American Economic Review 2018, 108(10): 3057–3103
https://doi.org/10.1257/aer.20170299

Evaluating Strategic Forecasters†

By Rahul Deb, Mallesh M. Pai, and Maher Said*

Motivated by the question of how one should evaluate professional election forecasters, we study a novel dynamic mechanism design problem without transfers. A principal who wishes to hire only high-quality forecasters is faced with an agent of unknown quality. The agent privately observes signals about a publicly observable future event, and may strategically misrepresent information to inflate the principal's perception of his quality. We show that the optimal deterministic mechanism is simple and easy to implement in practice: it evaluates a single, optimally timed prediction. We study the generality of this result and its robustness to randomization and noncommitment. (*JEL* C53, D72, D82)

> *A foolish consistency is the hobgoblin of little minds, adored by little statesmen and philosophers and divines.*
>
> —Ralph Waldo Emerson

Forecasting is an important industry whose experts' services are utilized in a variety of different fields, including politics, sports, meteorology, banking, finance, and economics. Forecasters differ based on the quality of their predictions which, in turn, is determined by the accuracy of their information and their ability to process it. The career prospects of an expert depend on public perceptions of his ability, and hence a strategic forecaster may make predictions designed to inflate those perceptions. In this paper, we study the dynamic mechanism design problem of a principal who uses an expert's predictions to determine whether that expert is worth hiring. In a nutshell, we are interested in determining the optimal method of screening strategic forecasters.

*Deb: Department of Economics, University of Toronto, 150 St. George Street, Toronto, ON M5S 3G7, Canada (email: rahul.deb@utoronto.ca); Pai: Department of Economics MS-22, Rice University, 6100 Main Street, Houston, TX 77005 (email: mallesh.pai@rice.edu); Said: Department of Economics, Stern School of Business, New York University, 44 West Fourth Street, New York, NY 10012 (email: maher.said@nyu.edu). This paper was accepted to the *AER* under the guidance of Jeff Ely, Coeditor. We are very grateful to Ben Brooks, Yeon-Koo Che, Ying Chen, Martin Cripps, Alex Frankel, Hanna Hałaburda, Johannes Hörner, Matt Jackson, Rohit Lamba, Elliot Lipnowski, Alessandro Lizzeri, Marco Ottaviani, Andrea Prat, Larry Samuelson, Vasiliki Skreta, Peter Sørensen, Ennio Stacchetti, Colin Stewart, Lucy White, three anonymous referees, and various seminar and conference participants for insightful comments and valuable feedback that helped greatly improve the paper. We would also like to thank the Study Center Gerzensee for its hospitality. Deb would like to thank Rice University and the Indian Statistical Institute for their hospitality and the Social Sciences and Humanities Research Council of Canada for their generous financial support. Pai would like to thank the Cowles Foundation for its hospitality, and the National Science Foundation (grant CCF-1101389) for their generous support. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

†Go to https://doi.org/10.1257/aer.20170299 to visit the article page for additional materials and author disclosure statement(s).

For example, consider a governor or senator who is contemplating a presidential run in the next electoral cycle. She would like to hire a professional election forecaster to help her accurately determine the viability of her future candidacy. To evaluate the forecaster's ability, she observes his predictions at various points in the current electoral cycle; she also eventually observes the current electoral outcome. What is the best way for her to determine whether the forecaster is worth hiring? The important factors that the politician needs to incorporate into her hiring decision are that (i) the forecaster's information and the election outcome are both noisy signals of the underlying preferences of the electorate; (ii) the forecaster learns about those preferences more precisely as the election nears; and (iii) his predictions are strategically chosen to make himself appear to be of higher quality as he anticipates the career implications of his performance.

We develop a novel dynamic model to study these issues. In its simplest symmetric, and binary form (on which the bulk of the paper focuses), the framework can be described as follows. There is a persistent, unknown state of the world governing the data-generating process. This unobserved state takes one of two values with equal probability. A binary public outcome, which is a potentially noisy signal of the underlying state, occurs at time $T+1$. Leading up to that outcome, the agent (forecaster) privately learns about the state (and therefore the expected outcome) via a sequence of T noisy signals. These binary signals are correlated with the state but are otherwise conditionally independent and identically distributed. The agent is equally likely to be either a "good" or a "bad" type, where a good type observes more precise information. At each point of time, the agent strategically reports his signal. After the outcome has been realized, the principal decides whether to hire the agent based on a mechanism that is announced (and committed to) at the beginning of the game. A mechanism in this context is a deterministic mapping from the history of reported signals and the eventual outcome to a hiring decision. Both parties care only about this hiring decision, as both the underlying state and the agent's signals are payoff irrelevant. Their incentives diverge, however: the principal only wants to hire the good type, while the agent always wants to be hired, regardless of his private type.

The critical modeling assumptions of our environment are supported by the disparate literatures that study forecasting in psychology, statistics, economics, and finance. Our underlying information structure, an unknown data-generating process that the forecaster learns over time, is a standard (albeit simplified) feature of statistical models of forecasting (an up-to-date survey is Elliott and Timmermann 2016; recent empirical evidence on learning by professional forecasters can be found in Lahiri and Sheng 2008 and the papers that follow). Psychologists have shown that experts differ in their forecasting abilities and that better forecasters are consistently more accurate (see, for instance, the work described in Tetlock 2005 and Tetlock and Gardner 2015). Trueman (1994), Ottaviani and Sørensen (2006c), and others have argued that experts who differ by ability choose their forecasts with the intention of influencing clients' assessments of that ability. At a high level, the key departure of this paper from this latter economics literature is that we incorporate a strategic principal (as opposed to a passive market) who optimally chooses her method of evaluating such strategic forecasters.

To understand the role played by incentives in this environment, it is instructive to examine the benchmark case where the principal does not know the agent's type but can observe his signals. Here, the principal can screen the agent, even in the absence of a public outcome, by using the *variance* of the observed signals: since the good type receives more precise information, he is more likely to receive signal profiles with a large proportion of identical signals (or, equivalently, profiles with lower signal variance). The public outcome provides another instrument for screening, as the good type's signals are also more likely to match that outcome. The principal's optimal hiring decision in this benchmark therefore combines *accuracy* and *consistency* by using two thresholds (Theorem 1). The agent is hired either when he is consistently accurate and receives enough signals that match the outcome (a simple majority of matches is not sufficient) or when he is consistently inaccurate and receives enough signals that do *not* match the outcome (a majority of matches is not necessary).[1] Hence, when the agent cannot strategically misreport his signals, the principal screens using *both* the accuracy and the consistency of the agent's information. Note that, given a profile of received signals, the order in which the signals arrive plays no role as they are generated from a conditionally i.i.d. process.

An immediate and important economic insight is that when the agent is free to report his signals strategically, the optimal mechanism does not screen using consistency. The reason is quite intuitive: it is always possible for the agent to report consistent signals regardless of the actual information he receives. Instead, we show that it is optimal for the principal to screen using a combination of the *accuracy* of the agent's signals along with the *order* in which they arrive. Specifically, our main result shows that the optimal deterministic mechanism takes the very simple form of a *prediction mechanism*: the principal optimally chooses a time period $T^* \leq T$ to solicit a *single* prediction of the final outcome, and the agent is hired if and only if that prediction matches that outcome (Theorem 3).[2] The principal utilizes the order in which the signals arrive by ignoring information that arrives after T^*.

This result has a number of features that are worth emphasizing. It may be surprising to some readers (perhaps in light of the "testing experts" literature we discuss below) that screening is possible at all in this strategic environment, especially when the principal's only screening instrument is a coarse hiring decision. Unlike the benchmark case, screening a strategic agent is not possible in the absence of a public outcome as the bad type is free to follow any reporting strategy. However, with a public outcome, screening becomes possible: since the good type receives more precise information, his prediction (if truthful) of the outcome in any period is more likely to be correct than the bad type's. Thus, a hiring rule where the agent is picked if and only if his prediction in a given period turns out to be accurate is more likely (compared to the initial belief) to result in the hiring of the good type. Moreover, focusing on a single period's prediction (and ignoring the agent's reports

[1] The most transparent demonstration of why *two* cutoffs are optimal and the agent may be hired when sufficiently inaccurate is the corner case where the good type's signals are perfectly informative while both the public outcome and the bad type's signals are completely uninformative. The likelihood that the bad type repeatedly receives the same signal is sufficiently small to ensure that the principal is happy to hire an agent with a perfectly consistent signal profile, regardless of whether it matches the (uninformative) outcome.

[2] It is worth stressing that this mechanism is optimal within the full class of deterministic direct revelation mechanisms that, in addition to the signals, also ask the agent to report his initial private type. Since such type reporting is not observed in practice and, as Lemma 1 shows, there is no loss of generality in dispensing with it, we deliberately focus on mechanisms that do not solicit this information. As we discuss later in Section VI, our results also generalize to the case where the agent has no initial private information.

at all other periods) also ensures that it is optimal for the strategic agent to sincerely predict the outcome he believes to be more likely.

As mentioned above, the optimal mechanism uses the order of signals for screening by discarding information that the agent receives after period T^*. To see why doing so might help the principal, suppose she instead always chooses to solicit predictions at the end of period T after the agent has acquired all possible information. When T is large, both types of the agent learn the underlying state with high probability, which makes screening by predictions ineffective. Instead, the principal can choose to screen at an intermediate time period when the learning advantage for the good type (from receiving more precise information) is at its highest. An insight from the main result is that the principal is unable to improve screening in a deterministic mechanism (over and above soliciting a prediction) by using any information that arrives after period T^*.

A strength of the optimal mechanism is that it is very easy to describe and implement in practice. Moreover, we show that the same optimal outcome can be achieved even *without commitment* (Theorem 4), thereby making our results applicable in settings where the principal has little or no commitment power. This is another novel aspect of our framework as it is quite unusual for commitment power to not benefit the principal in a dynamic mechanism design environment.

In Section VI, we discuss the scope of the main insight driving our result by showing that it also applies to very general environments (Theorem 7). We show that the key assumption we need for the optimality of prediction mechanisms is that the public outcome is binary.[3] As long as this assumption holds, prediction mechanisms remain optimal even if the agent's type is drawn from a general space and the information he receives is generated from a general time-varying signal process. Additionally, even in this general environment, commitment is not required to implement the optimal mechanism. The simplicity of optimal mechanisms in so general a setting opens the door to further research on even richer models which have strategic forecasting as a component (and we discuss a few avenues for future research in our concluding remarks).

Finally, while our focus on deterministic hiring rules is driven by their suitability for our motivating applications, randomization plays an interesting theoretical role in our environment. This is most easily demonstrated in the optimal *stochastic* mechanism for the special case of $T = 3$ periods (Theorem 5). Here, we show how the principal fine-tunes her screening by hiring the agent with different (strictly positive) probabilities that depend on the order of signal arrivals in addition to the overall composition of the signal profile. Finally, we show that a sufficient condition for the optimality of randomization is that the time horizon is long enough (Theorem 6).

Related Literature.—Expert forecasting is an important industry and the input of forecasters is often solicited for numerous decisions made by firms and policymakers alike. While the statistical work on evaluating forecasting models is well developed (see, for instance, the aforementioned survey Elliott and Timmermann

[3] We more carefully discuss the role this binary-outcome assumption plays in Section VIIC.

2016), there is relatively less research examining the incentives of strategic experts and how these incentives influence their forecasts (a recent survey of this work is Marinovic, Ottaviani, and Sørensen 2013). The theoretical work in this latter literature (see, for instance, Ottaviani and Sørensen 2006a, c) differs in that forecasters are evaluated by a rational, but otherwise passive, market and that the environment is static. This paper differs in that we consider a dynamic environment in which a strategic principal can alter the incentives of the forecaster by choosing her evaluation criterion.

The literature on testing experts (starting with Foster and Vohra 1998; a recent survey is Olszewski 2015) shares a similar motivation. Here an abstract dynamic environment is considered and the focus is on determining the existence of a test which (i) cannot be passed by a strategic forecaster without knowledge of the true data-generating process and (ii) can be passed almost surely when the forecaster knows the process. Both our model and overall objective differ in that we allow the agent to be imperfectly informed about the data-generating process and that the principal's goal is to design a mechanism that maximally separates the good forecaster from the bad, even if that screening is imperfect.

Since we consider a setting where a principal can commit to her hiring policy (based on sequential information received from the agent), our results are related to those in the literature on dynamic mechanism design. The binary private signals in our simplified model are a key feature of Battaglini (2005) and Boleslavsky and Said (2013), of which the latter also features private information about the signal process. These papers differ not only in their reliance on transfers but also in terms of the payoffs, the structure of the stochastic process governing signal evolution, and (as a result) the applications to which their models apply. Though it also differs along these latter dimensions, Guo and Hörner (2018) is more closely related as it also examines a dynamic mechanism design problem in a binary environment without transfers. In another strand of this literature, Aghion and Jackson (2016) show that tenure schemes can provide incentives for an agent to take actions that reveal his competence. However, their setting yields distinct economic insights as (among other differences) they rely on having multiple opportunities to learn about the agent's competence as well as on principal preferences that depend on the agent's actions instead of his underlying type. We will further discuss the relation of our results to the dynamic mechanism design literature in more detail in Section VIIA.

Finally, since we also examine the dynamic cheap talk setting where the principal cannot commit, this paper is related to the literature studying how an agent with a privately known type builds reputation via dynamic communication. The key difference between our setting and this literature (in addition to the different applications modeled and the fact that we also characterize the full commitment optimum) is that our principal dynamically screens across types with differential rates of learning of a fixed underlying state. This is in contrast with Ottaviani and Sørensen (2006b) and Li (2007), where the agent is evaluated by a competitive market and so his payoff is simply the posterior belief about his type. Alternatively, Morris (2001) considers a repeated, two-period setting where a principal makes a decision in each period based on the agent's report. While both the principal and the agent in his setting have very different preferences from ours, an important distinction is that our principal makes a single decision after cheap talk has ended.[4]

I. Model

We consider a T-period, discrete time, finite horizon framework in which a principal determines whether to hire an agent who is an expert forecaster. To make the main insights transparent, we define a simplified, symmetric version of the model on which the majority of the paper focuses. We discuss the full generality of the results in Section VI.

A. *The Environment*

State.—The forecaster is being judged on his ability to learn about an unknown state of the world ω. This state, which governs the data-generating process, is equally likely to be either high (h) or low (l), so the commonly known prior distribution of states is $\Pr(\omega = h) = \Pr(\omega = l) = 1/2$.

Agent's Private Information.—There is a single forecaster whose privately known type (his forecasting ability) θ can either be good (g) or bad (b) with equal likelihood; thus, the commonly known prior distribution of ability is $\Pr(\theta = g) = \Pr(\theta = b) = 1/2$.[5]

In each period $t = 1, \ldots, T$, the forecaster privately observes a binary signal $s_t \in \{h, l\}$ about the unknown state ω. The accuracy of these signals (that is, the probability that each signal "matches" the true state) is

$$\alpha_\theta \coloneqq \Pr(s_t = \omega | \theta).$$

We assume that $1/2 < \alpha_b < \alpha_g < 1$, so that the type-$g$ agent's signals are more precise than the type-b agent's.[6] We write $s^t \coloneqq (s_1, \ldots, s_t)$ to denote a sequence of t signals.

Outcome.—At the end of period T, a publicly observed binary outcome $r \in \{h, l\}$ is realized. This outcome is correlated with the true state ω; we denote by $\gamma \in [1/2, 1]$ the probability with which the outcome r "matches" the true state ω, so

$$\gamma \coloneqq \Pr(r = \omega).$$

[4] This aspect is also reminiscent of Krishna and Morgan (2004) (and the papers that follow), where an additional long communication protocol is added to the canonical model of Crawford and Sobel (1982).

[5] Note that we do not require symmetry in either the state or type distributions for any of the results in the binary model. This assumption merely allows us to simplify the notation and shorten the proofs without compromising our main economic insights.

[6] We exclude the corner cases $\alpha_g = 1$ and $\alpha_b = 1/2$ to simplify our exposition, though our results continue to hold.

The corner case where $\gamma = 1$ corresponds to situations where the public outcome fully reveals the underlying state, while $\gamma < 1$ reflects environments where that outcome is only a noisy signal.[7]

B. *The Game*

In each period t, the agent strategically reports his signal $\tilde{s}_t \in \{h, l\}$, possibly as the realization of a mixed strategy (we will discuss implementations where the agent makes predictions instead of reporting signals in Section IIIC). Our main focus will be on the case where the principal has full commitment, but we will also examine what happens in the absence of commitment power.

Histories.—At the beginning of any period t, $h_t^A = (s^t, \tilde{s}^{t-1})$ denotes the agent's private history. This contains the t privately observed signals s^t and the $t-1$ reports $\tilde{s}^{t-1}$ made prior to period t. We use $\mathcal{H}^A = \cup_{t=1}^T \left(\{h, l\}^t \times \{h, l\}^{t-1}\right)$ to denote the set of all histories for the agent.

The relevant history for the principal $h^P = (\tilde{s}^T, r)$ at which she makes a hiring decision contains the entire sequence of reports made by the agent in all T periods and the final outcome. We use $\mathcal{H}^P = \{h, l\}^{T+1}$ to denote the set of all such public histories.

Agent's Strategy.—The type-θ agent's strategy $\sigma^\theta : \mathcal{H}^A \to \Delta\{h, l\}$ determines the distribution of signal reports at each history. We will use the signal subscript $\sigma_s^\theta(h_t^A)$ to denote the probability that the agent reports signal $s \in \{h, l\}$.

Principal's Strategy.—We denote by $x_r(\tilde{s}^T) \in \{0, 1\}$ the principal's strategy at history $(\tilde{s}^T, r) \in \mathcal{H}^P$. It determines the probability with which she hires the agent as a function of the T reported signals $\tilde{s}^T$ and the outcome r. We focus on deterministic hiring decisions for the principal as we feel that this is the more natural modeling assumption for the applications we consider. That said, randomization plays an interesting theoretical role in our model that we discuss in Section V.

When the principal has commitment power, we sometimes refer to x as a mechanism, although it does *not* correspond to a direct revelation mechanism (since x does not condition on the agent's private type θ). We restrict attention to this game as it more closely mirrors the applications of our model (forecasters do not typically report their types in practice); note, however, that this restriction is without loss of generality (as we show in Lemma 1).

Payoffs.—The payoffs only depend on the agent's type and the hiring decision. The principal receives a payoff of 1 if she hires a good ($\theta = g$) forecaster, a payoff of -1 if she hires a bad ($\theta = b$) forecaster, and a payoff of 0 otherwise. Essentially,

[7] Election outcomes are often affected by last-minute events uncorrelated with the electorate's underlying preferences. For example, bad weather on election day can significantly reduce voter turnout: see Gomez, Hansford, and Krause (2007) for more. Similarly, unanticipated in-game injuries often lead to upsets of the "better" team in a sporting event.

the principal's goal is to maximize the difference in likelihood between hiring the two types.

The agent's preferences are type independent: both types want to maximize the probability with which they are hired. To capture this, we assume that the agent receives a payoff of 1 if she is hired and a payoff of 0 if not.

Timing.—For easy reference, the following flow chart summarizes the game.

Nature draws unobserved state $\omega \in \{h, l\}$; Agent privately learns type $\theta \in \{g, b\}$;	$\rightarrow$	Agent observes private signal $s_1 \in \{h, l\}$ and reports $\tilde{s}_1 \in \{h, l\}$;	$\rightarrow \cdots \rightarrow$	Agent observes private signal $s_T \in \{h, l\}$ and reports $\tilde{s}_T \in \{h, l\}$;	$\rightarrow$	Outcome $r \in \{h, l\}$ publicly realized;	$\rightarrow$	Principal makes hiring decision $x_r(\tilde{s}^T)$;	$\rightarrow$	Payoffs are realized.		

II. Benchmark: Publicly Observed Signals

Before we analyze the game, we consider a simple benchmark in which the agent's signals are publicly observed. Here, the agent is passive, and the only private information is his initial type. This benchmark helps highlight the issues inherent in trying to attain this "first-best" payoff for the principal when the agent must be incentivized to truthfully reveal his private signals.

A consequence of the payoff structure is that the principal's ex ante expected payoff from any hiring decision x can be written as

$$\begin{aligned}\Pi &:= \sum_{r\in\{h,l\}} \sum_{s^T\in\{h,l\}^T} \Pr(r, s^T)\big[\Pr(\theta = g | r, s^T) - \Pr(\theta = b | r, s^T)\big]\, x_r(s^T) \\ &= \frac{1}{2} \sum_{r\in\{h,l\}} \sum_{s^T\in\{h,l\}^T} \big[\Pr(r, s^T | \theta = g) - \Pr(r, s^T | \theta = b)\big]\, x_r(s^T),\end{aligned}$$

which is the difference in the expected probabilities that the g and b types of the agent are hired. Therefore, the optimal hiring decision in this benchmark is given by

$$x_r^{FB}(s^T) = \begin{cases} 1 & \text{if } \Pr(r, s^T | \theta = g) \geq \Pr(r, s^T | \theta = b) \\ 0 & \text{otherwise} \end{cases}. \tag{1}$$

Observe that the principal cannot benefit from randomizing her hiring decision in this first-best benchmark. In addition, the probabilities that determine the first-best hiring policy can be readily expressed in terms of the model primitives. In particular, since signals are conditionally i.i.d., only their frequencies (and not the specific order in which signals arrive) play a role. With this in mind, the probability of type θ observing a signal profile s^T in which n signals that match the outcome is

$$\Pr\left(\sum_{t=1}^{T} \mathbf{1}_r(s_t) = n \,|\, \theta\right) = \binom{T}{n} \beta_{n,T,\theta},$$

where

$$\beta_{n,T,\theta} \coloneqq \gamma\alpha_\theta^n(1-\alpha_\theta)^{T-n} + (1-\gamma)\alpha_\theta^{T-n}(1-\alpha_\theta)^n$$

and we define $\mathbf{1}_r(s_t)$ to be the indicator function that takes the value 1 if the period-t signal matches the outcome ($s_t = r$) and 0 otherwise.

The first term of $\beta_{n,T,\theta}$ corresponds to the cases where the outcome matches the underlying state (that is, when $r = \omega$), while the second term corresponds to the complementary cases where the outcome does not match the state (that is, when $r \neq \omega$). Since $\Pr(r, s^T)$ is constant across all signal profiles s^T with the same number of signals matching the outcome r, the first-best is then easy to state: hire an agent who receives exactly n signals that match the outcome if and only if the agent is more likely to be of type g than type b, so that

$$\Delta_{n,T} \coloneqq \beta_{n,T,g} - \beta_{n,T,b} \geq 0.$$

To make the incentive issues in implementing x^{FB} explicit, we now provide a qualitative characterization of the first-best hiring policy in (1).

THEOREM 1: *In the benchmark with publicly observable signals, the first-best hiring policy x^{FB} can be characterized by two cutoffs $\overline{n}$ and $\underline{n}$ with $T/2 < \overline{n} \leq T$ and $\underline{n} \leq T - \overline{n}$ such that*

$$x_r^{FB}(s^T) = \begin{cases} 1 & \text{if} \sum_{t=1}^T \mathbf{1}_r(s_t) \geq \overline{n} \text{ or } \sum_{t=1}^T \mathbf{1}_r(s_t) \leq \underline{n} \\ 0 & \text{otherwise} \end{cases}.$$

Theorem 1 shows that it is not enough for a majority of the agent's signals to correctly match the outcome; in general, this is neither necessary nor sufficient for being hired. Instead, that the principal screens using a combination of *accuracy* and *consistency*.

- The agent is hired if he is consistently accurate, with at least $\overline{n}$ signals that match the outcome. When this *match threshold* is $\overline{n} > \lceil (T+1)/2 \rceil$, the agent may not be hired even when the majority of his signals match the outcome.
- The agent is hired if he is consistently inaccurate, with at least $T - \underline{n}$ signals that do not match the outcome. When this *mismatch threshold* is $\underline{n} \geq 0$, the agent may be hired even when the majority of his signals mismatch the outcome.
- The mismatch threshold is more stringent than the match threshold, as $\underline{n} \leq T - \overline{n}$. Thus, greater consistency is required to compensate for inaccuracy.

For some intuition on why the first-best utilizes the mismatch cutoff $\underline{n}$, consider the case where $\gamma = 1/2$, so that the outcome is uninformative about the true underlying state. Here, accuracy provides no information with which to evaluate the forecaster. Instead, the principal must exploit the fact that the good type is more consistent than the bad type: his more precise information is more highly correlated with the true state, and so his signal profile is likely to have lower variance. Thus, the first-best here relies only on consistency, and the two cutoffs are symmetric (i.e., $\underline{n} = T - \overline{n}$).

At the other extreme, when $\gamma = 1$, so that the outcome perfectly reveals the true underlying state, accuracy is meaningful as incorrect signals are now an indication that the agent is the bad type, and many incorrect signals are an even stronger indication, implying that $\underline{n} < 0$ and the agent is not hired when consistently inaccurate.[8] Consistency continues to be an important part of the first-best, however, especially when the signal precisions α_g and α_b are relatively large: separation of the two types can be improved by imposing a *super*-majority threshold of matching signals $\overline{n} > \lceil (T+1)/2 \rceil$ so that a mere majority of correct signals is not sufficient.[9]

Observe that an immediate consequence of Theorem 1 is that strategic behavior by the agent precludes the implementation of the first-best policy when $\overline{n} > \lceil (T+1)/2 \rceil$ or when $\underline{n} \geq 0$. In the former case, the agent has an incentive to misreport signals after histories where he has already observed and truthfully reported $T - \overline{n}$ signals of each type. Unless each of his $2\overline{n} - T$ remaining reports is identical (which is not guaranteed with truthful reporting), the agent has no chance of meeting the match threshold $\overline{n}$ and being hired. In the latter case, the agent is always hired for sure when he reports the same signal in all periods (regardless of whether it matches the outcome), and such consistency can always be mimicked. As a result, the principal achieves no separation whatsoever (both types will be hired with probability 1) in this case when faced with a strategic agent. Nontrivial screening is, however, always possible using the simple class of mechanisms (that are easy to implement in practice) that we present in the next section.

III. The Optimal Mechanism with Commitment

In this section, we consider the case where the principal can commit in advance to the mechanism x. We begin by describing a simple class of mechanisms via which screening can always be achieved. As we will argue, the optimal mechanism also belongs to this class.

A. *Prediction Mechanisms*

A hiring policy x is a *period-t prediction mechanism* if the agent is hired whenever the outcome r matches the state that is most likely, given the signals s^t reported through period t. Put differently, the agent is asked to predict the final outcome at period t, and is hired whenever this prediction matches the outcome. Formally, a period-t prediction mechanism can be implemented as a function of the reported signals as follows:

$$x_r(\tilde{s}^T) = \begin{cases} 1 & \text{if } \sum_{t' \leq t} \mathbf{1}_r(\tilde{s}_{t'}) > t/2 \\ 1 & \text{if } \sum_{t' \leq t} \mathbf{1}_r(\tilde{s}_{t'}) = t/2 \text{ and } \tilde{s}_t = r \cdot \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

[8] For $\gamma < 1$, it is always possible to find α_g and α_b such that $\underline{n} \geq 0$, regardless of how close γ is to 1.

[9] In general, as either type's signal precision α_θ rises and incorrect signals become less likely, the match threshold $\overline{n}$ increases and the first-best becomes less forgiving of mistakes. Meanwhile, the mismatch threshold $\underline{n}$ may be nonmonotone in α_θ: increasing the quality of an agent's information decreases the variance of his signals while increasing their accuracy, and these two effects have countervailing impacts on the likelihood of signals consistently contradicting the public outcome.

In words, this mechanism hires the agent for sure when a strict majority of his reported signals up to t match the outcome. When the agent reports an equal number of h and l signals through period t (so both states are equally likely), then the period-t report serves as a tie breaker and determines the hiring decision.[10]

It is straightforward to argue that truth-telling is optimal for *both* types g and b in response to the mechanism defined in (2). To see this, note that the agent does not have an incentive to misreport before period t even if he could pick a reporting strategy $\tilde{s}^t$ after observing all t signals s^t (instead of having to report them sequentially), as the majority of the signals corresponds to the outcome that is more likely to arise. Additionally, since the signals reported after period t do not affect the hiring decision, it is trivially optimal to report them truthfully. Finally, since the good type is always more likely to observe a majority of signals corresponding to the underlying state, he will be hired with greater probability than the bad type, and hence this mechanism always achieves nontrivial screening.

The principal can optimize within this class of mechanisms by choosing the period in which she solicits a prediction. The next result shows that it may not be optimal for the principal to wait until the final period, but instead should limit the information observed by the agent.

THEOREM 2: *There exists a $\hat{T} \geq 1$ such that the principal's payoff from a period-t prediction mechanism is increasing in t for all $t \leq \hat{T}$ and decreasing in t for all $t \geq \hat{T}$.*

The intuition for the nonmonotonicity of payoffs in t is simple. As t grows larger, both types learn about the underlying state more precisely. But in the limit as t becomes arbitrarily large, both types learn the state perfectly and thus make the same prediction. As a result, screening becomes less effective, and for sufficiently long time horizons T, the principal prefers to solicit a prediction at an intermediate time when the learning advantage of the type-g agent over his type-b counterpart is at its highest. In what follows, we will use $\hat{T}$ to denote the optimal period for the principal to solicit the prediction.[11]

B. *The Optimal Mechanism*

In this section, we will argue that the optimal mechanism is a prediction mechanism. To begin, it is worth reiterating that the class of mechanisms we consider (functions of the reported signals alone) is a strict subset of the set of direct revelation mechanisms. This is because the mechanisms x do not condition on the agent's initial private type. While a restriction to such mechanisms can be justified by appealing to their realism, we now argue that this restriction is also without loss of generality: the principal can achieve the same payoff by maximizing over the class of (indirect) mechanisms x as she can from the optimal direct revelation mechanism.

[10] Note that we can use other tie-breaking rules to implement the same outcome when $\sum_{t' \leq t} \mathbf{1}_r(\tilde{s}_{t'}) = t/2$. For instance, we can choose an arbitrary period $t' \leq t$ and hire the agent whenever $\tilde{s}_{t'} = r$.

[11] It is straightforward to show that $\hat{T}$ is decreasing in both α_g and α_b. Note that as α_g grows larger, the type-g agent's information becomes more precise quickly and his relative advantage over type b peaks sooner; meanwhile, as α_b grows larger, the type-b agent is able to "catch up" quickly to the type-g agent.

A *direct revelation mechanism* $\chi_r(\theta, s^T) \in \{0,1\}$ (or direct mechanism for short) determines the probability that the agent is hired as a function of his reported initial type θ, his profile of reported signals s^T, and the final outcome r. Note that the revelation principle applies in this environment, so it is without loss to consider the message space $\{g,b\} \times \{h,l\}^T$.[12] The next result states that the principal cannot attain a higher payoff from using this larger class of direct mechanisms.

LEMMA 1: *There is an optimal direct mechanism that does not depend on the reported type. Specifically, for any incentive-compatible direct mechanism χ, there is an indirect mechanism x with the following properties:*

(*i*) *The principal's payoff from x is* (*weakly*) *higher than her payoff from χ*;

(*ii*) *The type-g agent has an incentive to report his signals truthfully*; *and*

(*iii*) *The type-b agent reports his signals optimally*.

The intuition for this lemma is transparent. Fix any incentive-compatible direct mechanism χ that depends on the agent's reported type. Incentive compatibility implies that it is optimal for the agent to report his initial type truthfully; in particular, the type-*b* agent receives a lower payoff from initially misreporting his type as *g* and then optimally reporting his signals. So consider the indirect mechanism $x_r(\cdot) := \chi_r(g, \cdot)$. By definition, it is optimal for the good type to report his signals truthfully (property (ii)). Additionally, incentive compatibility of χ ensures that the bad type's payoff is lower from x than from χ. Since the principal's payoff is increasing in type *g*'s payoff and decreasing in type *b*'s, this indirect mechanism will be no worse for her (property (i)). Finally, note that since x is not a direct mechanism, we cannot a priori restrict attention to mechanisms where *both* types report their signals truthfully (made explicit by property (iii)).

Henceforth, when we refer to an *optimal* mechanism x, this will correspond to a mechanism that yields the highest payoff that the principal can achieve in the *full space* of direct mechanisms χ. The next theorem characterizes the optimal mechanism.

THEOREM 3: *Let $T^* := \min\{T, \hat{T}\}$. A period-$T^*$ prediction mechanism is an optimal mechanism.*

There are several aspects of the result above that are worth emphasizing. First, the optimal mechanism takes a very simple form that is easy to implement in practice, as it is both easy to time when predictions are solicited and to institute a hiring policy that depends on the accuracy of the predictions. Second, observe that the optimal mechanism has the property that truth-telling is optimal for *both* types of the agent. This property finds support from the empirical evidence on anonymous

[12] Strausz (2003) shows that the revelation principle does not always apply when the principal is restricted to deterministic mechanisms. However, it does apply in single-agent settings such as ours; for reference, we present a formal statement and proof in the online Appendix.

analyst surveys; for example, Marinovic, Ottaviani, and Sørensen (2013, p. 716) point out that, "According to industry experts, forecasters often seem to submit to the anonymous surveys the same forecasts they have already prepared for public (i.e., non-anonymous) release." This is suggestive evidence for the fact that strategic forecasters in the real world predict truthfully as they have no reason to lie in anonymous surveys.

Third, while the optimal period at which to solicit the prediction T^* depends on the underlying parameters, it does not do so in a fine-grained way. Put differently, the optimality of the period-T^* prediction mechanism will be robust to "small" inaccuracies in the principal's beliefs about the underlying parameters. Finally, while we will show the full generality of the insight driving the result above (in Section VI), it is worth mentioning that it is easy to incorporate asymmetries (in the prior belief regarding the state, the agent's type distribution, and the principal's payoffs from hiring the good or the bad type) within this simplified version of the model. The only change to the result above is that a trivial decision (to always or never hire the agent) may become optimal under some model parameters.

It might seem surprising that the optimal mechanism does not involve more elaborate screening. The reason is that the principal has limited instruments at her disposal and, as a result, incentive compatibility significantly restricts the set of effective mechanisms the principal can utilize. The characterization of the set of mechanisms that induce the good type-g to report truthfully is the crucial step in the proof of Theorem 3 and is described in the following lemma.

LEMMA 2: *A mechanism x induces truthful signal reporting from the type-g agent if and only if it is one of the following mechanisms:*

(*i*) *A trivial mechanism: the principal's hiring decision does not depend on the agent's reports, so that* $x_r(s^T) = x_r(\hat{s}^T)$ for $r = h, l$ *and all* $s^T, \hat{s}^T \in \{h, l\}^T$; *or*

(*ii*) *A period-t prediction mechanism for some* $1 \leq t \leq T$.

Consequently, a mechanism that induces truthful reporting by the type-g agent also induces truthful reporting by the type-b agent.

This lemma shows that incentive compatibility for type g (which, by Lemma 1, is a property of an optimal mechanism) implies that the only nontrivial mechanisms at the principal's disposal are prediction mechanisms. Combined with Theorem 2's payoff single-peakedness result, this implies Theorem 3. As we will argue further in Section VI, the insight in this lemma is remarkably general, applying immediately to substantially generalized versions of the model.

The following example is useful to develop intuition for the lemma. Suppose that a type-g incentive-compatible mechanism x is such that, at some period-T history, the hiring decision is a nontrivial function of the final period-T report s_T and the outcome r; that is, there is a sequence of signals s^{T-1} such that the hiring decisions $(x_h(s^{T-1}, h), x_l(s^{T-1}, h))$ after a report $\tilde{s}_T = h$ differ from the hiring decisions $(x_h(s^{T-1}, l), x_l(s^{T-1}, l))$ after a report $\tilde{s}_T = l$. Since the hiring rule is deterministic, this is only possible when $(x_h(s^{T-1}, \tilde{s}_T), x_l(s^{T-1}, \tilde{s}_T))$ equals either $(1, 0)$ or

$(0,1)$.[13] Incentive compatibility implies that, if $(x_h(s^{T-1}, s_T), x_l(s^{T-1}, s_T)) = (1,0)$, then the agent must believe the outcome $r = h$ is more likely, as he could instead report $s'_T \neq s_T$ and face $(x_h(s^{T-1}, s'_T), x_l(s^{T-1}, s'_T)) = (0,1)$. However, this essentially implies that the agent is hired if and only if the outcome he believes to be more likely is realized; in other words, the mechanism is effectively soliciting a prediction at this history and then hiring based on its accuracy. The proof of Lemma 2 generalizes this argument to all histories.

C. *Alternative Implementations and Interpretations of the Optimal Mechanism*

In this section, we revisit our leading example (a political candidate seeking to hire a forecaster) to discuss alternate ways in which the optimal mechanism can be implemented. This discussion also allows us to demonstrate the flexibility of our framework to capture the various different forms that political forecasts often take. The period-t prediction mechanism as defined in (2) captures the case where the agent reports his signal in each period: this can be interpreted as a pollster sequentially releasing the predicted outcomes from each poll he conducts. (As we will argue in Section VI, the model can be generalized to allow for signals with a continuous support in which case this will correspond to releasing the poll results as a percentage instead of a prediction.) Here, the prediction from a period-t poll corresponds to his period-t signal and not the cumulative information he has acquired.

Alternatively, political punditry often takes the form of an expert predicting who he thinks is more likely to win in each period after aggregating all his past information. In this case, each report $\tilde{s}_t$ can be interpreted as a prediction of the final outcome and not a signal report. When the agent reports this way, the period-t prediction mechanism simply becomes

$$x_r(\tilde{s}^T) = \begin{cases} 1 & \text{if } \tilde{s}_t = r \\ 0 & \text{otherwise} \end{cases}. \tag{3}$$

In words, this mechanism asks the agent to predict the final outcome at each period and hires the agent if and only if the period-t prediction matches the outcome (all other reports are ignored). A strategic agent facing this mechanism will predict at period t whichever outcome he believes is more likely to eventually arise. This will simply be the outcome for which the agent has received more signals up to period t (and he will be indifferent if he has received an equal number of h and l signals). This implementation clearly achieves the identical payoff to the principal as that in (2).

Finally, our framework is flexible enough to allow for the forecaster to make predictions on the odds of the likely winner (in the form of a percentage). Such predictions are made by forecasters like The Upshot of the *New York Times* or FiveThirtyEight by Nate Silver who aggregate information from various polls (and

[13] If $x_h(\hat{s}^T) = x_l(\hat{s}^T) = 1$ for some sequence $\hat{s}^T$, then the agent can guarantee he is hired with probability 1 by *always* reporting $\hat{s}^T$, regardless of his true signals. Since this potential deviation remains unused (as the type-g agent is willing to report truthfully), the principal must therefore always (trivially) hire the agent. An analogous argument applies if there is some $\hat{s}^T$ with $x_h(\hat{s}^T) = x_l(\hat{s}^T) = 0$.

their type determines the accuracy of this aggregation). To model this, we can simply alter the message space so that the agent is asked to make a percentage prediction (of course, the revelation principle implies that enlarging the message space in this way does not alter the optimal mechanism). Here the period-t prediction mechanism can be implemented by hiring the agent if and only if the outcome he predicts is more likely (in a percentage sense) in period t ends up occurring.

IV. The Optimal Mechanism without Commitment

In this section, we derive the equilibrium that maximizes the principal's payoff when she cannot commit to her hiring policy x. Of course, the principal is always weakly better off with commitment power as she can always choose to commit to whatever strategy she can play in its absence. We show that the principal can achieve the same payoff as in Theorem 3 even when she does not have commitment power. We view this as further support for the optimal mechanism in Theorem 3 as, in practice, the level of commitment possessed by principals may vary.

In the absence of commitment, our setting constitutes a dynamic cheap talk game. Here, the agent (the sender) can costlessly make either report in every period t. The reports $\tilde{s}^T$ themselves are not payoff-relevant; instead, their only purpose is to inform the principal's (the receiver) decision. The principal's payoff-relevant information is, instead, the agent's private type, and (as in the standard cheap talk setting) the principal and agent have divergent preferences over the former's action choice as a function of this type. Finally, rather than consider alternative message spaces or games, we will directly show that the principal can achieve the same payoff both with and without commitment power.

THEOREM 4: *There is a sequential equilibrium of the game without commitment that yields the principal the same payoff as a period-t prediction mechanism. In particular, this implies that the principal can achieve the same payoff as in the optimal mechanism with commitment.*

Due to the simple structure of the optimal mechanism under full commitment, this result is remarkably straightforward to show. We now describe equilibrium strategies that replicate the outcome of a period-t prediction mechanism. The principal's strategy is to ignore all reports of the agent except that in period t, and she hires the agent if and only if his period-t report matches the outcome. In response, both types of the agent babble in all periods except period t, where they report the signal corresponding to the outcome that they consider more likely to arise.

Formally, the principal's strategy is

$$x_r(s^T) = \begin{cases} 1 & \text{if } s_t = r \\ 0 & \text{otherwise} \end{cases}.$$

For any $t' \neq t$, the agent's strategy is

$$\sigma_h^\theta(h_{t'}^A) = 1 - \sigma_l^\theta(h_{t'}^A) = \frac{1}{2}$$

for all period-t' agent histories; that is, he mixes both reports with equal probability. For a period t history $h_t^A = (s^t, \tilde{s}^{t-1})$ (recalling that s^t denotes the t observed signals and $\tilde{s}^{t-1}$ denotes the $t-1$ reports made prior to period t), the agent's strategy is

$$\sigma_h^\theta(s^t, \tilde{s}^{t-1}) = 1 - \sigma_l^\theta(s^t, \tilde{s}^{t-1}) = \begin{cases} 1 & \text{if } \sum_{t' \leq t} \mathbf{1}_h(s_{t'}) > t/2 \\ \frac{1}{2} & \text{if } \sum_{t' \leq t} \mathbf{1}_h(s_{t'}) = t/2 \cdot \\ 0 & \text{otherwise} \end{cases}$$

It is straightforward to see that these strategies constitute an equilibrium. Since the principal ignores the reports in all periods except t, the agent is indifferent at all such histories; in particular, babbling is therefore a best response. In addition, he is hired only if his period-t report matches the outcome, so it is a best response for him to report whichever signal he has seen more often (and is indifferent if he has seen an equal number of h and l signals). Conversely, since the agent is babbling at all periods except t, it is a best response for the principal to ignore these reports. Finally, since the type-g agent is more likely to correctly predict the outcome, it is optimal for the principal to hire the agent when his period-t report matches the outcome.

Note that all possible signal reports are on-path in the agent's strategy above. Thus, as in the canonical cheap talk setting of Crawford and Sobel (1982), standard refinements have no bite as there is no need to discipline off-path behavior. In particular, the equilibrium constructed above is a sequential equilibrium. To the best of our knowledge, there are no accepted refinements of dynamic cheap talk games; moreover, since our setting is quite different from the canonical setting, it is not clear how to extend the refinements designed specifically for the static Crawford and Sobel (1982) environment (most notably Chen, Kartik, and Sobel 2008). The design of such refinements for dynamic cheap talk games is an important topic of research but is beyond the scope of this paper.

V. Stochastic Mechanisms

In this section, we describe how the principal can utilize randomization to fine-tune screening. Formally, the principal's strategy, which we refer to as a *stochastic mechanism* when she has commitment, now has the entire unit interval as its range. We will use the same notation as before: $x_r(\tilde{s}^T) \in [0,1]$ denotes the probability with which she hires the agent as a function of the T reported signals $\tilde{s}^T$ and the outcome r. For brevity, we will sometimes drop the additional "stochastic" qualifier in this section when it is clear that we are referring to a stochastic mechanism.

The optimal mechanism is difficult to derive for arbitrary time horizons T. This is primarily because the set of incentive-compatible stochastic mechanisms is much larger and harder to characterize than in the deterministic mechanism case. Similar issues are also encountered in dynamic mechanism design environments with transfers (hence the restriction to deterministic mechanisms in Courty and Li 2000 or Krähmer and Strausz 2011, for instance). The main aim of this section is to show that the screening is more subtle with randomization for which the restriction to the special case of $T = 3$ suffices. That said, we also provide a simple sufficient condition in for when randomization is a feature of the optimal stochastic mechanism (Theorem 6).

A. *The Role of Randomization When $T = 3$*

In this subsection, we describe the optimal stochastic mechanism for the $T = 3$ period case and discuss its qualitative properties. This special case is convenient to highlight the role played by randomization as the optimal stochastic mechanism can be characterized and is easy to describe.

We begin by describing the first-best mechanism x^{FB} for the case where the agent's signals (but not his initial type θ) are also observed by the principal. The following characterization of the set of possible match and mismatch thresholds (corresponding to Theorem 1) that can arise in x^{FB} is instructive as a point of contrast with the optimal mechanism.

LEMMA 3: *Suppose $T = 3$. Then the first-best mechanism x^{FB} is one of the following:*

(*i*) *Hire the agent if and only if all three of his signals are accurate* (*so $\overline{n} = 3$ and $\underline{n} = -1$*);

(*ii*) *Hire the agent if and only if all three of his signals are consistent* (so $\overline{n} = 3$ *and* $\underline{n} = 0$); *or*

(*iii*) *Hire the agent if and only if a majority of his signals match the outcome* (*so $\overline{n} = 2$ and $\underline{n} = -1$*).

Observe that the first-best mechanism in case (iii) is simply a period-3 prediction mechanism, and is therefore implementable: the agent will predict the outcome corresponding to the majority of his signals. This case arises when $\Delta_{2,3} \geq 0$ and the type-g agent is more likely to observe exactly two matches than the type-b agent. When this is not the case and $\Delta_{2,3} < 0$, however, the first-best payoffs corresponding to cases (i) and (ii) cannot be achieved as a strategic agent can easily feign consistency by simply "cascading" on his first signal. In such circumstances, the optimal mechanism (characterized in the following theorem) is distorted away from the first-best.

THEOREM 5: *Suppose $T = 3$. When $\Delta_{2,3} \geq 0$, the period-3 prediction mechanism is an optimal stochastic mechanism. Conversely, when $\Delta_{2,3} < 0$, the optimal stochastic mechanism is given by*

$$x_r(s^3) = \begin{cases} 1 & \text{if } s_1 = s_2 = r \\ \dfrac{1}{2(\gamma\alpha_b + (1-\gamma)(1-\alpha_b))} & \text{if } s_1 \neq s_2 \text{ and } s_3 = r. \\ 0 & \text{otherwise} \end{cases}$$

Faced with this mechanism, it is optimal for both types of the agent to truthfully report all signals.

The optimal stochastic mechanism when $T = 3$ hires the agent only if a majority of his reported signals match the outcome. Moreover, when the type-b agent is more likely to match exactly two of three signals than the type-g agent (that is, when $\Delta_{2,3} < 0$), the *order* of reported signals influences the hiring decision. Specifically, the optimal mechanism rewards early accuracy: in profiles where exactly two of the three reports match the outcome, the agent is hired with higher probability when the first two reports are correct than when one of them mismatches.

Intuitively, when $\Delta_{2,3} < 0$, the principal would prefer not to hire the agent at histories where he truthfully reports only two signals matching the outcome (as such profiles are more likely for the type-b agent). But as we have seen, deterministic mechanisms compel the principal to hire the agent at such profiles whereas, when the principal can randomize, she can reduce the hiring probability at such profiles without violating incentive compatibility.

To better understand how randomization permits such a reduction, it is helpful to reinterpret the hiring rule in Theorem 5 as an option mechanism: in the second period, the agent is offered the opportunity to make a prediction immediately or to delay his prediction to period 3. A correct prediction in period 2 is rewarded by hiring the agent for sure, while a correct prediction in period 3 is rewarded by hiring the agent with a probability strictly less than 1. (An incorrect forecaster is never hired, regardless of the timing of his prediction.) Faced with this option, an agent who has observed two identical signals will always make a prediction in period 2; no matter what he observes in the third period, the agent's prediction will remain unchanged but his probability of being hired is lower. However, an agent who has observed contradictory signals is uncertain about the underlying state, and therefore benefits from delaying his prediction by a period. Indeed, the reduced probability of being hired in period 3 after mixed signals is chosen precisely to ensure that the type-b agent is indifferent about delaying his prediction, while type g's better information gives him a strict incentive to wait for an additional signal.[14] Of course, since the type-g agent is more likely to observe two matching signals in the first two periods, he is correspondingly more likely to make an early prediction (with a larger hiring probability), compounding his pure informational advantage over the type-b agent.

Interestingly, it is possible to implement the optimal stochastic mechanism when $\Delta_{2,3} < 0$ in a sequential equilibrium of the game without commitment. In particular, the type-b agent's indifference after contradictory signals (that is, after observing $s_2 \neq s_1$) permits the appropriate mixed strategy that rationalizes the principal's randomization.[15] It remains an open question, however, whether this property generalizes beyond the special case of $T = 3$.

It is instructive to briefly contrast the proof strategy for Theorem 5 with that for Theorem 3. First, observe that Lemma 1 also applies to stochastic mechanisms, so it is without loss to consider stochastic mechanisms where the type-g agent reports truthfully while type b is allowed to optimally misreport. Unlike with deterministic

[14] Note that the randomization necessary to generate this indifference for the type-b agent relies on the informativeness γ of the public outcome. This is in contrast to the case of deterministic mechanisms where, as long as the public outcome is equally accurate in both states of the world, the optimal prediction mechanism does not depend on γ.

[15] Details of the equilibrium construction are available from the authors on request.

mechanisms, which showed that incentive compatibility for type g implies incentive compatibility for type b, incentive constraints in a stochastic mechanism may be less restrictive. In particular, there are stochastic mechanisms where truthful reporting of signals is incentive compatible for type g but not for type b. Therefore, it is difficult to formulate a tractable version of the principal's optimization problem. Our proof instead relies on an auxiliary problem that is both easier to solve and yields the principal a greater payoff; we then show that resulting solution is in fact feasible in the original problem.

B. *When Is Randomization Optimal?*

We now provide a simple sufficient condition for the optimality of randomization.

THEOREM 6: *The principal's payoff from the optimal stochastic mechanism is strictly higher than that from the optimal (deterministic) mechanism when* $T > \hat{T} + 1$.

This result states that the principal strictly benefits from using randomization for sufficiently long time horizons. Intuitively, recall that the optimal (deterministic) mechanism for $T > \hat{T}$ is a period-$\hat{T}$ prediction mechanism: in this mechanism, the principal ignores reports after $\hat{T}$. In the optimal stochastic mechanism for $T = 3$, when the agent in period 2 has conflicting signals (and therefore thinks both outcomes are equally likely), the principal can fine-tune screening by lowering the hiring probability (which is beneficial since type b's higher signal variance implies he is more likely to receive an equal number of h and l signals) without destroying incentive compatibility. The principal can similarly lower the hiring probability at profiles in which the agent reports the same number of h and l signals (or in which the difference between h and l signals is one) by conditioning the mechanism on reports after period $\hat{T}$. The agent prefers such a mechanism as it allows him the chance to better learn the underlying state.

It is hard to fully characterize the optimal stochastic mechanism for $T > 3$ since incentive compatibility for type g alone is no longer sufficient to pin down type b's reporting strategy. As a result, the derivation of the optimal mechanism must account for optimal misreporting, which makes the problem intractable. In the $T = 3$ period case, it is possible to identify and individually account for histories at which the type-b agent might have an incentive to misreport; when $T > 3$, however, the set of such histories becomes large, and this approach is no longer feasible.

VI. A More General Model

In this section, we show the generality of our main insight that prediction mechanisms are the optimal way to screen strategic forecasters. As we will argue, the critical assumption driving our result is that the outcome r that is being predicted is binary; every other assumption can be substantially generalized. We describe the key components of the model in their full generality below and deliberately overload the notation to make the generalization of each assumption explicit. The timing of the model remains unchanged.

State.—There is an unknown underlying state ω drawn from an arbitrary set Ω that drives the data-generating process. State ω is distributed according to a commonly known probability measure $p_0 \in \Delta(\Omega)$. Notice that Ω need not be binary; this permits the analysis, for instance, of environments where the state ω is the realized sample path of a general stochastic process.

Agent's Private Information.—The agent's type θ is drawn from an arbitrary (again, not necessarily binary) set Θ. The commonly known prior distribution of θ is given by $\mu_0 \in \Delta(\Theta)$.

In this general setting, we allow for the possibility that the agent does not perfectly observe his initial type, but instead learns about his forecasting ability over time.[16] We model this by adding an additional signal: formally, in period 0, the forecaster observes a single private signal $\lambda \in \Lambda$, where the set Λ is arbitrary. This signal λ is drawn from a (commonly known) measure $\mu_\theta \in \Delta(\Lambda)$ that may vary by type θ. Thus, the case of a perfectly informed agent corresponds to $\Lambda = \Theta$ and $\mu_\theta(\{\theta\}) = 1$ for all $\theta \in \Theta$. On the other hand, the case where $\mu_\theta = \mu_{\theta'}$ for all $\theta, \theta' \in \Theta$, so the distribution of λ does not vary by type, corresponds to a "signal jamming" version of our model (similar to the career concerns literature following Holmström 1999) where both the principal and the agent start with the same information.

In each period, $t = 1, \ldots, T$, the forecaster privately observes a noisy but informative signal s_t, drawn from an arbitrary signal space S_t, about the unknown state ω. Signals are conditionally independent given the underlying state ω and the agent's type θ, and s_t is drawn from a distribution $\alpha_{\omega,\theta,t} \in \Delta(S_t)$. Note that both the signal spaces and distributions may vary over time.

Outcome.—As in the simplified model, a binary outcome $r \in \{h, l\}$ is publicly realized at the end of period T. We denote by $\gamma_\omega \in [0,1]$ the probability of outcome h arising when the true state is ω. Note that the outcome remains a potentially noisy signal of the state, but the joint distribution is not restricted in any way.

Payoffs.—The agent of type θ now receives a payoff $u_\theta > 0$ from being hired, and a payoff of 0 if he is not. Note that this does not change the agent's incentives compared to the simplified model in Section I as his objective is still to maximize the probability of being hired.

Finally, the principal's payoff can also be made type-dependent: she receives a payoff $\pi_\theta \in \mathbb{R}$ if she hires an agent of type θ, and a payoff of 0 from not hiring the agent.

Appropriate definitions of strategies and mechanisms generalize to this richer environment in the obvious way. The next result shows that prediction mechanisms remain optimal even in this very general environment. Note that the definition of a prediction mechanism as a function of reported signals will differ from that in (2) as

[16] In discussing the important directions for future research on strategic forecasters, Marinovic, Ottaviani, and Sørensen (2013, p. 717) state that a "key challenge lies in finding a tractable and sufficiently general multi-period environment with learning about the precision as well as about the state." Our general model takes a step in this direction.

the environment is no longer symmetric and the signal space is not binary. Instead, we make use of the alternative definition in (3).

THEOREM 7: *In the general model, one of the following mechanisms is optimal:*

(*i*) *A trivial mechanism: the principal's hiring decision does not depend on the agent's reports*; *or*

(*ii*) *A period-t prediction mechanism for appropriately chosen t.*

Additionally, the principal can implement the same outcome and thereby achieve the same payoff in a sequential equilibrium of the game without commitment.

As in the case of simplified model (Theorem 3), prediction mechanisms are optimal within the full class of direct revelation mechanisms. As we allow for more than two types, it is no longer possible to directly argue (as in Lemma 1) that the principal cannot benefit by asking the agent to report his type in a direct revelation mechanism. Instead, our proof characterizes incentive-compatible direct mechanisms in this general setting. Effectively, we show that the only nontrivial incentive-compatible direct mechanisms are prediction mechanisms and thus, finding the optimal mechanism only involves choosing the time at which to solicit the prediction. Of course, additional structure is necessary to fully characterize the optimal prediction period as a function of model parameters.

Essential to Theorem 7's characterization of incentive compatibility is the assumption that the publicly observable outcome (that is, the information available to the principal when evaluating a forecast) is binary. Enriching the set of possible outcomes yields the principal a substantially more complex set of instruments: agents could, for instance, be asked to make predictions about nested partitions of possible outcomes. This increased dimensionality of the set of possible mechanisms precludes a characterization of the principal's optimal mechanism.

On the other hand, the result does not require us to take a stand on the relationship between the principal's payoff and the information of the agent. For instance, we do not need to assume that "good" types (for which $\pi_\theta > 0$) receive better information than "bad" types (for which $\pi_\theta < 0$). Of course, some additional structure (like that we impose in our simplified model) is desirable to capture specific applications.

The critical assumptions of the general model in this section are supported by our leading application. The richness of the time-varying signal space captures the myriad different sources of information that are available to political pundits. Importantly, this application satisfies the key driving assumption of our model: political predictions are always about the eventual winner which is a binary outcome in the (effectively) two-party US political system. As we argued in Section IIIC, our model is flexible enough to capture the variety of different forms that political predictions come in. It is worth noting that an election result taken as the difference in vote counts can be considered as a continuous outcome variable; however, to the best of our knowledge, political forecasters always predict, and are judged on, election winners' identities and not their margins of victory.

VII. Discussion

In this section, we address a few important structural assumptions of the model. For ease of exposition, the discussion will employ the simplified setup of Section I.

A. *The Role of Sequential Reporting*

In the canonical dynamic mechanism design environment with transfers (see, for instance, Courty and Li 2000 or Pavan, Segal, and Toikka 2014), the fact that the agent receives his private information sequentially plays an important role for the tractability of the model. Because the agent has single dimensional private information at the time of contracting, incentive compatibility is easier to characterize than in static, multidimensional mechanism design environments where the agent has acquired all his private information before contracting (Eső and Szentes 2017 demonstrate the generality of this technique). An important underlying economic insight is that the principal benefits from being able to contract with an agent when her informational disadvantage is at its lowest as the agent has not acquired the entirety of his private information.

We now isolate the role played by sequential reporting in our model by drawing a contrast with the optimal stochastic mechanism in the static multidimensional version of our environment. In the static game, the agent's strategy σ^θ is defined as follows: he first observes his T signals s^T and his strategy $\sigma^\theta(s^T) \in \Delta(\{h,l\}^T)$ determines the distribution over T-vectors $\tilde{s}^T \in \{h,l\}^T$ of signal reports. As before, the principal's strategy $x(\tilde{s}^T, r) \in [0,1]$ depends on the vector of reported signals and the outcome but observe that we also allow the principal to randomize. We refer to the principal's strategy when she can commit as a *static stochastic mechanism.*

THEOREM 8: *The optimal static stochastic mechanism yields the principal the same payoff as that from a period-T prediction mechanism in the dynamic game with sequential reporting.*[17]

The optimal static mechanism is equivalent to the period-T prediction mechanism in the dynamic environment. There are two aspects of Theorem 8 that are worth highlighting. The first is that, with long time horizons $T \geq \hat{T}$, the principal cannot prevent the agent from using the information he receives after period $\hat{T}$. This is in contrast with Theorem 3 where the principal chooses to ignore reports after $\hat{T}$. Recall also, that for the case of deterministic hiring policies, the principal does not even need commitment to maximize her payoff (Theorem 4). Thus, in our main setting of interest (deterministic mechanisms), the dynamics of agent learning plays a greater role than commitment.

Secondly, observe that it is not optimal for the principal to employ randomization in the static setting. This is in contrast with the optimality of randomization when the time horizon is long (Theorem 6). This latter aspect is also a feature of the sequential screening setting of Courty and Li (2000). There too, the principal may

[17] The proof of this result can be found in the online Appendix.

employ randomization with dynamic reporting but will not if restricted to using a static mechanism after the agent has acquired all his private information. This similarity is captured by the Myersonian (1981) approach that we take in the proof of Theorem 8.

To summarize, simple mechanisms are optimal in our model unless the principal can randomize *and* the environment is dynamic. Put differently, *both* these aspects must be present simultaneously in order for the optimal mechanism to take a form more complex than a prediction mechanism.

B. *Transfers*

While our setting without transfers is appropriate for the applications we have in mind, it is natural to explore the theoretical implications of permitting them. We begin by discussing the optimal *direct* mechanism but, as we will argue below, it is also possible to implement this direct mechanism by with an indirect mechanisms that does not condition on the agent's type. A *direct mechanism with transfers* consists of two functions

$$\chi_r(\theta, s^T) \in \{0,1\} \quad \text{and} \quad \tau_r(\theta, s^T) \in \mathbb{R},$$

where χ is (as before) the hiring decision and τ is a transfer that also depends on the reported type, signals, and outcome. Both types receive (arbitrary) strictly positive utility from being hired and zero utility if they are not.

Since the signal distributions of both types are correlated, we can use the insight of Crémer and McLean (1988) to induce the agent to reveal his type with a zero expected transfer, thereby ensuring that the principal only hires the type-g agent.[18] To see this, consider the mechanism

$$\chi_r(\theta, s^T) = \begin{cases} 1 & \text{if } \theta = g \\ 0 & \text{if } \theta = b \end{cases}, \quad \text{and} \quad \tau_r(\theta, s^T) = \begin{cases} \kappa\overline{\beta} & \text{if } \theta = g \text{ and } s_1 = r \\ \kappa\underline{\beta} & \text{if } \theta = g \text{ and } s_1 \neq r, \\ 0 & \text{if } \theta = b \end{cases}$$

where $\overline{\beta}, \underline{\beta}, \kappa > 0$ and

$$[\alpha_g \gamma + (1-\alpha_g)(1-\gamma)]\kappa\overline{\beta} - [\alpha_g(1-\gamma) + (1-\alpha_g)\gamma]\kappa\underline{\beta} = 0.$$

In words, this mechanism hires the agent only if he reports type g, and the transfer depends on the reported period-1 signal. Moreover, this transfer is such that, if the agent reports type g, he receives $\kappa\overline{\beta}$ if the first signal matches the outcome, makes a payment of $\kappa\underline{\beta}$ to the principal if it does not, and has an expected payment from reporting truthfully (for a type-g agent) of 0.

Now observe that reporting truthfully is optimal for type g. He has no incentive to report his initial type as b, as he is hired with an expected transfer of zero if he is truthful. Moreover, he has no incentive to misreport his period-1 signal s_1 as he

[18] Olszewski and Peski (2011) apply a general version of this insight to the "testing experts" problem discussed earlier.

receives a positive transfer when the report matches the outcome (and a negative transfer when it does not). Type b receives zero utility from truthful reporting. If, instead, the type-b agent misreports his type, he will then find it optimal to report his period-1 signal s_1 truthfully (for the same reason that type g does). However, type b will now have to make a strictly positive expected payment to the principal (as $\alpha_b < \alpha_g$). Note that κ can always be chosen to be large enough so that this payment will be greater than the utility that type b gets from being hired. Thus, this mechanism achieves the best possible outcome for the principal.

Finally note that we can implement a similar outcome using the class of mechanisms that does not depend on the explicit announcement of the type (due to the presence of transfers, Lemma 1 does not apply here). The principal can always use the first signal report to proxy for the type announcement (for instance, interpreting $s_1 = h$ as an announcement that the agent is type g and vice versa). Having solicited this information, the principal can choose her hiring rule and construct similar transfer lotteries as above using the signal reports s_t from periods $t > 1$ to ensure that only type g is hired at a zero expected transfer.

C. *More than Two Outcomes*

As we discussed in Section VI, the key assumption driving our results is that the eventual outcome being predicted is binary. Aside from our leading example of an election with two candidates, the binary outcome assumption is appropriate for a variety of other environments. For instance, binary forecasts are critical for defense and intelligence decisions made by governments (Does a given "rogue state" possess nuclear arms capabilities or not? Will protests and civil unrest in some country lead to regime change or not?), as well as for a broad range of economic decisions (Will a trade treaty be ratified or not? Will a proposed merger be approved by antitrust authorities or not?). That said, there is a broad range of important nonbinary events (election outcomes with more than two viable candidates, for instance, or continuous economic variables like the rate of GDP growth) for which professional forecasts are indispensable. In this section, we argue that prediction mechanisms are no longer generally optimal in such environments, and discuss some of the richness that arises from multiple outcomes.

In a sense, Theorem 5 (characterizing the optimal stochastic mechanism when $T = 3$) already demonstrates that prediction mechanisms are not generally optimal deterministic mechanisms in environments with rich outcome spaces. To see why, notice that a stochastic mechanism can be interpreted as a deterministic mechanism with a *public* randomization device. For instance, consider an environment with the binary state and signal structure from Section I, but an enriched outcome space: outcome $r = h$ is replaced with a draw from the uniform distribution on $(0,1)$, while outcome $r = l$ is replaced with a draw from the uniform distribution on $(-1,0)$. Since the signal structure is unchanged, the agent continues to learn only about the underlying state and, as a result, can only predict whether the final outcome is likely to be positive or negative. By conditioning on the specific realized value of the public outcome (and not just its sign), an optimal deterministic mechanism with this richer outcome space can perfectly replicate the optimal stochastic mechanism described in Theorem 5.

It is worth highlighting that this argument does not require an outcome space as rich as that described above; indeed, adding only a single "noise" outcome can result in the suboptimality of prediction mechanisms, as the following example demonstrates.

Example: Consider the basic binary state and signal structure from Section I, but with a public outcome $r \in \{h, m, l\}$ whose distribution is given by

$$\Pr(r \mid \omega) = \begin{cases} \varepsilon & \text{if } r = m \\ \gamma(1-\varepsilon) & \text{if } r = \omega \\ (1-\gamma)(1-\varepsilon) & \text{otherwise} \end{cases}.$$

In words, m is a noise outcome that occurs with probability $\varepsilon > 0$, regardless of the underlying state; conditional on not observing outcome m, the outcome distribution is as in our baseline model, and the outcome matches the true state with probability $\gamma \in [1/2, 1]$. Although the agent's signals provide no information about it, this additional outcome can be used to offer dynamic options to the agent that improve upon a simple prediction mechanism. For instance, consider the mechanism that offers the agent the option of making a prediction in period 2 or postponing until period 3: the agent is hired (with certainty) if the prediction is made in period 2 and is not contradicted by the eventual outcome (i.e., if the outcome matches the prediction *or* if the noise outcome m is realized), or if the prediction is made in period 3 and exactly matches the eventual outcome. Thus, an early prediction is rewarded by "lenience" in the set of acceptable outcomes after which the agent is hired, whereas a delayed prediction leaves the agent less leeway. Even for arbitrarily small $\varepsilon > 0$, such an option mechanism can strictly improve on a "pure" prediction mechanism in precisely the same manner that randomization raises the principal's payoff in Theorem 5.

The discussion above focuses on environments where the noise in the final outcome r is rich enough, relative to the underlying state space, to play an additional role as an explicit randomization device. Even when this is not the case and the outcome perfectly reveals the state, however, screening with simple prediction mechanisms may fail to be optimal if the outcome is not binary.

Example: Consider an environment with three possible states $\omega \in \{h, m, l\}$, each with equal probability, and a final outcome $r \in \{h, m, l\}$ that perfectly reveals the state, so $\Pr(r = \omega) = 1$. The agent is one of two types $\theta \in \{g, b\}$, and preferences are as in the basic model from Section I. In each period t, a conditionally independent signal $s_t \in \{h, m, l\}$ is drawn from the distribution

$$\Pr(s_t \mid \omega, \theta) = \begin{cases} \alpha_\theta & \text{if } s_t = \omega \\ \frac{1-\alpha_\theta}{2} & \text{if } s_t \neq \omega \end{cases},$$

where $\alpha_g = 0$ and $\alpha_b = 1/2$. Thus, type g observes "negative" signals that never correspond to the underlying state, so the agent knows with certainty that the outcome cannot be any signal he observes. In particular, after observing two distinct signals, the type-g agent can predict the outcome correctly with probability 1. Meanwhile, the type-b agent observes informative "positive" signals that are likely to match the true state.

It is straightforward to compute that, when $T = 3$, the optimal prediction mechanism in this setting asks the agent to make a prediction in the final period. However, the principal can improve on such a mechanism by using a two-stage mechanism: in period 1, the agent must rule out an outcome; in period 3, the agent then chooses one of the two remaining outcomes as his final prediction; and, finally, the agent is hired (deterministically) if this final prediction is correct.

Observe that the type-g agent is indifferent between this two-stage mechanism and the period-3 prediction mechanism, as his period-3 prediction is the same in both mechanisms (he never predicts s_1, which he rules out in the two-stage mechanism). The type-b agent, on the other hand, strictly prefers the period-3 prediction mechanism. After observing s_1, the type-b agent eliminates an outcome $r' \neq s_1$; however, there is a strictly positive probability that $s_2 = s_3 = r'$ and the outcome that is ex post most likely is no longer a permissible prediction in period 3. Thus, type b's (constrained) prediction in the two-stage mechanism is more likely to be incorrect than his (unconstrained) prediction in the period-3 prediction mechanism. Therefore, this two-stage mechanism yields greater separation than the period-3 prediction mechanism; in fact, it is possible to show that it is the optimal deterministic mechanism in this environment.

The two-stage mechanism in this example leverages an important aspect of nonbinary dynamic learning environments: when there are many states and many outcomes, the agent's beliefs follow a multidimensional stochastic process, and so different types' posteriors can take very different paths while converging toward correct beliefs. In the example above, although both types eventually learn the underlying state, the type-g agent can quickly learn enough to eliminate an individual outcome while the type-b agent cannot. The optimal mechanism thus takes advantage of that early separation in the two type's posteriors by accounting for the "direction" of learning.

With a binary outcome, on the other hand, beliefs are single-dimensional, and so learning cannot generate the same diversity of possible belief process paths. (There is no difference between "positive" and "negative" signals in a binary environment, for instance.) This permits the general optimality result in Theorem 7 for binary-outcome environments.

VIII. Concluding Remarks

In this paper, we introduce the problem of evaluating a strategic forecaster based on the dynamics of the reports he makes about an upcoming event. In doing so, we bring two novel aspects to the study of evaluating forecasters that differ from the existing literature in economics and psychology: prediction dynamics and mechanism design by the evaluator. In a very general setting, we derive the optimal deterministic dynamic mechanism for the principal and show that it takes a very simple and easy-to-implement form. The simplicity of the optimal mechanism, combined with the fact that commitment is not necessary to implement it, implies that it can serve as a simple guideline for hiring forecasters.

More generally, our main economic insight is that optimal screening relies on a combination of accuracy and the speed of learning. Unlike in the first-best benchmark (where the agent's signals are publicly observed), strategic reporting by the agent prevents the principal from screening more finely by conditioning her decision on other features of the agent's information (such as the consistency of signals in the simplified model we analyze). This intuition can be captured with a simple analogy.[19] Suppose a teacher is trying to determine the ability (high or low) of a student by using a test consisting of a single difficult true-false question, and is free to determine the duration of the exam. If the teacher allows insufficient time, the student will essentially be forced to guess randomly, while if she allows too much time, both types will be able to answer correctly. Hence, to maximally distinguish between the two types, the optimal duration must be an intermediate time. In principle, the teacher might have solicited the student's level of confidence in his answer for her evaluation; indeed, this would be first-best. But this is not incentive compatible, as a strategic student would always report the level of confidence that maximizes his chances of passing.

We conclude by noting two natural avenues of inquiry that merit further investigation. As we discussed in Section VIIC, the first is to examine different environments with richer (nonbinary) outcome spaces. A second natural generalization considers optimal contest design for multiple forecasters. We hope to investigate these questions in future research.

Appendix

PROOF OF THEOREM 1:

In the first-best, the principal observes the agent's signals s^T, but not the agent's type. Therefore, the first-best optimal mechanism must solve

$$\max_{x_h(\cdot),x_l(\cdot)}\left\{\sum_{r\in\{h,l\}}\sum_{s^T\in\{h,l\}^T}\left[\frac{1}{2}\Pr(r,s^T\,|\,\theta=g)-\frac{1}{2}\Pr(r,s^T\,|\,\theta=b)\right]x_r(s^T)\right\}.$$

Note, however, that $\Pr(r,s^T\,|\,\theta)$ is constant across all s^T with the same number of matching signals $s_t = r$, regardless of the order of those signals. Therefore, with slight abuse of notation, we can write any solution $x_r(s^T)$ to the principal's problem as $x(n)$, where $n = \sum_t \mathbf{1}_r(s_t)$.

Therefore, with slight abuse of the notation from the main text, we write

$$\beta_{n,\theta,\gamma} := \gamma\alpha_\theta^n(1-\alpha_\theta)^{T-n}+(1-\gamma)\,\alpha_\theta^{T-n}(1-\alpha_\theta)^n \quad \text{and} \quad \Delta_{n,\gamma} := \beta_{n,g,\gamma}-\beta_{n,b,\gamma}$$

[19] We are grateful to a referee for suggesting this metaphor.

for all $n \in [0,T]$. (Recall that $\binom{T}{n}\beta_{n,\theta,\gamma}$ is the probability that exactly n of the agent's T signals with precision α_θ match the precision-γ realized outcome.) We can then write the principal's observable-signal problem as

$$\max_{x(\cdot)}\left\{\frac{1}{2}\sum_{n=0}^{T}\binom{T}{n}\Delta_{n,\gamma}x(n)\right\}.$$

It is trivial to see that the solution of this linear program depends entirely on the signs of the $\Delta_{n,\gamma}$ coefficients: we have $x^{FB}(n) = 1$ if $\Delta_{n,\gamma} > 0$, and $x^{FB}(n) = 0$ if $\Delta_{n,\gamma} < 0$.

CLAIM: *Suppose* $\Delta_{n,\gamma} \geq 0$. *Then* $\frac{\partial^2}{\partial n^2}\Delta_{n,\gamma} > 0$.

PROOF OF CLAIM:

Note first that

$$\frac{\partial}{\partial n}\Delta_{n,\gamma} = \left[\ln\left(\frac{\alpha}{1-\alpha}\right)\left(\gamma\alpha^n(1-\alpha)^{T-n} - (1-\gamma)\alpha^{T-n}(1-\alpha)^n\right)\right]_{\alpha_b}^{\alpha_g},$$

implying that

$$\begin{aligned}\frac{\partial^2}{\partial n^2}\Delta_{n,\gamma} &= \left[\ln^2\left(\frac{\alpha}{1-\alpha}\right)\left(\gamma\alpha^n(1-\alpha)^{T-n} + (1-\gamma)\alpha^{T-n}(1-\alpha)^n\right)\right]_{\alpha_b}^{\alpha_g} \\ &= \left[\ln^2\left(\frac{\alpha}{1-\alpha}\right)\right]_{\alpha_b}^{\alpha_g}\left(\gamma\alpha_g^n(1-\alpha_g)^{T-n} + (1-\gamma)\alpha_g^{T-n}(1-\alpha_g)^n\right) \\ &\quad + \ln^2\left(\frac{\alpha_b}{1-\alpha_b}\right)\left[\gamma\alpha^n(1-\alpha)^{T-n} + (1-\gamma)\alpha^{T-n}(1-\alpha)^n\right]_{\alpha_b}^{\alpha_g} \\ &= \left[\ln^2\left(\frac{\alpha}{1-\alpha}\right)\right]_{\alpha_b}^{\alpha_g}\beta_{n,g,\gamma} + \ln^2\left(\frac{\alpha_b}{1-\alpha_b}\right)\Delta_{n,\gamma}.\end{aligned}$$

Since $\ln\left(\frac{\alpha}{1-\alpha}\right)$ is strictly positive and increasing on $(1/2, 1)$ and $\beta_{n,g,\gamma} > 0$, the assumption that $\Delta_{n,\gamma} \geq 0$ implies that the expression above is strictly positive. ∎

Thus, $\Delta_{n,\gamma}$ is strictly convex on a neighborhood of any $m \in [0,T]$ at which $\Delta_{m,\gamma} \geq 0$. Therefore, if there exists some $\underline{n} \in [0,T]$ with $\Delta_{\underline{n},\gamma} = 0$ and $\frac{\partial}{\partial n}\Delta_{\underline{n},\gamma} \leq 0$, then $\frac{\partial}{\partial n}\Delta_{m,\gamma} < 0$ for all $m < \underline{n}$. This implies that $\Delta_{m,\gamma} > 0$ for all $m \in [0,\underline{n})$. Similarly, if there exists some $\overline{n} \in [0,T]$ such that $\Delta_{\overline{n},\gamma} = 0$ and $\frac{\partial}{\partial n}\Delta_{\overline{n},\gamma} \geq 0$, then $\frac{\partial}{\partial n}\Delta_{m,\gamma} > 0$ for all $m > \overline{n}$. This implies that $\Delta_{m,\gamma} > 0$ for all $m \in (\overline{n}, T]$.

Hence, we may conclude that the function $\Delta_{n,\gamma}$ has at most two zeros in $[0,T]$.

CLAIM: *There exists a unique* $\overline{n} \in \left(\frac{T}{2}, T\right)$ *such that* $\Delta_{\overline{n},\gamma} = 0$.

PROOF OF CLAIM:

Note first that $\Delta_{T,1} = \alpha_g^T - \alpha_b^T > 0$ since $\alpha_g > \alpha_b$. In addition, note that $\Delta_{T,\frac{1}{2}} = \frac{1}{2}\left[\alpha^T + (1-\alpha)^T\right]_{\alpha_b}^{\alpha_g}$. However,

$$\frac{\partial}{\partial\alpha}\left[\alpha^T + (1-\alpha)^T\right] = T\left[\alpha^{T-1} - (1-\alpha)^{T-1}\right] > 0 \quad \text{for all } \alpha > \frac{1}{2}.$$

Therefore, $\Delta_{T,\frac{1}{2}} > 0$. But since $\Delta_{T,\gamma}$ is linear in γ, this implies that $\Delta_{T,\gamma} > 0$ for all $\gamma \in [1/2, 1]$.

Next, consider

$$\Delta_{\frac{T}{2},\gamma} = \left[\gamma\alpha^{\frac{T}{2}}(1-\alpha)^{T-\frac{T}{2}} + (1-\gamma)\alpha^{T-\frac{T}{2}}(1-\alpha)^{\frac{T}{2}}\right]_{\alpha_b}^{\alpha_g} = \left[(\alpha(1-\alpha))^{\frac{T}{2}}\right]_{\alpha_b}^{\alpha_g}.$$

Since $\alpha(1-\alpha)$ is strictly decreasing on $(1/2, 1)$, we have $\Delta_{\frac{T}{2},\gamma} < 0$ for all $\gamma \in [1/2, 1]$.

Finally, because $\Delta_{n,\gamma}$ is continuous in n, there must exist some $\overline{n} \in (T/2, T)$ such that $\Delta_{\overline{n},\gamma} = 0$. Moreover, the convexity argument above implies that this $\overline{n}$ is the unique zero in $(T/2, T)$. ∎

The existence of a second zero is not guaranteed; in particular, there exists some $n \in [0, T/2)$ with $\Delta_{\underline{n},\gamma} = 0$ if and only if $\Delta_{0,\gamma} \geq 0$. (Note that $\underline{n} = 0$ in the boundary case where $\Delta_{0,\gamma} = 0$.) Again, the convexity argument above implies that this is the unique zero below $T/2$.

CLAIM: *Suppose there exists some* $\underline{n} < \frac{T}{2}$ *with* $\Delta_{\underline{n},\gamma} = 0$. *Then* $\underline{n} < T - \overline{n}$.

PROOF OF CLAIM:

We can write

$$\Delta_{\underline{n},\gamma} = \gamma\Delta_{\underline{n},1} + (1-\gamma)\Delta_{T-\underline{n},1}, \quad \text{where} \quad \Delta_{\underline{n},1} = \left[(\alpha^{\underline{n}}(1-\alpha))^{T-\underline{n}}\right]_{\alpha_b}^{\alpha_g}.$$

Note, however, that

$$\begin{aligned}\frac{\partial}{\partial\alpha}\left[\alpha^{\underline{n}}(1-\alpha)^{T-\underline{n}}\right] &= \underline{n}\alpha^{\underline{n}-1}(1-\alpha)^{T-\underline{n}} - (T-\underline{n})\alpha^{\underline{n}}(1-\alpha)^{T-\underline{n}-1} \\ &= (\underline{n} - \alpha T)\alpha^{\underline{n}-1}(1-\alpha)^{T-\underline{n}-1}.\end{aligned}$$

Since $\underline{n} < T/2$, this expression is strictly negative whenever $\alpha \in (1/2, 1)$; since $1 > \alpha_g > \alpha_b > 1/2$, this implies that $\Delta_{\underline{n},1} < 0$.

Thus, $\Delta_{T-\underline{n},1} > 0$ since $\Delta_{\underline{n},\gamma} = \gamma\Delta_{\underline{n},1} + (1-\gamma)\Delta_{T-\underline{n},1} = 0$. But since $\gamma \in [1/2, 1]$, this implies that $\Delta_{T-\underline{n},\gamma} = \gamma\Delta_{T-\underline{n},1} + (1-\gamma)\Delta_{\underline{n},1} > 0$, which is only possible if $T - \underline{n} > \overline{n}$. ∎

Thus, there exist $\overline{n} \in (T/2, T)$ and $\underline{n} < T - \overline{n}$ (where $\underline{n} < 0$ if $\Delta_{0,\gamma} < 0$) such that, for all $n \in [0, T]$,

$$\Delta_{n,\gamma} \begin{cases} > 0 & \text{if } n > \overline{n} \text{ or } n < \underline{n} \\ = 0 & \text{if } n = \overline{n} \text{ or } n = \underline{n}. \\ < 0 & \text{if } \overline{n} > n > \underline{n} \end{cases}$$

The first-best policy x^{FB} described in the theorem follows immediately. ∎

PROOF OF THEOREM 2:

When the principal uses a period-t prediction mechanism, her payoff is simply the difference in prediction-matching probabilities between the type-g and type-b agents. To that end, recall the following notation:

$$\beta_{m,n,\theta} := \gamma \alpha_\theta^m (1 - \alpha_\theta)^{n-m} + (1 - \gamma) \alpha_\theta^{n-m} (1 - \alpha_\theta)^m \quad \text{and} \quad \Delta_{m,n} := \beta_{m,n,g} - \beta_{m,n,b}.$$

Note that $\binom{n}{m} \beta_{m,n,\theta}$ is the probability that exactly m out of n signals with precision α_θ match the public precision-γ outcome. Thus, for any $k \geq 0$, we can write the principal's payoff from using a period-$2k$ or -$(2k+1)$ prediction mechanism as

$$\Pi(2k) := \frac{1}{2}\binom{2k}{k} \Delta_{k,2k} + \sum_{j=k+1}^{2k} \binom{2k}{j} \Delta_{j,2k} \quad \text{and} \quad \Pi(2k+1) := \sum_{j=k+1}^{2k+1} \binom{2k+1}{j} \Delta_{j,2k+1},$$

respectively. Finally, define $\delta(n) := \Pi(n) - \Pi(n-1)$. Note that since $\Pi(0) = 0$ and $\Pi(1) > 0$, we know that $\delta(1) > 0$.

CLAIM: *For any $k \geq 1$, both the principal and agent (of either type) are indifferent between the $(2k-1)$-period and $2k$-period prediction mechanisms.*

PROOF OF CLAIM:

In the $(2k-1)$-period prediction mechanism, a type-θ agent is hired if k or more signals match the outcome. Partitioning that event into the case where exactly k signals match and the case where at least $k+1$ signals match, we can write the probability of hiring a type-θ agent in the $(2k-1)$-period prediction mechanism as

$$\binom{2k-1}{k} \beta_{k,2k-1,\theta} + \sum_{j=k+1}^{2k-1} \binom{2k-1}{j} \beta_{j,2k-1,\theta}.$$

In the $2k$-period prediction mechanism, on the other hand, a type-θ agent is hired with probability $1/2$ if exactly k signals match the principal's, and with certainty if $k+1$ or more signals match. Focusing on the first $2k-1$ periods, this implies that three events may lead to the agent being hired:

- at least $k+1$ of the first $2k-1$ signals match the public outcome, in which case the agent is hired regardless of the realization of the $2k$th signal;

- exactly k of the first $2k-1$ signals match, in which case the agent is hired with probability 1 if the $2k$th signal matches, and with probability $1/2$ if it does not; and
- exactly $k-1$ of the first $2k-1$ signals match, in which case the agent is hired with probability $1/2$ if the $2k$th signal matches, and is not hired otherwise.

Therefore, the probability of hiring a type-θ agent in the $2k$-period prediction mechanism is

$$\sum_{j=k+1}^{2k-1}\binom{2k-1}{j}\beta_{j,2k-1,\theta}+\binom{2k-1}{k}\Big(\tfrac{1}{2}\beta_{k,2k,\theta}+\beta_{k+1,2k,\theta}\Big)+\binom{2k-1}{k-1}\Big(\tfrac{1}{2}\beta_{k,2k,\theta}\Big)$$

$$=\sum_{j=k+1}^{2k-1}\binom{2k-1}{j}\beta_{j,2k-1,\theta}+\binom{2k-1}{k}\big(\beta_{k,2k,\theta}+\beta_{k+1,2k,\theta}\big)$$

$$=\sum_{j=k+1}^{2k-1}\binom{2k-1}{j}\beta_{j,2k-1,\theta}+\binom{2k-1}{k}\beta_{k,2k-1,\theta},$$

where the first equality follows from the fact that $\binom{2k-1}{k}=\binom{2k-1}{k-1}$, and the second from the observation that $\beta_{k,\,2k,\,\theta}+\beta_{k+1,\,2k,\,\theta}=\beta_{k,\,2k-1,\,\theta}$.

Thus, a type-θ agent is hired with exactly the same probability in the $2k$- and $(2k-1)$-period prediction mechanisms, and so is indifferent between the two; this also implies that $\delta(2k)=0$. ∎

CLAIM: *For any* $k>0$, $\delta(2k+1)=\binom{2k}{k}\Big[\alpha^k(1-\alpha)^k\Big(\gamma\alpha+(1-\gamma)(1-\alpha)-\frac{1}{2}\Big)\Big]_{\alpha_b}^{\alpha_g}$.

PROOF OF CLAIM:

In the $(2k+1)$-period prediction mechanism, a type-θ agent is hired if $k+1$ or more signals match the public outcome. Focusing on the first $2k$ periods, this implies that two events may lead to the agent being hired:

- at least $k+1$ of the first $2k$ signals match the outcome, in which case the agent is hired regardless of the realization of the $(2k+1)$th signal; or
- exactly k of the first $2k$ signals match, in which case the agent is hired with probability 1 if the $(2k+1)$th signal matches, and is not hired otherwise.

Therefore, the probability of hiring a type-θ agent in the $(2k+1)$-period prediction mechanism is

$$\sum_{j=k+1}^{2k}\binom{2k}{j}\beta_{j,2k,\theta}+\binom{2k}{k}\beta_{k+1,2k+1,\theta}.$$

In the $2k$-period prediction mechanism, an agent is hired with probability $1/2$ if exactly k signals match the outcome, and with certainty if at least $k+1$ match, so the probability of hiring type θ is

$$\frac{1}{2}\binom{2k}{k}\beta_{k,2k,\theta} + \sum_{j=k+1}^{2k}\binom{2k}{j}\beta_{j,2k,\theta}.$$

Thus, the difference between these two probabilities is

$$\binom{2k}{k}\beta_{k+1,2k+1,\theta} - \frac{1}{2}\binom{2k}{k}\beta_{k,2k,\theta} = \binom{2k}{k}\alpha_\theta^k(1-\alpha_\theta)^k\Big(\gamma\alpha_\theta + (1-\gamma)(1-\alpha_\theta) - \frac{1}{2}\Big).$$

Since the principal's payoff is the difference between the type-g and type-b agents' payoffs, this yields the desired result. ∎

The result above therefore implies that $\delta(2k+1)$ is proportional to

$$\begin{aligned} z(k) &:= \Big[\alpha^k(1-\alpha)^k\Big(\gamma\alpha + (1-\gamma)(1-\alpha) - \frac{1}{2}\Big)\Big]_{\alpha_b}^{\alpha_g} \\ &= \Big[\frac{1}{2}\alpha^k(1-\alpha)^k(2\gamma-1)(2\alpha-1)\Big]_{\alpha_b}^{\alpha_g}. \end{aligned}$$

There is a unique k^* such that $z(k^*) = 0$; expanding the expression above and taking logs yields

$$k^* = \ln\Big(\frac{2\alpha_b - 1}{2\alpha_g - 1}\Big)\Big/\ln\Big(\frac{\alpha_g(1-\alpha_g)}{\alpha_b(1-\alpha_b)}\Big).$$

Furthermore, note that

$$z'(k) = \Big[\frac{1}{2}\alpha^k(1-\alpha)^k\ln(\alpha(1-\alpha))(2\gamma-1)(2\alpha-1)\Big]_{\alpha_b}^{\alpha_g}.$$

Since $\alpha_g > \alpha_b > 1/2$, we must have $\alpha_g(1-\alpha_g) < \alpha_b(1-\alpha_b)$, implying that $z'(k) < 0$. By continuity and the fact that $z(\cdot)$ has a unique root, we must have $z(k) > 0$ for all $k < k^*$ and $z(k) < 0$ for all $k > k^*$. Of course, this implies that $\delta(2k+1) > 0$ for all $k < k^*$ and $\delta(2k+1) < 0$ for all $k > k^*$. ∎

PROOF OF LEMMA 1:

Fix any incentive-compatible direct mechanism $\{\chi_h(\theta, s^T), \chi_l(\theta, s^T)\}$ with payoff Π, and define the alternative mechanism $\{x_h(s^T), x_l(s^T)\}$ by

$$x_r(s^T) := \chi_r(g, s^T) \quad \text{for all } r \in \{h,l\} \text{ and all } s^T \in \{h,l\}^T.$$

Denote by $\mu(\tilde{s}^T | s^T, \sigma)$ the probability that an agent who observes signals s^T and follows strategy $\sigma \in \Sigma$ reports the sequence $\tilde{s}^T$, where Σ is the set of all dynamic reporting strategies adapted to the signal process (as defined in Section IB). The principal's payoff from $\{x_h(\,\cdot\,), x_l(\,\cdot\,)\}$ is then

$$\Pi' = \frac{1}{2}\sup_{\sigma^g \in \Sigma}\left\{\sum_{(r,s^T)} \Pr(r, s^T \mid g) \sum_{\tilde{s}^T} \mu(\tilde{s}^T \mid s^T, \sigma^g)\, \chi_r(g, \tilde{s}^T)\right\}$$

$$- \frac{1}{2}\sup_{\sigma^b \in \Sigma}\left\{\sum_{(r,s^T)} \Pr(r, s^T \mid b) \sum_{\tilde{s}^T} \mu(\tilde{s}^T \mid s^T, \sigma^b)\, \chi_r(g, \tilde{s}^T)\right\}.$$

Note, however, that incentive compatibility of the original mechanism implies that the type-g agent finds truthful reporting of signals to be optimal, implying that

$$\Pi' = \frac{1}{2}\sum_{(r,s^T)} \Pr(r, s^T \mid g)\, \chi_r(g, s^T) - \frac{1}{2}\sup_{\sigma^b \in \Sigma}\left\{\sum_{(r,s^T)} \Pr(r, s^T \mid b) \sum_{\tilde{s}^T} \mu(\tilde{s}^T \mid s^T, \sigma^b)\, \chi_r(g, \tilde{s}^T)\right\}.$$

In addition, incentive compatibility of the original mechanism implies that forcing the type-b agent to misreport his initial type and then reoptimize reduces his expected utility; this implies that

$$\Pi' \geq \frac{1}{2}\sum_{(r,s^T)} \Pr(r, s^T \mid g)\, \chi_r(g, s^T) - \frac{1}{2}\sum_{(r,s^T)} \Pr(r, s^T \mid b)\, \chi_r(b, s^T) \;=:\; \Pi.$$

Thus, since the principal's objective is *decreasing* in the utility of the type-b agent, the new mechanism $\{x_h(\,\cdot\,), x_l(\,\cdot\,)\}$ improves the principal's payoff. As $\{\chi_h(\theta,\,\cdot\,), \chi_l(\theta,\,\cdot\,)\}$ was an arbitrary incentive-compatible mechanism, it is without loss to restrict attention to mechanisms that solicit only the agent's signals and in which the type-g agent is incentivized to report truthfully. ∎

PROOF OF LEMMA 2:

Trivial contracts are trivially incentive compatible: if the hiring decision does not depend on the agent's reports, then there is no incentive for the agent (of either type) to misreport any of his signals.

So fix any nontrivial deterministic and incentive-compatible contract $x_h, x_l : \{h, l\}^T \to \{0, 1\}$. Incentive compatibility and nontriviality of this contract imply that there is no sequence of signals $s^T \in \{h, l\}^T$ such that $x_h(s^T) = x_l(s^T) = 1$; if there were such a sequence, then the agent would always have an incentive to report it and guarantee his hiring (unless the contract were an "always hire" trivial contract). Similarly, there is no sequence $s^T \in \{h, l\}^T$ such that $x_h(s^T) = x_l(s^T) = 0$; if there were such a sequence, then agent would never be willing to report it truthfully (unless the contract were a "never hire" trivial contract).

Note that, by backward induction, there must be some latest period $T' \leq T$ and history of reports $\hat{s}^{T'-1} \in \{h, l\}^{T'-1}$ such that the agent's period-T' report is pivotal; that is,

$$(x_h(\hat{s}^{T'-1}, h,\,\cdot\,), x_l(\hat{s}^{T'-1}, h,\cdot\,)) \neq (x_h(\hat{s}^{T'-1}, l,\,\cdot\,), x_l(\hat{s}^{T'-1}, l,\,\cdot\,)).$$

To see this, start in the period T. If there is no such $\hat{s}^{T-1}$, then the final-period report *never* affects the principal's hiring decision, which must then depend only on the reports from the first $T-1$ periods. Proceeding in this manner yields T' and an $\hat{s}^{T'-1}$. (Note that $T' \geq 1$ since otherwise the contract does not depend on the agent's report, contradicting the assumption that it is nontrivial.) Since periods $T'+1$ through T do not affect the hiring decision, we can without loss overload notation and describe the contract as two functions $x_h, x_l : \{h,l\}^{T'} \to \{0,1\}$.

As argued above, nontriviality and incentive compatibility imply that

$$(x_h(\hat{s}^{T'-1},h), x_l(\hat{s}^{T'-1},h)), (x_h(\hat{s}^{T'-1},l), x_l(\hat{s}^{T'-1},l)) \in \{(1,0),(0,1)\}.$$

Since, by construction, we know that $(x_h(\hat{s}^{T'-1},h), x_l(\hat{s}^{T'-1},h)) \neq (x_h(\hat{s}^{T'-1},l), x_l(\hat{s}^{T'-1},l))$, it must then be the case that

$$(x_h(\hat{s}^{T'-1},h), x_l(\hat{s}^{T'-1},h)) = (1,0) \quad \text{and} \quad (x_h(\hat{s}^{T'-1},l), x_l(\hat{s}^{T'-1},l)) = (0,1).$$

This follows from incentive compatibility, and the fact that the agent's posterior beliefs are such that $\Pr(r = h \mid \hat{s}^{T'-1}, h) > \Pr(r = h \mid \hat{s}^{T'-1}, l)$. Further, these beliefs must be such that

$$\Pr(r = h \mid \hat{s}^{T'-1}, h) \geq \frac{1}{2} \quad \text{and} \quad \Pr(r = h \mid \hat{s}^{T'-1}, l) \leq \frac{1}{2},$$

as otherwise the pivotality of the period-T' report following history $\hat{s}^{T'-1}$ would lead to a violation of incentive compatibility.

Now consider any other history $\tilde{s}^{T'} \in \{h,l\}^{T'}$. Nontriviality and incentive compatibility again imply that $(x_h(\tilde{s}^{T'}), x_l(\tilde{s}^{T'})) \in \{(1,0),(0,1)\}$. We claim that we must have $(x_h(\tilde{s}^{T'}), x_l(\tilde{s}^{T'})) = (1,0)$ whenever $\Pr(r = h \mid \tilde{s}^{T'}) > 1/2$ and $(x_h(\tilde{s}^{T'}), x_l(\tilde{s}^{T'})) = (0,1)$ whenever $\Pr(r = h \mid \tilde{s}^{T'}) < 1/2$. To see why this must be true, suppose the contrary and note that this must yield a violation of incentive compatibility. In particular, consider the alternative agent strategy of always reporting $\hat{s}^{T'-1}$ in the first $T'-1$ periods regardless of his true signals, and then choosing a period-T' report that matches his posterior belief; that is, he reports h if $\Pr(r = h \mid \tilde{s}^{T'}) > 1/2$, l if $\Pr(r = h \mid \tilde{s}^{T'}) < 1/2$, and chooses arbitrarily if $\Pr(r = h \mid \tilde{s}^{T'}) = 1/2$. Such a strategy increases the agent's payoff over truthful reporting as it guarantees that the agent is hired precisely at the outcome he thinks more likely (whereas truthful reporting may lead to being hired only in the less likely outcome).

Finally, note that $\Pr(r = h \mid \tilde{s}^{T'}) > 1/2$ if and only if $\sum_{\tau \leq T'} \mathbf{1}_h(s_\tau) > T'/2$. Thus, the (arbitrarily-chosen) nontrivial deterministic and incentive-compatible contract x_h, x_l is equivalent to a period-T' prediction mechanism. Therefore, *any* deterministic nontrivial and incentive-compatible contract is a period-t prediction mechanism for some $1 \leq t \leq T$. ∎

PROOF OF THEOREM 3:

Recall that Lemma 1 establishes that it is without loss to consider only mechanisms that induce the type-g agent to report her signals truthfully (and allowing the type-b agent to optimally misreport). Therefore, greatly simplifies the class of mechanisms over which the principal must optimize. In particular, the principal must either abandon screening entirely (that is, employ a trivial mechanism, which yields a payoff of zero) or employ a period-t prediction mechanism for some $1 \leq t \leq T$.

Of course, Theorem 2 showed that the principal's payoff, within this class of mechanisms, is increasing in t until reaching a maximum at some $\hat{T}$. Therefore, the optimal deterministic mechanism is a period-T^* prediction mechanism, where $T^* := \min\{T, \hat{T}\}$. ∎

PROOF OF THEOREM 4:

The result follows immediately from the argument in the main text. ∎

PROOF OF LEMMA 3:

Recall from the proof of Theorem 1 that we can write the principal's problem when the agent's signals are observable as

$$\max_{x(\cdot)} \left\{ \frac{1}{2} \sum_{k=0}^{T} \binom{T}{k} \Delta_{k,T} x(k) \right\},$$

where $x(k)$ denote the principal's hiring decision when she observes k signals that match the public outcome, and where

$$\Delta_{k,T} := \left[\gamma \alpha^k (1-\alpha)^{T-k} + (1-\gamma)\, \alpha^{T-k} (1-\alpha)^k\right]_{\alpha_b}^{\alpha_g}.$$

The solution to this linear program depends entirely on the signs of $\Delta_{k,T}$. We now focus on signing these terms when $T = 3$:

- $\Delta_{0,3}|_{\gamma=1} = \left[(1-\alpha)^3\right]_{\alpha_b}^{\alpha_g} < 0$ and $\Delta_{0,3}|_{\gamma=\frac{1}{2}} = \left[\frac{1}{2}(\alpha^3 + (1-\alpha)^3)\right]_{\alpha_b}^{\alpha_g} > 0$;
- $\Delta_{1,3}|_{\gamma=1} = \left[\alpha(1-\alpha)^2\right]_{\alpha_b}^{\alpha_g} < 0$ and $\Delta_{1,3}|_{\gamma=\frac{1}{2}} = \left[\frac{1}{2}\alpha(1-\alpha)\right]_{\alpha_b}^{\alpha_g} < 0$;
- $\Delta_{2,3}|_{\gamma=1} = \left[\alpha^2(1-\alpha)\right]_{\alpha_b}^{\alpha_g}$ is ambiguously signed (it may be positive or negative), while $\Delta_{2,3}|_{\gamma=\frac{1}{2}} = \left[\frac{1}{2}\alpha(1-\alpha)\right]_{\alpha_b}^{\alpha_g} < 0$; and
- $\Delta_{3,3}|_{\gamma=1} = \left[\alpha^3\right]_{\alpha_b}^{\alpha_g} > 0$ and $\Delta_{3,3}|_{\gamma=\frac{1}{2}} = \left[\frac{1}{2}(\alpha^3 + (1-\alpha)^3)\right]_{\alpha_b}^{\alpha_g} > 0$.

Since $\Delta_{k,T}$ is linear in γ, we can unambiguously sign $\Delta_{1,3} < 0$ and $\Delta_{3,3} > 0$; therefore, we must have $x^{FB}(3) = 1$ and $x^{FB}(1) = 0$; the principal always hires the agent when all three of his signals match the realized outcome, and never hires the agent when only one of his signals matches the realized outcome.

By the same logic, it is *not* possible to unambiguously sign $\Delta_{0,3}$ and $\Delta_{2,3}$; however, we can characterize the solution x^{FB} for the various feasible sign combinations:

- If $\Delta_{0,3} < 0$ and $\Delta_{2,3} < 0$, then the solution must be such that $x^{FB}(0) = x^{FB}(1) = x^{FB}(2) = 0$ and $x^{FB}(3) = 1$; that is, the principal hires the agent if and only if all three of her signals are accurate (so $\overline{n} = 3$ and $\underline{n} = -1$).

- If $\Delta_{0,3} \geq 0$ and $\Delta_{2,3} < 0$, then the solution must be such that $x^{FB}(0) = x^{FB}(3) = 1$ and $x^{FB}(1) = x^{FB}(2) = 0$; that is, the principal hires the agent if and only if all three of her signals are consistent (so $\overline{n} = 3$ and $\underline{n} = 0$).
- If $\Delta_{0,3} < 0$ and $\Delta_{2,3} \geq 0$, then the solution must be such that $x^{FB}(0) = x^{FB}(1) = 0$ and $x^{FB}(2) = x^{FB}(3) = 1$; that is, the principal hires the agent if and only if a majority (at least two out of three) of her signals are accurate (so $\overline{n} = 2$ and $\underline{n} = -1$).

Note that the fourth possible sign combination (both $\Delta_{0,3} \geq 0$ and $\Delta_{2,3} \geq 0$) is not feasible. To see why, suppose that α_g and α_b are such that $\Delta_{2,3}|_{\gamma=1} \geq 0$ (otherwise, $\Delta_{2,3} < 0$ for all γ and we are done). Thus, as γ goes from $1/2$ to 1, $\Delta_{0,3}$ crosses from positive to negative while $\Delta_{2,3}$ goes from negative to positive. Let γ^* be such that $\Delta_{0,3}|_{\gamma=\gamma^*} = 0$; that is,

$$\gamma^* = \frac{\left[\alpha^3\right]_{\alpha_b}^{\alpha_g}}{\left[(\alpha^3 - (1-\alpha)^3)\right]_{\alpha_b}^{\alpha_g}}.$$

Then

$$\Delta_{2,3}|_{\gamma=\gamma^*} = \gamma^*\left[\alpha^2(1-\alpha)\right]_{\alpha_b}^{\alpha_g} + (1-\gamma^*)\left[\alpha(1-\alpha)^2\right]_{\alpha_b}^{\alpha_g}$$

$$= -\frac{(\alpha_g - \alpha_b)(\alpha_g + \alpha_b - 1)[(2\alpha_g - 1)^2 + (2\alpha_b - 1)^2 + (2\alpha_g - 1)(2\alpha_b - 1)]}{(\alpha_g + \alpha_b - 1)^2 + (1-\alpha_g)^2 + (1-\alpha_b)^2 + \alpha_g + \alpha_b} < 0,$$

where the inequality follows from the fact that $1 > \alpha_g > \alpha_b > 1/2$. Therefore, whenever $\Delta_{0,3} \geq 0$ (that is, whenever $\gamma \leq \gamma^*$), we must have $\Delta_{2,3} < 0$.

Thus, the first-best mechanism when $T = 3$ takes on one of the three desired forms. ∎

PROOF OF THEOREM 5:

We begin by recalling that Lemma 1 shows that it is without loss of generality for the principal to offer a contract of the form $x_r : \{h, l\}^T \to [0,1]$, $r \in \{h, l\}$, such that the type-g agent is incentivized to report her signals truthfully while the type-b agent is free to misreport optimally. Therefore, letting Σ denote the set of all dynamic reporting strategies that are adapted to the signal process and $\mu(\tilde{s}^T | s^T, \sigma)$ the probability that an agent who observes signals s^T and follows strategy $\sigma \in \Sigma$ reports the sequence $\tilde{s}^T$, we can write the principal's problem as

$$(\mathcal{P}) \quad \max_{x_h, x_l}\left\{\frac{1}{2}\sum_{(r,s^T)} \Pr(r, s^T | \theta = g)\, x_r(s^T) - \frac{1}{2}\sup_{\sigma^b \in \Sigma}\left\{\sum_{(r,s^T)} \Pr(r, s^T | \theta = b) \sum_{\tilde{s}^T} \mu(\tilde{s}^T | s^T, \sigma^b)\, x_r(\tilde{s}^T)\right\}\right\}$$

$$\text{subject to} \sum_{(r,s^T)} \Pr(r, s^T | \theta = g)\, x_r(s^T) \geq \sum_{(r,s^T)} \Pr(r, s^T | \theta = g) \sum_{\tilde{s}^T} \mu(\tilde{s}^T | s^T, \sigma')\, x_r(\tilde{s}^T) \quad \text{for all } \sigma' \in \Sigma.$$

Note that the constraint is simply the type-g agent's incentive-compatibility condition, whereas the type-b agent's optimal reporting strategy has been incorporated into the objective function.

We will proceed to the solution of problem $(\mathcal{P})$ as follows:

- We define a relaxed problem with a restricted set of strategies available to the type-b agent.
- We will then argue that the solution to this relaxed problem features truthful reporting at certain histories by the type-b agent.
- We then incorporate the corresponding incentive-compatibility constraints into a further relaxation of the problem, which we then solve.
- Finally, we demonstrate that our proposed solution is indeed feasible in the original problem, in the sense that the strategy we impose on the type-b agent's behavior in the relaxed problem is optimal given the identified solution.

We begin by restricting the set of possible misreports of the type-b agent. Denote by $\hat{\Sigma} \subset \Sigma$ the set of strategies where, for all $s^3 \in \{h, l\}^3$ and any $s_2', s_3' \in \{h, l\}$,

$$\mu(\tilde{s}^3 \mid s^3, \sigma) > 0 \quad \text{if and only if} \quad \begin{cases} \tilde{s}^T = (s_1, s_2, s_3) \text{ and } s_1 = s_2 = s_3, \\ \tilde{s}^T = (s_1, s_2, s_3') \text{ and } s_1 = s_2 \neq s_3, \text{ or} \\ \tilde{s}^T = (s_1, s_2', s_3) \text{ and } s_1 \neq s_2. \end{cases}$$

Thus, any strategy $\sigma \in \hat{\Sigma}$ reports truthfully at all histories except possibly those where the agent first observes a contradictory signal. With this in hand, define the relaxed problem

$$(\mathcal{R}) \quad \max_{x_h, x_l} \left\{ \frac{1}{2} \sum_{(r, s^T)} \Pr(r, s^T \mid \theta = g) x_r(s^T) - \frac{1}{2} \sup_{\sigma^b \in \hat{\Sigma}} \left\{ \sum_{(r, s^T)} \Pr(r, s^T \mid \theta = b) \sum_{\tilde{s}^T} \mu(\tilde{s}^T \mid s^T, \sigma^b) x_r(\tilde{s}^T) \right\} \right\}$$

$$\text{subject to} \sum_{(r, s^T)} \Pr(r, s^T \mid \theta = g) x_r(s^T) \geq \sum_{(r, s^T)} \Pr(r, s^T \mid \theta = g) \sum_{\tilde{s}^T} \mu(\tilde{s}^T \mid s^T, \sigma') x_r(\tilde{s}^T) \quad \text{for all } \sigma' \in \Sigma.$$

CLAIM: *The solution to the relaxed problem* $(\mathcal{R})$ *yields the principal a higher payoff than the original problem* $(\mathcal{P})$.

PROOF OF CLAIM:

Consider any solution x_r^* to problem $(\mathcal{P})$. Since $\hat{\Sigma} \subset \Sigma$, we must have

$$\sup_{\sigma^b \in \hat{\Sigma}} \left\{ \sum_{(r, s^T)} \Pr(r, s^T \mid \theta = b) \sum_{\tilde{s}^T} \mu(\tilde{s}^T \mid s^T, \sigma^b) x_r^*(\tilde{s}^T) \right\}$$
$$\leq \sup_{\sigma^b \in \Sigma} \left\{ \sum_{(r, s^T)} \Pr(r, s^T \mid \theta = b) \sum_{\tilde{s}^T} \mu(\tilde{s}^T \mid s^T, \sigma^b) x_r^*(\tilde{s}^T) \right\},$$

which implies that the maximal payoff from $(\mathcal{P})$ is achievable in $(\mathcal{R})$. ■

Now further relax the problem by dropping the incentive-compatibility constraints for the type-g agent; that is, consider the problem

$$(\mathcal{R}') \quad \max_{x_h, x_l} \left\{ \frac{1}{2} \sum_{(r, s^T)} \Pr(r, s^T \mid \theta = g) x_r(s^T) - \frac{1}{2} \sup_{\sigma^b \in \hat{\Sigma}} \left\{ \sum_{(r, s^T)} \Pr(r, s^T \mid \theta = b) \sum_{\tilde{s}^T} \mu(\tilde{s}^T \mid s^T, \sigma^b) x_r(\tilde{s}^T) \right\} \right\}$$

and note that (since it is less constrained) the solution to $(\mathcal{R}')$ yields the principal a higher payoff than that of $(\mathcal{R})$.

CLAIM: *There is a solution to* $(\mathcal{R}')$ *such that the type-b agent reports his signals truthfully at all histories.*

PROOF OF CLAIM:

Suppose, by way of contradiction, that there is a solution x_r^* to $(\mathcal{R}')$ in which the type-b agent who has observed $s^2 = (i,j)$ strictly prefers to misreport $s_2 = j$ as $\tilde{s}_2 = i$ for some $i,j \in \{h,l\}$ with $i \neq j$.

Since the preference is strict, it must be the case that the expected probability of being hired after reporting one of the sequences (i,j,i) or (i,j,j) is strictly less than 1 (otherwise, the agent would optimally report the second signal j truthfully). This implies, however, that only the type-g agent (who always reports truthfully in $(\mathcal{R}')$) ever reports sequences (i,j,i) and (i,j,j). Therefore, the alternative hiring rule x_r^{**} defined by

$$x_r^{**}(\hat{s}^T) \coloneqq \min\left\{x_r^*(\hat{s}^T) + \varepsilon \mathbf{1}_{\{(i,j,i),(i,j,j)\}}(\hat{s}^T),\, 1\right\}$$

for sufficiently small $\varepsilon > 0$ strictly increases the probability that the principal hires the type-g agent without influencing the strategy of the type-b agent. This, of course, increases the principal's payoff, contradicting the assumption that x_r^* solves $(\mathcal{R}')$.

An identical argument applies when $s^3 = (i,i,j)$. (Note that this argument can be applied separately across these two types of sequences since compound misreports are ruled out in $\hat{\Sigma}$.) ∎

This argument implies that, instead of incorporating the type-b agent's problem into the objective function as in $(\mathcal{R}')$, we can instead incorporate the *solution* (truthful reporting) to that problem while also imposing the requisite incentive-compatibility constraints. Thus, $(\mathcal{R}')$ is equivalent to

$$(\mathcal{R}'') \quad \max_{x_h, x_l} \left\{\frac{1}{2}\sum_{(r,s^T)} \Pr(r, s^T \mid \theta = g)\, x_r(s^T) - \frac{1}{2}\sum_{(r,s^T)} \Pr(r, s^T \mid \theta = b)\, x_r(s^T)\right\}$$

$$\text{subject to} \sum_{(r,s^T)} \Pr(r, s^T \mid \theta = b)\, x_r(s^T) \geq \sum_{(r,s^T)} \Pr(r, s^T \mid \theta = b) \sum_{\tilde{s}^T} \mu(\tilde{s}^T \mid s^T, \sigma')\, x_r(\tilde{s}^T) \quad \text{for all } \sigma' \in \hat{\Sigma}.$$

Since this relaxed problem is separable in histories conditioned on the agent's first signal (as we have assumed truthful reporting of the first signal), we can solve the problem separately for each of the two cases $s_1 \in \{h,l\}$. Formally, when the first signal is $s_1 = h$, we can write $(\mathcal{R}'')$ as

$$(\mathcal{R}_h'') \quad \max_{x_r} \big\{\Delta_{3,3} x_h(h,h,h) + \Delta_{2,3}[x_h(h,h,l) + x_h(h,l,h) + x_l(h,l,l)]$$
$$+ \Delta_{1,3}[x_h(h,l,l) + x_l(h,h,l) + x_l(h,l,h)] + \Delta_{0,3} x_l(h,h,h)\big\}$$

subject to

$$\beta_{2,3,b}x_h(h,l,h) + \beta_{1,3,b}x_l(h,l,h) + \beta_{1,3,b}x_h(h,l,l) + \beta_{2,3,b}x_l(h,l,l)$$
$$\geq \beta_{2,3,b}x_h(h,h,h) + \beta_{1,3,b}x_l(h,h,h) + \beta_{1,3,b}x_h(h,h,l) + \beta_{2,3,b}x_l(h,h,l),$$

$$\beta_{2,3,b}x_h(h,h,l) + \beta_{1,3,b}x_l(h,h,l) \geq \beta_{2,3,b}x_h(h,h,h) + \beta_{1,3,b}x_l(h,h,h).$$

CLAIM: *Suppose* $\Delta_{2,3} \geq 0$*. Then the solution to* $(\mathcal{R}_h'')$ *is given by*

$$x_h(h,h,h) = x_h(h,h,l) = 1, \quad x_l(h,h,h) = x_l(h,h,l) = 0, \tag{A4}$$
$$x_h(h,l,h) = x_l(h,l,l) = 1, \quad x_l(h,l,h) = x_h(h,l,l) = 0.$$

PROOF OF CLAIM:

We will proceed by showing that there exist multipliers λ and μ corresponding to the two incentive-compatibility constraints in $(\mathcal{R}_h'')$ such that the conjectured solution (A4) satisfies the Karush-Kuhn-Tucker conditions. These conditions may be written as

$$x_h(h,h,h): \quad \Delta_{3,3} - \lambda\beta_{2,3,b} - \mu\beta_{2,3,b} \geq 0, \tag{A5}$$

$$x_h(h,h,l): \quad \Delta_{2,3} - \lambda\beta_{1,3,b} + \mu\beta_{2,3,b} \geq 0, \tag{A6}$$

$$x_l(h,h,h): \quad \Delta_{0,3} - \lambda\beta_{1,3,b} - \mu\beta_{1,3,b} \leq 0, \tag{A7}$$

$$x_l(h,h,l): \quad \Delta_{1,3} - \lambda\beta_{2,3,b} + \mu\beta_{1,3,b} \leq 0, \tag{A8}$$

$$x_h(h,l,h), x_l(h,l,l): \quad \Delta_{2,3} + \lambda\beta_{2,3,b} \geq 0, \tag{A9}$$

$$x_l(h,l,h), x_h(h,l,l): \quad \Delta_{1,3} + \lambda\beta_{2,1,b} \leq 0. \tag{A10}$$

The directions of the inequalities above are determined by the feasibility constraint that each variable $x_r(\,\cdot\,)$ lies between 0 and 1.

Note that, at the conjectured solution, the first constraint (corresponding to period-2 incentive compatibility) reduces to

$$\beta_{2,3,b} \geq \beta_{1,3,b}.$$

Of course, this inequality holds strictly, and so the constraint is slack. Therefore, we must have

$$\lambda = 0.$$

In addition, recall (from the proof of Lemma 3), that $\Delta_{3,3} > 0 > \Delta_{1,3}$ and that $\Delta_{0,3} < 0$ whenever $\Delta_{2,3} \geq 0$ (as was assumed). Therefore, it is easy to see that choosing

$$\mu = 0$$

leads to the satisfaction of all the KKT conditions above: which are, of course, both necessary and sufficient for the linear program $(\mathcal{R}_h'')$. ■

CLAIM: *Suppose $\Delta_{2,3} < 0$. Then the solution to $(\mathcal{R}_h'')$ is given by*

$$\text{(A11)} \qquad x_h(h,h,h) = x_h(h,h,l) = 1, \quad x_l(h,h,h) = x_l(h,h,l) = 0,$$

$$x_h(h,l,h) = x_l(h,l,l) = \frac{\beta_{1,3,b} + \beta_{2,3,b}}{2\beta_{2,3,b}}, \quad x_l(h,l,h) = x_h(h,l,l) = 0.$$

PROOF OF CLAIM:

We will proceed by showing that there exist multipliers λ and μ corresponding to the two incentive-compatibility constraints in $(\mathcal{R}_h'')$ such that the conjectured solution (A11) satisfies the Karush-Kuhn-Tucker conditions. These conditions may be written as

$$\text{(A12)} \qquad x_h(h,h,h): \qquad \Delta_{3,3} - \lambda\beta_{2,3,b} - \mu\beta_{2,3,b} \geq 0,$$

$$\text{(A13)} \qquad x_h(h,h,l): \qquad \Delta_{2,3} - \lambda\beta_{1,3,b} + \mu\beta_{2,3,b} \geq 0,$$

$$\text{(A14)} \qquad x_l(h,h,h): \qquad \Delta_{0,3} - \lambda\beta_{1,3,b} - \mu\beta_{1,3,b} \leq 0,$$

$$\text{(A15)} \qquad x_l(h,h,l): \qquad \Delta_{1,3} - \lambda\beta_{2,3,b} + \mu\beta_{1,3,b} \leq 0,$$

$$\text{(A16)} \qquad x_h(h,l,h), x_l(h,l,l): \qquad \Delta_{2,3} + \lambda\beta_{2,3,b} = 0,$$

$$\text{(A17)} \qquad x_l(h,l,h), x_h(h,l,l): \qquad \Delta_{1,3} + \lambda\beta_{1,3,b} \leq 0.$$

The directions of the inequalities above are determined by the feasibility constraint that each variable $x_r(\cdot)$ lies between 0 and 1.

Note first that (A16) implies that (since $\Delta_{2,3} < 0$) we must have

$$\lambda = -\frac{\Delta_{2,3}}{\beta_{2,3,b}} > 0.$$

Substituting this value into (A17) yields

$$\beta_{2,3,b}\Delta_{1,3} \leq \beta_{1,3,b}\Delta_{2,3},$$

which is easily verified to hold. In addition, we can rewrite (A12) and (A15) as

$$\mu \leq \frac{\Delta_{3,3} + \Delta_{2,3}}{\beta_{2,3,b}} \quad \text{and} \quad \mu \leq -\frac{\Delta_{2,3} + \Delta_{1,3}}{\beta_{1,3,b}},$$

respectively. Clearly, choosing

$$\mu = \min\left\{\frac{\Delta_{3,3} + \Delta_{2,3}}{\beta_{2,3,b}}, -\frac{\Delta_{2,3} + \Delta_{1,3}}{\beta_{1,3,b}}\right\}$$

satisfies both of these conditions. It remains to be shown that this choice of μ satisfies (A13) and (A14).

So suppose first that $\mu = \frac{\Delta_{3,3} + \Delta_{2,3}}{\beta_{2,3,b}} \leq -\frac{\Delta_{2,3} + \Delta_{1,3}}{\beta_{1,3,b}}$. Then we can rewrite (A13) as

$$0 \leq \Delta_{2,3} + \frac{\Delta_{2,3}}{\beta_{2,3,b}}\beta_{1,3,b} + \frac{\Delta_{3,3} + \Delta_{2,3}}{\beta_{2,3,b}}\beta_{2,3,b} = \frac{(\alpha_g - \alpha_b)(2\gamma - 1)(1 - \gamma + (2\gamma - 1)\alpha_b + \alpha_g(1 - \alpha_g))}{1 - \gamma + (2\gamma - 1)\alpha_b},$$

and likewise rewrite (A14) as

$$0 \geq \Delta_{0,3} + \frac{\Delta_{2,3}}{\beta_{2,3,b}}\beta_{1,3,b} + \frac{\Delta_{3,3} + \Delta_{2,3}}{\beta_{2,3,b}}\beta_{1,3,b} = -\frac{(\alpha_g - \alpha_b)(2\gamma - 1)((\alpha_g + \alpha_b)^2 + 3\alpha_b(1 - \alpha_b))}{1 - \gamma + (2\gamma - 1)\alpha_b}.$$

It is straightforward to see that both of these inequalities hold as $1 \geq \alpha_g > \alpha_b \geq 1/2$ and $1 > \gamma > 1/2$.

On the other hand, suppose that $\mu = -\frac{\Delta_{2,3} + \Delta_{1,3}}{\beta_{1,3,b}} \leq \frac{\Delta_{3,3} + \Delta_{2,3}}{\beta_{2,3,b}}$. Then we can rewrite (A13) as

$$-\frac{\Delta_{2,3} + \frac{\beta_{1,3,b}}{\beta_{2,3,b}}\Delta_{2,3}}{\beta_{2,3,b}} \leq -\frac{\Delta_{2,3} + \Delta_{1,3}}{\beta_{1,3,b}}.$$

Note, however, that (A17) implies that $-\frac{\beta_{1,3,b}}{\beta_{2,3,b}}\Delta_{2,3} \leq -\Delta_{1,3}$. Therefore, since (as is simple to verify) $\beta_{1,3,b} < \beta_{2,3,b}$, this inequality is satisfied. Finally, we can write (A14) as

$$0 \geq \Delta_{0,3} + \frac{\Delta_{2,3}}{\beta_{2,3,b}}\beta_{1,3,b} + \frac{\Delta_{2,3} + \Delta_{1,3}}{\beta_{1,3,b}}\beta_{1,3,b} = \frac{(\alpha_b - \alpha_g)(2\gamma - 1)((1 - \alpha_g)^2 + (\alpha_g + \alpha_b) + \gamma(2\alpha_b - 1))}{1 - \gamma + (2\gamma - 1)\alpha_b}.$$

Again, the inequality is satisfied since $1 > \alpha_g > \alpha_b > 1/2$ and $1 > \gamma > 1/2$.

Thus, the conjectured solution, along with λ and μ as defined above, satisfy the KKT conditions. Of course, these conditions are both necessary and sufficient for the linear program $(\mathcal{R}_h'')$. ■

Finally, it remains to be shown that the conjectured solutions to $(\mathcal{R}_h'')$ above solve the unrelaxed problem $(\mathcal{P})$. Note that the original problem $(\mathcal{P})$ imposes incentive-compatibility constraints on the type-g agent while the relaxed problem assumed truthful reporting; likewise, the original problem allowed the type-b agent to optimally misreport while the relaxed problem imposed incentive-compatibility constraints on two histories and assumed truthful reporting at the others. Therefore, it suffices to show that the conjectured behavior in the relaxed problem is indeed optimal in the unrelaxed one.

CLAIM: *Suppose the principal chooses either of the mechanisms described in* (A4) *or* (A11). *Then it is optimal for the agent to always report her private signals truthfully.*

PROOF OF CLAIM:

We begin by noting that the solution in (A4) corresponds to a period-3 prediction mechanism, as it deterministically hires the agent if a majority of his reported signals match the eventual outcome. Lemma 2 then immediately implies that this mechanism induces truthful reporting for both the type-g and type-b agents.

We now turn to the solution in (A11), which can be implemented by offering the agent the option in period 2 to either make a prediction immediately (and be hired, if correct, with probability 1) or to make a prediction in period 3 (and be hired, if correct, with probability $\rho := \frac{\beta_{1,3,b} + \beta_{2,3,b}}{2\beta_{2,3,b}} < 1$). Note that there is an onto mapping from the set of signal-reporting strategies to the set of prediction strategies in this option implementation. In particular, truthful reporting of signals in (A11) corresponds to making a sincere prediction in period 2 if both signals match, and otherwise making a sincere prediction in period 3. Hence, showing that this conjectured behavior is optimal for the agent is sufficient for showing the optimality of truthful signal reporting in (A11).

To see why this behavior is optimal for the agent, note first that observing two matching signals in periods 1 and 2 yields the agent enough information to make a prediction in period 3: regardless of whether the third signal matches or not, he will make the same prediction. Since $\rho < 1$, a period-2 prediction yields the agent a strictly higher payoff than postponing. On the other hand, suppose that the agent has observed a pair of mismatched signals in the first two periods, leaving him with a uniform posterior over states. This implies that an early prediction (of either h or l) yields the type-θ agent an expected payoff of

$$\frac{1}{2}\gamma + \frac{1}{2}(1-\gamma) = \frac{1}{2}.$$

Postponing the prediction to period 3 (and then making a sincere prediction that follows the third private signal) yields the type-θ agent an expected payoff of

$$(\gamma\alpha_\theta + (1-\gamma)(1-\alpha_\theta))\rho = \frac{1}{2}\left(\frac{\gamma\alpha_\theta + (1-\gamma)(1-\alpha_\theta)}{\gamma\alpha_b + (1-\gamma)(1-\alpha_b)}\right).$$

Clearly, a type-b agent with mixed signals in period 2 is indifferent about delay, whereas a type-g agent with mixed signals in period 2 strictly prefers to delay his prediction since $\alpha_g > \alpha_b$. This implies that the mechanism in (A11) is incentive compatible for both types of the agent.[20] ∎

Thus, the assumed behavior for the agent in the relaxed problem $(\mathcal{R}'')$ is in fact a best response to the principal's proposed mechanism. This implies that the conjectured solution indeed solves the original problem $(\mathcal{P})$. ∎

PROOF OF THEOREM 6:

Recall from the proof of Theorem 2 that both the principal and agent (of either type) are indifferent between the $(2k-1)$-period and $2k$-period prediction mechanisms; therefore, assume without loss that $\hat{T}$ is odd, and let $\overline{k}$ be such that $\hat{T} = 2\overline{k}+1$ (and therefore, since $T > \hat{T}+1$, we have $T \geq 2\overline{k}+3$).[21]

CLAIM: $\Delta_{\overline{k}+2,\,2\overline{k}+3} < 0$.

PROOF OF CLAIM:

Recall from the proof of Theorem 2 that we defined $\delta(n)$ to be the difference between the principal's expected payoff from an n-period and an $(n-1)$-period prediction mechanism. Since $\hat{T}$ is the optimal length for a prediction mechanism, Theorem 2 implies that $0 > \delta(\hat{T}+2) = \delta(2\overline{k}+3) = \delta(2(\overline{k}+1)+1)$.

However, the second claim in that proof showed that

$$\begin{aligned}
\delta(2(\overline{k}+1)+1) &= \binom{2(\overline{k}+1)}{\overline{k}+1}\Big[\alpha^{\overline{k}+1}(1-\alpha)^{\overline{k}+1}\Big(\gamma\alpha + (1-\gamma)(1-\alpha) - \frac{1}{2}\Big)\Big]_{\alpha_b}^{\alpha_g} \\
&= \binom{2(\overline{k}+1)}{\overline{k}+1}\Big[\gamma\alpha^{\overline{k}+2}(1-\alpha)^{\overline{k}+1} + (1-\gamma)\alpha^{\overline{k}+1}(1-\alpha)^{\overline{k}+2}\Big]_{\alpha_b}^{\alpha_g} \\
&\quad - \binom{2(\overline{k}+1)}{\overline{k}+1}\frac{1}{2}\Big[\gamma\alpha^{\overline{k}+1}(1-\alpha)^{\overline{k}+1} + (1-\gamma)\alpha^{\overline{k}+1}(1-\alpha)^{\overline{k}+1}\Big]_{\alpha_b}^{\alpha_g} \\
&= \binom{2(\overline{k}+1)}{\overline{k}+1}\Big(\Delta_{\overline{k}+2,2\overline{k}+3} - \frac{1}{2}\Delta_{\overline{k}+1,2\overline{k}+2}\Big).
\end{aligned}$$

But $\Delta_{\overline{k}+1,\,2\overline{k}+2} = (\alpha_g(1-\alpha_g))^{\overline{k}+1} - (\alpha_b(1-\alpha_b))^{\overline{k}+1} < 0$ since $\alpha_g > \alpha_b > 1/2$. Therefore, to avoid contradicting the fact that $\delta(\hat{T}+2) < 0$, we must have $\Delta_{\overline{k}+2,\,2\overline{k}+3} < 0$. ∎

[20] Since the private signals and the public outcome are (positively) correlated with the underlying state, insincere predictions (that is, those that contradict the agent's private signals) are clearly dominated.

[21] Note that the argument that follows applies immediately to $\hat{T}$ even, so the proposed bound $T > \hat{T}+1$ continues to be sufficient for the optimality of randomization in that case.

Now consider the alternative mechanism defined by

$$\hat{x}_r(s^T) := \begin{cases} 1 & \text{if } \sum_{\tau=1}^{2\bar{k}+2} \mathbf{1}_r(s_\tau) \geq \bar{k}+2 \\ \rho & \text{if } \sum_{\tau=1}^{2\bar{k}+2} \mathbf{1}_r(s_\tau) = \bar{k}+1 \text{ and } s_{2\bar{k}+3} = r, \\ 0 & \text{otherwise} \end{cases}$$

where

$$\rho := \frac{\beta_{1,3,b} + \beta_{2,3,b}}{2\beta_{2,3,b}} = \frac{1}{2(\gamma\alpha_b + (1-\gamma)(1-\alpha_b))}$$

is the same probability as in the three-period optimal stochastic mechanism described in Theorem 5. Essentially, $\hat{x}_r(\cdot)$ does not solicit any information from the agent until period $2\bar{k}+2$. At that point, it offers the agent the option of either making an immediate prediction in period $2\bar{k}+2$ or waiting one period until $2\bar{k}+3$ to make a prediction. The agent is hired with probability 1 if his early prediction is correct, with probability ρ if the late prediction is correct, and with probability 0 if his prediction is incorrect.

Clearly, if the agent has at least $\bar{k}+2$ identical signals in the first $2\bar{k}+2$ periods, he will continue to have a strict majority of that signal in period $2\bar{k}+3$; therefore, his recommendation will be the same in both periods, but delaying lowers the probability of being hired if the recommendation is correct. Therefore, such an agent will choose to make an immediate prediction in period $2\bar{k}+2$.

On the other hand, an agent with exactly $\bar{k}+1$ of each signal in period $2\bar{k}+2$ would prefer to wait until the next period before making a prediction. Notice that ρ is chosen to leave the type-b agent indifferent between guessing immediately and waiting for one additional signal, while (since $\alpha_g > \alpha_b$) the type-g agent's more informative signal gives him a strict incentive to delay.

Thus, it remains to be shown that the stochastic mechanism $\hat{x}_r(\cdot)$ defined above yields the principal a higher payoff than the period-$\hat{T}$ prediction mechanism.

As shown in the proof of Theorem 2, the principal's payoff of the period-$\hat{T}$ prediction mechanism (for $\hat{T} = 2\bar{k}+1$) equals that of the period-$(2\bar{k}+2)$ prediction mechanism. In that latter mechanism, the agent is hired with probability 1 when he observes at least $\bar{k}+2$ signals that match the outcome, with probability $1/2$ when he observes exactly $\bar{k}+1$ signals that match the outcome, and with probability 0 otherwise.

Therefore, the difference in the principal's payoff between $\hat{x}_r$ and the period-$\hat{T}$ prediction mechanism arises precisely from the situation where the agent has observed exactly $\bar{k}+1$ of each signal by period $2\bar{k}+2$, and therefore chooses to postpone predicting under $\hat{x}_r$. This leads to a payoff differential of

$$\binom{2\bar{k}+2}{\bar{k}+1}\left(\rho\Delta_{\bar{k}+2,\,2\bar{k}+3} - \frac{1}{2}\Delta_{\bar{k}+1,\,2\bar{k}+2}\right),$$

since there are exactly $\binom{2\bar{k}+2}{\bar{k}+1}$ signal sequences that lead the agent to be exactly tied in $2\bar{k}+2$ periods. The deterministic prediction mechanism will hire the agent with probability $1/2$ if he is exactly tied (due to the agent mixing when indifferent), whereas $\hat{x}_r$ hires the agent with probability ρ if the final signal matches (yielding a net payoff of $\rho\,\Delta_{\overline{k}+2,2\overline{k}+3}$).

Note, however, that

$$\begin{aligned}\rho\,\Delta_{\overline{k}+2,2\overline{k}+3}-\frac{1}{2}\Delta_{\overline{k}+1,2\overline{k}+2} &= \frac{\Delta_{\overline{k}+2,2\overline{k}+3}}{2(\gamma\alpha_b+(1-\gamma)(1-\alpha_b))}-\frac{1}{2}\Delta_{\overline{k}+1,2\overline{k}+2}\\ &= \frac{\left[\gamma\alpha^{\overline{k}+2}(1-\alpha)^{\overline{k}+1}+(1-\gamma)\alpha^{\overline{k}+1}(1-\alpha)^{\overline{k}+2}\right]_{\alpha_b}^{\alpha_g}}{2(\gamma\alpha_b+(1-\gamma)(1-\alpha_b))}-\frac{1}{2}\Delta_{\overline{k}+1,2\overline{k}+2}\\ &= \frac{\left[(\alpha(1-\alpha))^{\overline{k}+1}(\gamma\alpha+(1-\gamma)(1-\alpha))\right]_{\alpha_b}^{\alpha_g}}{2(\gamma\alpha_b+(1-\gamma)(1-\alpha_b))}-\frac{1}{2}\left[(\alpha(1-\alpha))^{\overline{k}+1}\right]_{\alpha_b}^{\alpha_g}\\ &= \frac{1}{2}(\alpha_g(1-\alpha_g))^{\overline{k}+1}\left(\frac{\gamma\alpha_g+(1-\gamma)(1-\alpha_g)}{\gamma\alpha_b+(1-\gamma)(1-\alpha_b)}-1\right) > 0.\end{aligned}$$

Therefore, the mechanism $\hat{x}_r$ defined above, which nontrivially randomizes in $2\overline{k}+3$ periods, achieves a strictly higher revenue than the optimal deterministic mechanism, a period-$(2\overline{k}+1)$ recommendation mechanism. ∎

PROOF OF THEOREM 7:

Note first that the revelation principle (see the online Appendix) implies that, when the principal has commitment power, it is without loss of generality to restrict attention to incentive-compatible direct mechanisms $\chi:\Lambda\times S^T\to\{0,1\}^2$, where we write $\chi(\cdot)=(\chi_h(\cdot),\chi_l(\cdot))\in\{0,1\}^2$ for the principal's hiring decision given outcomes h and l, respectively.

So fix any nontrivial and incentive-compatible direct mechanism χ, and note that we must have $\chi(\lambda,s^T)\in\{(1,0),(0,1)\}$ for all $(\lambda,s^T)\in\Lambda\times S^T$. Note that if $\chi(\hat{\lambda},\hat{s}^T)=(1,1)$ for some reports $\hat{\lambda},\hat{s}^T$, then incentive compatibility requires that the agent is *always* hired, regardless reports (otherwise he would deviate by always reporting $\hat{\lambda},\hat{s}^T$). Similarly, if $\chi(\hat{\lambda},\hat{s}^T)=(0,0)$ for some $(\hat{\lambda},\hat{s}^T)$, then incentive compatibility requires that the agent is *never* hired, regardless of his reports (otherwise he would deviate by never reporting $\hat{\lambda},\hat{s}^T$).

Now fix any $\lambda\in\Lambda$, and let t_λ be the largest period such that $\chi(\lambda,\cdot)$ is measurable with respect to the first t_λ reports; that is, $\chi(\lambda,s^T)=\chi(\lambda,\hat{s}^T)$ for all $s^T=(s^{t_\lambda},s_{t_\lambda+1},\ldots,s_T)$ and $\hat{s}^T=(s^{t_\lambda},\hat{s}_{t_\lambda+1},\ldots,\hat{s}_T)$ that coincide in their first t_λ periods. Since periods $t_\lambda+1$ through T do not affect the hiring decision given an initial period report of λ, we abuse notation somewhat and write $\chi(\lambda,s^{t_\lambda})$ to denote the principal's hiring rule.

The definition of t_λ as the final period in which the agent's reported signal potentially changes the hiring decision as a function of the ultimate outcomes implies the existence of $\hat{s},\hat{s}'\in S_{t_\lambda}$ and $\hat{s}^{t_\lambda-1}\in\prod_{\tau=1}^{t_\lambda}S_\tau$ such that

$$\chi(\lambda,\hat{s}^{t_\lambda-1},\hat{s})=(1,0)\neq(0,1)=\chi(\lambda,\hat{s}^{t_\lambda-1},\hat{s}').$$

Moreover, we must have

$$\Pr(r = h|\lambda, \hat{s}^{t_\lambda - 1}, \hat{s}) \geq \frac{1}{2} \geq \Pr(r = h|\lambda, \hat{s}^{t_\lambda - 1}, \hat{s}');$$

if this did not hold, the pivotality of the period-t_λ report following history $(\lambda, \hat{s}^{t_\lambda - 1})$ would lead to a violation of incentive compatibility.

This implies that the agent who initially observes signal $\lambda \in \Lambda$ has a strategy which guarantees that he is always hired at the outcome he thinks more likely in period t_λ: simply report $\hat{s}^{t_\lambda - 1}$ regardless of signals seen in the first $t_\lambda - 1$ periods, and then report either $\hat{s}$ or $\hat{s}'$ in period t_λ based on his true signals and his posterior expectation of the most likely outcome. Therefore, incentive compatibility implies that the continuation mechanism $\chi(\lambda, \cdot)$ must be payoff equivalent (for an agent who initially observes signal $\lambda \in \Lambda$) to making a prediction at period t_λ.

Finally, note that the signal structure is such that the agent, regardless of his initial private signal, weakly prefers to make a prediction as late as possible. Therefore, by incentive compatibility of the initial signal report, it must be the case that the agent observing $\lambda \in \Lambda$ is always ex ante indifferent between being asked to make a prediction in t_λ or in $t^* := \max_{\lambda' \in \Lambda}\{t_{\lambda'}\}$. As a result, the principal is indifferent between offering the direct mechanism χ or a period-t^* prediction mechanism.

To see that this outcome (and hence payoffs) remains implementable in the game without commitment, note that if the principal ignores all reports of the agent except that in period t^* (hiring if and only if the period-t^* prediction matches the ultimate outcome), it is a best response by the agent to babble in all periods except t^*. Of course, this babbling justifies the principal ignoring the reports in those periods. Meanwhile, hiring the agent after a correct period-t^* prediction is also sequentially rational for the principal; if it were not, then the mechanism's payoff in the full commitment model would be negative, contradicting its optimality. Thus, as in Theorem 4 for the baseline model, the lack of commitment does not change the outcomes or payoffs. ∎

REFERENCES

Aghion, Philippe, and Matthew O. Jackson. 2016. "Inducing Leaders to Take Risky Decisions: Dismissal, Tenure, and Term Limits." *American Economic Journal: Microeconomics* 8 (3): 1–38.

Battaglini, Marco. 2005. "Long-Term Contracting with Markovian Consumers." *American Economic Review* 95 (3): 637–58.

Boleslavsky, Raphael, and Maher Said. 2013. "Progressive Screening: Long-Term Contracting with a Privately Known Stochastic Process." *Review of Economic Studies* 80 (1): 1–34.

Chen, Ying, Navin Kartik, and Joel Sobel. 2008. "Selecting Cheap-Talk Equilibria." *Econometrica* 76 (1): 117–36.

Courty, Pascal, and Hao Li. 2000. "Sequential Screening." *Review of Economic Studies* 67 (4): 697–717.

Crawford, Vincent P., and Joel Sobel. 1982. "Strategic Information Transmission." *Econometrica* 50 (6): 1431–51.

Crémer, Jacques, and Richard P. McLean. 1988. "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions." *Econometrica* 56 (6): 1247–57.

Elliott, Graham, and Allan Timmermann. 2016. "Forecasting in Economics and Finance." *Annual Review of Economics* 8 (1): 81–110.

Eső, Péter, and Balázs Szentes. 2017. "Dynamic Contracting: An Irrelevance Theorem." *Theoretical Economics* 12 (1): 109–39.

Foster, Dean P., and Rakesh V. Vohra. 1998. "Asymptotic Calibration." *Biometrika* 85 (2): 379–90.

Gomez, Brad T., Thomas G. Hansford, and George A. Krause. 2007. "The Republicans Should Pray for Rain: Weather, Turnout, and Voting in U.S. Presidential Elections." *Journal of Politics* 69 (3): 649–63.

Guo, Yingni, and Johannes Hörner. 2017. "Dynamic Allocation without Money." Unpublished.

Holmström, Bengt. 1999. "Managerial Incentive Problems: A Dynamic Perspective." *Review of Economic Studies* 66 (1): 169–82.

Krähmer, Daniel, and Roland Strausz. 2011. "Optimal Procurement Contracts with Pre-Project Planning." *Review of Economic Studies* 78 (3): 1015–41.

Krishna, Vijay, and John Morgan. 2004. "The Art of Conversation: Eliciting Information from Experts through Multi-Stage Communication." *Journal of Economic Theory* 117 (2): 147–79.

Lahiri, Kajal, and Xuguang Sheng. 2008. "Evolution of Forecast Disagreement in a Bayesian Learning Model." *Journal of Econometrics* 144 (2): 325–40.

Li, Wei. 2007. "Changing One's Mind When the Facts Change: Incentives of Experts and the Design of Reporting Protocols." *Review of Economic Studies* 74 (4): 1175–94.

Marinovic, Iván, Marco Ottaviani, and Peter Norman Sørensen. 2013. "Forecasters' Objectives and Strategies." In *Handbook of Economic Forecasting,* Vol. 2B, edited by Graham Elliott and Allan Timmermann, 689–720. Amsterdam: Elsevier.

Morris, Stephen. 2001. "Political Correctness." *Journal of Political Economy* 109 (2): 231–65.

Myerson, Roger B. 1981. "Optimal Auction Design." *Journal of Mathematics of Operations Research* 6 (1): 58–73.

Olszewski, Wojcieh. 2015. "Calibration and Expert Testing." In *Handbook of Game Theory with Economic Applications,* Vol. 4, edited by H. Peyton Young and Schmuel Zamir, 949–84. New York: Elsevier.

Olszewski, Wojciech, and Marcin Peski. 2011. "The Principal-Agent Approach to Testing Experts." *American Economic Journal: Microeconomics* 3 (2): 89–113.

Ottaviani, Marco, and Peter Norman Sørensen. 2006a. "Professional Advice." *Journal of Economic Theory* 126 (1): 120–42.

Ottaviani, Marco, and Peter Norman Sørensen. 2006b. "Reputational Cheap Talk." *RAND Journal of Economics* 37 (1): 155–75.

Ottaviani, Marco, and Peter Norman Sørensen. 2006c. "The Strategy of Professional Forecasting." *Journal of Financial Economics* 81 (2): 441–66.

Pavan, Alessandro, Ilya Segal, and Juuso Toikka. 2014. "Dynamic Mechanism Design: A Myersonian Approach." *Econometrica* 82 (2): 601–53.

Strausz, Roland. 2003. "Deterministic Mechanisms and the Revelation Principle." *Economics Letters* 79 (3): 333–37.

Tetlock, Philip E. 2005. *Expert Political Judgment: How Good Is It? How Can We Know?* Princeton, NJ: Princeton University Press.

Tetlock, Philip E., and David Gardner. 2015. *Superforecasting: The Art and Science of Prediction.* New York: Crown Publishing Group.

Trueman, Brett. 1994. "Analyst Forecasts and Herding Behavior." *Review of Financial Studies* 7 (1): 97–124.