Naive Calibration*

Yair Antler[†] and Benjamin Bachi[‡]

October 16, 2024

Abstract

We develop a model of non-Bayesian decision-making in which an agent obtains a signal about a relevant economic fundamental and subsequently takes an action. To interpret the signal, the agent calibrates a simple prediction rule based on a dataset that consists of previous signals and state realizations. Her subsequent action affects the probability with which the current signal and the corresponding state realization will be observed and recorded in the dataset that will be used in future decisions. We show that this procedure converges to a steady state and that it results in a seemingly pessimistic behavior that is exacerbated by feedback loops. We apply our model to project selection problems and second-price IPV auctions.

1 Introduction

Imagine that you are a manager of a company that has to make an important decision whose outcome depends on an economic variable, and that you have a noisy estimate of the variable's value. You look in the company's records and find that in similar decision problems faced by your predecessors, on average, the estimates at their disposal were

*We thank Daniel Bird, Duarte Gonçalves, Philippe Jehiel, Botond Kőszegi, Dotan Persitz, Ariel Rubinstein, and Rani Spiegler for helpful comments. We also thank seminar audiences at BRIC 2022, ESEM-EEA 2022, Hebrew University, NASM 2022, NYU Abu Dhabi, Tel Aviv University, UAB, UCL, and the University of Haifa for useful comments. Antler gratefully acknowledges financial support from the Coller Foundation and the Henry Crown Institute of Business Research in Israel.

[†]Coller School of Management, Tel Aviv University. *E-mail:* yair.an@gmail.com.

[‡]Department of Economics, University of Haifa. *E-mail:* bbachi@econ.haifa.ac.il

10% higher than the actual outcomes. How would you interpret the current estimate in light of this finding?

A natural way to account for the apparent bias in the company's records is to adjust the estimate and discount it by 10%. This type of adjustment is called *reference class forecasting* and its use is advocated by researchers, government agencies, and professional associations as a way to account for optimism bias and to prevent cost overruns in construction projects (Kahneman and Lovallo, 1993; HM Treasury, 2003a,b; AACE International, 2012; Flyvbjerg, 2008).[1] In particular, the UK Department for Transport has employed it in the appraisal process of the cost and the construction duration of large transportation projects (Flyvbjerg et al., 2004; UK Department for Transport and Oxford Global Projects, 2020). Similar procedures are used in the oil and gas extraction industries (Nesvold and Bratvold, 2022).[2]

One problem with this approach is that systematic differences between estimates and ex-post outcomes in the dataset may emerge due to selection bias, and not biased estimates. For example, the company's records may depend on one's predecessors' decisions or the organizational memory. It is well known that when failing to properly account for selection bias, one may end up reaching an erroneous conclusion. But what happens when the person who reaches the erroneous conclusion is not an outside observer but rather a decision maker whose actions affect the selection bias in the data she is using, as in the examples in the previous paragraph?

In this paper, we address this question by developing a model where decision-makers attempt to naively correct for a selection bias they themselves have unknowingly contributed to. Our framework extends existing research on decision-making under uncertainty and selection bias by highlighting the recursive nature of selection bias in various decision problems. To illustrate our conceptual framework and some of our findings, consider the next example, which is inspired by Jehiel's (2018) model of investment decisions.

Example 1 *A risk-neutral entrepreneur decides whether to implement risky projects based on their estimated returns. Implementing a project costs $c < 0.5$ and yields a revenue of $\theta \sim U[0,1]$. The estimated return, s, equals θ with probability $p < 1$.*

[1] Typically, projects are sorted into classes according to their type (e.g., tunnel project, bridge project, railway project). Within each class, estimates are adjusted based on the differences between estimates and realizations in all of the projects in the class that were implemented in the past.

[2] Oil extraction projects are often selected based on their estimated reserves and, in ex-post audits, these reserves are typically lower than initially estimated (Brashear et al., 2001; Chen and Dyer, 2009).

Otherwise, s is independently drawn from $U[0,1]$.

Our entrepreneur does not know the joint distribution of revenues and estimates, and so she cannot use the standard Bayesian methodology. Instead, she has access to a large dataset $\mathcal{D}$ that consists of estimates and realized returns of similar projects implemented in the past. She uses this dataset to calibrate a prediction rule that she later uses to (point) predict the project's revenue. Specifically, we assume that her prediction rule takes the simple form $\hat{\theta}(s) = s - b$, where b is a debiasing factor that reflects the possibility that the estimates might be biased.[3] *She calibrates her rule by choosing a debiasing factor that minimizes the quadratic loss function over the dataset, and so it is equal to the average difference between the estimated and actual returns in the dataset, i.e., $b_{\mathcal{D}} = E[s - \theta | (s, \theta) \in \mathcal{D}]$. The entrepreneur implements the project if and only if $\hat{\theta}(s) \geq c$.*

The entrepreneur's dataset is endogenous as it includes only data from implemented projects. Specifically, if the entrepreneur believes that there is a bias b and chooses the optimal implementation cutoff given this belief, that is, $b + c$, then in the long run she will obtain the dataset $\mathcal{D}_b := \{(s, \theta) | s \geq b + c\}$. Note that

$$E[s | s \geq b + c] = 0.5(c + b + 1)$$

and

$$E[\theta | s \geq b + c] = 0.5p(c + b + 1) + 0.5(1 - p),$$

which implies that

$$E[s - \theta | s \geq b + c] = 0.5(1 - p)(c + b).$$

Since $b_{\mathcal{D}}$ and $\mathcal{D}_b$ depend on each other, we take a steady-state approach, which we shall refer to as an equilibrium. *An equilibrium is a perceived bias $b^\star$ and a dataset $\mathcal{D}^\star$ such that the entrepreneur implements projects if and only if $s \geq c + b^\star$, the dataset $D^\star$ consists of these projects, i.e., $\mathcal{D}^\star = \{(s, \theta) | s \geq c + b^\star\}$, and*

$$b^\star = E[s - \theta | (s, \theta) \in \mathcal{D}^\star].$$

[3] In the example (and in the model), the debiasing is additive for the sake of tractability. In Section 6, we discuss other forms of debiasing and show that the main insights of the model hold when the de-biasing factor is multiplicative.

Hence, $b^\star$ must satisfy:

$$b^\star = 0.5(1-p)(c+b^\star). \tag{1}$$

The solution to (1) is $b^\star = c \times \frac{1-p}{1+p}$ and an induced implementation cutoff of $s^\star = c \times \frac{2}{1+p}$. Figure 1 illustrates the predictions and implementation cutoff of our entrepreneur in equilibrium relative to the cutoff she would have were she to calibrate her prediction rule using an uncensored dataset (in which case, her perceived bias would be null as $E[s] = E[\theta]$) and the cutoff of a Bayesian entrepreneur.

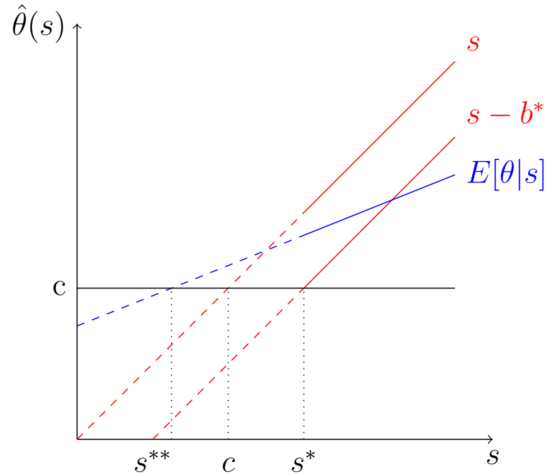

Figure 1: The implementation cutoffs of our entrepreneur (s^*), a naive entrepreneur who uses an uncensored dataset (c), and a Bayesian entrepreneur (s^{**}) in Example 1. The dashed lines represent the observations that are censored from our entrepreneur's dataset.

The entrepreneur's behavior in Example 1 reveals two main insights. First, in a steady state, the entrepreneur holds an incorrect belief that estimates are upwardly biased.[4] The flawed belief is a result of selection bias and model misspecification. Selection bias arises since she calibrates her prediction rule using a censored dataset that includes only implemented projects, i.e., projects for which s is relatively high.

[4]Despite the fact that this belief is incorrect, the selected data the entrepreneur observes is, on average, consistent with this belief given her misspecified model of the world. The idea that agents maintain confidence in their worldview as long as it is not contradicted by the data they gather has its roots in models of conjectural equilibria (Battigalli and Guaitoli, 1997) and self-confirming equilibria (Fudenberg and Levine, 1993) and is by now common in the bounded rationality literature (e.g., Spiegler, 2016).

In these instances, the estimated returns are, on average, higher than the actual returns as $E[s-\theta|s]$ is increasing in s. Were the entrepreneur to use an uncensored dataset that includes all projects, she would reach a conclusion that the estimates are unbiased as $E[\theta]=E[s]$. The misspecification contributes to this flawed belief since the entrepreneur uses a (point) prediction rule $\hat{\theta}(s)=s-b^{\star}$ when in reality $E[\theta|s]=ps+0.5(1-p)$. Without the restriction to this family of prediction rules, the entrepreneur could calculate $E[\theta|s]$ for every signal and reach optimal decisions despite the selection in the data. The combination of these two assumptions yields a perceived bias of

$$b^{\star}=E[s-\theta|s>c+b^{\star}]>E_s[E[s-\theta|s]]=0.$$

Second, the entrepreneur's decisions are too conservative, in the sense that she sets a higher implementation cutoff relative to a Bayesian entrepreneur.[5] A Bayesian entrepreneur with an estimate s would form a correct belief $s-(s-E[\theta|s])$. The marginal Bayesian entrepreneur, who is indifferent whether to implement or not, holds an estimate $s^{\star\star}$ such that $s^{\star\star}-(s^{\star\star}-E[\theta|s^{\star\star}])=c$. Hence, his perceived bias is $b^{\star\star}=s^{\star\star}-E[\theta|s^{\star\star}]$. To gain intuition on why our entrepreneur sets a higher implementation cutoff, suppose that she uses the same implementation cutoff $s^{\star\star}$. This would result in a dataset $\{(s,\theta)|s\geq s^{\star\star}\}$ in the long run. Due to her misspecified prediction rule, she would reach the conclusion that the estimates are biased by $E[s-\theta|s\geq s^{\star\star}]$. Since $E[s-\theta|s]$ is increasing, $E[s-\theta|s\geq s^{\star\star}]>E[s-\theta|s=s^{\star\star}]=b^{\star\star}$. In other words, our entrepreneur pools together all the observations in which $s\geq s^{\star\star}$, which leads to a perceived bias higher than the bias of the marginal Bayesian entrepreneur, who considers only observations in which $s=s^{\star\star}$.

The entrepreneur in Example 1 attempts, but ultimately fails, to properly correct the estimate at her disposal by drawing on past experience. Despite having access to an infinitely large dataset, her ability to properly calibrate her prediction rule is hampered by a feedback loop created by her own decisions, leading to persistent errors in future judgments.

We now introduce a more general decision-making framework that captures these dynamics. In this generalized model, a DM faces a sequence of similar decision problems. In each problem, the DM observes an independent signal $s\in\mathbb{R}$ about the state of a relevant variable $\theta\in\mathbb{R}$ and then takes an action. One interpretation of the signal

[5] To calculate the Bayesian entrepreneur's implementation cutoff, note that $E[\theta|s=\frac{2c+p-1}{2p}]=c$. Hence, the Bayesian implementation cutoff is $max\{\frac{2c+p-1}{2p},0\}<\frac{2c}{1+p}$.

is a noisy estimate of the state's value. We assume that the higher the value of the signal is, the higher the expected value of the state, and that the higher the expected value of the state is, the higher the DM's optimal action.

The DM does not know the joint distribution of signals and states. Instead, she uses a restricted dataset that consists of past signals and state realizations to calibrate a simple prediction rule. Specifically, according to the DM's prediction rule, the signal s reflects the true state θ up to a constant bias b, and so she considers prediction rules of the form $\hat{\theta}(s) = s - b$. The DM's model is misspecified as, in reality, $E[\theta|s]$ may take a different form (e.g., in Example 1, $ps + 0.5(1 - p)$). The DM calculates b by minimizing the quadratic loss function over her dataset, which implies that b is the average difference between signals and state realizations in the dataset, as in Example 1. When facing a new decision, the DM uses the current signal s and the prediction rule to form a point estimate $\hat{\theta}(s)$, and chooses an action as if the state were $\hat{\theta}(s)$.

The DM's data is endogenous. After taking an action, the DM observes the state's realization with a probability that is increasing in her action. This monotonicity assumption is naturally satisfied in the investment decision setting. Another example is a second-price IPV auction in which the DM relies on a signal about the object's value. The higher the signal she observes, the higher the expected value of the object and the optimal bid are, and therefore the more likely the DM is to win the object and learn its actual value. The DM records all the state's realizations she observes and their respective signals in the dataset.

Since the DM's actions affect the data she observes and vice versa, we take an equilibrium approach to characterizing the steady state of this system. An equilibrium consists of a strategy (a mapping from signals to actions) and a perceived bias b, such that the strategy is a best response given the point prediction $\hat{\theta}(s) = s - b$, and the perceived bias b minimizes the quadratic loss function over the dataset generated by the DM's strategy. We show that the DM's behavior converges to an equilibrium under mild conditions. Moreover, in equilibrium, the perceived bias is always greater than $E[s] - E[\theta]$, which we refer to as the *inherent bias*. This leads to a conservative interpretation of the signals as illustrated in Example 1.

We refer to the mapping from the DM's actions to the probability that she observes the actual realizations ex post as a *feedback function*. In Example 1, the feedback function assigns a probability of 1 to observing the returns of implemented projects and a probability of 0 to observing unimplemented ones. More generally, this function reflects basic properties of the environment in which the DM operates; for instance, different feedback functions may represent different types of organizational memory. We derive a tight condition that enables us to rank different feedback functions in terms of the perceived bias that they induce in equilibrium. Essentially, a feedback function ϕ induces a higher perceived bias than the feedback function ϕ' for every objective distribution of states and signals if and only if ϕ dominates ϕ' in the likelihood ratio sense. In Section 5 we apply this condition to the setting of a second-price IPV auction in which bidders rely on a signal to bid and observe the actual value of the object only if they win. We use the fact that the feedback function depends on the number of bidders to show that the feedback-ranking condition implies that the bidders' perceived bias is increasing in the number of bidders for any information structure.

We then analyze the externalities that DMs impose on one another. In this analysis, we interpret the DM as a sequence of DMs facing similar problems and assume that a fraction of the DMs use our misspecified calibration heuristic while the others are Bayesian in the sense that they know the joint distribution of states and signals and apply the Bayesian methodology. Since they know the joint distribution, Bayesian DMs ignore the dataset when taking an action. Nonetheless, their actions select state realizations into the dataset, thereby affecting the information naive DMs rely on. We find that the presence of Bayesian DMs mitigates the externalities naive DMs impose on their successors through the data. Interestingly, the naive DMs' perceived bias does not vanish even when the share of Bayesian DMs approaches 1.

Overall, our results indicate that calibrating a prediction rule in an attempt to account for the discrepancy in the data can lead to a seemingly pessimistic behavior. This may explain well-documented phenomena such as the hurdle rate premium puzzle, which is the tendency of firms to set investment hurdle rates that are substantially higher than the cost of capital (Poterba and Summers, 1995; Meier and Tarhan, 2007). This microfoundation for irrational pessimism in equilibrium is a contribution to the literature on misspecified beliefs. While the literature provides much evidence and many models of overoptimism, scholars have devoted little attention to irrational pessimism, which is an equally important topic: just like optimism, pessimism may lead individuals to erroneous conclusions and suboptimal choices.

Related literature

The apparent pessimism in our model is related to Compte's (2002) and Compte and Postlewaite's (2018) *cautious behavior*. While the apparent pessimism in our model follows from constraints on the DM's prediction rules (i.e., on her beliefs), the cautious behavior in Compte and Postlewaite (2018) follows from constraints on the agents' strategy space (e.g., restricting bidders to using a fixed shading factor across different signals). Essentially, our DM has misspecified beliefs whereas their agents are restricted to using coarse strategies. Moreover, under both modeling approaches, the agents optimize an objective function given a certain dataset, while biases in the dataset can have an effect on their decision making. However, a key difference between the two approaches in this respect is that our DM chooses the prediction rule (from a constrained set of rules) that has the best fit with the data, whereas their agents choose the strategy that maximizes their payoff (from a constrained set of strategies). Importantly, the data in our model is endogenous and affected by the DM's actions, which plays a key role in magnifying the DM's biased predictions. By contrast, in Compte and Postlewaite (2018) the dataset is exogenous.

The effects we find are also reminiscent of choice-driven optimism (Van den Steen, 2004) and the optimizer's curse (Smith and Winkler, 2006). These authors consider the perspective of an outside observer. By contrast, we consider the perspective of a DM who tries to account for selection bias while her actions affect the selection of state realizations into the dataset, which requires an equilibrium approach that is absent from the aforementioned papers.

This paper contributes to the literature on naive learning from endogenously selected data. Esponda and Vespa (2018) and Barron et al. (2024) provide empirical evidence that DMs tend to neglect selection and extrapolate naively from endogenous data. Esponda (2008) shows that selection neglect can exacerbate adverse-selection problems and Jehiel (2018) lays the equilibrium microfoundations of overoptimism when there is selection in an investment decision setting similar to that of Example 1. In Section 5, we explain why Jehiel's procedure leads to opposite results from ours.

This paper belongs to a growing literature on decision making and strategic interaction when agents hold misspecified models of the world (Piccione and Rubinstein, 2003; Eyster and Rabin, 2005; Jehiel, 2005; Esponda, 2008; Esponda and Pouzo, 2016; Spiegler, 2016; Heidhues et al., 2018, 2023).[6] For excellent reviews on this topic see Jehiel (2020) and Spiegler (2020).

The calibration approach in this paper is related to Esponda and Pouzo's (2016) approach to learning with misspecified models. They propose a solution concept, called

[6]There is ample evidence that economic agents depart from the Bayesian methodology (for a comprehensive review, see Benjamin, 2019).

the Berk–Nash equilibrium, in which the DM's subjective beliefs minimize the Kullback–Leibler divergence from the true distribution. The main difference between the solution concepts in the two models is that, in our model, the DM's equilibrium beliefs minimize a quadratic loss function over the dataset rather than minimizing the Kullback–Leibler divergence as in Esponda and Pouzo (2016). In general, the two methods yield different results. Furthermore, the convergence to an equilibrium differs in the two models. In Esponda and Pouzo (2016) the DM has (potentially misspecified) prior beliefs that are updated each period using the Bayesian methodology. By contrast, in our model the DM has no prior beliefs; her periodic beliefs minimize a quadratic loss function over the data gathered up to that period.

The paper proceeds as follows. We present the model in Section 2 and analyze it in Section 3. In Section 4 we show that the DM's behavior converges to an equilibrium. In Section 5 we use our results to study second-price IPV auctions and investment decision problems. Section 6 discusses alternative families of prediction rules, and Section 7 concludes. All proofs are relegated to the appendix.

2 The Model

The environment is composed of two random variables, namely, the state of nature θ and a signal s, which are distributed according to a bivariate probability distribution function $F(\theta, s)$ defined over $\mathbb{R}^2$. We assume that the (marginal) distribution over signals has a log-concave density and denote it by $f(s)$. Denote by $k := E[s] - E[\theta]$ the *inherent bias* of the signal. We assume that $E[\theta|s]$ and $s - E[\theta|s]$ are nondecreasing in s. When the support of the distribution of signals is unbounded, we require that $\lim_{s' \to \infty} E[\theta|s \geq s']$ be unbounded as well. These assumptions are satisfied by many prevalent information structures. In particular, the signal can be the sum of the state and a noise term.[7]

A DM receives a signal s and chooses an action. A strategy is a mapping $a : \mathbb{R} \to A$ from signals to actions, where $A = [\underline{a}, \overline{a}] \subset \mathbb{R}$. The DM's payoff $\pi(\theta, a)$ is a function of the state and her action. Let $a^\star(\theta)$ be an action that maximizes the DM's payoff given a state θ. We assume that $a^\star(\theta)$ is weakly increasing in θ.

The DM has access to an infinitely large dataset $\mathcal{D}$ that consists of signals and state

[7]For example, let $s = \theta + \epsilon$, where ϵ and θ are independently drawn from log-concave density functions and $E[\epsilon] = 0$. As per Efron (1965), $s - E[\theta|s]$ and $E[\theta|s]$ are nondecreasing in s.

realizations from past decisions. She uses this dataset to calibrate a (point) prediction rule $\hat{\theta}(s)$, which she then uses to interpret the signal at her disposal when facing a new decision problem. She restricts attention to a family of simple prediction rules from which she chooses the one that best fits the data. Specifically, the DM considers rules of the form $\hat{\theta}(s) = s - b$, where $b \in \mathbb{R}$, and chooses the rule that minimizes the quadratic loss function over the dataset

$$E[(\theta - \hat{\theta}(s))^2 | (s, \theta) \in \mathcal{D}]. \tag{2}$$

Minimizing (2) yields

$$b_{\mathcal{D}} := \arg\min_{b \in \mathbb{R}} E[(\theta - (s - b))^2 | (s, \theta) \in \mathcal{D}] = E[s - \theta | (s, \theta) \in \mathcal{D}]. \tag{3}$$

We refer to $b_{\mathcal{D}}$ as the DM's *perceived bias.*

The dataset $\mathcal{D}$ is formed endogenously. The probability that each pair (s, θ) is recorded in $\mathcal{D}$ depends on the action the DM chooses given s. Formally, a pair (s, θ) is recorded with probability $\phi(a(s))$, where $\phi : A \to [0, 1]$ is a feedback function and $a(s)$ is the action the DM chooses given s. We assume that $\phi(\cdot)$ is nondecreasing. If the DM's strategy is constant over time and the dataset is not empty (i.e., when $\phi(a(s)) > 0$ for a strictly positive measure of signals), then the proportion of each signal s in $\mathcal{D}$ in the long run is

$$\frac{f(s)\phi(a(s))}{\int_{-\infty}^{\infty} f(s)\phi(a(s))ds}. \tag{4}$$

Thus, we can write (3) as a function of the strategy $a(\cdot)$:

$$b_{\mathcal{D}} = \frac{\int_{-\infty}^{\infty} f(s)\phi(a(s))(s - E[\theta|s])ds}{\int_{-\infty}^{\infty} f(s)\phi(a(s))ds}. \tag{5}$$

In order to treat also cases in which the dataset is empty, we rewrite (5) as

$$\int_{-\infty}^{\infty} f(s)\phi(a(s))(s - E[\theta|s] - b_{\mathcal{D}})ds = 0. \tag{6}$$

Note that when the dataset is empty, any perceived bias is consistent with the dataset.

We now present our equilibrium notion. The first requirement in Definition 1 is that the dataset be generated by the DM's strategy, and that the prediction rule minimize the quadratic loss function. The second requirement is that the DM's strategy be optimal given her prediction rule.

Definition 1 *A strategy $a(\cdot)$ and a bias b constitute an equilibrium if the following conditions are met:*

1. *b and $a(\cdot)$ satisfy Equation (6).*

2. *$a(s) = a^{\star}(s - b)$ for every signal s.*

Note that the definition does not preclude the possibility that $\mathcal{D} = \emptyset$ in equilibrium. This may occur when $\phi(\underline{a}) = 0$ and there exists a large enough bias b such that $a^{\star}(s - b) = \underline{a}$ for every signal s in the support of f. In such a case, every $b' \geq b$ is consistent with (6) and is part of an equilibrium. We refer to equilibria in which $\mathcal{D} = \emptyset$ as *corner equilibria* and to equilibria in which $\mathcal{D} \neq \emptyset$ as *interior equilibria*.

3 Analysis

In this section, we show that in an interior equilibrium, the perceived bias is always greater than the inherent bias. We then study how the naive DM's behavior changes with the feedback function, and compare the equilibrium behavior of our naive DM to the behavior of a Bayesian DM. Finally, we investigate the externalities imposed by the naive DM's behavior. We start by showing that an equilibrium exists.

Proposition 1 *An equilibrium exists.*

To gain intuition for this result, note that a perceived bias b pins down the DM's strategy $a^{\star}(s - b)$, which pins down a dataset $\mathcal{D}_b$. If there exists a bias b such that $\mathcal{D}_b = \emptyset$, then the data is consistent with any "perception" of the bias and, in particular, with b. In this case, b is clearly an equilibrium. The proof shows that if there is no such bias, then there exists a bias $b^{\star}$ such that

$$b^{\star} = E[s - \theta | (s, \theta) \in \mathcal{D}_{b^{\star}}].$$

If there exists more than one interior equilibrium, then we often focus on the two interior equilibria with the minimal and maximal biases. Denote the minimal and maximal (interior) equilibrium biases by $\overline{b}$ and $\underline{b}$, respectively.

Having established that an equilibrium exists, we now turn to study its properties. We start with the perceived bias.

Proposition 2 *In every interior equilibrium, the perceived bias is greater than the inherent bias, i.e.,* $b \geq k$.

To understand this result, note that if the DM were to observe all signals and their respective state realizations, the perceived bias would be equal to the inherent bias. However, the DM's dataset includes only a selected sample of such pairs. In particular, since $\phi(\cdot)$ and $a^{\star}(\cdot)$ are nondecreasing, the dataset contains disproportionately more cases in which the signal is high. The assumption that $s - E[\theta|s]$ is nondecreasing in s implies that the dataset also contains disproportionately more cases in which the difference $s - \theta$ is high, which in turn leads to a perceived bias greater than the inherent bias.

We now provide a definition that allows us to rank different feedback functions according to the degree of selection they induce in the data. We then use this definition to show how the degree of selection affects the perceived bias in equilibrium.

Definition 2 *The feedback function* ϕ *dominates the feedback function* $\tilde{\phi}$ *in the likelihood ratio sense if* $\frac{\phi(a)}{\tilde{\phi}(a)}$ *is nondecreasing in* a.

To illustrate this definition, note that the feedback function $\phi(a) = a^n$ dominates the feedback function $\tilde{\phi}(a) = a^m$ for $m < n$ and $a > 0$. The next result establishes that if a feedback function dominates another feedback function in the likelihood ratio sense, then in an interior equilibrium it induces a higher perceived bias. Let us denote by $\underline{b}_{\phi}$ and $\overline{b}_{\phi}$ the minimal and maximal bias in an interior equilibrium given the feedback function ϕ, when such an equilibrium exists.

Proposition 3 *If* ϕ *dominates* $\tilde{\phi}$ *in the likelihood ratio sense and* $\underline{b}_{\phi}$ *and* $\underline{b}_{\tilde{\phi}}$ *exist, then* $\underline{b}_{\phi} \geq \underline{b}_{\tilde{\phi}}$ *and* $\overline{b}_{\phi} \geq \overline{b}_{\tilde{\phi}}$.

The dominant feedback function ϕ assigns relatively more weight to high actions. Thus, intuitively, it assigns more weight to high signals and, as a result, more weight to instances in which the signal is high relative to the actual state realization. The more weight a feedback function assigns to these instances relative to instances in which signals are low, the higher the DM's perceived bias.

Likelihood ratio dominance is a tight condition in the sense that, if ϕ does not dominate $\tilde{\phi}$ and both functions are continuous, then there exist a distribution $F(\cdot, \cdot)$ and a payoff function $\pi(\cdot, \cdot)$ such that $\tilde{\phi}$ induces a higher perceived bias than ϕ in equilibrium. The intuition for this tightness is that if ϕ does not dominate $\tilde{\phi}$ in the likelihood ratio sense, then there is some interval $[a_l, a_h]$ on which $\phi|_{[a_l,a_h]}$ is dominated by $\tilde{\phi}|_{[a_l,a_h]}$ in the likelihood ratio sense. It is possible to find a distribution $F(\cdot,\cdot)$ that is concentrated on that interval and a payoff function $\pi(\cdot,\cdot)$ such that the result of Proposition 3 is reversed. The following corollary summarizes this discussion.

Corollary 1 *If ϕ does not dominate $\tilde{\phi}$ in the likelihood ratio sense and both functions are continuous, then there exist a distribution $F(\cdot,\cdot)$ and a payoff function $\pi(\cdot,\cdot)$, such that $\underline{b}_\phi < \underline{b}_{\tilde{\phi}}$ and $\overline{b}_\phi < \overline{b}_{\tilde{\phi}}$.*

At this point, it is worth comparing the prediction of our naive DM, $\hat{\theta}(s) = s - b$, to the prediction of a Bayesian DM, $E[\theta|s]$. Our DM's prediction is higher if and only if $s - b > E[\theta|s]$. Recall that, by assumption, $s - E[\theta|s]$ is nondecreasing in s. In an interior equilibrium, b is a weighted average of $s - E[\theta|s]$ in the dataset and, therefore, $b \in (\inf(s - E[\theta|s]), \sup(s - E[\theta|s]))$. Thus, in an interior equilibrium, there is a signal $\hat{s}$ such that our DM's prediction is higher than the Bayesian DM's prediction if and only if $s > \hat{s}$.

While our DM's predictions are higher than a Bayesian DM's predictions for high signals, they are, on average, lower than the Bayesian DM's predictions. This follows from the equilibrium bias being greater than the inherent bias (Proposition 2). Our DM's average prediction is $\int_{-\infty}^{\infty} f(s)(s-b)ds = E(s) - b = E[\theta] + k - b$, whereas the Bayesian DM's average prediction is $\int_{-\infty}^{\infty} f(s)E[\theta|s]ds = E(\theta)$. Corollary 2 summarizes this discussion.

Corollary 2 *In an interior equilibrium, there exists a signal $\hat{s}$ such that the naive DM's prediction $s - b$ is higher than a Bayesian DM's prediction $E[\theta|s]$ if and only if $s > \hat{s}$. Moreover, on average, the naive DM's prediction is weakly lower than the Bayesian DM's prediction.*

So far, we have assumed that the dataset on which the DM bases her decisions is generated by her own actions or by the actions of other DMs who use the same heuristic. However, in some situations DMs are more sophisticated and know the joint distribution of states and signals. Such DMs can apply the Bayesian methodology to make use of the signals at their disposal. Since their actions select different signals into the dataset, they affect the discrepancy in the data and the perceived bias. In turn, this bias affects the naive DMs' behavior and the signals they select into the dataset.

We now introduce Bayesian DMs into their model and vary their share to examine the externalities they impose on our naive DM.

We incorporate this idea into our model by assuming that a share $\alpha > 0$ of the DMs are Bayesian. Since Bayesian DMs are unaffected by the dataset (as they know the joint distribution of signals and states), varying their share enables us to study the externalities they impose on naive DMs without worrying that their own behavior is being affected by the presence of the naive DMs. In particular, it allows us to explore the implications for the naive DMs' equilibrium bias.

We assume that a Bayesian DM plays the strategy[8] $a^\star(E[\theta|s])$. By similar arguments to those used in the proofs of Propositions 1 and 2, an equilibrium exists and the perceived bias is greater than the inherent bias for any $\alpha \in [0, 1]$. When an interior equilibrium exists, we denote the maximal and minimal equilibrium biases by $\bar{b}_\alpha$ and $\underline{b}_\alpha$, respectively. Next, we show that the presence of Bayesian DMs mitigates the discrepancy in the data and lowers the perceived bias: the more Bayesian DMs there are, the lower the perceived bias is in an interior equilibrium.

Proposition 4 *Suppose that an interior equilibrium exists for both α and $\alpha' > \alpha$. Then $\bar{b}_\alpha \geq \bar{b}_{\alpha'}$ and $\underline{b}_\alpha \geq \underline{b}_{\alpha'}$.*

Relative to our naive DMs, Bayesian DMs assign a lower weight to the value of the signal as they take their prior beliefs into account. Therefore, as established in Corollary 2, Bayesian DMs play lower (resp., higher) actions when the signal is high (resp., low). Since $\phi(\cdot)$ and $a^\star(\cdot)$ are nondecreasing, it follows that actions taken by a Bayesian DM generate less (resp., more) feedback in situations in which the signal is high (resp., low). As a result, the average bias in the dataset is lower when there are more Bayesian DMs. Thus, the presence of Bayesian DMs imposes an externality on naive DMs that leads them to choose higher actions. In turn, this further decreases the bias in the data they rely on.

When the share of naive DMs vanishes, the additional bias in the data, $b - k$, does not disappear. This is because Bayesian DMs also take higher actions given higher signals, which implies that higher signals are more likely to be recorded in the dataset. Thus, decision making based on data generated by Bayesian DMs results in an additional bias, albeit lower than when DMs use the naive calibration procedure.

[8]This assumption holds in various settings and, in particular, in both of our applications, namely, investment decisions and second-price IPV auctions.

4 Convergence

Our analysis so far has focused on the steady state of the DM's learning process. We now describe the learning process in more detail and show that indeed it converges to a steady state.

We assume that, in each period $t \in \mathbb{N}^+$, the DM receives a signal s_t and forms a point prediction $\hat{\theta}_{\mathcal{D}_t}(s_t)$ based on the signal and the data obtained up to period t, $\mathcal{D}_t$. She then chooses the optimal action given the predicted state, $a^\star(\hat{\theta}_{\mathcal{D}_t}(s_t))$. At the end of the period, with probability $\phi(a^\star(\hat{\theta}_{\mathcal{D}_t}(s_t)))$, the signal s_t and the realization θ_t are recorded in the dataset that will be available in the future. Let $\mathbb{1}_j$ be an indicator that equals 1 if (s_j, θ_j) are recorded in the dataset and 0 otherwise. For every period t, let

$$b_t = \frac{\sum_{j=1}^{t} \mathbb{1}_j (s_j - \theta_j)}{|\mathcal{D}_t|}$$

if $\mathcal{D}_t \neq \emptyset$, and let $b_t = 0$ otherwise. As in the baseline model, $\hat{\theta}_{\mathcal{D}_t}(s) = s - b_t$. Throughout this section, we assume that the support of $F(\cdot, \cdot)$ is bounded.

Let

$$T(b_t) := \frac{\int_{-\infty}^{\infty} f(s)\phi(a^\star(s - b_t))[s - E[\theta|s]]ds}{\int_{-\infty}^{\infty} f(s)\phi(a^\star(s - b_t))ds} \tag{7}$$

when $\int_{-\infty}^{\infty} f(s)\phi(a^\star(s - b_t))ds > 0$ and $T(b_t) = b_t$ otherwise. When the RHS of (7) is well defined, $T(b_t)$ is the expected difference between s_{t+1} and θ_{t+1}, conditional on (s_{t+1}, θ_{t+1}) being recorded in the dataset. Note that b_t is part of an equilibrium of our model if, and only if, $b_t = T(b_t)$.

Proposition 5 *The sequence b_t converges almost surely to a bias b such that $T(b) = b$.*

The proof shows that $b_t : t \geq 0$ is essentially a *stochastic approximation process.* As such, with probability 1 it visits any segment $[a, b]$ in which $T(b) - b$ is bounded away from zero only a finite number of times (Pemantle, 2007, Lemma 2.6). Since each of these segments is almost surely excluded from the limit set of b_t, the process almost surely converges to points in which $T(b) = b$ (Pemantle, 2007, Corollary 2.7), namely, to an equilibrium of our model.

The proposition guarantees that the process converges to an equilibrium. However, it does not guarantee that it converges to every possible equilibrium with strictly positive probability. Due to the stochasticity of the process, there are *unstable* equilibria from which the process almost surely deviates. This happens when there exists an interval $(b^\star - \epsilon, b^\star + \epsilon)$ in the neighborhood of an equilibrium bias $b^\star$ such that $T(b) < b$ for any $b \in (b^\star - \epsilon, b^\star)$ and $T(b) > b$ for any $b \in (b^\star, b^\star + \epsilon)$. This observation follows immediately from Theorem 2.9 in Pemantle (2007). The proof of the next result is omitted as it is easy to verify that the conditions for Theorem 2.9 in Pemantle (2007) are met.

Proposition 6 *Suppose that for some $b^\star$ and $\epsilon > 0$, $sign(T(b) - b) = sign(b - b^\star)$ for all $b \in (b^\star - \epsilon, b^\star + \epsilon)$. Then, $\mathcal{P}(b_t \to b^\star) = 0$.*

5 Applications

We now apply our results to two settings: investment decisions and second-price IPV auctions. In both applications, our naive calibration procedure leads to conservative behavior relative to the behavior of a Bayesian DM: rejection of marginally good projects in investment decisions and underbidding in auctions.

5.1 Investment Decisions

An entrepreneur selects which projects to implement based on their predicted revenue. Denote the revenue by θ and its estimate by s. Implementing a project entails a cost of c. Denote a decision to implement a project by $a = 1$ and a decision to forgo it by $a = 0$. The entrepreneur wishes to implement a project if and only if $\theta \geq c$, and so $a^\star(\theta) = 1$ if $\theta \geq c$ and $a^\star(\theta) = 0$ otherwise.

The entrepreneur bases her decisions on a dataset that includes signals and actual revenues of implemented projects, i.e., $\phi(a) = a$. These projects were implemented by a set of entrepreneurs of which a share α are Bayesian and a share $1-\alpha$ use our heuristic, where $0 \leq \alpha \leq 1$. Denote the entrepreneur's perceived bias in an interior equilibrium by b_α when it exists (we show below that the interior equilibrium is unique). The naive entrepreneur predicts a revenue of $s - b_\alpha$ and, therefore, implements projects whose signals are higher than $c+b_\alpha$. Hence, in equilibrium, she uses an implementation cutoff of

$$s_\alpha = c + b_\alpha. \tag{8}$$

As a benchmark, note that a Bayesian entrepreneur implements a project if and only if $E[\theta|s] \geq c$. Since $E[\theta|s]$ is nondecreasing in s, there is a cutoff s_B such that a Bayesian entrepreneur implements a project if and only if $s \geq s_B$, where $E[\theta|s = s_B] = c$.

The next result characterizes the interior equilibrium bias as a function of[9] α.

Claim 1 *There exists at most one interior equilibrium. In such an equilibrium, (i) $s_\alpha \geq c + k$, (ii) $s_\alpha \geq s_B$, and (iii) s_α is decreasing in α.*

Parts (i) and (iii) follow from Propositions 2 and 4, respectively. Part (ii) shows that the naive entrepreneur's cutoff is higher than the Bayesian entrepreneur's cutoff. To obtain intuition for this effect, note that since $E[\theta|s = s_B] = c$, the Bayesian entrepreneur's correction at the cutoff is $E[s-\theta|s = s_B]$. Were the naive entrepreneur to use the Bayesian cutoff s_B (or, equivalently, use a dataset that Bayesian entrepreneurs generate), then her perceived bias would be $E[s-\theta|s \geq s_B]$, which is greater than $E[s-\theta|s = s_B]$ as $E[s - \theta|s]$ is nondecreasing. Proposition 4 implies that the entrepreneur's bias in equilibrium is even higher, which results in a more conservative implementation cutoff.

The result of Claim 1 makes clear that our entrepreneur's perceived bias in equilibrium, $E[s-\theta|s \geq s_\alpha]$, is higher than the optimal bias, $E[s-\theta|s_B]$. A related question is whether the entrepreneur's perceived bias is payoff-enhancing relative to taking signals at face value. Clearly, naive calibration can be better than taking signals at face value when $E[s] \neq E[\theta]$. However, it can also be payoff-enhancing when $E[s] = E[\theta]$, as we now illustrate. Suppose that $c = 1$, $\alpha = 0$, and $s = \theta + \epsilon$, where θ and ϵ are drawn independently from the standard normal distribution. A DM who takes signals at face value would set a cutoff of 1 while our DM would set a cutoff slightly lower than 2.34 in equilibrium. On average, projects whose estimated returns are between 1 and 2.34 yield less than 1, and so the entrepreneur's perceived bias is payoff-enhancing in this case.[10]

The results in this section run counter to the results obtained in Jehiel's (2018) model of investment decisions even though the setting is similar. In particular, in the equilibrium of his model, the entrepreneur uses an implementation cutoff that is too low relative to the cutoff of a Bayesian entrepreneur.

In Jehiel's model, the cutoff $s^\star$ generates a dataset $\mathcal{D}_{s^\star}$ of implemented projects. The dataset consists of the same projects as the dataset in our model given a cutoff

[9] Clearly, in every corner equilibrium, all projects are forgone.

[10] A Bayesian DM would set a cutoff of 2 under these parameters.

$s^\star$; however, the dataset in Jehiel (2018) does not include the signals that were used to select the projects. Rather, it includes only revenues. Therefore, Jehiel's entrepreneur samples a new signal $s'(\theta)$ for every project $\theta \in \mathcal{D}_{s^\star}$. To use the signal $s^\star$ at her disposal, the marginal entrepreneur considers a subset of projects $\{\theta \in \mathcal{D}_{s^\star} | s'(\theta) = s^\star\}$ (i.e., projects that are similar to the current project in the sense that their new signal is equal to $s^\star$). She then expects that the new project will result in a revenue of $E_{\mathcal{D}_{s^\star}}[\theta | s'(\theta) = s^\star]$. Since every project in $\mathcal{D}_{s^\star}$ was selected based on a signal $s \geq s^\star$ and $E[\theta|s]$ is increasing in s, it holds that $E_{\mathcal{D}_{s^\star}}[\theta | s'(\theta) = s^\star] \geq E[\theta|s^\star]$. Thus, the marginal entrepreneur in Jehiel's model is overoptimistic about the project's expected revenue relative to a conventional Bayesian entrepreneur. Therefore, she sets a lower implementation cutoff than such an entrepreneur would.

Other applications of our model: Medical referrals, recommendation systems, and credit provision

While we have been using investment decision terminology, the analysis in this section is relevant in other contexts as well. For example, the DM might be a physician who refers patients for a diagnostic test based on the result of a screening test (e.g., an Antigen test followed by a PCR test for Covid 19).[11] To evaluate the screening test score, she uses data that includes results (of both tests) of past patients. Alternatively, the DM might be an individual who uses a recommendation system to decide which products and services to consume and naively calibrates the recommendation she receives based on her actual enjoyment in previous situations in which she consumed these products and services. Finally, the DM might be a credit officer who approves credit applications based on a credit score that is calibrated based on the return rate of previous successful applications but not of unsuccessful ones. In all of these situations our results imply a seemingly pessimistic behavior: setting the bar too high for medical referrals and credit provision, and downgrading positive recommendations from an algorithm or a friend.

5.2 Second-Price IPV Auctions

While our baseline model considers the decision of a single DM (or a sequence of such DMs), its framework can be extended to strategic situations. In general, this requires

[11] A screening test provides information about the risk of a certain disorder or condition. It is typically less costly and so it is used by a larger group of patients. A diagnostic test establishes the existence of a condition or a disease.

extending the payoff and feedback functions, and making assumptions on how agents reason about other agents' behavior. However, when a game is dominance-solvable, as in a second-price IPV auction, such assumptions become moot.

We assume that there are $n \geq 2$ bidders who bid for an object, each of whom receives a signal s_i about the value she will derive from the object, θ_i. The value and its signal for each bidder are independently drawn from $F(\theta, s)$. Recall that, in a second-price IPV auction, bidding one's value is a dominant strategy, i.e., $a^\star(\theta_i) = \theta_i$. To bypass the problem of potentially negative bids, we assume that $\text{supp}(F) = [1, 2]^2$. Denote bidder i's perceived bias by b_i. Bidder i's predicted value is $\hat{\theta}(s_i) = s_i - b_i$. We assume that each bidder i uses the bidding function $a(s_i) = a^\star(\hat{\theta}(s_i)) = a^\star(s_i - b_i)$. Finally, a bidder learns her true valuation of the object if and only if she wins the object. Thus, $b_i = E[s_i - \theta_i | (s_i, \theta_i) \in \mathcal{D}_{win}]$, where $\mathcal{D}_{win}$ is the data collected when she wins. Thus, essentially, each bidder i is assumed to use the bidding rule $a(s_i) = s_i - E[s_i - \theta_i | (s_i, \theta_i) \in \mathcal{D}_{win}]$.

Following is the formal definition of an equilibrium in this game, which extends Definition 1.

Definition 3 *An equilibrium in a second-price IPV auction is a profile of bidding functions such that (i) the entire profile constitutes a Nash equilibrium in undominated strategies and (ii) each bidder's bidding function is part of an equilibrium at the individual level according to Definition 1.*

Next, we establish that the equilibrium is unique and that its induced bias is strictly positive. Moreover, it provides comparative statics with respect to the number of bidders.

Claim 2 *There exists a unique symmetric interior equilibrium. In such an equilibrium it holds that $b > k$. Moreover, b is increasing in n.*

In a symmetric (interior) equilibrium, a bidder obtains feedback (wins the object) if and only if her signal is the highest; i.e., she receives feedback with probability $F(s_i)^{n-1}$. This feedback function is independent of the bidders' perceived bias, and so there is a unique perceived bias that is consistent with our calibration procedure. Furthermore, the perceived bias in equilibrium is strictly greater than the inherent bias. This follows from the strict monotonicity of the bidding function. As a result, different signals lead to different frequencies of observing the actual realization, which precludes the possibility of a solution in which the additional bias, $b - k$, is null. Finally, the comparative statics with respect to the number of bidders follow directly from Proposition 3. To see this, recall that the feedback probability when there are n bidders is $F(s)^{n-1}$. This function is dominated in the likelihood ratio sense by $F(s)^{m-1}$ for $m > n$ bidders.

As in a second-price IPV auction with Bayesian bidders, the bidder with the highest signal wins the object. Thus, the equilibrium outcome is efficient. However, the naive calibration procedure affects the bidding strategy and, therefore, the division of surplus. We now turn to the auctioneer's perspective and compare her expected revenue in the case in which bidders are Bayesian to the case in which they use our heuristic.

Claim 3 *The auctioneer's revenue when bidders use the naive calibration procedure is lower than her revenue when bidders are Bayesian.*

This comparison is not obvious ex ante as our bidders' bids can be higher or lower than the ones submitted in a second-price IPV auction with Bayesian bidders (Corollary 2). To see why, denote the highest and second-highest signals by $s_{(n)}$ and $s_{(n-1)}$, respectively. When bidders are naive, the winner pays $s_{(n-1)} - b$, where the perceived bias b is the average difference between the highest signal and the expected value conditional on receiving the highest signal. Thus, a naive winner pays, on average, $E[s_{(n-1)}] - (E[s_{(n)}] - E[\theta|s_{(n)}])$. When bidders are Bayesian, the winner pays the expected value of the object given the second-highest signal, $E[\theta|s_{(n-1)}]$. Since $s - E[\theta|s]$ is increasing in s, and the distribution of $s_{(n)}$ first-order stochastically dominates the distribution of $s_{(n-1)}$, it holds that $E[s_{(n-1)} - E[\theta|s_{(n-1)}]] < E[s_{(n)} - E[\theta|s_{(n)}]]$. Therefore, on average, the naive winner bids less than the Bayesian one.

A corollary of Claim 3 is that, in expectation, if all bidders are naive, they obtain a higher payoff than they would obtain if they were all Bayesian. However, at an individual level (where the strategies of the other bidders are held fixed), the equilibrium prediction rule need not be optimal, even within the class of prediction rules of the form $\hat{\theta}(s) = s - b$. Note also that the restriction to additive prediction rules entails a loss. The optimal bid given complete knowledge of the joint distribution of signals and states is $E[\theta|s]$. However, a constant bias b that satisfies $s - b = E[\theta|s]$ for every s does not exist.

Our analysis of second-price IPV auctions is related to the winner's curse in IPV auctions identified by Compte (2002). In his model, bidders in a procurement auction rely on a noisy estimate of the cost. Due to selection bias, the estimate is likely to be lower than the actual cost conditional on winning. Compte analyzes the model under an assumption that bidders are constrained to use a coarse bidding function: they choose a fixed markup and add it to their estimated cost to correct for the selection bias. The bidders maximize their net payoff subject to the above constraint. In equilibrium, the markup is positive (to correct for selection) and increasing in the number of bidders. The intuition for the latter effect is that, as in our model, competition exacerbates the selection bias (i.e., the expected value of the noise term associated with the highest signal increases in the number of bidders)

There are two main differences between our model and the model in Compte (2002). First, bidders in the two models maximize different objective functions. While our bidders maximize the fit of their (constrained) prediction rule to the data, the bidders in Compte (2002) maximize their payoff subject to a fixed markup constraint. Second, in our model the dataset over which this objective is maximized is endogenous and (in equilibrium) biased toward high realizations of s. By contrast, in Compte (2002), bidders' markup is a standard (constrained) best response to the other bidders' behavior.

It is possible to apply our model to procurement auctions and compare the the optimal markup in Compte (2002) to the markup in our model, which equals $-b$ (since the bid in our model is $s - b$, the markup is $-b$). Proposition 4 in Compte (2002) establishes that the equilibrium markup in his model is $-E[(s_i - \theta_i)|s_i = \underline{s}_{-i}]$, where $\underline{s}_i$ is the smallest estimate among all bidders but i. In our model, the markup is $-E[s_i - \theta_i|s_i \leq \underline{s}_{-i}]$. Since $s_i - E(\theta_i|s_i)$ is nondecreasing, the markup is our model is weakly greater than the one in Compte (2002).[12]

6 Alternative Prediction Procedures

Throughout the analysis we examined the behavior of a DM who calibrates an additive prediction rule of the form $\hat{\theta}(s) = s - b$, such that b is the average difference between signals and state realizations in her dataset. We now examine two alternative prediction procedures. The first is motivated by *reference class forecasting*, which we discussed in

[12]This observation is not general. For instance, consider the feedback function $\tilde{\phi} = \epsilon\phi + (1-\epsilon)$, where ϕ is the feedback function used throughout this section. When ϵ is close to zero, our bidders obtain feedback almost every time they participate in the auction, and so the perceived bias in equilibrium is close to the inherent bias (zero, in this case). In this case, the equilibrium markup in our model is smaller than in Compte (2002), as his model is unaffected by the feedback function.

the Introduction. In this procedure, the DM calibrates a multiplicative prediction rule $\hat{\theta}(s) = \eta s$, such that η is the ratio between the expected state and the expected signal in her dataset. In the second procedure, the DM estimates a linear model $\hat{\theta}(s) = \beta_0 + \beta_1 s$ using a standard OLS regression. Both procedures are relatively simple and are widely used by academic researchers and practitioners in various areas.

Multiplicative prediction rule

In this procedure, when the DM receives a signal s, she makes a point prediction $\hat{\theta}(s) = \eta s$. She calibrates η to fit the ratio between the average realization and the average signal in her dataset, namely,

$$\eta_{\mathcal{D}} := \frac{E[\theta|(s,\theta) \in \mathcal{D}]}{E[s|(s,\theta) \in \mathcal{D}]}.$$

In order to show that the paper's main insight, namely, that in an interior equilibrium, on average, the naive DM's prediction is lower than the prediction of a Bayesian DM, we slightly modify one of the baseline modeling assumptions. Specifically, we replace the assumption that $s - E[\theta|s]$ is nondecreasing by an assumption that $s\frac{E[\theta]}{E[s]} - E[\theta|s]$ is nondecreasing. This assumption is satisfied, for instance, when $s = \theta + \epsilon$, and $E[\epsilon] = 0$. Due to the multiplicative structure of this procedure, it might be tempting to assume instead that the ratio $\frac{s}{E[\theta|s]}$ is nondecreasing. However, the fact that $\frac{s}{E[\theta|s]}$ is nondecreasing implies that $\frac{sE[\theta]}{E[s]} - E[\theta|s]$ is nondecreasing as well.

Let $m := \frac{E[\theta]}{E[s]}$. Note that $\eta \leq m$ is analogues to $b \geq k$ in the baseline model. We can interpret $1-m$ as the inherent bias and $1-\eta$ as the perceived bias in this case. The next result establishes that, in interior equilibria, the perceived bias is greater than the inherent bias. Furthermore, it suggests that the reference class forecasting procedure yields equilibrium beliefs that are pessimistic on average relative to Bayesian beliefs, as in our baseline model.

Claim 4 *In every interior equilibrium, $\eta \leq m$ and $E[\hat{\theta}] \leq E[\theta]$.*

Linear prediction rule

We now modify the baseline model by assuming that $\hat{\theta}(s) = \beta_0 + \beta_1 s$. We recall the baseline model's assumptions; that is, in order to estimate β_0 and β_1, the DM minimizes the quadratic errors over her endogenous dataset. Thus, given a dataset $\mathcal{D}$, the parameters that minimize the quadratic loss function (i.e., the OLS estimators) satisfy

$$\beta_0 = E[\theta|(s,\theta) \in \mathcal{D}] - \beta_1 E[s|(s,\theta) \in \mathcal{D}] \quad and \quad \beta_1 = \frac{cov(s,\theta|(s,\theta) \in \mathcal{D})}{var(s|(s,\theta) \in \mathcal{D})}.$$

The model studied in this section allows the DM to believe that some signals are upwardly biased while others are downwardly biased, unlike the previous two models that, by construction, impose that the DM either believes that all signals are upwardly biased or that all signals are downwardly biased. Therefore, we examine the average perceived bias by comparing the the naive DM's average prediction $E[\hat{\theta}]$ to a Bayesian DM's average prediction, $E[\theta]$. We show that whether the naive DM's average prediction is lower or higher than the Bayesian DM's average prediction depends on the particular shape of $E[\theta|s]$. To do so, we use the investment decisions setting of Section 5 and establish that if $E[\theta|s]$ is convex (resp., concave), then the naive DM's prediction is lower (resp., higher) than the Bayesian DM's prediction on average. Furthermore, we show that when $E[\theta|s]$ is convex (resp., concave), she sets an implementation cutoff that is too high (resp., low).

Claim 5 *Consider the investment decisions application of Section 5. If $E[\theta|s]$ is convex (resp., concave) then $E[\hat{\theta}] \leq E[\theta]$ (resp., $E[\hat{\theta}] \geq E[\theta]$), and the DM chooses an excessively high (resp., low) implementation cutoff.*

7 Concluding Remarks

We studied a model in which an agent has a misspecified model of the world that she calibrates based on a dataset that suffers from selection bias. The agent inadvertently contributes to this selection bias by taking actions that affect the data collection procedure. We show that the naive calibration procedure can generate substantial biases and exacerbate the misspecification errors. Our findings indicate that these errors are consistently in the same direction and result in a seemingly conservative behavior.

The naive calibration procedure may result in a lower payoff relative to no calibration at all. For instance, in the investment decision setting of Example 1, a DM who takes the estimates at face value implements more projects than our naive calibrator (i.e., sets a lower cutoff), and all these additional projects are profitable in expectation. If one interprets the attempt to calibrate the signal rather than taking it at face value as an indication of sophistication, then this example illustrates that a higher degree of sophistication may actually lead to a worse outcome. In recent years, it has been shown that in strategic interactions, players who are more sophisticated may obtain lower payoffs (e.g., Ettinger and Jehiel, 2010; Eyster and Piccione, 2013). We contribute to this literature by showing that a higher degree of sophistication may worsen outcomes in decision problems.

On the other hand, the naive calibration procedure can improve the DM's welfare even relative to a Bayesian DM in situations where the signals are provided by a strategic agent. As an illustration, consider a buyer who receives noisy information about the suitability or quality of various products advertised by a strategic seller. The latter might have an incentive to send biased signals and add noise to the transmitted information. However, a buyer who uses the naive calibration procedure not only corrects for the bias, but also interprets the noise as an additional systematic upward bias and corrects for that as well, which lowers her willingness to pay. Thus, a strategic seller who takes the DM's calibration procedure into account may have an incentive to provide the DM with the most precise information possible.

Our analysis has implications for real-world procedures that correct for optimism bias, such as reference class forecasting. These procedures typically consider the cost overruns in similar past projects and raise new cost forecasts accordingly. However, these procedures ignore the selection bias problem that is at the heart of our analysis. As long as the cost of such projects affects the decision to carry them out, our results imply that these adjustments may yield excessively high cost forecasts. Thus, the attempt to fully correct for cost overruns based on past projects may lead decision makers who rely on such forecasts to make suboptimal choices such as forgoing socially beneficial projects.

References

AACE International (2012): *Total Cost Management Framework: An Integrated Approach to Portfolio, Program, and Project Management.*

Bagnoli, M. and T. Bergstrom (2005): "Log-concave probability and its applications," *Economic Theory*, 26, 445–469.

Barron, K., S. Huck, and P. Jehiel (2024): "Everyday econometricians: Selection neglect and overoptimism when learning from others," *American Economic Journal: Microeconomics*, 162–198.

Battigalli, P. and D. Guaitoli (1997): "Conjectural equilibria and rationalizability in a game with incomplete information," *Decisions, Games, and Markets,*, 97–124.

Benjamin, D. J. (2019): "Chapter 2 - Errors in probabilistic reasoning and judgment biases," in *Handbook of Behavioral Economics - Foundations and Applications 2*, ed. by B. D. Bernheim, S. DellaVigna, and D. Laibson, North-Holland, vol. 2 of *Handbook of Behavioral Economics: Applications and Foundations 1*, 69–186.

Brashear, J., A. Becker, and D. Faulder (2001): "Where have all the profits gone?" *Journal of petroleum technology*, 53, 20–73.

Chen, M. and J. Dyer (2009): "Inevitable disappointment in projects selected based on forecasts," SPE Journal, vol. 14, 216–221.

Compte, O. (2002): "The winner's curse with independent private values," *Working Paper*.

Compte, O. and A. Postlewaite (2018): *Ignorance and Uncertainty*, Cambridge University Press.

Efron, B. (1965): "Increasing properties of Polya frequency function," *The Annals of Mathematical Statistics*, 36, 272–279.

Esponda, I. (2008): "Behavioral equilibrium in economies with adverse selection," *American Economic Review*, 98, 1269–1291.

Esponda, I. and D. Pouzo (2016): "Berk–Nash equilibrium: A framework for modeling agents with misspecified models," *Econometrica*, 84, 1093–1130.

Esponda, I. and E. Vespa (2018): "Endogenous sample selection: A laboratory study," *Quantitative Economics*, 9, 183–216.

Ettinger, D. and P. Jehiel (2010): "A theory of deception," *American Economic Journal: Microeconomics*, 2, 1–20.

Eyster, E. and M. Piccione (2013): "An approach to asset pricing under incomplete and diverse perceptions," *Econometrica*, 81, 1483–1506.

Eyster, E. and M. Rabin (2005): "Cursed equilibrium," *Econometrica*, 73, 1623–1672.

Flyvbjerg, B. (2008): "Curbing optimism bias and strategic misrepresentation in planning: Reference class forecasting in practice," *European planning studies*, 16, 3–21.

Flyvbjerg, B., C. Glenting, and A. Rønnest (2004): "Procedures for dealing with optimism bias in transport planning," *London: The British Department for Transport, Guidance Document*.

Fudenberg, D. and D. K. Levine (1993): "Self-confirming equilibrium," *Econometrica*, 523–545.

Heidhues, P., B. Kőoszegi, and P. Strack (2023): "Misinterpreting Yourself," *Available at SSRN 4325160*.

Heidhues, P., B. Kőszegi, and P. Strack (2018): "Unrealistic expectations and misguided learning," *Econometrica*, 86, 1159–1214.

HM Treasury (2003a): *The Green Book: appraisal and evaluation in central government: Treasury guidance*, HM Treasury.

——— (2003b): "Optimism Bias: Supplementary Green Book Guidance," .

Jehiel, P. (2005): "Analogy-based expectation equilibrium," *Journal of Economic Theory*, 123, 81–104.

——— (2018): "Investment strategy and selection bias: An equilibrium perspective on overoptimism," *American Economic Review*, 108, 1582–1297.

——— (2020): "Analogy-based expectation equilibrium and related concepts: Theory, applications, and beyond," Working Paper.

Kahneman, D. and D. Lovallo (1993): "Timid choices and bold forecasts: A cognitive perspective on risk taking," *Management Science*, 39, 17–31.

Meier, I. and V. Tarhan (2007): "Corporate investment decision practices and the hurdle rate premium puzzle," *Available at SSRN 960161*.

Nesvold, E. and R. B. Bratvold (2022): "Debiasing probabilistic oil production forecasts," *Energy*, 258, 124744.

Pemantle, R. (2007): "A survey of random processes with reinforcement," *Probability Surveys*, 4, 1–79.

Piccione, M. and A. Rubinstein (2003): "Modeling the economic interaction of agents with diverse abilities to recognize equilibrium patterns," *Journal of the European Economic Association*, 1, 212–223.

Poterba, J. M. and L. H. Summers (1995): "A CEO survey of US companies' time horizons and hurdle rates," *MIT Sloan Management Review*, 37, 43.

Smith, J. E. and R. L. Winkler (2006): "The optimizer's curse: Skepticism and postdecision surprise in decision analysis," *Management Science*, 52, 311–322.

Spiegler, R. (2016): "Bayesian networks and boundedly rational expectations," *The Quarterly Journal of Economics*, 131, 1243–1290.

——— (2020): "Behavioral implications of causal misperceptions," *Annual Review of Economics*, 12, 81–106.

UK Department for Transport and Oxford Global Projects (2020): *Updating the Evidence Behind the Optimism Bias Uplifts for Transport Appraisals*.

Van den Steen, E. (2004): "Rational overoptimism (and other biases)," *American Economic Review*, 94, 1141–1151.

Appendix: Proofs

Proof of Proposition 1. Suppose that the DM plays the strategy $a^\star(s-b)$. If there exists a bias b such that $\phi(a^\star(s-b)) = 0$ for every signal s, then there exists an equilibrium in which the bias is b and $D = \emptyset$. Otherwise, the DM's dataset is nonempty and the DM's strategy induces a perceived bias of

$$T(b) = \frac{\int_{-\infty}^{\infty} f(s)\phi(a^\star(s-b))[s - E[\theta|s]]ds}{\int_{-\infty}^{\infty} f(s)\phi(a^\star(s-b))ds}. \tag{9}$$

To establish the existence of an equilibrium, we show that there exists a bias b for which $T(b) = b$. Note that $\lim_{b\to-\infty} \phi(a^\star(s-b)) = \phi(\overline{a}) \in (0,1]$ for every s, and so

$$\lim_{b\to-\infty} T(b) - b = \frac{\int_{-\infty}^{\infty} f(s)[s - E[\theta|s]]ds}{\int_{-\infty}^{\infty} f(s)ds} - b = E[s] - E[\theta] - b = \infty.$$

Since $s - E[\theta|s]$ is nondecreasing, $T(b)$ is no higher than

$$T^x(b) := \frac{\int_{x+b}^{\infty} f(s)\phi(a^\star(s-b))[s - E[\theta|s]]ds}{\int_{x+b}^{\infty} f(s)\phi(a^\star(s-b))ds}. \tag{10}$$

Moreover, $T^x(b)$ is nondecreasing in x. Fix a small $\epsilon > 0$. There exists a sufficiently large x such that for every pair $s, s' > x$, it holds that $|\phi(a^\star(s)) - \phi(a^\star(s'))| < \epsilon$. Hence, for such x it holds that $T^x(b)$ is arbitrarily close to $E[s - \theta|s \geq x + b]$ for every b.

Log-concavity of the signals' distribution implies that $E[s|s \geq x + b] - x - b$ is decreasing in b (Bagnoli and Bergstrom, 2005). Hence, for a sufficiently large b we have that $E[s|s \geq x+b] - b - E[\theta|s \geq x+b] < 0$. Thus, $E[s-\theta|s \geq x+b] < b$ and, therefore, $T(b) < T^x(b) < b$. By the intermediate value theorem there exists b for which $T(b) = b$.

Proof of Proposition 2. Consider an interior equilibrium and recall that the equilibrium bias b satisfies $b = T(b)$ (see (9)). Let $s^\star$ denote an arbitrary signal that satisfies $s^\star - E[\theta|s^\star] = k$. It follows that

$$\int_{-\infty}^{s^\star} f(s)\phi(a^\star(s-b))[s - E(\theta|s) - k]ds \geq \int_{-\infty}^{s^\star} f(s)\phi(a^\star(s^\star-b))[s - E(\theta|s) - k]ds$$

and

$$\int_{s^\star}^{\infty} f(s)\phi(a^\star(s-b))[s - E(\theta|s) - k]ds \geq \int_{s^\star}^{\infty} f(s)\phi(a^\star(s^\star-b))[s - E(\theta|s) - k]ds.$$

The sum of the RHS of the two inequalities is 0 as $E(s) - E(\theta) = k$. Thus,

$$\frac{\int_{-\infty}^{\infty} f(s)\phi(a^\star(s-b))[s - E(\theta|s)]ds}{\int_{-\infty}^{\infty} f(s)\phi(a^\star(s-b))ds} \geq \frac{k\int_{-\infty}^{\infty} f(s)\phi(a^\star(s-b))ds}{\int_{-\infty}^{\infty} f(s)\phi(a^\star(s-b))ds} = k.$$

To conclude the proof, note that the LHS is equal to $T(b)$ in an interior equilibrium.

Proof of Proposition 3. To prove this result, we show that $T_\phi(b) \geq T_{\tilde{\phi}}(b)$ for every b, where T_ϕ (resp., $T_{\tilde{\phi}}$) denotes the operator T when the feedback function is ϕ (resp., $\tilde{\phi}$). Observe that $T_\phi(b) \geq T_{\tilde{\phi}}(b)$ if and only if

$$\text{(11)} \quad \frac{\int_{-\infty}^{\infty} f(s)\phi(a^\star(s-b))[s - E[\theta|s]]ds}{\int_{-\infty}^{\infty} f(s)\phi(a^\star(s-b))ds} \geq \frac{\int_{-\infty}^{\infty} f(s)\tilde{\phi}(a^\star(s-b))[s - E[\theta|s]]ds}{\int_{-\infty}^{\infty} f(s)\tilde{\phi}(a^\star(s-b))ds}.$$

Without loss of generality, we can assume that the feedback functions satisfy

$$\text{(12)} \qquad \int_{-\infty}^{\infty} f(s)\phi(a^\star(s-b))ds = \int_{-\infty}^{\infty} f(s)\tilde{\phi}(a^\star(s-b))ds.$$

Since $\frac{\phi(a^\star(s-b))}{\tilde{\phi}(a^\star(s-b))}$ is nondecreasing in s, (12) implies that there exists $s^\star$ such that $\phi(a^\star(s^\star - b)) \geq \tilde{\phi}(a^\star(s^\star - b))$ for $s > s^\star$ and the inequality is reversed for $s < s^\star$. Using the normalization in (12), we can write (11) as

$$\text{(13)} \qquad \int_{-\infty}^{s^\star} f(s)\left(\phi(a^\star(s-b)) - \tilde{\phi}(a^\star(s-b))\right)[s - E[\theta|s] - k]ds + \int_{s^\star}^{\infty} f(s)\left(\phi(a^\star(s-b)) - \tilde{\phi}(a^\star(s-b))\right)[s - E[\theta|s] - k]ds \geq 0.$$

Since $s - E[\theta|s]$ is nondecreasing in s, the LHS of (13) is higher than

$$\text{(14)} \qquad \int_{-\infty}^{s^\star} f(s)\left(\phi(a^\star(s-b)) - \tilde{\phi}(a^\star(s-b))\right)[s^\star - E[\theta|s^\star] - k]ds + \int_{s^\star}^{\infty} f(s)\left(\phi(a^\star(s-b)) - \tilde{\phi}(a^\star(s-b))\right)[s^\star - E[\theta|s^\star] - k]ds,$$

which, by (12), is equal to zero.

Proof of Proposition 4. In an interior equilibrium, $T_\alpha(b) = b$, where

$$T_\alpha(b) := \frac{\int_{-\infty}^{\infty} f(s)[(1-\alpha)\phi(a^\star(s-b)) + \alpha\phi(a^\star(E[\theta|s])](s - E[\theta|s])ds}{\int_{-\infty}^{\infty} f(s)[(1-\alpha)\phi(a^\star(s-b)) + \alpha\phi(a^\star(E[\theta|s])]ds}.$$

Note that $T_\alpha(b) = b$ if and only if

$$\int_{-\infty}^{\infty} f(s)g_\alpha(s)(s - b - E[\theta|s])ds = 0, \tag{15}$$

where

$$g_\alpha(s) = (1-\alpha)\phi(a^\star(s-b)) + \alpha\phi(a^\star(E[\theta|s])).$$

Assume that $T_\alpha(b^\star) = b^\star$, i.e., $b^\star$ is part of an interior equilibrium. Consider a signal $s^\star$ such that $s^\star - b^\star = E[\theta|s^\star]$. Note that $g_\alpha(s^\star)$ is independent of α. Moreover, by the monotonicity of $\phi(a(\cdot))$, it follows that $g_\alpha(s)$ is increasing (resp., decreasing) in α for every $s < s^\star$ (resp., $s > s^\star$). Since $s - b^\star - E[\theta|s] \leq 0$ if, and only if, $s \leq s^\star$, the LHS of (15) is nonpositive for $\alpha' > \alpha$. That is, $T_{\alpha'}(b^\star) \leq b^\star$ for $\alpha' > \alpha$. Similarly, $T_{\alpha'}(b^\star) \geq b^\star$ for $\alpha' < \alpha$.

Fix α and consider $\underline{b}_\alpha$. Let $\alpha' > \alpha$. Since $T_{\alpha'}(\underline{b}_\alpha) \leq \underline{b}_\alpha$, $T_{\alpha'}(k) \geq k$, and $T_{\alpha'}(\underline{b}_{\alpha'}) = \underline{b}_{\alpha'}$ by continuity, it holds that $k \leq \underline{b}_{\alpha'} \leq \underline{b}_\alpha$. Next, consider $\overline{b}_\alpha$. By previous arguments, $T_\alpha(\overline{b}_{\alpha'}) \geq \overline{b}_{\alpha'}$. Note that $T_\alpha(b) < b$ for $b > \overline{b}_\alpha$ since $\lim_{b\to\infty} T_\alpha(b)$ is finite. Therefore, $\overline{b}_{\alpha'} \leq \overline{b}_\alpha$.

Proof of Proposition 5. We start by assuming that $\phi(a^\star(s)) > 0$ for every a. Under this assumption, (7) is well defined. For ease of notation, we shall assume that time advances if and only if new data is recorded in the dataset. Under that assumption, $|\mathcal{D}_t| = t$.

Let $\mathcal{F}(b_t) = T(b_t) - b_t$ and $\xi_{t+1} = s_{t+1} - \theta_{t+1} - T(b_t)$. Thus,

$$b_{t+1} - b_t = \frac{1}{t+1}(\mathcal{F}(b_t) + \xi_{t+1}),$$

where $E(\xi_{t+1}|b_t) = 0$. Thus, b_t is a stochastic approximation process according to the definition in Equation (2.6) in Pemantle (2007). Moreover, $E[\xi_{t+1}^2|b_t] \leq K$ for some finite K, since F is bounded. Furthermore, $\mathcal{F}$ is bounded. Hence, by Lemma 2.6 in Pemantle (2007), with probability 1, the process b_t visits any closed segment on which $\mathcal{F}$ is bounded sway from 0 only a finite number of times. Since $T(b_t)$ is continuous in b_t, by Corollary 2.7 in Pemantle (2007), the sequence of biases converges almost surely to the zero set of $\mathcal{F}$, namely, to a point b where $T(b) = b$, which is an interior equilibrium of our model.

Now, assume that $\phi(a^\star(s)) = 0$ for some s. Let $\tilde{\xi}_{t+1} = s_{t+1} - \theta_{t+1} - T(b_t)$ if $\int_{-\infty}^{\infty} f(s)\phi(a^\star(s - b_t))ds > 0$ and $\tilde{\xi}_{t+1} = 0$ otherwise.[13] Thus,

$$b_{t+1} - b_t = \frac{1}{t+1}(\mathcal{F}(b_t) + \tilde{\xi}_{t+1}).$$

Note that $T(b_t)$ is continuous everywhere except for $b_t = b'$, where

$$b' = \arg\min_b \int_{-\infty}^{\infty} f(s)\phi(a^\star(s - b))ds = 0.$$

Furthermore, $T(b_t)$ is bounded since F has a bounded support and $T(b_t)$ is a weighted average of $s - E[\theta|s]$. Moreover, $E[\tilde{\xi}^2_{t+1}|b_t] \leq K$ for some finite K, since F is bounded. The process b_t satisfies the requirements for Lemma 2.6 in Pemantle (2007), and converges almost surely to the zero set of $\mathcal{F}$.

Proof of Claim 1. Note that when $\alpha = 1$ it holds that $\mathcal{D} = \{(s,\theta)|s \geq s_B\}$. Moreover, there exists a unique b_1 that solves $b_1 = E[s - \theta|s \geq s_B]$. Since $E[s - \theta|s]$ is nondecreasing and $E[s] - E[\theta] = k$, it follows that $b_1 \geq k$. Since $c = s_B - E[s - \theta|s = s_B]$, it follows that

$$s_B = c + E[s - \theta|s = s_B] \leq c + E[s - \theta|s \geq s_B] = c + b_1 = s_1.$$

By Proposition 4, $\underline{b}_1 \leq \underline{b}_\alpha$ for any $\alpha < 1$. Hence, $c + \underline{b}_\alpha \geq c + k$ and $c + \underline{b}_\alpha \geq s_B$ for any α.

We now assume that b_α is part of an interior equilibrium and show that there cannot exist another interior equilibrium with a perceived bias $b'_\alpha > b_\alpha$. Let γ be the share of observations in the data that were induced by Bayesian DMs in the equilibrium in which the bias is b_α. That is,

$$\gamma_{b_\alpha} = \frac{\alpha(1 - F(s_B))}{\alpha(1 - F(s_B)) + (1 - \alpha)(1 - F(c + b_\alpha))}.$$

[13] We assume that if the process reaches some b_t such that $\int_{-\infty}^{\infty} f(s)\phi(a^\star(s - b_t))ds = 0$, then time progresses even though no new data is recorded. This allows us to apply the exact same argument as in the first part of the proof.

Note that $\gamma_{b'_\alpha} > \gamma_{b_\alpha}$ for $b'_\alpha > b_\alpha$. In an interior equilibrium,

$$T(b_\alpha) - b_\alpha = \gamma_{b_\alpha}(E[s-\theta|s \geq s_B] - b_\alpha)) + (1-\gamma_{b_\alpha})(E[s-\theta|s \geq c + b_\alpha] - b_\alpha) = 0.$$

Moreover, $E[s-\theta|s \geq c+b_\alpha] \geq E[s-\theta|s \geq s_B]$, and so

$$\gamma_{b'_\alpha}(E[s-\theta|s \geq s_B] - b_\alpha)) + (1-\gamma_{b'_\alpha})(E[s-\theta|s \geq c + b_\alpha] - b_\alpha) \leq 0.$$

Due to the log-concavity of the signals' distribution, $E[s-\theta|s \geq c+b_\alpha] - b_\alpha$ is decreasing in b_α. Clearly, $E[s-\theta|s \geq s_B] - b_\alpha$ is strictly decreasing in b_α. We conclude that $T(b'_\alpha) - b'_\alpha < 0$, in contradiction to the assumption that b'_α is part of an interior equilibrium.

Proof of Claim 2. Since agents' strategies are symmetric, we can write the bias as

$$b = \frac{\int_1^2 f(s)F(s)^{n-1}[s - E(\theta|s)]ds}{\int_1^2 f(s)F(s)^{n-1}ds}. \tag{16}$$

As the RHS of (16) is independent of b, it has a unique solution. Hence, there exists a unique interior equilibrium. Since (i) $E[s] = E[\theta] + k$, (ii) $s - E[\theta|s]$ is nondecreasing and nondegenerate, and (iii) $F(s)$ is *strictly* increasing, it follows that $b > k$ and, therefore, $a(s) = a^\star(s-b) = s - b < s - k$. Since $\phi_n(a^\star(s-b_n)) = F(s)^{n-1}$, it follows that $\frac{\phi_{n+1}(\cdot)}{\phi_n(\cdot)} = F(s)$ is increasing in s and, by Proposition 3, the proof is complete.

Proof of Claim 3. Let $s_1, s_2, ..., s_n$ be a random sample of signals of size n. Denote the k'th order statistic by $s_{(k)}$ and its distribution by $f_k(\cdot)$. Conditional on the bidder winning the auction, the expected values of the object and the signal are $\int_1^2 E[\theta|s]f_n(s)ds$ and $\int_1^2 sf_n(s)ds$, respectively. Thus, the equilibrium bias when bidders are naive is

$$b = \int_1^2 sf_n(s)ds - \int_1^2 E[\theta|s]f_n(s)ds. \tag{17}$$

Since the bidding function is $s - b$, the auctioneer's expected revenue is $E[s_{(n-1)}]$ net of the bias, that is,

$$\int_1^2 sf_{n-1}(s)ds - b. \tag{18}$$

A Bayesian bidder bids the expected value of the object given her signal. If all bidders were Bayesian, then the winner would pay the the expected value of θ given $s_{(n-1)}$,

$$\int_1^2 E[\theta|s] f_{n-1}(s) ds. \tag{19}$$

By (17), (18), and (19), the auctioneer's revenue is higher when agents are Bayesian if and only if

$$\int_1^2 (E[\theta|s] - s) f_{n-1}(s) ds > \int_1^2 (E[\theta|s] - s) f_n(s) ds. \tag{20}$$

Condition (20) holds since $s_{(n)}$ first-order stochastically dominates $s_{(n-1)}$, and $E[\theta|s]-s$ is nonincreasing in s.

Proof of Claim 4. Consider an interior equilibrium. Note that $\eta \leq m$ if

$$\frac{\int_{-\infty}^{\infty} f(s)\phi(a^\star(\eta s))E[\theta|s]ds}{\int_{-\infty}^{\infty} f(s)\phi(a^\star(\eta s))sds} \leq m.$$

Rearranging yields

$$\int_{-\infty}^{\infty} f(s)\phi(a^\star(\eta s))[sm - E[\theta|s]]ds \geq 0.$$

Observe that $\int_{-\infty}^{\infty} f(s)[sm - E[\theta|s]]ds = 0$. Since, by assumption, $\phi(\cdot)$, $a^\star(\cdot)$, and $s\frac{E[\theta]}{E[s]} - E[\theta|s]$ are weakly increasing, the above inequality holds. Thus,

$$\int_{-\infty}^{\infty} f(s)\hat{\theta}(s)ds \leq m \int_{-\infty}^{\infty} f(s)sds = mE[s] = E[\theta].$$

Proof of Claim 5. The DM uses a prediction rule $\hat{\theta}_{linear}(s) = \beta_0 + \beta_1 s$, and launches a project if and only if $\hat{\theta}(s) \geq c$. In equilibrium, there is a cutoff signal $s^\star$ such that the DM launches a project if and only if $s \geq s^\star$ and, therefore, her dataset includes only signals higher than $s^\star$.

Suppose that $E[\theta|s]$ is convex in s. The calibrated function $\hat{\theta}(s) = \beta_0 + \beta_1 s$ intersects with $E[\theta|s]$ at exactly two points, both of which are weakly higher than $s^\star$. To see this, note that if the two functions intersect at a single point, then rotating the calibrated function around the intersection point would decrease the squared errors (pointwise).

If the functions do not intersect at all, bringing them closer together by changing β_0 would clearly decrease the squared errors. Denote these intersection points by s_1 and s_2.

Since the calibrated function $\hat{\theta}(s) = \beta_0 + \beta_1 s$ intersects with $E[\theta|s]$ twice above the cutoff $s^\star$ and $E[\theta|s]$ is convex, we have that $E[\theta|s^\star] \geq \beta_0 + \beta_1 s^\star$. This means that while the DM expects a revenue of c at the cutoff, the actual revenue is higher. Thus, a Bayesian DM would choose a lower implementation cutoff. Moreover, $E[\theta|s] > \beta_0 + \beta_1 s$ for every $s < s^\star$. Hence, $\beta_0 + \beta_1 E[s|s < s^\star] < E[\theta|s < s^\star]$. By definition, the DM's prediction is correct in the selected sample, and so $\beta_0 + \beta_1 E[s|s \geq s^\star] = E[\theta|s > s^\star]$. Combining this with the out-of-sample prediction, we conclude that overall $E[\hat{\theta}] = \beta_0 + \beta_1 E[s] < E[\theta]$, as in our baseline model.

In an analogous manner, if $E[\theta|s]$ is concave, the same argument implies that $E[\hat{\theta}] = \beta_0 + \beta_1 E[s] > E[\theta]$ and that the DM's implementation cutoff is lower than the Bayesian cutoff.