrisome enough to sound the alarm about the possibility of a rare event. Sounding the alarm or not are the only actions available to Alice and Bob. In other words, as experts, they must make "bottom line" predictions.[3]

Ex ante, both Alice and Bob know that a rare event happens with some probability, but whether it is about to occur is not known to anyone. The complicated information of each expert is modeled as a signal, which is drawn from one distribution if the event is about to occur and from another distribution if it is not. Some signals are more worrisome than others, and the signals of the two experts are independent conditional on the realization of the event (i.e., whether it is about to occur).

After observing their signals, both Alice and Bob must choose whether to predict the event or not. Alice's payoff depends on her prediction and whether the bad event ends up occurring. The same holds for Bob.[4] Alice faces the usual statistical tradeoff between

[3]In Section 6 we extend our model to allow experts to make finer-grained predictions.

[4]In the general model we will also permit an expert's payoff to depend on the predictions of the other

sounding the alarm when she should not have (making a Type I error) and failing to sound it when she should have (making a Type II error). So does Bob. Each of them is a rational agent, optimizing against the uncertainty inherent in the environment.

Suppose that Alice and Bob start to get systematically more informative signals. That is, they gain the ability to make predictions that have both a lower chance of Type I error and a lower chance of a Type II error than any prediction they could make before. This may happen because they get access to new data or because they figure out a way to process existing data better. Society now has the potential to be better-informed than before.

Nevertheless, an improvement in the quality of Alice's and Bob's information may result in a perverse effect. We identify a condition under which an improvement in the quality of Alice and Bob's information can make their collective predictions less informative. This condition requires both Alice and Bob to 'envy' each other's error rates before their information improvements.

Each of Alice and Bob will always make either fewer Type I errors, or fewer Type II errors following an information improvement, but neither of them has to improve on both dimensions. Suppose that initially Bob makes fewer Type I errors than Alice but more Type II errors. Following an improvement in his information, Bob, envious of Alice, may choose to reduce his Type II error rate but increase his Type I error rate . Alice, envious of Bob, may choose to make fewer Type I errors but more Type II errors following an improvement in her information. If the reduction in Bob's Type II errors is dominated by Alice's increase in Type II errors, while Alice's reduction in Type I errors is dominated by Bob's increase in Type I errors, then Bob and Alice's collective predictions can become less informative.[5] Information improvements that result in these perverse outcomes can themselves be fairly natural. For example, perverse outcomes are possible when the information improvements result from the experts learning each other's capabilities, e.g. via technology transfer (section 4.1).

Suppose the rare event being predicted by the experts is bad, but that society can take costly preemptive action. How can society best use the predictions of the experts to decide whether to take the preemptive action or not? Can society make a clever choice of a decision rule that mitigates the effect of a perverse information improvement? Unfortunately, we show that no decision rule can mitigate the effect of a perverse information improvement. After a perverse information improvement, every potential consumer of the experts' advice is simultaneously worse off, regardless of how that consumer trades off different types of errors.

expert.

[5]We note that perversity of information improvement is a group phenomenon. If Alice is the only expert and she gets a more informative signal, then she must do better with respect to at least one type of error.

Section 2 lays out the basic model, assumptions, and definitions, and discusses the existence of equilibrium. Section 3 defines perverse information improvements and provides a simple example of one. Section 4 identifies a necessary and sufficient condition for perverse information improvements to exist when experts' payoffs do not depend on each other's predictions. We then show that information improvements due to technology transfer can be perverse. These results are extended first to environments with strategic experts (whose payoffs *do* depend on each other's actions) in Section 5, and then in Section 6 to permit the experts to also make finer predictions. Section 7 concludes.

1.1 Related Literature

Our paper is most closely related to the literature on committees.[6,7] That literature focuses on how committees can make informed decisions or give informed advice when individual members have private information and only partially aligned incentives. The problem of committee members collating their information to make a decision is therefore similar to the problem of a third party (principal) interpreting the predictions of multiple experts to make a decision.

The typical setup in the committees literature is as follows: (i) the committee must make a binary decision; (ii) committee members have private information; (iii) all members have the broadly aligned objective of wanting the committee to avoid Type I and Type II errors; and (iv) although all committee members dislike both Type I and Type II errors, their relative preferences over avoiding Type I versus Type II errors can differ. Committee members, knowing each other's preferences, may then strategically adjust their reports so that the committee's decision trades off the possibility of making Type I and Type II errors in a way that aligns with their preferences. These strategic manipulations result in private information becoming garbled and the committee making less well informed decisions than

[6]There are also some connections with the literature on testing experts. In this paper we assume that experts truly have expertise in their field and restrict them to make a coarse and easily falsifiable prediction about a single event. However, when experts' can make more nuanced predictions about many events, the issue of separating true experts from false experts is raised. There are then important philosophical questions about what constitutes a "scientific" prediction, and whether true experts can be distinguished from false but strategic experts. See Sandroni (2003), (2006), Olszewski and Sandroni (2008; 2010), Fortnow and Vohra (2009), Al-Najjar and Weinstein (2008), and Feinberg and Stewart (2008).

[7]A second related literature is the one on herding, starting with the seminal papers of Banerjee (1992) and Bikhchandani, Hirschleifer, and Welch (1992). One interpretation of herding models is that "experts" receive many coarse and dependent signals (the actions choices of previous experts) before making a coarse "recommendation". It is the coarseness and dependence in *observed* signals that restricts information aggregation. In our model observed signals drawn from a continuous distribution and only recommendations are coarse.

it could if it shared information completely (Li, Rosen and Suen (2001)).

We differ from the committees literature primarily in focusing on different phenomena and asking different questions. In the committees literature, a primary concern is how information should be transmitted to mitigate strategic distortions. Solutions involve coarse information being transmitted in equilibrium, in a way that is reminiscent of cheap talk models (Crawford and Sobel (1982)). A primary question is then how fine-grained the information communicated is. We instead fix the coarseness of information transmission and then consider the informativeness of experts' joint predictions as the quality of their information varies.[8] In particular, we focus on the joint informativeness of the experts' advice before and after their information has improved. While in the committees literature more informed experts are typically beneficial (Li, Rosen and Suen (2001), Li and Suen (2004), Suen (2004)), we show that experts can be collectively less informative following an information improvement.

One exception to the tendency of better information to lead to better predictions is Ottaviani and Sørensen (2006). In that work, experts care about their perceived ability and have different (privately known) abilities, modeled through the quality of the signals they receive. Experts then distort their predictions in order to appear to be high ability and endogenously only binary recommendations can be sent in informative equilibria. However, there may be multiple such equilibria and it is possible for a better-informed expert to provide a recommendation that is of no use to a *specific* principal, while a less well informed expert sometimes provides a useful recommendation. Our model is very different insofar as our experts do not seek to affect a principal's estimate of their unobserved information type; indeed, in our model all information structures are commonly known. We also get a stronger perversity result. Following improvements in their information, *all* potential principals simultaneously can receive strictly less informative recommendations, rather than just a given principal with specific preferences over Type I and Type II errors.

2 Predictions Game

We first study the model with binary predictions. An extension in which experts can make n predictions is discussed in Section 6.

[8]This is also an important way in which we differ from the delegation literature. The delegation literature, starting with Holmstrom (1983), broadly studies how a principal should optimally restrict an agent's freedom of choice when delegating a decision to that agent. By assumption, in our model, each expert can only make coarse predictions.

2.1 The Environment

Players There are two experts indexed indexed $i = 1, 2$.

Actions Each expert chooses whether to predict that an event will occur. We denote by $x_i \in \{1, 0\}$ the prediction of expert i, with $x_i = 1$ corresponding to the positive prediction.

Timing and Information Structure

1. Nature draws a Bernoulli random variable X with distribution

$$X = \begin{cases} 1 & \text{with probability } p \\ 0 & \text{with probability } 1 - p. \end{cases}$$

   We define the event E to be $\{X = 1\}$; its complement $\{X = 0\}$ is denoted by $\overline{E}$.

2. For each expert i, nature draws a private signal $S_i \in \mathcal{S} = [0, 1]$ whose cumulative distribution function is: G_i if $X = 1$ and F_i otherwise. These signals are conditionally independent, where the conditioning is on E.

3. Simultaneously[9], each expert predicts $x_i \in \{1, 0\}$.

Strategies The strategy of each expert is a Borel measurable function $\sigma_i : [0, 1] \to [0, 1]$, with $\sigma_i(s)$ describing the probability with which i plays $x_i = 1$ upon observing the signal $S_i = s$.

Payoffs Each expert's payoff depends on her prediction, the other expert's prediction and whether the event E occurs. If i makes prediction x_i and j makes prediction x_j then i's payoff if E occurs is $\pi_i(x_i, x_j) \equiv \pi_i(x_i, x_j \,|\, E)$ and $\overline{\pi}_i(x_i, x_j) \equiv \pi_i(x_i, x_j \,|\, \overline{E})$ if E does not occur.

The experts simultaneously attempt to correctly predict the occurrence of the event E. They predict the event E by choosing the action 1 (and conversely predict the event $\overline{E}$ by choosing the action 0). We assume that the experts like making correct predictions and dislike making incorrect predictions: $\pi_i(1, x_j) > \pi_i(0, x_j)$ and $\overline{\pi}_i(0, x_j) > \overline{\pi}_i(1, x_j)$ for either prediction x_j that j can make. An expert who knew the event E had occurred would have a

[9] See Ottaviani and Sørensen (2001) for a comparison of sequential and simultaneous predictions when experts wish to maintain their reputation for being well informed.

dominant strategy to play 1, and similarly, one who knew that $\overline{E}$ had occurred would have a dominant strategy to play 0. However, the experts receive only a noisy signal of E. Given the signal they receive they update their beliefs[10] before choosing their action $x_i \in \{1, 0\}$. Expert j's prediction may or may not affect expert i's payoff.

2.2 Signals

We assume that F_i has a density f_i and G_i has a density g_i, each defined on $[0, 1]$, such that $f_i(s) > 0$ and $g_i(s) > 0$ for all $s \in [0, 1]$ and for $i \in \{1, 2\}$. No signal $s_i \in (0, 1)$ is fully informative and so the quantity $g_i(s)/f_i(s)$ is well-defined in this domain. This quantity corresponds to expert i's odds ratio on the event E (relative to its complement $\overline{E}$) after observing a signal $S_i = s$. That is,

$$\frac{g_i(s)}{f_i(s)} = \frac{\mathbf{P}_i(E \mid S_i = s)}{1 - \mathbf{P}_i(E \mid S_i = s)}.$$

We assume that the signals in the domain $[0, 1]$ are ordered such that this odds ratio is strictly increasing as a function of s for each expert i and that the odds ratio is lowest when $s = 0$. Higher signals indicate the event E is more likely to occur. This is the *monotone likelihood ratio property*. Finally, we assume for sufficiently high signals, E becomes very likely and that for sufficiently low signals E becomes very unlikely: as $s_i \to 1$, we have $\frac{g_i(s)}{f_i(s)} \to \infty$ and as $s_i \to 0$, we have $\frac{g_i(s)}{f_i(s)} \to 0$.

Via a change of variables, we may assume without loss of generality that $F_i(s) = s$ for all $s \in [0, 1]$: that is, we may assume that S_i is uniformly distributed when the event E does not occur. This makes $f_i(s)$ the uniform density. It also makes $g_i(s)$ equal to the odds ratio. One immediate implication of the monotone likelihood ratio property is that $g_i(s)$ is strictly increasing in s.

2.3 Discussion of the Environment

We restrict the experts to make yes/no predictions to capture the idea that the "bottom-line" predictions experts make are much coarser than their information. In Section 6 we relax the assumption that predictions are binary by allowing the experts to choose from a finite number of predictions of different strengths. Our main results are easily extended to this setting.

[10] These are beliefs of expert i about the event E, about what the other expert will do, about the beliefs of the other expert regarding what i will do, and so on.

We also restrict the experts to receive conditionally independent signals. Including an additional signal observed by both experts, but not the consumers of their predictions, would not change results. The commonly observed signal would adjust the common prior of the experts but subsequent analysis would go through with this new common prior. A fully general model would permit for arbitrary correlations of experts' signals and model their joint distribution. However, this generality would come at the cost of substantial additional complexity, and so we focus on the case of conditionally independent signals.

2.4 Equilibrium and Errors

In Appendix A we show that cutoff strategies are always played in equilibrium, and that a Bayes-Nash equilibrium exists.[11] In a given equilibrium, without loss of generality we may assume there are numbers (γ_1^*, γ_2^*) such that expert i predicts the event E whenever $S_i \geq \gamma_i^*$, and we refer to $\gamma_i^* \in [0, 1]$ as i's cutoff. Given that cutoff strategies are being played, we can simply define some standard statistical quantities for each expert. In the event that E does not occur, the distribution of the signal S_i is uniform, so $1 - \gamma_i^*$ is the probability of expert i choosing $x_i = 1$ when E does not occur. This is the probability of expert i making a Type I error. Similarly, $G_i(\gamma_i^*)$ is the probability of expert i making a Type II error – that is, choosing $x_i = 0$ when E occurs. To summarize:

$1 - \gamma_i^*$	is i's probability of Type I error (when event E does *not* occur).
$G_i(\gamma_i^*)$	is i's probability of Type II error (when event E occurs).

3 Better Informed Experts Can Give Less Informative Advice

We now identify conditions under which "better" information becoming available to the experts can make their predictions less informative to any principal.

3.1 The Perturbation: Better Information

The first step is to formally define an improvement in information.

In Section 2.4 it was shown that the experts' cutoff strategies mapped directly into their probabilities of making the two kinds of errors (Type I and II). When choosing their cutoff

[11]We do this for the more general environment in which the experts can choose from n predictions to make.

strategies, experts are trading off a lower cutoff and a lower probability of a Type I error (a false positive) versus a higher cutoff and a lower probability of a Type II error (failing to predict the event, a false negative). This tradeoff can be represented as a *error avoidance possibility frontier* where the signal technology constrains their choice of Type I and Type II errors. This is shown in Figure 1.

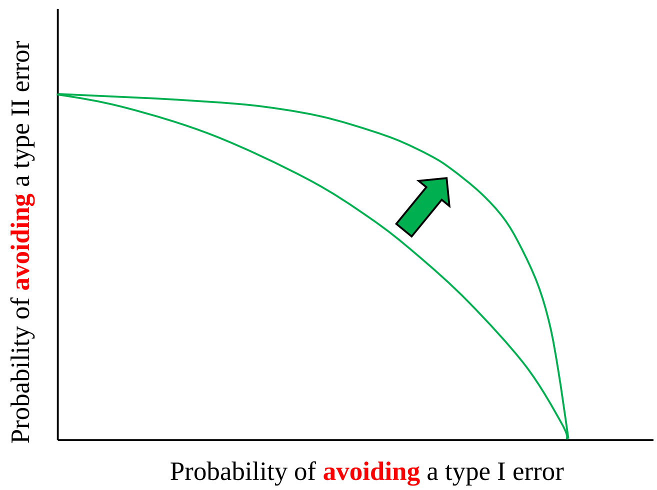


Figure 1: Better information: A shift outward of the error avoidance possibility frontier.

An expert i receives better information when her error avoidance possibility frontier shifts out. It is straightforward to show that this corresponds exactly to a first order stochastic dominance shift outwards in the CDF (G_i) of the signal under the event E, such that when the event E occurs, higher signals are received with higher probability. Indeed, the error avoidance possibility frontier is $s_i \mapsto 1 - G_i(s_i)$.[12] An expert with improved information can then reduce both her Type I and Type II error probabilites.

Definition 1. Formally, *an information improvement over* (G_1, G_2) is defined to be any pair of distributions $(\widetilde{G}_1, \widetilde{G}_2) \neq (G_1, G_2)$ such that $\widetilde{G}_i$ first-order stochastically dominates G_i for each $i \in \{1, 2\}$.

[12] As F_i is uniformly distributed on $s = [0, 1]$ the probability of i making a Type I error is $1 - \gamma_i^*$. The probability of making a Type II error is $G_i(\gamma_i^*)$. Plotting $1 - G_i(s_i)$ then identifies the trade-off an expert faces between avoiding a Type I error and a Type II error. As $g_i(s_i)$ is increasing in s_i (by the monotone likelihood ratio property), the function $1 - G_i(s_i)$ is concave.

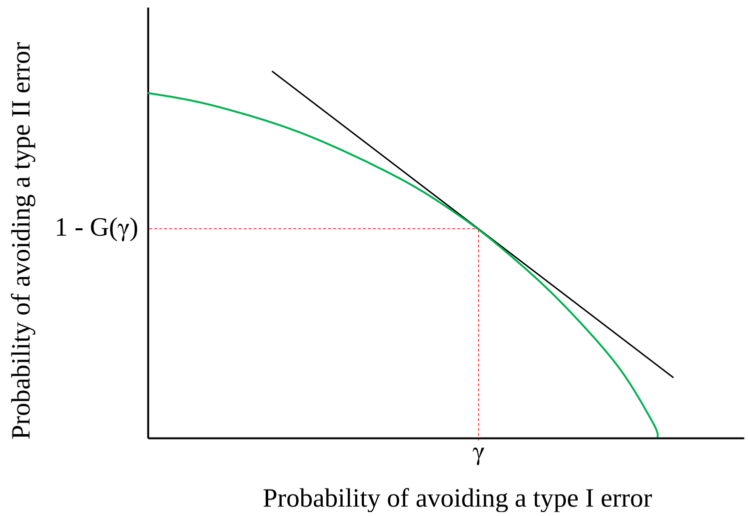


Figure 2: The slope of the line reflects the cost of a Type II error relative to a Type I error; along the line, an expert is indifferent. Setting it tangent to the curve described by $\gamma \mapsto 1 - G(\gamma)$ yields the optimal cutoff.

Looking at the tradeoff between different kinds of errors as a "production possibilities frontier" has another advantage. We can plot an indifference curve of an expert with a certain tradeoff between the two types of errors. Fixing the behavior of the other expert, this is a line, since the expert has expected utility preferences and her utility is thus linear in the probabilities of various types of errors. Let $\beta_i(\gamma_j)$ be the slope of i's indifference curve when expert j is choosing cutoff γ_j. This can be computed explicitly[13]. Finding a point of tangency between that line and the production frontier is a way to find the optimal cutoff (see Figure 2).

[13]Player i's expected payoff has the form:

$$\mathbb{E}[\pi_i|\gamma_j] = \Phi_0(\gamma_j) + \Phi_1(\gamma_j)(1 - G_i(\gamma_i)) + \Phi_2(\gamma_j)\gamma_i,$$

where $\Phi_0(\gamma_j)$, $\Phi_1(\gamma_j)$ and $\Phi_2(\gamma_j)$ are functions of γ_j (the other expert's cutoff choice) but not γ_i. Setting this expression equal to a constant defines this expert's (linear) indifference curve in the coordinates depicted in Figure 2, where a typical point is $(\gamma_i, 1 - G(\gamma_i))$: probabilities of avoiding the two types of errors. The slope of this indifference curve is then $\beta_i(\gamma_j) = -(\Phi_2(\gamma_j)/\Phi_1(\gamma_j))$. It can verified that

$$\Phi_1(\gamma_j) = p\left[(1 - G_j(\gamma_j))(\pi_i(1,1) - \pi_i(0,1)) + G_j(\gamma_j)(\pi_i(1,0) - \pi_i(0,0))\right],$$

and

$$\Phi_2(\gamma_j) = (1-p)\left[\gamma_j(\overline{\pi}_i(0,0) - \overline{\pi}_i(1,0)) + (1-\gamma_j)(\overline{\pi}_i(0,1) - \overline{\pi}_i(1,1))\right].$$

3.2 An Illustration of What Can Go Wrong

Improvements in experts' information do not necessarily have the intuitive effect of improving the quality of information available to an outside observer. Indeed, it is possible that experts' information strictly improves but the information available to be extracted becomes strictly worse in a sense that we now make precise. The following example illustrates how better-informed experts can make less informative predictions.

To emphasize that perverse information improvements are possible without any strategic effects (e.g., trying to "outguess" the other expert), suppose that both experts' payoffs are independent of each others' predictions. Let expert 1 start with the information depicted in Figure 3a and receive the information improvement depicted in Figure 3b. Following this information improvement expert 1 will choose to increase her cut-off, making more Type II errors but fewer Type I errors. Expert 1's rate of Type I and Type II errors therefore changes from the point 1 to the point 1'.

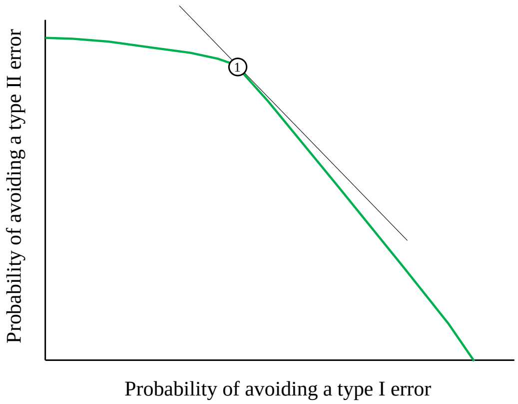
Expert 2 also receives an information improvement which is depicted as a move from Figure 3c to Figure 3d. After this information improvement expert 2 chooses to decrease her cut-off value, making fewer Type I errors but more Type II errors, and moving from point 2 to 2'.

After these information improvements, expert 1 makes more Type I errors *and* more Type II errors than expert 2 did before the information improvements. Similarly, expert 2 makes more Type II errors *and* more Type I errors than expert 1 did before the information improvements. This is shown in Figure 3e. Overall, the experts' collective predictions then become strictly less informative.

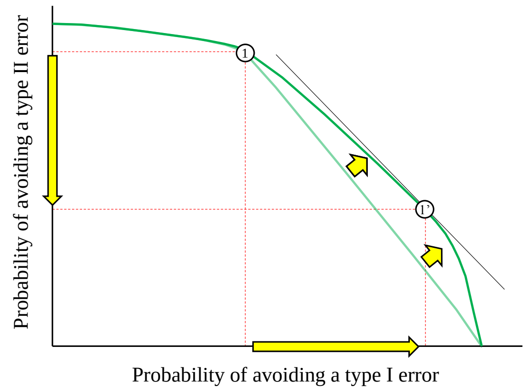
It is also interesting to consider the kind of information improvement that resulted in the perverse outcome. Comparing the experts' information before and after the information improvement in Figure 4a and 4b, it appears that, in some qualitative way, the experts' information is more similar after the information improvements. This observation forms the basis of Section 4.1 on how perversity can occur due to "technology r.

3.3 Utilizing the Experts' Predictions

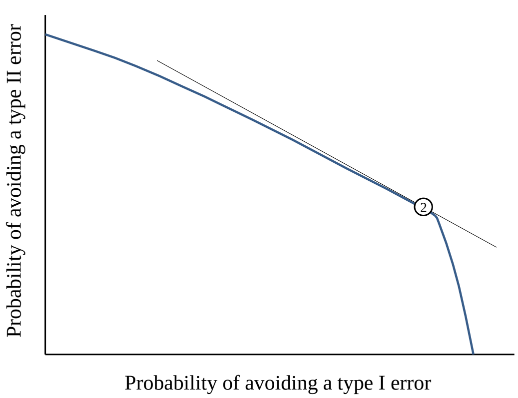
When do experts' collective predictions become less valuable to society? For the experts' predictions to have any impact there must exist decisions that are informed by those predic-



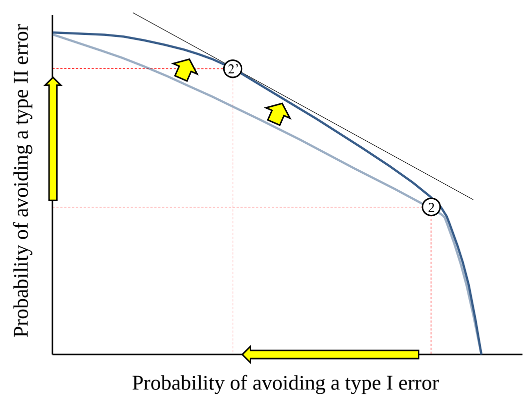
(a) 1's initial information.



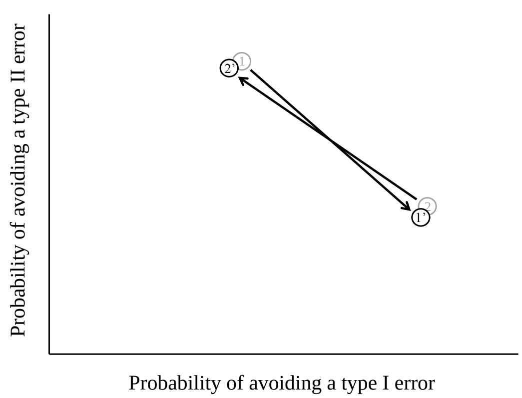
(b) 1's information improvement.



(c) 2's initial information.

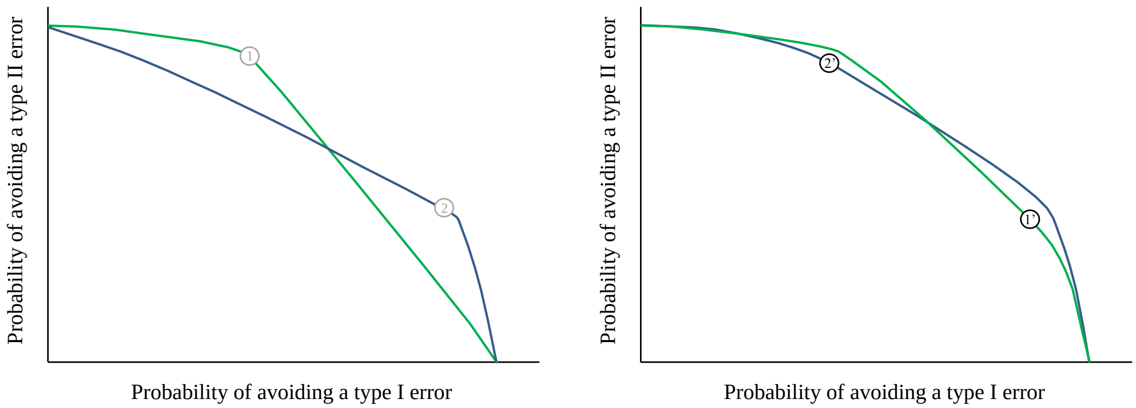


(d) 2's information improvement.



(e) Predictions become less informative.

Figure 3: Improvements in the experts' information leads them to make collectively less informative predictions.



(a) Old information. (b) New Information.

Figure 4: The experts' information looks (more) similar after the information improvement.

tions. To capture this idea, we suppose society observes the experts' predictions and must then decide whether to take costly preemptive action to mitigate the effect of the event E. We assume that society's collective payoff from taking the preemptive action is greater than not taking the preemptive action when the event E occurs, but lower when the event E does not occur. More formally, we introduce a principal to represent the interests of society and let the principal inherit society's preferences, as well as make decisions on behalf of society.[14] By choosing whether to take the preemptive action or not; the principal therefore trades off the possibility of making a Type I error (taking the preemptive action when the event E does not occur) and making a Type II error (not taking the preemptive action when the event E does occur).

The principal's payoffs are given by the following matrix, where y represents the principal's decision.

	E	$\overline{E}$
$y = 1$	a_0	$\overline{a}_0$
$y = 0$	b_0	$\overline{b}_0$

where, as just discussed, $a_0 > \overline{a}_0$ and $b_0 < \overline{b}_0$.

The principal can use the experts' predictions in any way when deciding whether to take the preemptive actions or not. The principal might take preemptive action only if both experts predict the event E, if only expert 1 predicts the event E, and so on. Let

[14] This formulation allows us to abstract from any preference aggregation issues, and focus on the informativeness of the experts' predictions.

$q(x_i, x_j) : \{0,1\}^2 \to [0,1]$ be a decision rule for the principal describing the probability with which the principal takes preemptive action following expert predictions x_1 and x_2.

From now on, we consider a larger game in which the principal is another player, who moves after the experts, choosing y, and whose payoffs are as above. "Equilibrium", will mean a Nash equilibrium of this larger game.

3.4 Perverse Information Improvements

As we are permitting the experts' payoffs to depend on each others predictions, experts' choices of predictions are modeled as a game and in general this game can have multiple equilibria. When defining perverse information improvements, the possibility of multiple equilibria needs to be accounted for. This leads us to introduce the concepts of *weakly perverse* information improvements (perverse for an equilibrium) and strongly perverse information improvements (perverse for all equilibria).

Fix an initial information structure (G_1, G_2) and initial equilibrium defined by cutoffs (γ_1^*, γ_2^*). We will assume throughout that in this initial equilibrium the principal uses the information provided by at least one expert – i.e. there exist a combination of recommendations the experts can make under which the principal chooses to take the preemptive action and other recommendations under which the principal chooses not to take the preemptive action.[15] An information improvement $(\widetilde{G}_1, \widetilde{G}_2)$ is *weakly perverse* at the equilibrium if and only if there exists an equilibrium of the predictions game under $(\widetilde{G}_1, \widetilde{G}_2)$ in which the principal is strictly worse off than before (i.e., at the original equilibrium (γ_1^*, γ_2^*) under the original information structure (G_1, G_2)), for *any* preferences the principal might have over Type I and Type II errors. Information improvements are *strongly perverse* if, in *all* equilibria of the predictions game under $(\widetilde{G}_1, \widetilde{G}_2)$, the principal is strictly worse off than before for *any* preferences the principal might have over Type I and Type II errors.[16]

The experts' recommendations (which can be thought of, from an outsider's or principal's perspective, as a pair of binary random variables). Denote by (X_1, X_2) and (X_1', X_2') these random variables before and after the information improvement, respectively. It can be

[15]These are the interesting cases. When the principal ignores both experts to begin with, no information improvement can make any principal worse off – the principal always has the option of continuing to ignore the experts, and by revealed preference, this was a best outcome for the principal before the change in information quality.

[16]A much weaker notion of perversity would require only that some principal is made worse off by the experts' information improvements. This would drop the requirement that all principals be worse off, regardless of their trade-off between different kinds of errors. For such a definition, it is straightforward to find "perverse" information improvements as long as the principal has different preferences over Type I and Type II errors from the experts.

verified that seeing (X_1', X_2') is less useful to every principal than seeing (X_1, X_2) if and only if (X_1', X_2') is a garbling of (X_1, X_2), in the sense of Blackwell (1953).

It will also be helpful to introduce the concept of "cross-dominated" information improvements. Let $\gamma_i^{*'}$ be i's cutoff in an equilibrium after the information improvement. This information improvement is *cross-dominated* for this equilibrium if $i = 1, 2$ makes both more Type I errors more Type II errors after the information improvement than $j \neq i$ did before: $1 - \gamma_i^{*'} > 1 - \gamma_j^*$ and $G_i(\gamma_i^{*'}) > G_j(\gamma_j^*)$, for $i, j \in \{1, 2\}$ and $i \neq j$. An information improvement is *strongly cross-dominated* if it is cross-dominated in all equilibria and *weakly cross-dominated* if it is cross-dominated in an equilibrium.

Lemma **1.** *The experts' information improvements are weakly (respectively, strongly) perverse if they are weakly (respectively, strongly) cross-dominated.*

The proof of Lemma 1 is in the Appendix. It can be shown, by example, that cross-dominated information improvements are not necessary for perverse information improvements – it is possible to have other perverse information improvements. In the next section we will focus on cross-dominated information improvements. Perversity is, however, more prevalent than this analysis alone suggests.

4 Perversity with Independent Experts

To identify conditions under which information improvements can be perverse, we consider first the case of *independent* experts, by which we mean experts whose payoffs do not depend on each other's predictions. We do this to emphasize that perversity is not an artifact of strategic interaction between experts.

It will be helpful to define a new condition on experts' preferences. Given a strategy profile of the game, an expert i is *envious* of j if i would strictly prefer to have j's Type I and Type II error rates over any error rates on i's own error avoidance frontier.

Theorem **1.** *Assume experts are independent and fix any equilibrium of the game. Then there exists a strongly cross-dominated information improvement at that equilibrium if and only if both experts are envious of each other. Consequently, if both experts are envious of each other, there exists a strongly perverse information improvement.*

The final sentence of the theorem follows by Lemma 1.

The proof of Theorem 1 is in the Appendix, but intuition can be gained from Figure 5. Figures 5a and 5b depict two envious experts. Expert 1 chooses point 1 on her error avoidance frontier but would prefer the error rates at point 2. Expert 2 chooses point 2 on her error avoidance frontier but would prefer the error rates at point 1.

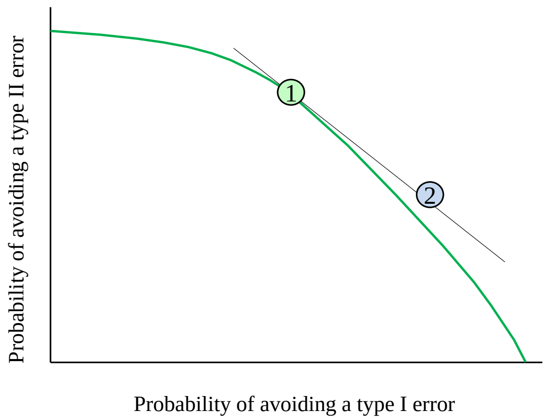
Information improvements for experts 1 and 2 can then be constructed so that, for each expert, the other expert's initial Type I and Type II error rate become just feasible. This is shown in Figures 5c and 5d. Expert 1's information improvement is constructed by finding the convex hull of 1's initial error avoidance frontier and 2's initial Type I and Type II error rates. As expert 1 was envious of expert 2, after her information improvement, expert 1 will choose expert 2's original point in the space of error rates. The analogous exercise can be done for expert 2. The experts therefore "switch places" given the information improvements shown in Figures 5b and 5c.

It is easy to adjust the constructed information improvements slightly so that the experts' optimal Type I and Type II error rates after the information improvements are cross-dominated – i.e., so that each expert after the information improvement is doing strictly worse than the other expert before the improvement. Strongly cross-dominated and therefore strongly perverse improvements have been constructed for experts who are envious of each other.

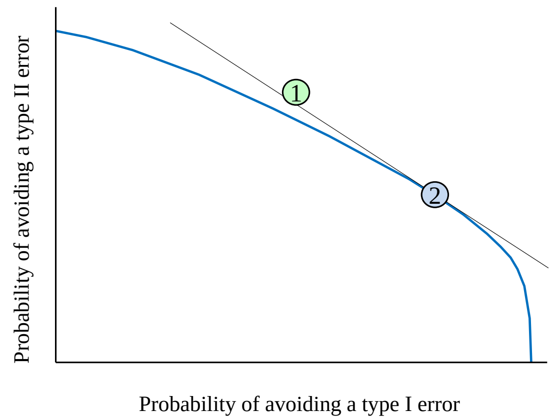
Cross-dominated information improvements with independent experts are only possible if the experts are both envious of each other. However, there do exist other types of information improvements that are also perverse. For this reason, independent experts being envious of each other is a sufficient but not necessary condition for a strongly perverse information improvement.

4.1 Information Improvements due to Technology Transfer

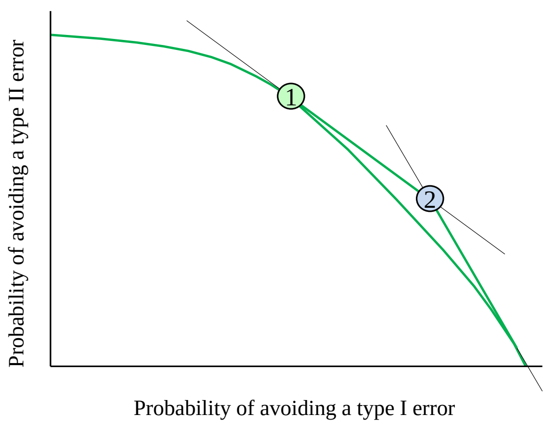
Theorem 1 only identifies conditions under which perverse information improvements are possible and is silent on which types of information improvements are perverse. A particular concern is that very specific and strange types of information improvements may be necessary for perversity. We show now that even natural, simple information improvements can be perverse. Consider two experts who have access to a fixed set of data but different technologies for analyzing the data. We will now consider information improvements that can result from the experts learning each other's technologies. We will say an information improvement *arises from technology transfer to expert i* if, loosely speaking, expert i's new



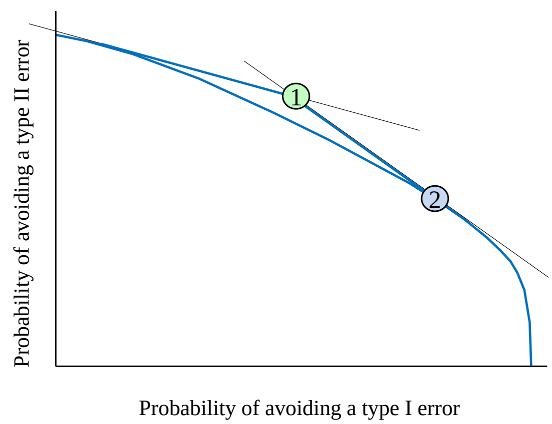
(a) 1 is envious of 2



(b) 2 is envious of 1



(c) 1's information improvement.



(d) 2's information improvement

Figure 5: The constructed improvements in the two envious experts' information results in them swapping positions – expert 1 makes Type I and Type II errors at the rate expert 2 did before the information improvements and expert 2 makes Type I and Type II errors at the rate expert 1 did before the information improvements.

error avoidance frontier lies within the convex hull[17] of her and the other expert's old frontiers. More formally, an information improvement arises from technology transfer to expert i if, for every point on expert i's new error avoidance frontier, there exists a point in the convex hull of i's own initial error avoidance frontier (the curve $\gamma \mapsto 1 - G_i(\gamma)$) and expert j's initial error avoidance frontier (the curve $\gamma \mapsto 1 - G_j(\gamma)$) with weakly fewer type I errors and weakly fewer type II errors. We say an information improvement *arises from technology*

[17] An expert that has access to two technologies for predicting the event can always achieve the convex hull of the error rates the two technologies can separately achieve – the expert simply randomizes between the predictions of one technology with one cut-off and the prediction of the other technology with a different cut-off.

transfer if it arises from technology transfer to either expert, or both.

Technology transfer may seem to be of limited value, since in some sense it does not add any new information, but merely shares information that was already present. Nevertheless, it is possible to characterize some circumstances in which such information improvements make it possible for both experts to reduce both their type I and type II errors – although they may not choose to do so.

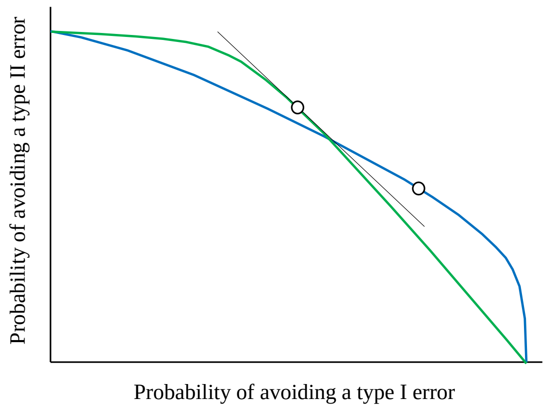
PROPOSITION **1.** *There exist information improvements that arise from technology transfer and permit* both *experts to lower their rates of type I* and *type II errors if and only if both experts are envious of each other.*

This result is straightforward and we omit a proof. The typical case is shown in Figures 6a and 6b, in which we start with two experts who strictly prefer each other's points in error space. Consider the convex hull of both experts' possible error combinations. As experts are envious of each other, they prefer each other's error rates; the indifference curve of each expert passes below the other expert's combination of error rates. As can be seen in Figures 6a and 6b, this implies that the convex hull of the experts' initial error avoidance frontiers will contain, in its interior, both expert's initial error rates.

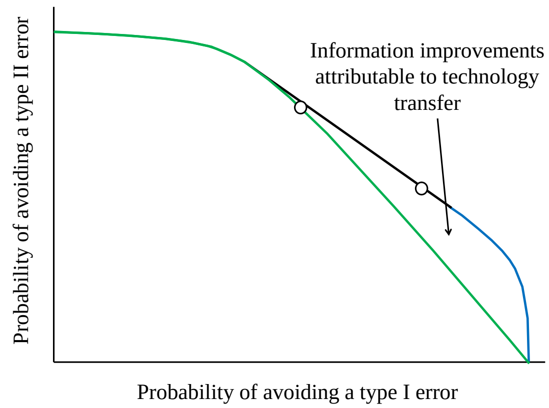
We now state a proposition describing how technology transfer can have unambiguously harmful effects.

PROPOSITION **2.** *If the experts are independent and envious of each other, then there exist information improvements which arise from technology transfer and are strongly perverse.*

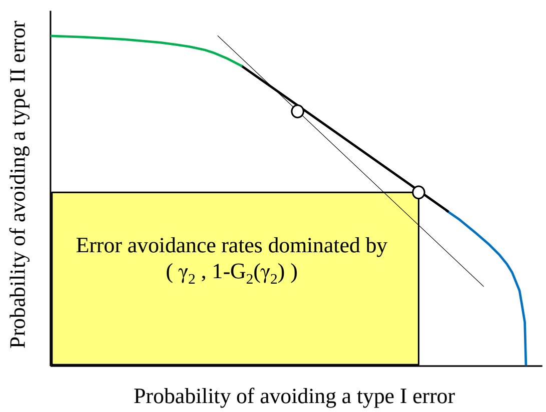
To gain intuition for this result, consider Figure 6. Figure 6a shows the initial error avoidance frontiers for experts 1 and 2. The convex hull of the curves $\gamma \mapsto 1 - G_i(\gamma)$ and $\gamma \mapsto 1 - G_j(\gamma)$ is then constructed in Figure 6b to illustrate the scope for information improvements for expert 1 that can be attributed to technology transfer. In Figure 6c the error avoidance rates dominated by expert 2's initial error avoidance rates of $(\gamma_2^*, 1 - G_2(\gamma_2^*))$ are shown. Figure 6d shows an information improvement that arises from technology transfer to expert 1 such that the new point of tangency between 1's indifference curves and 1's new error avoidance frontier is dominated by expert 2's initial error avoidance rates. Expert 1 will then choose these new error rates after her information improvement. Constructing a similar information improvement for expert 2 would make these information improvements perverse.



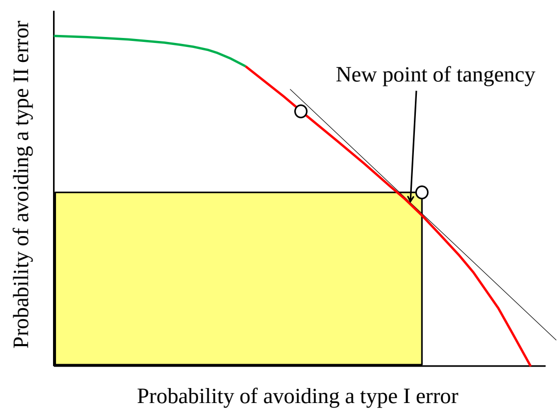
(a) Error avoidance frontiers.



(b) The convex hull.



(c) Region for new tangency point.



(d) A possible information improvement for 1.

Figure 6: An example of how to construct information improvements that can be attributed to technology transfer and are perverse. Here only the information improvement of expert 1 is shown. For perversity a similar information improvement is also required for expert 2.

5 Perversity with Dependent Expert Preferences

We now focus on extending the results above to situations where experts' payoffs are not necessarily independent – that is, where there is a direct payoff effect on one expert when the other changes her prediction. This can occur in a multitude of ways – for example, an expert correctly predicting a crisis when she was the only one to do so. There are then incentives to strategically manipulate predictions, depending on what the other expert is doing.

To extend Theorem 1 to permit strategic experts, a condition on the experts' initial equilibrium strategies called *playing to their strengths* is helpful. Expert i's prediction cannot be dominated by expert j if there is no cutoff strategy j can play that would result in j making both weakly fewer Type I errors and weakly fewer Type II errors than i. When, in a given equilibrium, an expert i is making predictions that cannot be dominated by expert j, we say that *expert i is playing to her strengths* in that equilibrium.

When the experts are playing to their strengths in an equilibrium (γ_1^*, γ_2^*), expert j's equilibrium avoidance of Type I and Type II errors $(\gamma_j^*, 1 - G_j(\gamma_j^*))$ will lie outside of i's error avoidance frontier. As in the proof of Theorem 1, we can then construct the convex hull of the point $(\gamma_j^*, 1 - G_j(\gamma_j^*))$ and i's error avoidance frontier (see Figure 5 throughout this discussion). Suppose for now that the error avoidance frontiers are continuously differentiable (see the proof of Proposition 3 for the general case, including more general definitions of $\widehat{\gamma}_{i,a}$ and $\widehat{\gamma}_{i,b}$). First construct lines tangent with i's error avoidance frontier and passing through the point $(\gamma_j^*, 1 - G_j(\gamma_j^*))$. Let these points of tangency be $(\widehat{\gamma}_{i,a}, 1 - G_i(\widehat{\gamma}_{i,a}))$ and $(\widehat{\gamma}_{i,b}, 1 - G_i(\widehat{\gamma}_{i,b}))$ where $\widehat{\gamma}_{i,a} < \gamma_j^* < \widehat{\gamma}_{i,b}$. In other words, the numbers $\widehat{\gamma}_{i,a}$ and $\widehat{\gamma}_{i,b}$ are defined to be the two values of x that solve $1 - G_i(x) + g_i(x)(x - \gamma_j^*) = 1 - G_j(\gamma_j^*)$.

Given her preferences, expert i may find that some of her possible cutoffs are dominated in the sense that she would not want to play them regardless of what expert j choose to do. Thus, we can define, in the usual way, the rationalizable cutoff strategies R_i and R_j by iteratively eliminating strictly dominated cutoff strategies for each expert (Bernheim, 1984).

Proposition **3.** *There exist weakly perverse information improvements over the initial information structure (G_1, G_2) at an initial equilibrium (γ_1^*, γ_2^*) if:*

(i) experts are playing to their strengths at (γ_i^*, γ_j^*); and

(ii) were the experts to swap error rates, each expert i's indifference curve would have a slope between the slopes of her frontier at the points of tangency[18]: for each i, we have[19] $\beta_i(\gamma_i^*) \in (-G_i'(\widehat{\gamma}_{i,b}), -G'(\widehat{\gamma}_{i,a}))$.

These information improvements are also strongly perverse if:

(iii) all $\gamma_i \leq \widehat{\gamma}_{i,a}$ and all $\gamma_i \geq \widehat{\gamma}_{i,b}$ are not rationalizable for i (i.e., all such γ_i are not in R_i), for each player $i \in \{1, 2\}$.

The proof of Proposition 3 is constructive and is in the Appendix.

Although the economic interpretation of condition (ii) in Proposition 3 is somewhat opaque, it is implied by a condition more closely related to the envious experts condition

[18]Recall from Section 3.1 that the slope of a given expert's indifference curve is a function of the *other* expert's cutoff; thus, the usual thing to write would be $\beta_i(\gamma_j^*)$. But here we are envisioning the experts switching places, with j playing γ_i^*, and thus the slope of interest is $\beta_i(\gamma_i^*)$.

[19]The negative signs arise because the frontier of expert i is $1 - G_i$.

from Theorem 1. The difficulty of extending the concept of envious experts to a setting in which experts' payoffs depend on each other's predictions is now each expert's preferences and trade-offs depend on the cut-off played by the other expert. The concept of an expert being envious of another expert must therefore specify what the other expert is doing. We will say an expert i is *robustly envious* of j if i would strictly prefer to have j's equilibrium Type I and Type II error rates (before the information improvements), over any error rates on i's error avoidance frontier (i) holding j's error rates constant and (ii) if j switched to making Type I and Type II errors at i's current error rates.

Corollary **1.** *There exist weakly perverse information improvements if both the experts are robustly envious of each other.*

The proof of Corollary 1 is omitted. It can be verified that the construction of perverse information improvements in the proof of Proposition 3 goes through with robustly envious experts. Without restricting the experts' preferences further, multiple equilibria are possible with experts whose payoffs depend on each other. This is why Corollary 1 only identifies conditions for weakly perverse information improvements.

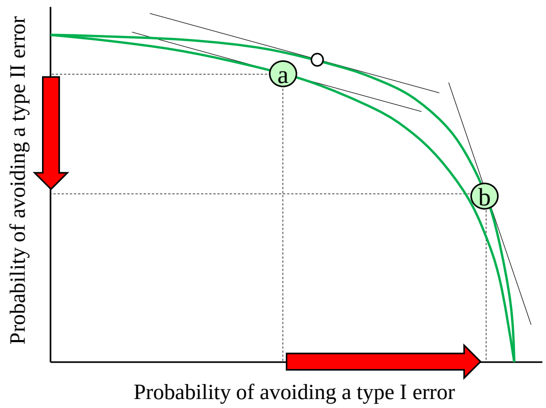
Proposition 2 can also be extended to environments with strategic experts.

Corollary **2.** *If both the experts are robustly envious of each other, then there exist information improvements that arise from technology transfer and are weakly perverse.*

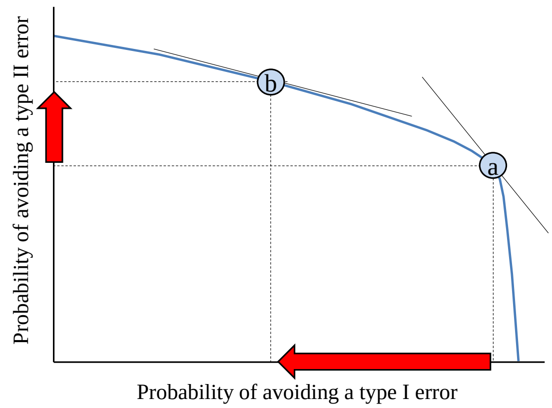
Corollary 2 follows immediately from Corollary 1 and the information improvements constructed in Proposition 3.

5.1 Cases in which Perversity Occurs *Only If* Experts' Payoffs are Dependent

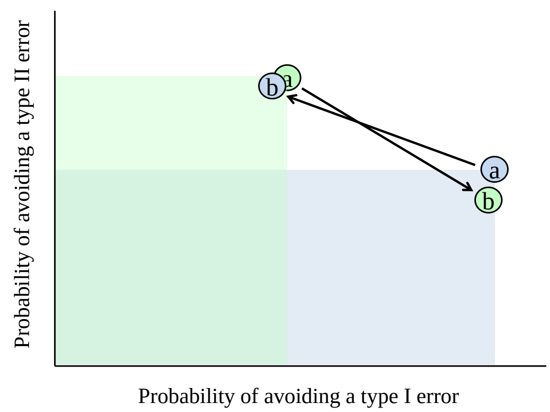
Given that we are able to give tighter (i.e., if and only if) conditions for perversity in the case where experts' payoffs are independent, it might be tempting to conclude that dependence of experts' payoffs on each others' predictions reduces the scope for perverse information improvements. This is not the case. In the following example, a perverse outcome is possible only when experts' payoffs depend on each other. Indeed, more generally, information improvements that affect only a single expert can be perverse only if the expert who does not receive the information improvement is strategic.



(a) Expert 1's information improves.



(b) Expert 2's error trade-off.



(c) Predictions become less informative.

Figure 7: An improvement in expert 1's information leads to less informative predictions.

Suppose expert 1 receives an information improvement as shown in Figure 7a. Without any strategic interaction expert 1 would then make fewer Type I errors and fewer Type II errors than before the improvement (moving from the point labeled "a" to the white circle). Her prediction would then be unambiguously more informative than her initial prediction. However, letting the experts' payoffs be dependent, each expert's optimal choice of cutoff can depend on what she expects the other expert's cutoff to be. If expert 2 expects expert 1 to increase her cutoff and make fewer Type I errors, then expert 2 may wish to differentiate herself from expert 1 by reducing her cutoff and making fewer Type II errors (but more Type I errors). Alternatively, expert 2 may want to herd with expert 1 and also increase her cutoff so she also makes fewer Type I errors (but more Type II errors).

In Figure 7 expert 1's movement out to the white circle changes expert 2's optimal choice of cutoff. At the new equilibrium, expert 2 moves along her error avoidance frontier from the point labeled "a" in Figure 7b to the point labeled "b", while expert 1 also moves along her new error avoidance frontier from the white circle (where she would be holding expert 2's cutoff fixed) to the point labeled "b" in Figure 7a.[20] In the new equilibrium, expert 1 chooses a higher cutoff and makes fewer Type I errors but more Type II errors. In contrast, expert 2 chooses a lower cutoff, so she makes fewer Type II errors but more Type I errors. Overall, the experts' collective predictions then become strictly less informative to the principal. If expert 2's payoffs did not depend on the other expert's, she would have kept chosing error rates at point labeled "a" in Figure 7b and expert 1 would have chosen error rates corresponding to the white circle in Figure 7a. Instead of being perverse, the information improvement would then have made any principal better off.

Although in this example a single information improvement was sufficient to generate a perverse outcome, perversity is, nevertheless, a multiple-expert phenomenon. If the principal listened to a single expert, then after an information improvement the expert would make fewer Type I errors or fewer Type II errors, or both. There would then exist some principal who traded off Type I errors and Type II errors in such a way as to be made better off by the information improvement, and the information improvement would not be perverse.

6 Less Coarse Recommendations

In this section, we extend the baseline binary prediction model by expanding the choice sets of the experts. Each expert can now choose one of n possible predictions, $x_i \in \{1, \ldots, n\}$. Letting Δ^n be the n-dimensional simplex, the strategy of each expert is a Borel-measurable function $\sigma_i : [0,1] \to \Delta^n$, with $\sigma_{i,k}(s)$ describing the probability with which i plays $x_i = k$, for $k \in \{1, \ldots, n\}$ upon observing the signal $S_i = s$. If expert i chooses a prediction x_i and the other expert chooses a prediction x_j, then i's payoff is denoted by $\pi_i(x_i, x_j) \equiv \pi_i(x_i, x_j|E)$ when E occurs and by $\overline{\pi}_i(x_i, x_j) \equiv \pi_i(x_i, x_j|\overline{E})$ when E does not occur. We assume that the principal interprets a prediction $k' > k''$ as a statement warning more strongly of the event E than prediction k'' and again assume that both experts dislike making Type I and Type II errors. It follows that if $k' > k''$ then $\pi_{i,k',k} > \pi_{i,k'',k}$ and $\overline{\pi}_{i,k',k} < \overline{\pi}_{i,k'',k}$.

[20]We omit the details of solving for an equilibrium here because the key conceptual point concerns the changes in error rates, not the details of how they come about.

6.1 Cutoff Strategies

If, in a given strategy, expert i chooses to make m different predictions with positive probability we will say that i's strategy has support of size m. For any strategy with support of size m we can label the predictions made with positive probability $x_1, \ldots, x_m$ while preserving their order. Suppose we can find a sequence $\gamma_{(x_1,x_2)}, \gamma_{(x_2,x_3)}, \ldots, \gamma_{(x_{m-1},x_m)}$ so that expert i makes the prediction x_k if and only if $s_1 \in [\gamma_{(x_{k-1},x_k)}, \gamma_{(x_k,x_{k+1})})$, where we interpret $\gamma_{(x_0,x_1)}$ to mean 0 and $\gamma_{(x_m,x_{m+1})}$ to mean 1. Then we will say cutoff strategies are played in equilibrium. It is shown in Appendix A.2 that equilibria exist and that in all equilibria cutoff strategies are played. In fact, something a little stronger is shown. Each expert will want to play a cutoff strategy regardless of the strategy chosen by the other expert.

6.2 Further Endogenous Coarsening

PROPOSITION **4.** *Under the maintained assumptions, there exist payoffs such that in all equilibria both experts only make predictions in the set* $\{1, n\}$.

To gain intuition for Proposition 4, consider the payoff expert 1 receives from making a prediction $h \in \{1, \ldots, n\}$ as a function of her signal. From Equation 3 in Appendix A.2, it follows that expert 1's payoff from choosing prediction h is linear in her expected probability of the event occurring conditional on her signal ($\mathbf{P}_1[E|s_1]$):

$$u_1(h) = \left[\sum_{k=1}^{n} \mathbf{P}[x_2 = k|\overline{E}]\overline{\pi}_1(h,k)\right] + \mathbf{P}_1[E|s_1]\left[\sum_{k=1}^{n} \left(\mathbf{P}[x_2 = k|E]\pi_1(h,k) - \mathbf{P}[x_2 = k|\overline{E}]\overline{\pi}_1(h,k)\right)\right]$$

To see how expert 1 will change her reports for different signals, we can plot equation 1 for expert 1's different possible predictions.[21] When $\mathbf{P}_1[E|s_1] = 0$, expert 1's highest possible payoff comes from making her weakest prediction of event E ($x_1 = 1$). When $\mathbf{P}_1[E|s_1] = 1$, expert 1's highest possible payoff comes from making her strongest prediction of event E ($x_1 = n$). An example of expert 1's possible prediction choices for different signals is shown in Figure 8 for $n = 4$.

[21] In Figure 8 we put $\mathbf{P}_1[E|s_1]$ on the x-axis, but as $\mathbf{P}_1[E|s_1]$ is monotonic in s we could instead have used s.

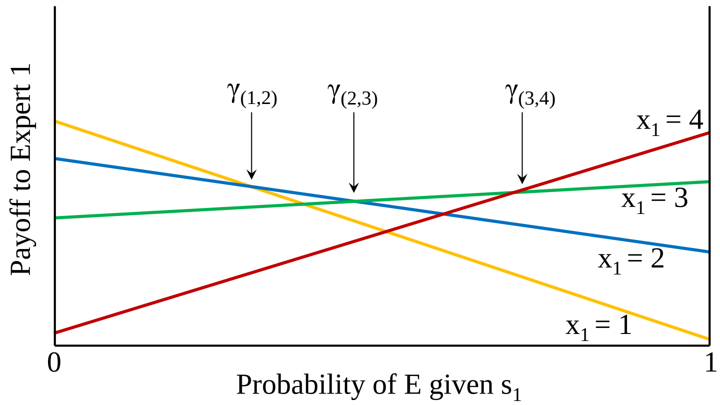


Figure 8: Expert 1's payoffs from different predictions.

In Figure 8, expert 1 uses each of her available predictions for some signals. However, this need not be the case. If the value to expert 1 of making the extreme predictions and being correct is increased, then expert 1 may endogenously choose only to ever report two predictions. Such a further endogenous coarsening of information is shown in Figure 9.

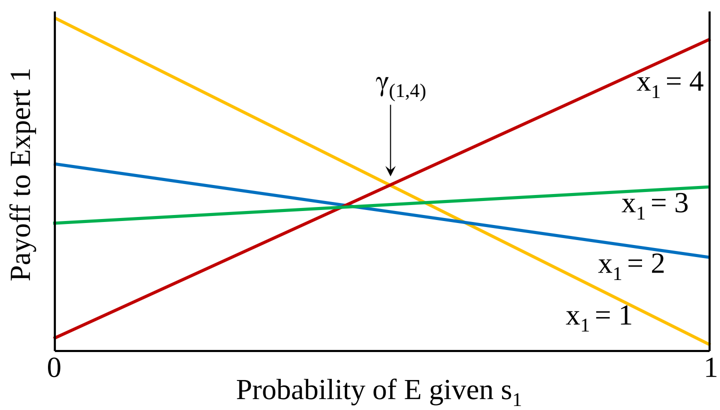
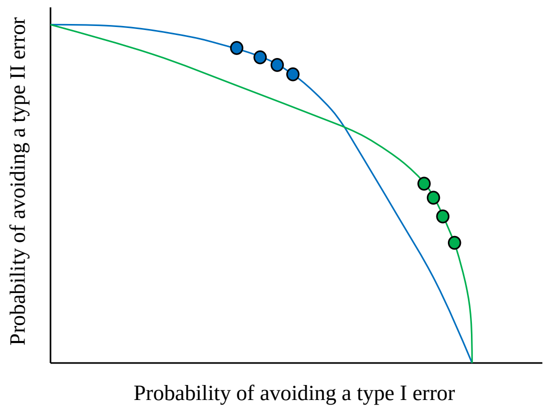


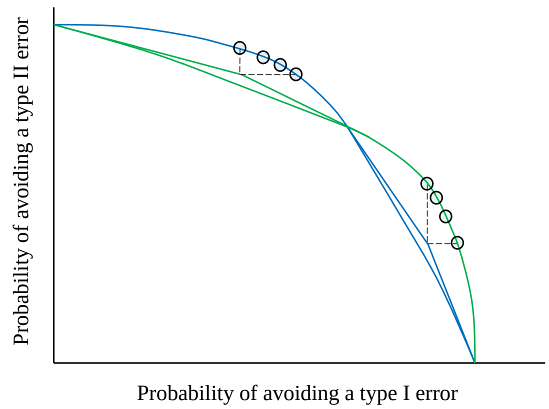
Figure 9: Expert 1 uses only her extreme predictions.

6.3 Perverse Information Improvements

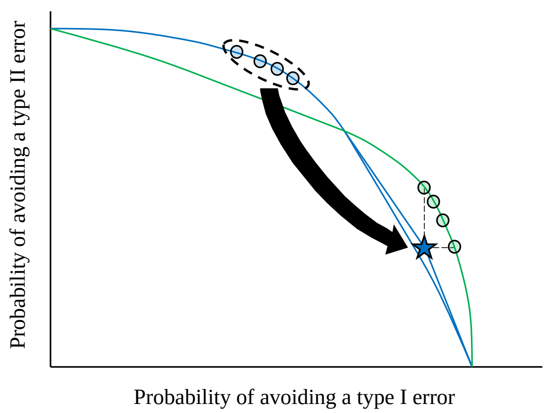
An immediate implication of Proposition 4 is that Proposition 3 extends to cases in which experts are able to make multiple possible predictions but choose to make only binary predictions. There are many other ways in which the sufficient conditions already identified for perverse information improvements can be extended to a setting of n predictions. However, in this section we focus on how information improvements can generate an endogenous coarsening of predictions that also results in information improvements being perverse. While this coarsening is not necessary for perverse information improvements, it is instructive to consider how coarsening can further inhibit a principal from extracting the potential benefits from information improvements. An example of perverse information improvements with endogenous coarsening is shown in Figure 10.



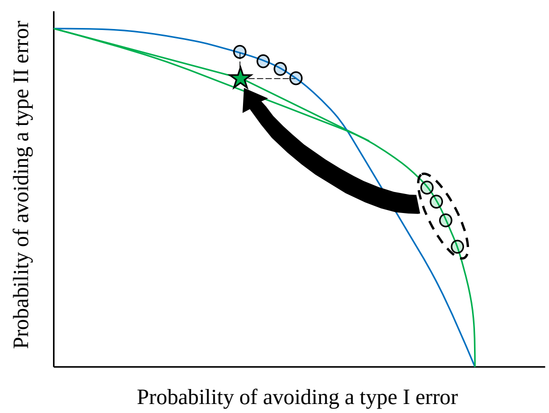
(a) Initial equilibrium (cutoffs shown as circles).



(b) Information improvements.



(c) New predictions for expert 1 (new single cutoff shown as star).



(d) New predictions for expert 2 (new single cutoff shown as star).

Figure 10: An improvement in the experts' information can result in both experts providing coarser and collectively less informative predictions (as shown by the stars, which represent the equilibrium cutoffs after an information improvement).

The example in Figure 10 works in much the same way as the proof of Theorem 3. The main difference is that after the information improvement, the kink in the experts' error avoidance frontier causes the experts to make just two predictions, with a cutoff point at the star (rather than the many predictions, and correspondingly many cutoffs, that were being used before). After this endogenous coarsening of their reports, with probability 1 both experts make one of the two extreme predictions.

7 Concluding Remarks

The problem of aggregating information from experts is increasingly important for society. With ever more data available, it is important to understand how this can best inform decisions, whether they are taken in a company, a government department, or in some other organization. We have provided some first steps towards a systematic analysis of situations in which information improvements can lead experts to make predictions that are less useful to every possible user of the predictions. While we have identified the existence of what we consider to be natural and simple information improvements that are perverse, starting from a wide range of parameter values, we have not systematically studied how "likely" perverse improvements are as a subset of all the possible improvements. There are substantial challenges to formalizing this question, but it may present an interesting avenue for further analysis.

Nevertheless, we believe the basic phenomena identified in the model have important substantive implications. Most importantly, understanding the impact of improvements in information quality requires not only an understanding of the information technology itself, but also of its effect on the optimizing behavior of the multiple experts who use it. Changes that are ambiguous when considering an individual expert (causing more of some kinds of errors and less of others) can result in experts becoming strictly less useful as a group. Because of this phenomenon, better econometric tools need not lead to better economic counsel; better diagnostic technology need not lead to better diagnoses; and better spying need not imply better military decisions. In each of these cases, the optimal use of the experts' recommendations can give society less value than society got before information improved.

Introducing a principal raises a natural question. Can society contract with the experts to eliminate perverse outcomes? While it is beyond the scope of this paper to undertake a full analysis of this problem, we note two difficulties with contracting. First, the value to society of correct predictions may be very large, which can make aligning the experts' incentives through a VCG type mechanism very expensive or else likely to run into limited liability problems. Second, experts are likely to have private information about their individual preferences over making Type I and Type II errors, as well as what signal they have received. Since this is a problem with multidimensional private information, the optimal contracting problem becomes quite difficult to solve, both for the principal and for the analyst.

A Equilibrium

We work in the setting of Section 6, where each expert can choose recommendations from the set $\{1, 2, \ldots, n\}$.

A.1 Existence

The solution concept we use is Bayes-Nash equilibrium.

PROPOSITION **5.** *A Bayes-Nash equilibrium of the game exists.*

REMARK **1.** This result uses no assumptions beyond the basic structure of the game: in particular all the signal distributions and payoffs are completely arbitrary.

Proof of Proposition 5

Proof. The strategy of each expert is a measurable function $\sigma_i : [0,1] \to \Delta^n$, where Δ^n is the n dimensional simplex and $\sigma_i(s)$ describes the probability with which i plays $x_i = k$ upon observing the signal $S_i = s$. The space of these functions is a convex, compact space and the utility function of each expert is an affine function of her own strategy, so the standard existence theorem applies (Fudenberg and Tirole, 1991, Theorem 1.2). □

A.2 Cutoff Strategies

Particularly natural strategies for the expert are cutoff strategies: they involve experts making lower predictions for lower signals and higher predictions for higher signals.

PROPOSITION **6.** *Cutoff strategies are played by both experts in all equilibria. More specifically, for each expert i and for any strategy σ_j played by the other expert, any strategy σ_i that is not a cutoff strategy is dominated by a cutoff strategy. Cutoffs are always interior.*

Proof of Proposition 6

Proof. Consider, without loss of generality, expert 1, and fix any strategy σ_2 for expert 2. Given $S_1 = s$, expert 1 assesses some conditional probability $r(s)$ of the event E,

$$r(s) = \mathbf{P}_1(E \mid S_1 = s) = \frac{p \cdot g(s)}{p \cdot g(s) + (1-p)}$$

Let the probability that $x_2 = k$ conditional on E be z_k and the probability that $x_2 = k$ conditional on $\overline{E}$ be $\overline{z}_k$:

$$z_k = \int_0^1 \sigma_{2,k}(s_2) g(s_2)\, ds_2 \tag{1}$$

$$\overline{z}_k = \int_0^1 \sigma_{2,k}(s_2)\, ds_2 \tag{2}$$

These definitions follow from the definition of $\sigma_{2,k}$ and the fact that S_2 is uniformly distributed under the null event $\overline{E}$. The expected utility to expert 1 of playing h is then:

$$u_i(h) = r(s) \left[\sum_{k=1}^{n} z_k \pi_1(h,k) \right] + (1 - r(s)) \left[\sum_{k=1}^{n} \overline{z}_k \overline{\pi}_1(h,k) \right]. \tag{3}$$

Thus,

$$\frac{\partial(u_i(h) - u_i(h-1))}{\partial s} = \sum_{k=1}^{n} r'(s) \left[z_k(\pi_1(h,k) - \pi_1(h-1,k)) + \overline{z}_k(\overline{\pi}_1(h-1,k) - \overline{\pi}_1(h,k)) \right]. \tag{4}$$

We know that $z_k \in [0,1]$ and $\overline{z}_k \in [0,1]$ for all k as they are probabilities. We also know that $\pi_1(h,k) > \pi_1(h-1,k)$ and $\overline{\pi}_1(h-1,k) > \overline{\pi}_1(h,k)$ because the experts dislike making errors. Finally, by our monotone likelihood ratio assumption, r is strictly increasing in s. Equation 4 is therefore always positive. It follows that expert 1 plays a cutoff strategy regardless of expert 2's strategy. The argument for expert 2 is symmetric.

The fact that cutoffs are interior follows from the fact that $g_i(\gamma) \to 0$ as $\gamma_i \to 0$ and $g_i(\gamma) \to \infty$ as $\gamma_i \to 1$, so the optimal actions near $S_i = 0$ and $S_i = 1$ are different. □

An important implication of Proposition 6 is that without loss of generality we can take the strategy space to be the set $[0,1]^{n-1}$, instead of the unwieldy space of functions $\sigma_i : [0,1] \to \Delta^n$. The only ambiguity this reduction leaves is what happens in the event $\{S_i = \gamma_i^*\}$, but since this event is of measure 0 and thus never affects expected payoffs or best responses, we ignore the issue.

It is also straightforward to establish that all equilibria are interior. As an expert's posterior probability of the event E approaches 1 for sufficiently high signals and approaches 0 for sufficiently low signals ($r(s) \to 1$ as $s \to 1$ and $r(s) \to 0$ as $s \to 0$), there exist sufficiently high signals for an expert to have a dominant strategy to make her strongest possible prediction of the event E by choosing $x_i = n$ and sufficiently low signals for the expert to make her weakest possible prediction $x_i = 1$. Thus, the predictions $x_i = n$ and

$x_i = 1$ are always in the support of the experts' strategies.

B Proofs of Main Results

B.1 Lemma 1

The experts' information improvements are weakly (respectively, strongly) perverse if they are weakly (respectively, weakly) cross-dominated.

Proof. Let U_q be the principal's utility under a decision rule q before the information improvement, and let U'_q be the principal's utility under decision rule q after the information improvement. Denote the principal's *optimal* decision rule after a cross-dominated information improvements by q^*. We now construct a decision rule, $\widehat{q}$ which behaves like q^*, but with the roles of the experts' predictions reversed: $\widehat{q}(x_1, x_2) = q^*(x_2, x_1)$. As the information improvement is cross-dominated, expert 1 makes both more Type I and Type II errors after the information improvement than expert 2 made before the information improvements, and expert 2 makes both more Type I and Type II errors after the information improvement than expert 1 made before. Therefore, $U_{\widehat{q}} > U'_{q^*}$: the principal is better off using $\widehat{q}$ before the information improvement than q^* after it. As the principal could have chosen the rule $\widehat{q}$ before the information improvement, the principal's equilibrium payoff before the improvement must have been even higher than U_{q^*}. □

B.2 Proposition 3

We present the proof of Proposition 3 here, since several of the subsequent results follow from it.

There exist weakly perverse information improvements over the initial information structure (G_1, G_2) at an initial equilibrium (γ_1^*, γ_2^*) if:

(i) experts are playing to their strengths at (γ_i^*, γ_j^*); and

(ii) were the experts to swap error rates, each expert i's indifference curve would have a slope between the slopes of her frontier at the points of tangency: for each i, we have $\beta_i(\gamma_i^*) \in (-G'_i(\widehat{\gamma}_{i,b}), -G'(\widehat{\gamma}_{i,a}))$.

These information improvements are also strongly perverse if:

(iii) all $\gamma_i \leq \widehat{\gamma}_{i,a}$ and all $\gamma_i \geq \widehat{\gamma}_{i,b}$ are not rationalizable for i (i.e., all such γ_i are not in R_i), for each player $i \in \{1, 2\}$.

Proof. **Weak Perversity:** We will construct an information improvement for each expert i by shifting out her error avoidance frontier. Figure 5 provides an example of what this information improvement might look like.

Fix a small $\delta > 0$ and for each i consider a pair of error avoidance probabilities, $\mathbf{z}_i = (\overline{\gamma}_i, \overline{\rho}_i)$, with $\overline{\gamma}_i \in (\gamma_j^* - \delta, \gamma_j^*]$ and $\overline{\rho}_i \in (1 - G_j(\gamma_j^*) - \delta, 1 - G_j(\gamma_j^*)]$. Our goal is to show that if δ is chosen small enough, then given any such $\mathbf{z}_1$ and $\mathbf{z}_2$, there exists an information improvement $(\widetilde{G}_1^{\mathbf{z}_1}, \widetilde{G}_2^{\mathbf{z}_2})$ such that $(\overline{\gamma}_1, \overline{\gamma}_2)$ is an equilibrium under the new information, where $(\widetilde{G}_1^{\mathbf{z}_1}, \widetilde{G}_2^{\mathbf{z}_2})$ is a pair of convex cumulative distribution functions. This clearly establishes the desired result.

Fix $(\mathbf{z}_1, \mathbf{z}_2)$ so that, for each i, we have $\mathbf{z}_i \in (\gamma_j^* - \delta, \gamma_j^*] \times (1 - G_j(\gamma_j^*) - \delta, 1 - G_j(\gamma_j^*)]$. Now fix i, and let j be the other agent.

Define[22]

$$H_i^{\mathbf{z}_i}(\gamma_i) := g_i(\gamma_i)(\gamma_i - \overline{\gamma}_i) - (G_i(\gamma_i) - \overline{\rho}_i)\,. \tag{5}$$

We will show that $H_i(\gamma_i)$ is positive when γ_i is near 0, negative at $\gamma_i = \gamma_j^*$, and positive when γ_i is near 1. Recall that as $\gamma_i \to 0$, we have both $G_i(\gamma_i) \to 0$ and (by assumption) $g_i(\gamma_i) \to 0$. As $\gamma_i \to 1$, we have both $G_i(\gamma_i) \to 1$ and (again by assumption) $g_i(\gamma_i) \to \infty$. Since experts are playing to their strengths, we have that $H_i^{(\gamma_j^*, 1-G_j(\gamma_j^*))}(\gamma_j^*) < 0$; otherwise, i could achieve j's equilibrium error rates. It follows by continuity of $H_i^{\mathbf{z}_i}$ in $\mathbf{z}_i$, that if δ is small enough, then for all $\mathbf{z}_i \in (\gamma_j^* - \delta, \gamma_j^*] \times (1 - G_j(\gamma_j^*) - \delta, 1 - G_j(\gamma_j^*)]$, we have $H_i^{\mathbf{z}_i}(\gamma_j^*) < 0$.

Set $\widehat{\gamma}_{i,a} = \inf\{\gamma_i : H_i^{\mathbf{z}_i}(\gamma_i) \leq 0\}$ and $\widehat{\gamma}_{i,b} = \sup\{\gamma_i : H_i^{\mathbf{z}_i}(\gamma_i) \leq 0\}$.

By the above argument, it follows that $\widehat{\gamma}_{i,a}$ and $\widehat{\gamma}_{i,b}$ are in $(0, 1)$, and so g_i is well-defined at these points.[23]

Using the numbers $\widehat{\gamma}_{i,a}$ and $\widehat{\gamma}_{i,b}$, we can construct a new error-avoidance frontier for i by setting:

$$1 - \widetilde{G}_i^{\mathbf{z}_i}(\gamma_i) \equiv \begin{cases} 1 - G_i(\gamma_i) & \text{if } \gamma_i \in [0, \widehat{\gamma}_{i,a}) \\ 1 - G_i(\widehat{\gamma}_{i,a}) - g_i(\widehat{\gamma}_{i,a})(\gamma_i - \widehat{\gamma}_{i,a}) & \text{if } \gamma_i \in [\widehat{\gamma}_{i,a}, \overline{\gamma}_i) \\ 1 - G_i(\widehat{\gamma}_{i,b}) + g_i(\widehat{\gamma}_{i,b})(\widehat{\gamma}_{i,b} - \gamma_i) & \text{if } \gamma_i \in [\overline{\gamma}_i, \widehat{\gamma}_{i,b}) \\ 1 - G_i(\gamma_i) & \text{if } \gamma_i \in [\widehat{\gamma}_{i,b}, 1] \end{cases}$$

We claim that, assuming δ was chosen small enough, $(\overline{\gamma}_i, \overline{\gamma}_j)$ is an equilibrium in cutoff strategies for the information structure given by $(\widetilde{G}_1^{\mathbf{z}_1}, \widetilde{G}_2^{\mathbf{z}_2})$. Suppose that j plays $\overline{\gamma}_j$. By the

[22]Heuristically, what we are doing in this part of the proof is showing that there are two uniquely determined lines intersecting the point $(\overline{\gamma}_i, \overline{\rho}_i)$ that are tangent to the curve $\gamma \mapsto 1 - G_i(\gamma)$; see Figure 5. These lines will be used to construct the information improvement.

[23]Recall that $g_i(1)$ is not well defined.

argument of Proposition 6, agent i has a cutoff strategy best response. If δ is small enough, we will argue that this best response cannot be anywhere except at $\overline{\gamma}_i$. For small enough δ, the slope $\beta_i(\overline{\gamma}_j)$ of i's indifference curve given j's strategy is, by continuity of β_i (recall Section 3.1, and footnote 13 in particular) arbitrarily close to $\beta_i(\gamma_i^*)$, which is in $(-G_i'(\widehat{\gamma}_{i,b}), -G'(\widehat{\gamma}_{i,a}))$ by assumption (ii). Thus, for sufficiently small δ, we have $\beta_i(\overline{\gamma}_j) \in (-G_i'(\widehat{\gamma}_{i,b}), -G'(\widehat{\gamma}_{i,a}))$ also. But at any $\gamma_i \in (\widehat{\gamma}_{i,a}, \overline{\gamma}_i)$, the curve $\gamma \mapsto 1 - \widetilde{G}_i^{\mathbf{z}_i}(\gamma)$ has a slope weakly less than $-G_i'(\widehat{\gamma}_{i,a})$, which is strictly shallower than $\beta_i(\overline{\gamma}_j)$ so there is a strict local improvement available by increasing γ_i slightly. Similarly, at any $\gamma_i \in (\overline{\gamma}_i, \widehat{\gamma}_{i,a})$, the curve $\gamma \mapsto 1 - \widetilde{G}_i^{\mathbf{z}_i}(\gamma)$ has a slope strictly steeper than $\beta_i(\overline{\gamma}_j)$, so there is a strict local improvement available by decreasing γ_i slightly. Thus, these cannot be best responses, and by the proof of Proposition 6, more extreme cutoffs cannot be best responses, either. This completes the proof.

Strong perversity:

If $\gamma_i \leq \widehat{\gamma}_{i,a}$ and $\gamma_i \geq \widehat{\gamma}_{i,b}$ are not rationalizable for expert i, then any equilibrium choice γ_i must lie in $(\widehat{\gamma}_{i,a}, \widehat{\gamma}_{i,b})$. By the last paragraph in the proof of the weak perversity part, this establishes that the equilibrium must be at $(\overline{\gamma}_i, \overline{\gamma}_j)$. □

B.3 Proposition 2

If the experts are independent and envious of each other, then there exist information improvements that arise from technology transfer and are strongly perverse.

Proof. Consider the information improvements constructed in the proof of Proposition 3. This information improvement for i involved taking the convex hull of the point $(\gamma_j^*, 1 - G_j(\gamma_j^*))$ and i's initial error avoidance frontier, the curve $\gamma \mapsto 1 - G_i(\gamma)$. By the proof of Theorem 1, these perverse information improvements can also be constructed for independent and envious experts. As the error avoidance rates $(\gamma_j^*, 1 - G_j(\gamma_j^*))$ are on j's initial error avoidance frontier and the experts are envious of each other, every point on i's new (constructed) error avoidance frontier is weakly dominated by some error avoidance rates on the convex hull of the curves $\gamma \mapsto 1 - G_i(\gamma)$ and $\gamma \mapsto 1 - G_j(\gamma)$. The perverse information improvements constructed in the proof of Proposition 3 can therefore arise from technology transfer. □

B.4 Theorem 1

If the experts are independent, there exists a strongly cross-dominated information improvement if and only if both experts are envious of each other.

Proof. **If:** This is a special case of Proposition 3. We show only that conditions (i), (ii) and (iii) of Proposition 3 are satisfied if both experts are envious of each other. First, if expert j is envious of expert i at the initial cutoff choices (γ_1^*, γ_2^*), then expert i must be playing to her strengths. Otherwise, by revealed preference, expert j would be choosing a suboptimal cutoff. To verify (ii): as i is envious of j, it follows that[24] $\widehat{\gamma}_{i,a} < \gamma_i^* < \widehat{\gamma}_{i,b}$, and so by the concavity of the error avoidance frontiers, we have $\beta_i(\gamma_i^*) \in (-G_i'(\widehat{\gamma}_{i,b}), -G_i'(\widehat{\gamma}_{i,a}))$. Finally, we verify (iii): since the experts are independent, β_i does not actually depend on its argument, so we may drip it. It follows immediately from $\beta_i \in (-G_i'(\widehat{\gamma}_{i,b}), -G_i'(\widehat{\gamma}_{i,a}))$ that all cutoff choices γ_i outside the interval $[\widehat{\gamma}_{i,a}, \widehat{\gamma}_{i,b}]$ are dominated.

Only if: Suppose expert i is not envious of expert j. This implies that i prefers error avoidance rates $(\gamma_i^*, 1-G_i(\gamma_i^*))$ to error avoidance rates $(\gamma_j^*, 1-G_j(\gamma_j^*))$. Consider any error avoidance rates (e_1, e_2) strictly dominated by $(\gamma_j^*, 1-G_j(\gamma_j^*))$ – that is, satisfying $e_1 < \gamma_j^*$ and $e_2 < 1-G_j(\gamma_j^*)$. We then have that $(\gamma_i^*, 1-G_i(\gamma_i^*)) \succeq_i (\gamma_j^*, 1-G_j(\gamma_j^*)) \succ_i (e_1, e_2)$. After her information improvement, by revealed preference, expert i must prefer her new error rates: $(\gamma_i^{*\prime}, 1-\tilde{G}_i(\gamma_i^{*\prime})) \succeq_i (\gamma_i^*, 1-G_i(\gamma_i^*))$. This establishes that $(\gamma_i^{*\prime}, 1-\tilde{G}_i(\gamma_i^{*\prime})) \succ_i (e_1, e_2)$ and so a cross-dominated information improvement is not possible. □

B.5 Proposition 4

Under the maintained assumptions, there exist payoffs such that in all equilibria both experts only make predictions in the set $\{1, n\}$.

Proof. Let $\gamma_{(1,n)}$ be the signal at which expert 1 would be indifferent between making prediction 1 and n. By playing this cutoff, expert 1 receives a payoff of

$$\begin{aligned} u_i(1) =& r(\gamma_{(1,n)}) \left[\sum_{k=1}^{n} q_k(\gamma_{(1,n)}) \pi_1(1,k) \right] + (1-r) \left[\sum_{k=1}^{n} q_k(\gamma_{(1,n)}) \overline{\pi}_1(1,k) \right] \\ =& r(\gamma_{(1,n)}) \left[\sum_{k=1}^{n} q_k(\gamma_{(1,n)}) \pi_1(n,k) \right] + (1-r) \left[\sum_{k=1}^{n} q_k(\gamma_{(1,n)}) \overline{\pi}_1(n,k) \right] = u_i(n) \end{aligned}$$

Suppose now that we increased $\pi_1(1,k)$ by the same constant for all k and $\overline{\pi}_1(n,k)$ by the same constant for all k such that the above equations continued to hold. This increase does not violate the assumption that experts dislike both Type I and Type II errors. This change increases the value of $u_i(1) = u_i(n)$ at $s_i = \gamma_{(1,n)}$ and does not change the payoff

[24] The numbers $\widehat{\gamma}_{i,a}$ and $\widehat{\gamma}_{i,b}$ are defined in Section 5.

of $u_i(k)$ for $k = 2, \ldots, n-1$ at any s_i. Furthermore, as $u_i(1)$ and $u_i(n)$ increase linearly in $\pi_1(1,k)$ and $\overline{\pi}_1(n,k)$ respectively for all k, there always exist such increases that make $u_i(1) = u_i(n) > u_i(k)$ at $s_i = \gamma_{(1,n)}$ for all $k \in \{2, \ldots, n-1\}$. Applying Proposition 6 this is sufficient to ensure that expert 1 only ever makes the predictions 1 and n. □

References

Al-Najjar, N. I. and J. Weinstein (2008): "Comparative Testing of Experts," *Econometrica*, 76, 541559.

Banerjee, A. (1992): "A simple model of herd behavior," *The Quarterly Journal of Economics*, 107, 797.

Bernheim, B. (1984): "Rationalizable strategic behavior," *Econometrica: Journal of the Econometric Society*, 1007–1028.

Bikhchandani, S., D. Hirshleifer, and I. Welch (1992): "A theory of fads, fashion, custom, and cultural change as informational cascades," *Journal of political Economy*, 992–1026.

Blackwell, D. (1953): "Equivalent Comparisons of Experiments," *The Annals of Mathematical Statistics*, 24, 265–272.

Council of Advisors of Science and Technology (2009): "Report to the president on U.S. preparations for 2009 - H1N1 influenza," .

Crawford, V. and J. Sobel (1982): "Strategic information transmission," *Econometrica: Journal of the Econometric Society*, 50, 1431–1451.

Dekel, E. and Y. Feinberg (2006): "Non-Bayesian Testing of a Stochastic Prediction," *Review of Economic Studies*, 73, 893–906.

Feinberg, Y. and C. Stewart (2008): "Testing Multiple Forecasters," *Econometrica*, 76, 561–582.

Fortnow, L. and R. V. Vohra (2009): "The Complexity of Forecast Testing," *Econometrica*, 77, 93105.

Fudenberg, D. and J. Tirole (1991): *Game Theory*, Cambridge, MA: MIT Press.

Holmstrom, B. (1983): *On the Theory of Delegation*, New York: North-Holland.

Li, H., S. Rosen, and W. Suen (2001): "Conflicts and Common Interests in Committees," *American Economic Review*, 91, 1478–1497.

Li, H. and W. Suen (2004): "Delegating decisions to experts," *Journal of Political Economy*, 112, 311–335.

Ottaviani, M. and P. Sørensen (2001): "Information aggregation in debate: who should speak first?" *Journal of Public Economics*, 81, 393–421.

——— (2006): "Professional advice," *Journal of Economic Theory*, 126, 120–142.

Raftery, J. and M. Chorozoglou (2011): "Possible net harms of breast cancer screening: updated modelling of Forrest report," *BMJ*, 343.

Sandroni, A. (2003): "The reproducible properties of correct forecasts," *International Journal of Game Theory*, 32, 151–159.

Sandroni, A. and W. Olszewski (2008): "Manipulability of Future-Independent Tests," *Econometrica*, 76, 1437–1466.

——— (2010): "Falsifiability," *American Economic Review*, forthcoming.

Select Committee Report (2004): "Report of the Select Committee on Intelligence on the U.S. Intelligence Community's Prewar Intelligence Assessments on Iraq," .

Suen, W. (2004): "The Self-Perpetuation of Biased Beliefs," *The Economic Journal*, 114, 377–396.