# The Strategy of Professional Forecasting*

**Marco Ottaviani**† **Peter Norman Sørensen**‡

April 2004

**Abstract**

This paper develops and compares two theories of strategic behavior of professional forecasters. The first theory posits that forecasters compete in a *forecasting contest* with pre-specified rules. In equilibrium of a winner-take-all contest, forecasts are excessively differentiated. According to the alternative *reputational cheap talk* theory, forecasters aim at convincing the market that they are well informed. The market evaluates their forecasting talent on the basis of the forecasts and the realized state. If the market expects forecaster honesty, forecasts are shaded toward the prior mean. With correct market expectations, equilibrium forecasts are imprecise but not shaded.

*Keywords:* Forecasting, contest, reputation, cheap talk.

*JEL Classification:* C72 (Noncooperative Games), D82 (Asymmetric and Private Information), D83 (Search, Learning, and Information).

---

*We thank Dean Croushore for essential feedback in the early stages of this project, Owen Lamont for generously giving us the *Business Week Investment Outlook* data along with very useful suggestions, Marco Battaglini for valuable discussion of the paper, Eric Zitzewitz for extremely helpful feedback, Andrew Chesher, Christian Dustmann, Douglas Gale, Peter R. Hansen, Harrison Hong, Boyan Jovanovic, Jon Levin, John Morgan, Kent Osband, Mike Peters, Ilya Segal, Daniel Seidmann, Hyun Shin, Gianluca Violante, and Jeffrey Zwiebel for useful input, and seminar audiences at Aarhus, American Economic Association (New Orleans), Bank of Italy, Bergen, Berlin, Bonn, Brussels, CEMFI (Madrid), Cambridge, Copenhagen, Danish Economic Society 2000, European Economic Association (Stockholm), ENTER 2001 (Mannheim), Erice Conference on Stochastic Methods in Decision and Game Theory with Applications, ESRC-PIPPS (Bristol), ESSET 2000 (Gerzensee), European Central Bank, EUI (Florence), Helsinki, INSEAD, Koç University, LBS, LSE, Munich, NYU Stern School, Paris (Roy Seminaire), Pompeu Fabra, Princeton, Royal Holloway, Southampton, Stanford, Surrey, and UCL for comments.

†Marco Ottaviani, Economics Subject Area, London Business School, Sussex Place, Regent's Park, London NW1 4SA, England. Phone: +44-20-7262-5050 ext. 3778. Fax: +44-20-7402-0718. E-mail: `mottaviani@london.edu`. Web: http://faculty.london.edu/mottaviani/.

‡Institute of Economics, University of Copenhagen, Studiestræde 6, DK–1455 Copenhagen K, Denmark. Phone: +45–3532–3056. Fax: +45–3532–3000. E-mail: `peter.sorensen@econ.ku.dk`. Web: http://www.econ.ku.dk/sorensen.

*The best is the enemy of the good. (Voltaire)*

# 1. Introduction

Professional forecasts guide market participants and inform their expectations about future economic conditions. Given the importance of this role and the potential rewards of accurate forecasting, one might expect that professional forecasters maximize their accuracy by truthfully releasing all their information.[1] As reported by Keane and Runkle (1998), "since financial analysts' livelihoods depend on the accuracy of their forecasts [...] we can safely argue that these numbers accurately measure the analysts' expectations." However, a number of commentators have suggested the contrary, arguing that forecasters might strategically misreport their information, even when they are not interested in manipulating the investment decisions of their target audience.

As argued for example by Croushore (1997), "some [survey] participants might shade their forecasts more toward the consensus (to avoid unfavorable publicity when wrong), while others might make unusually bold forecasts, hoping to stand out from the crowd." Following the recent evidence of the relevance of microeconomic incentives in forecasting (e.g., Graham (1999), Hong, Kubik and Solomon (2000), Lamont (2002), Welch (2000), Zitzewitz (2001a)), this paper develops the positive theory of strategic behavior of professional forecasters. As professional forecasts are often used as proxies for the unobservable expectations of market participants, our results have also implications for empirical tests of investment behavior.

To understand the basic ingredients of our model, consider Figure 1 depicting yearly GNP growth forecasts and realizations[2] during the period 1972-1993 from the *Business Week Investment Outlook*, as collected by Lamont.[3] The plot immediately reveals that there is substantial dispersion in the individual forecasts. Hence, our model assumes that forecasters are privately informed. Forecasts are more dispersed in some years, e.g., 1974 (in the aftermath of the oil shock), 1982-3, and 1991. To account for the variation in the forecast dispersion across years, the quality of private information and the precision of the prior are parameters in our model.

[1] Those that are successful in predicting the future are rewarded by markets and governments alike. For example, both Alan Greenspan and Laurence Meyer practiced economic forecasting before being appointed to the Board of Governors of the Federal Reserve System.

[2] A perennial problem in evaluating forecasts is that data on the realized values are revised over time. See Section 6.4 for more on this.

[3] This data is taken from a survey of professional forecasters run at the end of each year by *Business Week*. The economists belonging to the panel are asked to predict output, inflation, and unemployment for the subsequent year.

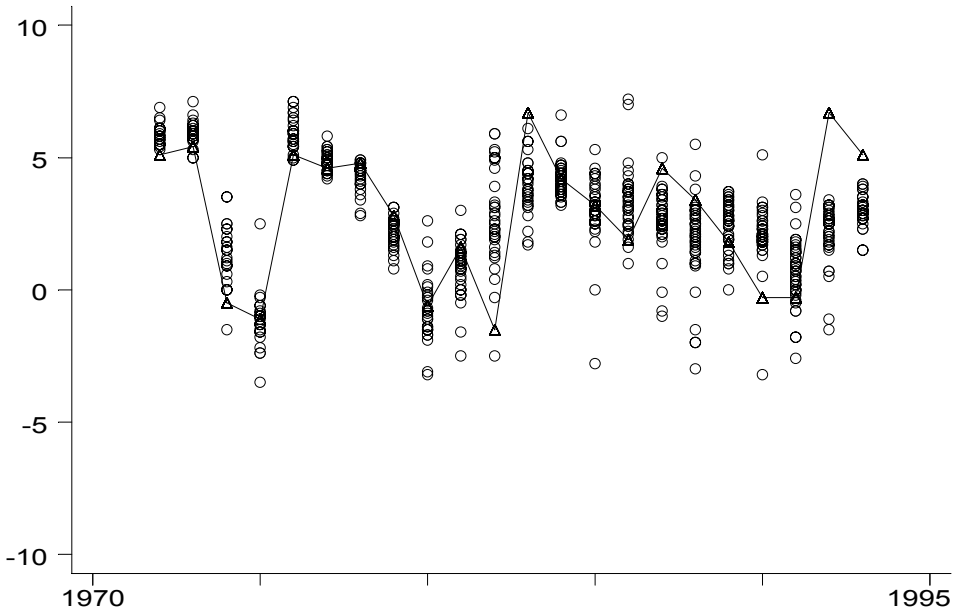


**Figure 1**. The circles represent individual forecasts of annual real GNP growth rate from the *Business Week Investment Outlook* survey for the period 1972-1993 and the connected triangles the realized values. Data source: Lamont (2002).

We formulate two distinct theories of strategic forecasting and contrast them with the benchmark case of non-strategic forecasting. In order to facilitate the comparison, we adopt a unified and tractable statistical model. The *state* has a normal prior distribution and the *signals* of the forecasters are normally distributed around the state. After the forecasters simultaneously release their forecasts, the state is publicly observed. In order to isolate the effect of the *professional* objectives of forecasters predicting the future evolution of economic or financial variables, we assume that these forecasters cannot affect the distribution of the state variable and do not care about the investment decisions taken on the basis of their forecasts. We therefore abstract from the additional strategic incentives relevant to partisan forecasters, whose payoff instead depends on the investment decisions made on the basis of their forecasts.[4]

Consider the benchmark case of a forecaster rewarded according to the *absolute* accuracy of the prediction. A non-strategic forecaster reports honestly the posterior expectation

[4]For empirical evidence of *partisan bias* of equity analysts see e.g. Michaely and Womack (1999) and Hong and Kubik (2003) and references therein and for theoretical investigations see e.g. Morgan and Stocken (2003). The study of the interaction of professional and partisan objectives is left to future research.

of the state, a weighted average of the signal and the prior mean. When the state turns out to be above the prior mean, the honest forecast tends to be lower than the realized state. This empirically documented negative correlation of the forecast errors with the realized state is often taken as evidence of herd behavior by the popular press. The academic empirical literature avoids this misconception and focuses instead on the correlation between forecasts and their errors.[5] This correlation is zero for "rational" forecasts in Muth's (1961) sense, because they are equal to conditional expectations.[6] The two theories of strategic forecasting developed here have different implications for this correlation.

Our first theory of strategic behavior posits that forecasters compete in a *forecasting contest* with pre-specified rules.[7] Forecasters are often ranked by their *relative* accuracy, see e.g. the semi-annual *Wall Street Journal Forecasting Survey* (macroeconomic variables), *WSJ All Star Analysts* (earnings) and *WSJ Best on the Street* (stock picking) competition.[8] We find that reporting the best predictor on the state (the posterior expectation conditional on the signal observed) is not an equilibrium in the contest. With an infinite number of forecasters, the equilibrium strikes a balance between two contrasting forces. First, an individual forecaster has an incentive to report the honest forecast, which is most likely to be on the mark. Second, one gains from moving away from the prior because the number of forecasters correctly guessing the state is lower the farther the state is from the prior mean. In equilibrium, forecasters differentiate their prediction from those of competitors by putting greater weight on their private signal than they would in an honest report of the posterior expectation. Yet, rational market participants can in principle invert the equilibrium strategy to recover the forecasters' information thereby constructing accurate expectations on the state.

According to the *reputational cheap talk* theory, forecasters wish to foster their reputation for being well informed.[9] In this second theory, forecasts and realized state are used by the market to evaluate forecasting talent.[10] Somewhat counter-intuitively, honest

[5] A number of studies have used professional forecasts to test the rational expectation hypothesis (e.g., Keane and Runkle (1990) and (1998)).

[6] Otherwise, a revised forecast would be more accurate.

[7] See Section 4 for a discussion of the connection with Laster, Bennett and Geoum (1999) and other related Hotelling location models.

[8] See Stekler (1987) for an early study of the relative accuracy of forecasts. A number of rankings are available on line. For example, Validea (www.validea.com), BigTipper.com (www.bigtipper.com) and BulldogResearch.com (www.bulldogresearch.com) track stock recommendations made by Wall Street professionals and then rank the analysts based on the performance of their selections. Forecasting contests are run also for non-economic variables (see e.g., the *National Collegiate Weather Forecasting Contest*).

[9] See Section 5 for a discussion of the connection with Holmström (1999), Scharfstein and Stein (1990), and other reputational models.

[10] Forecasters are assumed not to have private information about the talent prior to receiving the signal. The importance of this assumption is discussed at the end of Section 5.2.

forecasting cannot occur in equilibrium. Since the conditional expectation lies between the signal and the prior mean, the state is expected to be closer to the prior than the private signal is. But because the signals of more talented forecasters are on average closer to the state, there is an incentive to pretend to have received a signal equal to the posterior expectation of the state. When the market naively expects honest forecasting, the best deviation is to release a forecast even closer to the prior mean. This accords with the intuitive implication of career concerns suggested by recent empirical work. However, in equilibrium, the market must have rational expectations regarding the forecasters' behavior. The incentive to herd on the prior is self defeating. Equilibrium forecasts are not shaded, but are systematically less precise than if forecasters were not strategic. The analysts' desire to be perceived as good forecasters turns them into poor forecasters.[11]

We derive implications on how the forecast dispersion varies with the quality of private information and precision of the prior in the different models. In the symmetric equilibrium of our forecasting contest the forecast error is positively correlated with the forecast, while the correlation is negative in the reputational deviation. The reputational equilibrium forecast may be uncorrelated with its error, but is not efficient. The theory is also extended to allow for ex post innovation on the state. According to the empirical results of Zitzewitz (2001a), forecast errors exhibit a strong positive correlation with the forecast errors, consistently with our contest theory.

We believe that forecasting is a particularly apt laboratory for improving our understanding of strategic communication and positioning by non-partisan informed agents. Our theorizing is inspired and disciplined by the availability of data and the richness of institutional details, and can lead to further empirical testing. The insights gained in the analysis can be helpful in shedding light on other social and economic problems in which agents want to appear to be well informed, such as the choice of research topic by scientists.

The paper is organized as follows. Section 2 sets up the statistical model. Section 3 develops the benchmark case of honest forecasting. Section 4 introduces the forecasting contest theory, Section 5 the reputational cheap talk theory. Section 6 compares the predictions of the different theories and develops some extensions. Section 7 concludes by summarizing the results and discussing avenues for future research.

[11] Ironically, *The Economist* ("Dustmen as Economic Gurus," 3 June 1995) reports the surprisingly good performance of a sample of London garbagemen in forecasting key economic variables.

## 2. Information

Our model features $n$ forecasters who simultaneously issue forecasts on an uncertain state of the world. It is common prior belief that the state $x$ is normally distributed with mean $\mu$ and precision $\nu$, $x \sim N(\mu, 1/\nu)$, with p.d.f. $q(x) = \sqrt{\nu/2\pi} \exp\left(-\nu (x-\mu)^2/2\right)$. Each forecaster $i$ observes the private signal $s_i = x + \varepsilon_i$. Conditionally on state $x$, signals $s_i$ are assumed to be independently normally distributed with mean $x$ and precisions $\tau_i$, $s_i|x \sim N(x, 1/\tau_i)$, with the p.d.f. $g_i(s_i|x) = \sqrt{\tau_i/2\pi} \exp\left(-\tau_i (s_i - x)^2/2\right)$. Forecaster $i$'s observation of signal $s_i$ leads to a normal posterior belief on the state with mean $E[x|s_i] = (\tau_i s_i + \nu\mu)/(\tau_i + \nu)$ and precision $\tau_i + \nu$ (cf. DeGroot (1970)). The p.d.f. of this posterior distribution is denoted by $q_i(x|s_i)$. After the forecasts are released, the state of the world $x$ is realized and observed perfectly.[12]

It is natural to allow forecasters to have private information about the state, because actual forecasts are typically dispersed. Forecasters who share a common prior belief and are given the same (public) information without private information should make identical forecasts if they honestly report their expectations.[13] Forecast dispersion could also be due to heterogenous prior beliefs or different models. According to industry participants, forecasters seem to have access to the same pool of public data, but interpret it differently depending on their model.[14] In any case, forecasters have posterior beliefs that we assume to be private. These can be reinterpreted as deriving from private information of the model used to process the same public information. The observed forecast dispersion might also be the outcome of strategic behavior. In order not to introduce a bias against honest forecasting, we posit that forecasters are endowed with some private information about the state.[15]

For simplicity, we have assumed that forecasts have no impact on the distribution of the state or on the amount of information observed about its realization. Our model can be thus applied to situations in which the state cannot be affected by the forecasts and yet can be meaningfully forecasted and later observed.[16]

[12] Section 6.4 extends the model to allow for noisy observation of the state or, equivalently, ex-post innovation on the state.

[13] In their classic study on the rationality of forecasts using data from the NBER-ASA survey of professional forecasters (later to be called the Survey of Professional Forecasters), Keane and Runkle (1990) found that differences in individual forecasts cannot be explained by publicly available information. They inferred that differences in forecasts are due to asymmetric information, but this conclusion rests on the maintained assumption of honest forecasting.

[14] Indeed, Kandel and Pearson (1995) and Kandel and Zilberfarb (1999) have found empirical support for heterogeneous processing of public information.

[15] After all, the presence of heterogeneous private information is the very reason for the market to reward forecasters according to their (absolute or relative) accuracy.

[16] Additional strategic incentives would be present if forecasters were concerned with the effect of their

To simplify the presentation, our theories are developed under the assumption that the state is uni-dimensional, so that forecasters are evaluated on the basis of a single forecast of one variable. In some applications, forecasters are evaluated on the basis of several contemporaneous forecasts or a whole record of past forecasts. To better capture contests and evaluations that examine a number of different variables, our normal learning model can be extended to such a multi-variate setting. As explained in Sections 4 and 5, the strategic distortions identified here hold more generally with multi-dimensional states, signals and forecasts.

## 3. Honest Forecasting

Forecasters are presumed honest, unless proven strategic. As forcefully argued by Keane and Runkle (1990), "... professional forecasters ... have an economic incentive to be accurate. Because these professionals report to the survey the same forecasts that they sell on to the market, their survey responses provide a reasonably accurate measure of their expectations."

Our benchmark forecast is the honest report of the *Bayesian* posterior expectation,

$$h_i(s_i) = E\left[x|s_i\right] = \frac{\tau_i}{\tau_i + \nu} s_i + \frac{\nu}{\tau_i + \nu}\mu, \tag{3.1}$$

as assumed by most empirical investigations. In the normal model the posterior expectation minimizes the mean of any symmetric function of the forecast error, such as the mean squared error (cf. Bhattacharya and Pfleiderer (1985)).[17] The honest forecast incorporates all available evidence $s_i$, and can already offer some explanations for the data. Forecasts are dispersed due to private information. Forecasts are more dispersed and less accurate in years with relatively little public information.[18]

forecasts on the state, as is natural in a number of macroeconomic contexts. However, in the presence of many forecasters each of them individually would have a negligible influence on the state variable. The more general case in which forecasts affect the distribution of the state variable is left for future research.

[17] Motivated mainly by the needs of accurate probabilistic weather forecasts, a large literature in meteorology and statistics studies how to induce forecasters to form and state correct subjective probabilities (cf. Dawid's (1986) overview). Statisticians have developed scoring rules that elicit truthful information and avoid misrepresentation of a forecaster's beliefs (see e.g. de Finetti (1965) and Savage (1971)). In this paper we instead adopt a positive approach and characterize strategic manipulation of forecasts.

[18] This is consistent with a finding reported by Zarnowitz and Lambros (1987) on the ASA-NBER survey of professional forecasters. In addition to point forecasts, that survey initially asked forecasters to report probability distributions. Zarnowitz and Lambros documented that forecast dispersion is positively correlated with a measure of forecast uncertainty. Likewise, in the data of our Figure 1, a regression of the standard deviation of the forecasts on the absolute error of the consensus forecast, we find a coefficient of .145 with standard error .063. Thus, there is a significant negative correlation between accuracy and dispersion.

An important feature of the honest Bayesian forecast $h_i$ driving our theoretical analyses below, is that $h_i$ is a biased forecast in the statistical sense. With probability one we have $x \neq \mu$, and therefore $E[h_i|x] \neq x$. In fact, $E[h_i|x]$ always lies between $x$ and $\mu$.[19] Even with honest reporting, the forecast error $h_i - x$ is negatively correlated with $x$ (since $E[(h_i - x)x] = E[(\tau_i \varepsilon_i + \nu(\mu - x))x/(\tau_i + \nu)] = -1/(\tau_i + \nu) < 0$).

Nevertheless, the honest forecast $h_i(s_i)$ has the key statistical property of being uncorrelated with its forecast error $h_i(s_i) - x$: $E[E[x|s_i](E[x|s_i] - x)] = E[E[E[x|s_i](E[x|s_i] - x)|s_i]] = E[E[x|s_i]E[E[x|s_i] - x|s_i]] = 0$. This *orthogonality* property states that the honest forecast does not carry information about its forecast error. A large body of empirical literature on rational expectations (e.g., Keane and Runkle (1990) and (1998)) tests for orthogonality to see if forecasters are rational. Orthogonality may seem a necessary property of rational forecasts, but this is not the case. Asymmetric scoring rules generally result in forecasts violating the property, as noted by Granger (1969) and Zellner (1986).[20] In our model we instead maintain symmetry and examine whether rational players make non-orthogonal forecasts for strategic reasons.

As can also be verified with the data used in Figure 1, forecasts tend to be less volatile than realizations. This could suggest to some naive observers that forecasters herd. In particular, forecasters often fall short of the mark in years with extreme growth rates. But the realization $x = h_i + (x - h_i)$ is more volatile than any orthogonal forecast $h_i$ when the forecasters' information is noisy, $V[x] = V[h_i] + V[x - h_i] > V[h_i]$. This is almost always the case for macroeconomic, earnings, and weather forecasts.

Consider briefly *classical* statistics, i.e. forecasting in the absence of prior information. Forecaster $i$'s maximum likelihood estimator (MLE), maximizing $g_i(s_i|x_i)$ over $x_i$, would be $s_i$. Since $E[s_i|x] = x$, the MLE forecast is unbiased and the forecast error $\varepsilon_i$ is independent of $x$. However, the maximum likelihood forecast violates the orthogonality property, since $E[s_i(s_i - x)] = E[(x + \varepsilon_i)\varepsilon_i] = 1/\tau_i > 0$. Note that the MLE can be equivalently seen as resulting from Bayesian updating when the prior distribution on the state is the improper uniform distribution on the real line. We have chosen to posit a normal prior on $x$ to reflect that forecasters typically share some information about the variable to be predicted. The presence of prior information drives the results of the two theories developed

[19] In its widely circulated *World Economic Outlook*, even the International Monetary Fund (2001) laments that forecasters typically fall short of predicting changes in macroeconomic performance (cf. pages 6–8). In particular, forecasts are often too optimistic when the economy enters a recession, and pessimistic when the economy expands. It is misleading to interpret this as an indication that economic forecasters use their information inefficiently.

[20] Granger (1999) defines generalized forecast errors for any given loss function, and notes that these errors satisfy orthogonality.

below.[21]

## 4. Forecasting Contest Theory

Forecasters often participate in contests that reward the best performers.[22] It is natural to investigate whether the competitive pressure from other forecasters provides the right incentives to forecast honestly. This section develops a positive theory of a symmetric simultaneous winner-take-all contest with a large number of forecasters. In our forecasting contest, the participating forecasters simultaneously submit individual forecasts based on their private information. A prize is awarded to the forecaster whose forecast is closest to the realized state.

There is remarkably little previous work on forecasters' behavior in contests. Steele and Zidek (1980) were the first to study a *sequential* forecasting contest among two privately informed forecasters. They assumed away game-theoretic considerations by supposing truthful reporting by the first forecaster. After observing the first forecast, the second guesser faces a simple decision problem and has a clear advantage.[23] Indeed, Kutsoati and Bernhardt (2000) have empirically confirmed that the financial analysts who release late earning forecasts tend to overshoot the consensus forecast in the direction of their private information. Laster, Bennett and Geoum (1999) studied a winner-take-all *simultaneous* forecasting contest in which all forecasters share the same (public) information. We instead allow forecasters to have private information on the state and cast winner-take-all forecasting contests as problems of strategic location.[24] A forecaster's payoff is equal to the probability that the realized state is closer to her forecast than to any other forecast. This probability of falling closest to the state is equivalent to the market share or fraction of votes to be maximized in Hotelling's (1929) pure location game.[25]

[21] The importance of the role of public information is validated by a number of empirical studies. For example, Welch (2000) finds evidence that security analysts are strongly influenced by the prevailing consensus.

[22] In the much publicized Wall Street Journal semi-annual forecasting contest the most accurate forecaster over the previous six months is typically rewarded with a writeup. Even in the absence of monetary prizes, the publicity effect to the winner can be large.

[23] Steele and Zidek and some follow-up papers focused on the characterization of the second guesser's probability of winning.

[24] There would be no reason to reward accurate forecasters in the absence of heterogeneous private information. It turns out that the symmetric equilibrium in the forecasting contest is more appealing once private information is introduced, being in pure rather than mixed strategies.

[25] An extensive literature in economics and political science studies versions of this game without private information. As it is well known (cf. Osborne and Pitchik (1986)), equilibria in this classic game crucially depend on the number of players and often involve mixed strategies. Our forecasting contest theory extends Hotelling's simultaneous location game to scenarios where the forecasters (firms or politicians) have private information on the distribution of the state (location of consumers or voters).

### 4.1. Model

The forecasting contest game proceeds as follows. First, each forecaster observes the private signal $s_i$. Assume that the signals of all the forecasters are of common precision $\tau$.[26] Second, the forecasters simultaneously submit their forecasts $c_i$. Third, the true state $x$ is publicly observed and the forecaster whose forecast $c_i$ turns out to be closest to $x$ wins a prize proportional to the total number of forecasters participating in the contest. In the case of a tie, all winners share the prize evenly.

To play this game, forecasters form conjectures about the distribution of the opponents' forecasts and calculate their best response. Payoffs to forecaster $i$ can then be defined as follows. Suppose forecaster $i$ receives signal $s_i$ and conjectures that the opponents' forecasts are distributed according to the conditional density $\gamma(c|x, s_i)$ on the real line. By conditional independence of the signals, $x$ is sufficient for $s_i$ so that we can write $\gamma(c|x)$. Suppose forecaster $i$ submits the forecast $c_i$. The forecasts of the many opponents are dense on the support of $\gamma$, so forecaster $i$ wins if and only if $x = c_i$. This happens with chance $q_i(c_i|s_i)$. Conditional on winning, the size of the prize obtained by each player is $1/\gamma(c_i|c_i)$. In conclusion, the expected payoff to forecaster $i$ when releasing forecast $c_i$ is:

$$U_i(c_i|s_i) = \frac{q_i(c_i|s_i)}{\gamma(c_i|c_i)}. \tag{4.1}$$

A companion paper (Ottaviani and Sørensen (2002)) provides a formal derivation of this payoff function as the limit of payoff functions when the number of (identical) forecasters tends to infinity.[27]

### 4.2. Deviation

We now show that honest forecasting is not an equilibrium. Consider a single forecaster with signal $s$ competing against forecasters who are all reporting their honest forecasts. Without loss of generality, let $s > \mu$ as depicted in Figure 2. What is the best reply for such a forecaster?

According to (4.1), the best forecast maximizes the ratio between the probability of winning the first prize and the number of forecasters with whom this prize is shared. First, the probability of winning conditional on signal $s$ is equal to the posterior belief on the state $x|s$, the normal distribution centered at $E[x|s]$ and depicted on the right in Figure

[26] This symmetry assumption can be compatible with the fact that forecasters have unobserved heterogeneous forecasting abilities — see Section 5.

[27] Numerical simulations confirm that the equilibrium of the limit game is a good approximation of the equilibria in games with a large number of forecasters.

2. Second, the curve on the left in Figure 2 depicts $\gamma(x|x)$, the denominator of the ratio maximized by the forecaster. This represents how the mass of forecasters with *correct* forecasts changes as a function of the state $x$. The shape of $\gamma(x|x)$ depends on the weight assigned by the other forecasters to their signal. Since the other forecasters put a positive weight on the prior mean, $\gamma(x|x)$ is bell shaped around $\mu$.[28] As clear from the graphical illustration, the probability of winning is flat at the honest $E[x|s]$, while the frequency of correct forecasts is decreasing in the distance of the state $x$ from the prior mean $\mu$. At the posterior expectation it is then optimal to move away from prior mean toward the private signal, as the second-order loss resulting from lower probability of winning is more than compensated by the first-order gain due to reduced competition:[29]

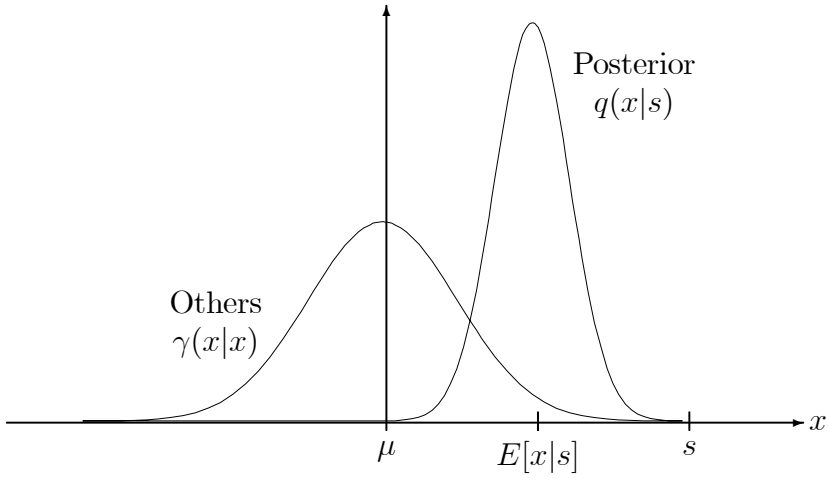


**Figure 2.** Optimal deviation in the forecasting contest model.

**Proposition 1 (Exaggeration in Contest Deviation)** *If all other forecasters are forecasting honestly by applying the strategy* $c(s) = (\tau s + \nu\mu) / (\tau + \nu)$, *the contest drives forecaster* $i$ *to exaggerate.*

The optimal deviation forecast is a weighted average of $s_i$ and $\mu$, but the weight on $\mu$ is lower than in the honest forecast. The contest deviation forecast is then positively correlated with its error: when $x$ is above $\mu$ the forecast is too high on average. More generally, it is optimal to bias the forecast against the prior mean whenever the other forecasters use a common strategy consisting of a convex combination of the prior and the signal received.

[28] It is worth noticing that $\gamma(x|x)$ is *not* a probability density function. For example, when the other forecasters put zero weight on the prior (for instance because they have perfectly informative signals), $\gamma(x|x)$ is constantly equal to 1.

[29] If $x$ and $s$ were multivariate Normal, a picture similar to Figure 2 would arise in the appropriate multidimensional space, with $\gamma(x|x)$ multivariate bell-shaped around $\mu$ and the posterior $q(x|s)$ multivariate bell-shaped around $E[x|s]$.Proposition 1 then continues to hold.

### 4.3. Equilibrium

Having established that honest forecasting is not compatible with equilibrium, we are now ready to characterize a symmetric Bayes-Nash equilibrium.[30] At such an equilibrium, each forecaster applies for every signal $s_i$ the best response to her conjecture about the opponents' distribution $\gamma(c|x)$, and this conjecture is correctly derived from the strategies actually used.

**Proposition 2 (Exaggeration in Contest Equilibrium)** *For any values of $\nu$ and $\tau > 0$ the contest has a unique symmetric linear equilibrium $c(s) = Cs + (1-C)\mu$ with $C \in [0,1]$. Forecasters put more weight on their private information than according to the conditional expectation: $C = \left(\sqrt{\tau^2 + 4\nu\tau} - \tau\right)/2\nu > \tau/(\nu+\tau)$.*

In the absence of private information ($\tau = 0$), the only symmetric equilibrium is in mixed strategies as in Laster, Bennett and Geoum (1999) and Osborne and Pitchik (1986), who find that with infinitely many symmetrically informed players the distribution of equilibrium locations replicates the common prior distribution about $x$. The addition of private information has the desirable effect of inducing a symmetric location equilibrium in pure rather than mixed strategies. Forecasters differentiate themselves from their competitors by putting excessive weight on their signals. As in the honest forecast, the weight on the signal is increasing in $\tau$ and decreasing in $\nu$. For all values of $\nu$ and $\tau$, this weight is larger than in the honest forecast, so the contest gives an incentive to move away from $\mu$.[31] The symmetric equilibrium strikes a balance: opponents disperse themselves to such an extent that forecaster $i$ is happy to reply precisely with the same dispersion. The equilibrium forecast is positively correlated with its error.

The contest equilibrium satisfies $C < 1$, so the forecast is not as extreme as the maximum likelihood estimate (MLE). However, the MLE results in the contest when the prior on the state $x$ is improper, i.e. uniform on the real line. If the opponents forecast $c = s$, their forecasts are normally distributed around $x$, and the term $\gamma(c|c)$ is constant in $c$. Forecaster $i$'s best reply will then be $c_i = s_i$, since this constant term does not distort the forecaster's problem. Truthtelling by all forecasters is then an equilibrium in the absence of public information. The contest distortion thus depends on the presence of prior information that anchors the forecasts of the opponents around $\mu$. The tendency of opponents to cluster around the prior mean drives forecasters away from it.

[30] We take a positive approach and do not search for asymmetric or non-linear equilibria.

[31] The equilibrium in linear strategies exists for all parameter values. This is slightly surprising since a best reply to honesty did not exist for all parameter values. Intuitively, with increased weight on their signal, the opponents are less concentrated around $\mu$, mitigating the incentive to move away from $\mu$.

## 5. Reputational Cheap Talk Theory

While in a forecasting contest, competition takes place according to rules clearly set out in advance, markets often reward successful performance in subtler and less structured ways. For example, the labor and financial markets perform informal (or subjective) evaluations of the forecasters' track record and performance. This section develops a positive theory of forecasting in which forecasters aim to impress the market with their talent. It is convenient to think of the market as performing the passive role of an evaluator. The market evaluates a forecaster's quality of information (or talent) by comparing the forecast with the ex-post realization of the state.[32] Instead of committing ex ante to a particular reward scheme, the market optimally evaluates ex post the forecasting talent based on all the information available. Forecasters with a better reputation can provide more valuable information and can therefore command higher compensation. To foster their careers, forecasters want to develop a good reputation for accuracy.

The first reputational cheap talk game was formulated in Scharfstein and Stein's (1990) seminal paper on reputational herding.[33] While Scharfstein and Stein assumed that better informed forecasters have conditionally more correlated signals, we focus on the case with conditionally independent signals and extend our model in Section 6.4 to allow for conditional signal correlation.[34] Reputational forecasting is a game of cheap talk (Crawford and Sobel (1982)), since a forecaster's payoff depends on the forecast released only indirectly through the market's evaluation. The forecaster plays the role of sender and the market is the receiver.

### 5.1. Model

The structure of our basic statistical model needs to be extended to introduce a latent parameter representing the unknown talent $t_i > 0$ of forecaster $i$. We assume that forecasters and the market share a common non-degenerate prior belief $p_i(t_i)$ on forecaster $i$'s talent, with all the talents $t_i$ and the state $x$ statistically independent. It remains the common prior that $x \sim N(\mu, 1/\nu)$. Conditionally on state $x$ and talent $t_i$, signal $s_i$

[32] In Zwiebel's (1995) model ability adds instead to the productivity of the agent, rather than parameterizing the quality of information about the state.

[33] Earlier, Holmström (1999) in the second part of his paper analyzed the behavior of an agent who aims at convincing the market that she is well informed in a situation in which the decision made affects whether the state is observed by the market or not. We instead follow Scharfstein and Stein by assuming that the forecasts influence neither the realization nor the observability of the state of the world.

[34] See Ottaviani and Sørensen (2000) on the connection of the reputational herding model of Scharfstein and Stein (1990) and Trueman (1994) with the statistical herding model of Bikhchandani, Hirshleifer and Welch (1992) and Banerjee (1992).

is distributed with p.d.f. $\tilde{g}(s_i|x,t_i)$. We assume that there exists a p.d.f. $\hat{g}$ on $[0,\infty)$, such that $\tilde{g}(s_i|x,t_i) = t_i\hat{g}(t_i|s_i - x|)/2$ — then we have a symmetric location experiment where $t_i$ is the scale parameter. Since we are extending our model, we keep the assumption that $s_i|x \sim N(x, 1/\tau)$. The primitives $\hat{g}$ and $p$ are thus restricted to satisfy $g_i(s_i|x) = \int_0^\infty \tilde{g}(s_i|x,t_i)\, p_i(t_i)\, dt_i$. We naturally assume that a greater forecasting talent is associated with smaller signal errors, i.e. that the likelihood ratio $\tilde{g}(s_i|x,t_i)/\tilde{g}(s_i|x,t_i')$ is increasing in $|s_i - x|$ when $t_i < t_i'$. As pointed out in Example 3.3 of Lehmann (1955), this is equivalent to the log-concavity in $a$ of $\hat{g}(e^a)$.[35] Straightforward algebra shows that both conditions reduce to monotonicity of the elasticity $(\varepsilon\hat{g}'(\varepsilon)/\hat{g}(\varepsilon)$ is decreasing in $\varepsilon)$.[36]

We further simplify the problem by eliminating any form of payoff interaction among the forecasters. First, conditionally on $x$ and $t_1,\ldots,t_n$, the $n$ forecasters' private signals $s_i$ are conditionally independent, so that forecaster $i$ cannot signal anything to the market about $t_j$ for $j \neq i$. Second, forecaster $i$'s objective depends solely on the posterior beliefs about $t_i$, regardless of the posterior beliefs about $t_j$.[37] With these assumptions there is no strategic interaction among forecasters and so the problem of each forecaster is separable from that of the others. For the remainder of this section, we will then focus on a single forecaster and remove the subscript $i$.

The reputational cheap talk game proceeds as follows. First, the forecaster observes the private signal $s$. Second, the forecaster issues the forecast (or message) $m$. Third, the true state $x$ is observed by the market which uses $(m,x)$ to update the belief $p(t)$ about the forecaster's talent $t$.[38] The forecast $m$ thus serves the role of a signal sent to the market about $s$, and the observation of additional information $x$ allows for sorting. The forecaster's goal is to obtain a favorable updating on the precision, with the understanding that the market rewards a good reputation.

To update the beliefs about the forecaster's talent, the evaluator applies a *conjecture* on the forecaster's strategy mapping $s$ into $m$ and derives the distribution of $m$ condi-

[35] As hinted by Lehmann (1955), if $\hat{g}$ is decreasing then the assumption is weaker than log-concavity of $\hat{g}$. Under this stronger property, Lehmann's (1988) Theorem 5.3 would guarantee that the talent $t_i$ parameterizes forecaster $i$'s Blackwell effectiveness in the class of monotone decision procedures.

[36] For an example of distributions satisfying our assumptions, let $\hat{g}(\varepsilon) = 2K_1 \exp\left(-\varepsilon^4/12\right)$ implying that $s|x,t$ has an exponential power distribution (Box and Tiao (1973), page 517) with p.d.f. $\tilde{g}(s|x,t) = K_1 t \exp\left(-t^4(s-x)^4/12\right)$, and let $t^{-4}$ follow a Gamma distribution such that $t > 0$ has p.d.f. $p(t) = K_2\tau^{3/2}t^{-4}\exp\left(-3\tau^2t^{-4}/4\right)$. Appealing to the well-known p.d.f. of the inverse Gaussian distribution, it is straightforward to find the normalizing constants $K_1$ and $K_2$ and verify that $g(s|x) = E[\tilde{g}(s|x,t)]$. Clearly, $\log(\hat{g}(\varepsilon)) = \log(2K_1) - \varepsilon^4/12$ is concave and decreasing in $\varepsilon$, implying also log-concavity of $\hat{g}(e^a)$.

[37] See Ottaviani and Sørensen (2003) for an extension of the model to allow for relative reputational concerns.

[38] In Prendergast and Stole's (1996) reputational signaling model instead the state is not observed. See Section 5.2 below for a more detailed discussion of the other differences with Prendergast and Stole.

tional on $x$ and $t$, denoted by $\varphi(m|x,t)$. Bayes' rule is then used by the evaluator to calculate the posterior reputation $p(t|m,x) = \varphi(m|t,x)\,p(t)/\varphi(m|x)$ where $\varphi(m|x) = \int_0^\infty \varphi(m|t,x)\,p(t)\,dt$. To model the forecaster's preferences over posterior reputations, we take a two-step von Neumann-Morgenstern formulation. We assume that the forecaster correctly knows the $\varphi$ function used by the evaluator, so that the forecaster can predict how the posterior reputation is calculated. The utility of reputation $p(t|m,x)$ is then given by its expected value of the Bernoulli function $u(t)$,[39]

$$W(m|x) \equiv \int_0^\infty u(t)p(t|m,x)\ dt. \tag{5.1}$$

Since a signal with higher $t$ is more valuable in all monotone decision problems, we assume that $u$ is strictly increasing. This implies that the market rewards more forecasters with a (first-order stochastically) better reputation. When reporting the message $m$, the forecaster does not yet know the state $x$, but believes it to be distributed according to $q(x|s)$. The forecaster then chooses the message $m$ which maximizes the expected $W$,

$$U(m|s) \equiv \int_{-\infty}^\infty W(m|x)q(x|s)\ dx. \tag{5.2}$$

Different messages might be differently appealing to a forecaster depending on the signal received.

### 5.2. Deviation

In order to show that truthtelling cannot be an equilibrium, consider what happens when the evaluator conjectures that the forecaster applies the benchmark honest strategy $h(s)$. A forecaster with signal $s > \mu$ believes that the state is concentrated around $E[x|s]$, a weighted average of the prior mean $\mu$ and the signal $s$ represented in Figure 3. If the forecaster were to honestly report $h(s) = E[x|s]$, by inverting the strategy $h(.)$ the market would infer that the signal $s$ is higher than $E[x|s]$. Note that the forecaster believes that the state is distributed symmetrically and unimodally around the posterior expectation. As proved formally below, signals closer to the state are better news about the talent of a forecaster and so result in higher expected reputational payoff. The forecaster then wants to be perceived as having a signal $\hat{s}$ equal to the posterior expectation $h(s) = E[x|s]$ on the state. We conclude that if the market naively believes that the

[39] This can be seen as a psychological game, in which the sender's payoff depends on the receiver's belief. The payoff is equal to the expected value of a function over types evaluated using the posterior belief about an individual's type based on the information signaled in equilibrium. A formally similar expected utility formulation has also been adopted in different contexts by Geanakoplos, Pearce and Stacchetti (1989) and Bernheim (1995).

forecast reflects truthfully the forecaster's posterior expectation, the forecaster deviates by reporting $d(s) = h(h(s)) = E[x|\hat{s} = E[x|s]]$.[40]

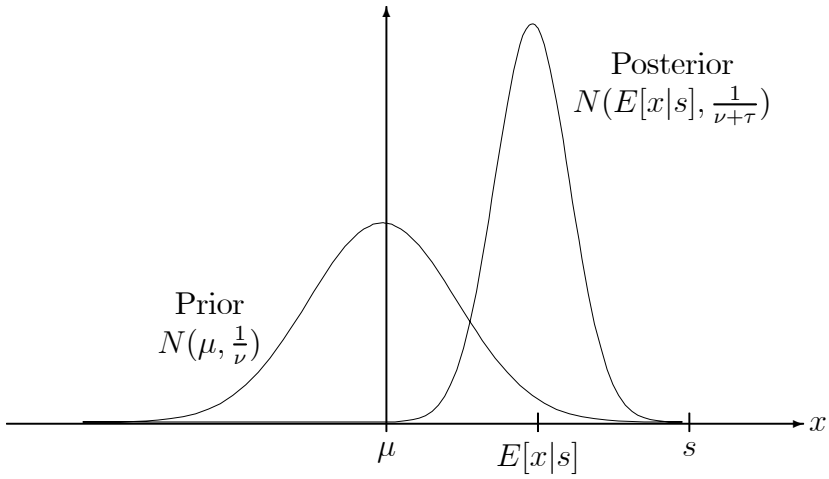


**Figure 3.** Optimal deviation in the reputational cheap talk model.

Another interpretation of the conservative deviation is based on the following logic. A forecaster who receives a signal $s$ above the prior mean $\mu$, concludes that the average forecast error $E[\varepsilon|s] = \nu(s-\mu)/(\tau+\nu)$ is positive. The forecaster then optimally deviates by removing this expected error from the true signal and so pretending to have signal $\hat{s} = s - E[\varepsilon|s] = E[x|s]$. In this deviation, forecasters are biased towards the prior mean:[41]

**Proposition 3 (Conservatism in Reputational Deviation)** *If the evaluator conjectures honest forecasting $h(s)$, the forecaster shades the forecast towards the prior mean by reporting*

$$d(s) = h(h(s)) = \left(\frac{\tau}{\tau+\nu}\right)^2 s + \left(1 - \left(\frac{\tau}{\tau+\nu}\right)^2\right)\mu.$$

Note that if the signal is perfectly informative or the prior is improper, the posterior expectation puts zero weight on the prior belief and so the inferred signal would be equal to $E[x|s]$. In both these cases truthtelling is an equilibrium. Likewise, this theory relies

[40] In the more general case with $x$ and $s_i$ multivariate and univariate talent $t$, we can extend the model by letting $\tilde{g}(s_i|x,t_i) = t_i\hat{g}(t_i||x-s_i||)/2$. The property that signals closer to the state are better news about the talent is retained. A picture similar to Figure 3 arises in the multi-dimensional space and Proposition 3 continues to hold.

[41] This result holds regardless of the forecaster's attitude toward risk. As seen in the proof of the proposition, posterior reputations depend only on the inferred absolute forecast error $|\hat{\varepsilon}|$ and are ordered in the sense of first order stochastic dominance. Equation (5.2) implies that forecasters are risk-neutral with respect to lotteries over posterior reputations. Any forecaster with an increasing $u$ therefore ranks messages the same way, regardless of the second derivative of $u$.

on the fact that the market *sequentially rationally* uses all the information available *ex post* to evaluate the forecaster. To understand this, suppose that instead the market were to *commit ex ante* to evaluate the forecaster by comparing the forecast $m$ with the realization $x$, according to the magnitude of the absolute distance $|m - x|$. In this case, the forecaster's optimal strategy is to honestly report $m = E[x|s]$. This is essentially the default case of honest forecasting also explained in Section 2. Notice that this could be reconciled with the reputational theory by positing that the market (incorrectly) believes that the forecaster's message is equal to her signal, $m = s$. It is perhaps quite natural that when forecasters use a separating strategy, the market treats them as if $m = s$. But this market conjecture remains incorrect in the presence of prior information, as the forecaster's best reply is $m = E[x|s]$.

According to the proposition, sophisticated forecasters who are taken at face value report conservative forecasts in order to fool the market into believing that they have more accurate signals. The full characterization of the deviation incentive is relevant for several purposes. Understanding the pressure to deviate from honesty provides intuition for the impossibility of truthtelling and sheds light on out-of-equilibrium forces. If the forecaster has mixed incentives, caring about both the reputation and the forecast accuracy, the incentive to deviate from honesty persists, and we can expect to find conservativeness in equilibrium. Finally, conservatism arises in real-world outcomes of communication when the evaluator is not fully rational.[42]

The conservatism of Proposition 3 is a new insight into herding (as a disequilibrium phenomenon). As the prior mean aggregates public information released previously by others, forecasters tend to herd. Herd behavior is here driven by concerns for absolute rather than relative accuracy. The content and logic of this result are different from the finding of Scharfstein and Stein (1990). Scharfstein and Stein argued that reputational herding requires the signals of better informed managers to be more correlated, conditionally on the state of the world. Our result instead does not rely on conditionally correlated signals.

At a superficial level, our result is reminiscent of Prendergast's (1993) "yes-men" effect, but is driven by different forces and essentially goes in the opposite direction.[43] While in

[42] For an example of reputational signaling with bounded rationality, see Zitzewitz's (2001b) model where the market uses a simple econometric technique to evaluate the quality of the information contained in the forecasts. In his model forecasters have information on their own ability. As discussed below, this introduces an incentive to exaggerate.

[43] Prendergast examines how to induce an agent to gather and report information. The agent has access to two private signals, one on the state of the world and the other on the principal's private signal on the state. If the principal commits to a reward scheme based on the difference between the agent's report and the principal's signal, the agent honestly reports her best estimate of the principal's private signal. Then

Prendergast's model the agent does not sufficiently move away from her signal about the principal's signal, in our model the agent does not move away from the prior mean. By identifying the principal's signal with the ex-post (noisy) realization of the state, it is seen that Prendergast's deviation report is biased toward the state, rather than the prior. As shown in the next section, also in our model the incentive to mis-report is self defeating. In both models, in equilibrium the agent cannot transmit all her information.

The incentive to deviate towards the prior relies on the assumption that forecasters do not know more than the evaluator about their forecasting talent (a.k.a. ability).[44] A forecaster who instead has private information about her own ability, will try to signal it. Since the posterior expectation with better information is further away from the prior mean, the incentive to signal ability generates an additional force pushing in the opposite direction to the one identified in Proposition 3.[45] The tendency to put excessive weight on the private signal is isolated by Prendergast and Stole's (1996) managerial reputational signaling model without ex-post information about the state. When evaluating forecasters, the market has instead access to additional information about the state, in the form of ex-post realization or contemporaneous forecasts of others. The addition of such ex-post information introduces the new conservatism effect isolated in Proposition 3.[46] Overall, concerns for absolute accuracy drive forecasters to be conservative if they do not know their ability, but to exaggerate if they know it well enough.[47]

### 5.3. Equilibrium

According to Proposition 3, honest forecasting is incompatible with equilibrium. Since this is a cheap talk game, ruling out truthtelling implies that there cannot be any fully separating equilibrium. By definition, in a fully separating equilibrium, the strategy mapping

the agent's report contains information from her two sources, and the principal can extract only part of the agent's direct signal about the state.

[44] This assumption allowed us to simplify the analysis, but is questionable in a dynamic setting since forecasters would learn more about their precision than the evaluator. Analysis of the resulting two-dimensional signaling problem is sensibly more involved. Preliminary investigations show that our conservatism result is robust to the introduction of small amounts of private information on own ability. See our companion paper (Ottaviani and Sørensen (2003)) for more on this.

[45] *Unconditionally* on the state, better informed agents have more variable posterior expectations. The variance of the honest forecast unconditional on the state, $V[h] = \tau / (\nu (\tau + \nu))$, increases in $\tau$.

[46] Exaggeration to signal own ability is also robust to the introduction of a small amount of ex-post information about the state. Trueman (1994) and Zitzewitz (2001b) exhibit exaggeration to signal own forecasting ability in the presence of ex-post information.

[47] As also suggested by Avery and Chevalier (1999), if younger managers have little private information about their own ability, they should have a tendency to be conservative; older managers would instead exaggerate, being more confident about their ability. Notice the contrast with Prendergast and Stole's (1996) prediction of impetuous youngsters and jaded old-timers when the same manager privately informed about own ability makes repeated observable decisions with a *constant* and unobserved state.

signals into forecasts can be inverted. As before, the evaluator infers the signal through inversion of the strategy, but the forecaster with signal $s$ then wishes to deviate to the forecast corresponding to signal $s' = E[x|s]$, which is different from $s$ whenever $s \neq \mu$.[48]

Non-existence of a fully separating equilibrium is not a particularly surprising finding in a cheap-talk setting. Another common property of cheap-talk games is that there exists an equilibrium with complete pooling. In such a babbling equilibrium, the forecaster issues the same message $m$ regardless of the signal received, and any message received by the evaluator is interpreted as carrying no information about the signal. More generally, not all information is conveyed to the evaluator. Equilibrium forecasting must necessarily involve some degree of pooling (or bunching) of signals into messages.

Due to the cheap talk nature of the game, the actual language used to send equilibrium messages is indeterminate. But the market can easily translate message $m$ into the best estimate conditionally on $m$, namely $E[x|m]$. So, the forecaster is effectively communicating $E[x|m]$ to the evaluator. Being a conditional expectation of $x$, this forecast is uncorrelated with its error. In this sense, the reputational equilibrium forecast satisfies the orthogonality property. We conclude:

**Proposition 4 (Coarseness in Reputational Equilibrium)** *There is no reputational cheap talk equilibrium with full revelation of information. Any equilibrium can be defined with a language such that the forecast has the orthogonality property.*

Because of the endogenous coarseness of the message, forecaster can communicate only part of the information about the state $x$. Additional information losses can result in dynamic extensions of the model, consistently with what observed by Welch (2000). Note that some information about the forecaster's talent $t$ is also lost. This in turn implies an additional loss of information about future states of the world, because more precise information on $t$ would be useful to assess the value of future forecasts from the same individual.

We now show by example that there is always a partially separating equilibrium in which some information is conveyed. This equilibrium involves a binary forecasting strategy, whereby the forecaster reports a high message $m_H$ whenever the signal $s$ weakly exceeds a threshold signal and a low message $m_L$ otherwise. The binary strategy in this

[48] In the presence of commitment or bounded rationality, the incentive to deviate can instead persist in equilibrium. The importance of commitment is illustrated by the predictions obtained in the model of Prendergast and Stole (1996), where the decision is delegated to the informed agent with reputational concerns. For an example of reputational signaling with bounded rationality, see Zitzewitz's (2001b) model where the market uses a simple econometric technique to evaluate the quality of the information contained in the forecasts.

equilibrium is symmetric, with threshold signal equal to $\mu$. Expressed in the natural language that identifies messages with their meaning understood in equilibrium, we have $m_L = E[x|s < \mu]$ and $m_H = E[x|s \geq \mu]$. A forecaster is indifferent between these two messages when receiving signal $\mu$. When observing a higher signal $s > \mu$, the forecaster expects high values of the state to be realized and so prefers to send message $m_H$ rather than $m_L$ in order to indicate a positive signal, implying smaller average errors:

**Proposition 5 (Binary Reputational Equilibrium)** *In the reputational cheap talk model there exists a symmetric binary equilibrium, and this is the unique equilibrium in binary strategies.*

The reputational cheap talk theory can be extended to allow for private information of forecasting ability, mixed objectives, and concern for relative reputation among forecasters. We refer to the companion paper Ottaviani and Sørensen (2003) for a broader theoretical analysis and discussion of the empirical literature. Rather than performing direct tests of reputational cheap talk, most of the existing empirical literature provides indirect evidence of reputational concerns based on heterogeneity across forecasters. Lamont (2002) finds that older forecasters tend to deviate more from the consensus. Chevalier and Ellison (1999) find that older mutual fund managers have bolder investment strategies. Hong, Kubik and Solomon (2000) conclude that the lower accuracy of older stock analysts is due the fact that they move earlier. Unfortunately, no one has so far attempted to model the endogenous timing of forecasts when the agents are concerned about their reputation or relative accuracy.[49]

# 6. Discussion

We now compare the empirical predictions of the different theories.

## 6.1. Forecast Variability

Except for the reputational equilibrium forecasts, we have found linear forecasting rules of the form $f_i(s_i) = F_i s_i + (1 - F_i)\mu$ for some constant weight $F_i$ between 0 and 1. The conditional distribution of the linear forecast is then

$$f_i|x \sim N\left(F_i x + (1 - F_i)\mu, F_i^2/\tau_i\right). \tag{6.1}$$

[49] In Gul and Lundholm (1995) forecasters care about the absolute accuracy as well as delay. A forecaster with a more extreme signal acts earlier at equilibrium.

We now describe the variance of the forecasts under the different theories. Observe that the prior variance $1/\nu$ of $x$ scales all variables of the model, while the weights $F_i$ depend on $\tau_i$ only through the relative signal precision $\rho_i \equiv \tau_i/\nu$. Apart from a factor $1/\nu$, all variances below can therefore be written as a function of $\rho_i$.

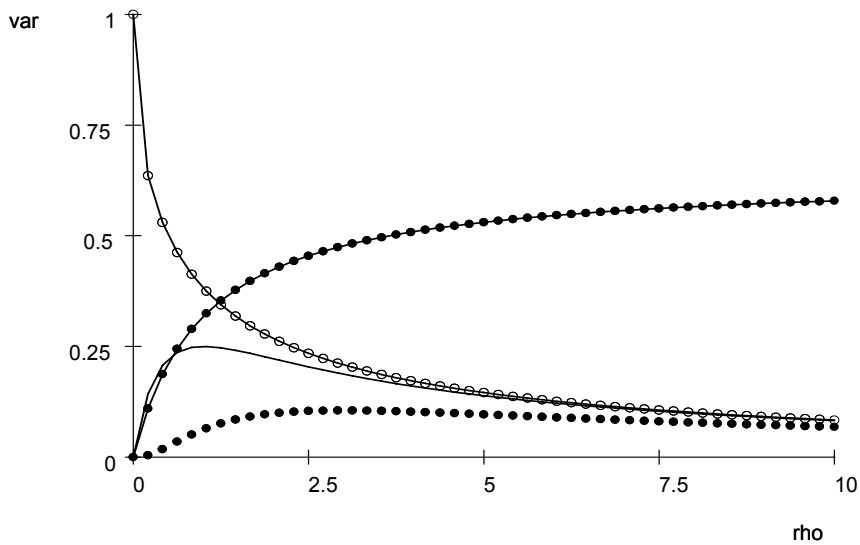


**Figure 4.** Conditional variances of forecasts as function of relative precision $\rho_i$, fixing $\nu = 1$. The solid line shows the conditional variance of the honest forecast $V[h|x]$, the dotted line the reputational deviation forecast $V[d|x]$, the line with full dots the reputational equilibrium forecast $V[r|x = \mu]$, and the line with empty dots the contest forecast $V[c|x]$.

**Honest Forecast.** The conditional variance of the honest forecast $V[h_i|x] = (1/\nu)\rho_i/(1+\rho_i^2)$ is increasing in the signal precision if the forecaster is imprecise enough ($\rho_i < 1$), but decreasing if precise ($\rho_i > 1$). When the signal is poorly informative ($\rho_i \approx 0$), the honest forecast is concentrated on the prior mean and so the conditional variance is also near 0. When $\rho_i$ is very large, the conditional variance is again near 0 because a perfectly informative signal gives an honest forecast concentrated on the true state.

**Contest Equilibrium Forecast.** The conditional variance of the forecast in the symmetric equilibrium of our symmetric contest (with $\tau_i = \tau$ and so $\rho_i = \rho$) is $C^2/\tau = (1/\nu)\left(2 + \rho - \sqrt{\rho^2 + 4\rho}\right)/2$. In the limit as the private signals become uninformative $\rho \to 0$, the distribution of equilibrium locations replicates the common prior distribution about the state and its conditional variance converges to $1/\nu$. This limit result is consistent with the findings of Osborne and Pitchik (1986) and Laster, Bennett and Geoum

(1999). Note that the conditional variance of the distribution of the equilibrium contest forecasts decreases in $\rho$ and converges to 0 as $\rho \to \infty$. We conclude that the addition of private information decreases the conditional variance of the contest forecasts, but this variance is consistently higher than the one resulting from honest forecasting.[50] This is in sharp contrast to the non-monotonicity of the conditional variance of the honest forecast. In cases with imprecise private signals, one could check empirically whether forecasts are very widely dispersed as in the contest, or quite close together as in the case of honest forecasting.

**Reputational Deviation Forecast.** The conditional variance of the reputational deviation forecast is $(1/\nu)\,\rho_i^3/\left(1+\rho_i\right)^4$ with variance first increasing in $\rho_i$, maximal at $\rho_i = 3$, and then decreasing in $\rho_i$. Compared to the forecasts under truthtelling and in the forecasting contest, the reputational deviation forecast puts more weight on the prior mean, and is therefore less variable.[51]

**Reputational Equilibrium Forecast.** The reputational equilibrium forecast $r_i$ of Proposition 5 is binomially distributed and therefore not directly comparable with the normally distributed forecasts that result in the other cases described above. The following characterization allows a partial comparison when the forecasters use the natural language $m_H = E\left[x|s > \mu\right]$ and $m_L = E\left[x|s \leq \mu\right]$.

**Proposition 6 (Distribution of the Binary Reputational Equilibrium)** *In the binary equilibrium of the reputational model* $m_H - \mu = \mu - m_L = \sqrt{2\rho_i/\pi\nu\left(1+\rho_i\right)}$ *and* $V[r_i|x] \leq (1/\nu)\,2\rho_i/\left(\pi\left(1+\rho_i\right)\right)$.

The amount by which the message moves the prior beliefs on the state increases in the relative precision $\rho_i$ of the forecaster's signal. To compare the conditional variance of the reputational binary equilibrium forecast with that of the honest forecast, observe that $2\rho_i/\left(\pi\left(1+\rho_i\right)\right) \leq \rho_i/\left(1+\rho_i\right)^2$ if and only if $\rho_i \leq \left(\pi-2\right)/2$. If the signal is not

[50] Morris and Shin (2002) consider a model in which informed agents are interested in being close to the state as well as to the average action taken by the others. In their symmetric equilibrium, agents underreact to private information. Revelation of public information might then result in an increase of action variability and so reduce welfare. Forecasters want instead to be far from the others and so overreact to private information. It can be shown that the variance of the contest equilibrium forecast conditional on the state is decreasing in the precision of the prior. More public information in our contest model unambiguously decreases forecast variability.

[51] Graham (1999) was the first to develop comparative statics predictions in a reputational cheap talk model. He considered a sequential setting with binary signals and he looked at how the second mover's *incentives for deviation* from a separating equilibrium are affected by changes in prior reputation, forecast ability and conditional signal correlation.

very precise, the conditional variance is uniformly higher under honesty than in the binary equilibrium. But the inequality is reversed when the signal has high precision and $x$ is close to $\mu$. In that case, the signal and thus the honest forecast are highly concentrated near $\mu$, while the reputational forecast is highly variable with even chance of a positive and negative update of amount $\sqrt{2\nu\rho_i/\pi\,(1+\rho_i)}$.

**Comparison of Variances.** Figure 4 plots the conditional variances as a function of $\rho_i$ when $\nu = 1$ for four forecasts: honest $h$, contest $c$, reputational deviation $d$, reputational binary equilibrium $r$. The graph makes clear that herding or exaggeration can be inferred from forecast dispersion only after controlling for the quality of the forecaster's information. This point is also emphasized by Zitzewitz (2001a).

### 6.2. Properties of Forecast Errors

The linear forecasts (6.1) with weight $F_i < 1$ are not unbiased, since $E[f_i|x] > x$ when $x < \mu$ and $E[f_i|x] < x$ when $x > \mu$. The conditional mean of the linear forecast only goes part of the way from the prior $\mu$ to the true state $x$. We see that all these forecasts, including the honest one, possess the oft-lamented property that forecasters fail to predict extreme values. This property is therefore not evidence of inefficient conservativeness on the part of the forecasters.

In a similar vein, the forecast error $f_i - x$ is negatively correlated with $x$ (since $E[(f_i - x)x] = E[(F_i\varepsilon_i + (1-F_i)(\mu - x))\,x] = -(1-F_i)/\nu < 0$). When $x$ is high (low) the error tends to be negative (positive). The forecast error can be predicted after the state $x$ has been observed, even if forecasters are not conservative. Furthermore, even though their signals are conditionally independent, there is correlation among the forecast errors of any pair $i, j$ of forecasters. The covariance of $f_i - x = (F_i\varepsilon_i + (1-F_i)(\mu - x))$ with $f_j - x$ is $(1-F_i)(1-F_j)/\nu > 0$. This covariance is positive because forecasters tend to make equal-signed errors of opposite sign of $x$.

To remove this correlation, one could alternatively study forecaster $i$'s shock as $f_i - E[f_i|x] = F_i\varepsilon_i$. This error follows a normal distribution with mean 0 and variance $F_i^2/\tau_i$, and is uncorrelated with $x$ and with the errors of other forecasters. This makes them useful observations for regression analysis. Empirically, $E[f_i|x]$ may not be known by the data analyst even when $x$ has been realized. It is a common approach to estimate $E[f_i|x]$ using the consensus forecast, to which we now turn.

When $n \geq 2$ forecasters have issued their forecasts, it is simple to calculate the unweighted average forecast $\bar{f} = \sum_{i=1}^{n} f_i/n$, often referred to as the consensus forecast. In

general, this is not the optimal forecast given the $n$ signals.[52] For example, when the individual errors $\varepsilon_i$ are statistically independent, the correlation of the honest consensus forecast with its error is negative, since the forecast tends to be too low when $x$ exceeds $\mu$ and too high when $x$ is below $\mu$.[53] When all forecasters have equal precision and independent errors, we see that $\bar{f}$ converges almost surely to $E[f|x]$ as $n \to \infty$ by the strong law of large numbers. Thus, asymptotically the shocks relative to the consensus $f_i - \bar{f}$ have the desirable zero-correlation properties.

### 6.3. Orthogonality

Consider linear forecasting rules of the form $f_i(s_i) = F_i s_i + (1 - F_i)\mu$. We have already noted that under honesty the forecast is uncorrelated with its error $f_i - x$ when $F_i = \tau_i / (\nu + \tau_i)$. More generally, the correlation is

$$E[(f_i - x) f_i] = E[(F_i \varepsilon_i + (1 - F_i)(\mu - x))(F_i (x + \varepsilon_i) + (1 - F_i)\mu)] = F_i \left( \frac{F_i}{\tau_i} - \frac{1 - F_i}{\nu} \right),$$

and so has the same sign as $F_i - \tau_i / (\tau_i + \nu)$. There is positive correlation when $F_i$ is larger as in the contest, and negative correlation when $F_i$ is smaller as in the reputational deviation. The reputational equilibrium forecast satisfies orthogonality but is not efficient.

As noted above, all types of forecast had errors negatively correlated with the state $x$. Thus, after knowing $x$ the sign of the errors could be predicted. However, this is an unreasonably strong test of the forecasters' abilities since $x$ is still unknown when the forecasts are released. If they report honestly, their error cannot be predicted from the forecast. The contest forecast and the reputational deviation forecast fail instead to inherit this property, so that once a forecast has been released the sign of its error can be predicted.

A typical empirical test for the hypothesis that the forecasts are conditional expectations ($E[x|I_i]$ for some information set $I_i$) is based on regressing the realized forecast error on the forecasts. Most studies report a positive correlation of the forecast and its

[52] An extensive literature in operations research studies the problem of how to optimally combine forecasts obtained with different methods or incorporating different information sets. See e.g. the early contributions by Bates and Granger (1969) and Bunn (1975). That literature does not, however, consider the possibility of strategic behavior by the forecasters.

[53] The weight that the conditional expectation $E[x|s_1, \ldots, s_n] = (\nu\mu + \tau_1 s_1 + \cdots + \tau_n s_n)/(\nu + \tau_1 + \cdots + \tau_n)$ attaches to $s_i$ is $\tau_i/\tau_j$ times the weight to $s_j$. In the consensus forecast, the ratio of weights is instead $\tau_i(\tau_j + \nu)/\tau_j(\tau_i + \nu)$, so that too much weight is given to the least precise signals. Even when all forecasters are equally precise, the weight accorded to the prior mean $\mu$ is too large and the consensus forecast fails to inherit the orthogonality property from the individual forecasts. In this case the consensus honest forecast is $\bar{h} = (n\nu\mu + \tau \sum_{i=1}^n s_i) / (n\nu + n\tau) = (n\nu\mu + n\tau x + \tau \sum_{i=1}^n \varepsilon_i) / (n\nu + n\tau)$ and the error is $\bar{h} - x = (n\nu(\mu - x) + \tau \sum_{i=1}^n \varepsilon_i) / (n\nu + n\tau)$, so that the covariance is always negative: $E[\bar{h}(\bar{h} - x)] = -(n-1)\tau / n(\nu + \tau)^2 < 0$ for $n > 1$. Kim, Lim and Shaw (1998) suggest methods to correct for the loss of information in the consensus forecast.

error, consistent with the prediction of our contest theory. For example, Batchelor and Dua (1992) find that forecasters put too little weight on the forecasts previously released by other forecasters (or, equivalently in our model, on the prior mean). However, Keane and Runkle (1990, 1998) cannot support any bias, as they note that the tests are not as powerful as is usually assumed once the correlation between forecast errors across forecasters is properly taken into account (see also Section 6.4 below). Recently, Zitzewitz (2001a) has perfected the orthogonality methodology to test for the presence of herding or exaggeration.[54] Zitzewitz applies the test to I/B/E/S analysts and finds significant and strong exaggeration.

### 6.4. Model Extension: Common Error

As stressed by Keane and Runkle (1998), the significant positive correlation among the residuals in the orthogonality regression indicates the presence of a common error in the forecasts. Since the forecasts are released well in advance, there are often unpredictable changes to the variables after the forecasts are submitted.[55] In this section we extend the model to account for this correlation found in the data. We find that our results are robust to the introduction of ex-post innovations in the state.

Suppose that each forecaster receives signal $s_i = y + \varepsilon_i$, that the variables $y, \varepsilon_1, \ldots, \varepsilon_n$ are stochastically independent as before, and that the state later observed is $x = y + \varepsilon_0$. Forecaster $i$ observes $s_i = x - \varepsilon_0 + \varepsilon_i$, so that the error $\varepsilon_0$ plays the role of a common error in the signals about the observed state $x$. We naturally assume that the the innovation is unpredictable, i.e., $\varepsilon_0 \sim N(0, 1/\tau_0)$ is independent of $y$ and the other errors. The honest forecast of $x$ is then the same as the honest forecast of $y$. Indeed, $E[x|s_i] = E[y + \varepsilon_0|s_i] = E[y|s_i]$ by the independence assumption. However, the posterior beliefs about $x$ are less precise than the posterior beliefs about $y$ due to the added error term. The variance of the normally distributed $x|s_i$ is $1/(\tau_i + \nu) + 1/\tau_0$. Let $\tilde{q}_i(x|s_i)$ denote the p.d.f. of this posterior belief.

[54] In our above theoretical derivation of the correlation of the forecast and its error, we treated the prior mean $\mu$ as a parameter. Data analysis must control for the prior mean in order to correctly identify when a realized state is "high" or "low." Zitzewitz' key innovation with respect to Keane and Runkle is to modify the regression equation by subtracting the expectation containing all publicly available information at the moment of forecasting from both the realization on the left hand side and the forecast on the right hand side.

[55] In addition, often realizations are observed with noise. For example, national statistics data are revised and become more accurate in later releases. In our model the evaluation of forecasters is based on one of these realizations.

**Contest Theory.** Let forecaster $i$ conjecture that the opponents' messages are distributed according to the conditional p.d.f. $\gamma(c|x, s_i)$. His expected payoff from forecast $c_i$ is $U_i(c_i|s_i) = \tilde{q}_i(x|s_i)/\gamma(c_i|c_i, s_i)$. Where we previously found that $\gamma(c_i|c_i)$ peaked at $\mu$, we now derive that $\gamma(c_i|c_i, s_i)$ peaks between $\mu$ and $h(s_i)$. A value of $s_i$ above $\mu$ suggests that $y > \mu$ and therefore that the opponents will issue relatively high forecasts — relative to the situation before, this moves up the state wherein the greatest mass of opponents guess correctly. Extending the analysis of Section 4, an optimal reply to honesty still involves exaggeration. Now there is no symmetric linear Nash equilibrium when $\tau/\nu$ is large relative to $\tau_0/\nu$. In this case, the forecasters have poor information on the location of $x$, and $\log(\tilde{q}_i(c_i|s_i))$ is not very concave. Still, they have good information about $y$ and about the other forecasters' signals, resulting in the convexity of $\log(\gamma(c_i|c_i, s_i))$ being too large when the other forecasters use best replies. When instead $\tau/\nu$ is small relative to $\tau_0/\nu$, there is a linear equilibrium with similar features to the one studied in the benchmark model. We conclude that the results of Section 4 hold in the presence of enough ex-post information about the state:

**Proposition 7** *In the forecasting contest with common error, there is exaggeration in the best reply to honesty. If the noise in the common error is small enough, there exists a pure-strategy symmetric equilibrium with exaggeration.*

**Reputational Theory.** Since errors $\varepsilon_1, \ldots, \varepsilon_n$ are correlated conditionally on $x$, in the reputational model we must explictly consider the interaction of the $n$ forecasters. In order to update beliefs about the precision of forecaster $i$, the evaluator uses the information on the location of $y$ contained in the realized $x$ as well as in the $n$ forecasts. We find that the results of Section 5 are robust to the addition of ex-post noise:

**Proposition 8** *In the reputational model with common error, the best reply to honesty of any other fully separable strategy is conservative. There exists an equilibrium in which every forecaster uses the binary reporting strategy of Proposition 5.*

A forecaster with signal $s_i = \mu$ regards the two possible messages as symmetric, and is thus indifferent. A forecaster with signal $s_i > \mu$ thinks it more likely that other forecasters report that they saw $s_j > \mu$, and thinks it more likely that the market observes $x > \mu$ — alas forecaster $i$ considers it more likely that the market's posterior beliefs on $y$ are shifted upwards. Forecaster $i$ then prefers to send the message that signals $s_i > \mu$.

Since the evaluator does not have access to $y$, even honest revelation of $s_i$ does not allow for the calculation of the individual $\varepsilon_i$. The evaluator modifies the Bayesian procedure

in this context, averaging over the possible values of $y$ given the available information, including the messages from the other forecasters. Recall that under a linear strategy $f_i - E[f_i|y] = F_i\varepsilon_i$ is proportional to the error and that in the case of many forecasters, the consensus forecast provides a good approximation for $E[f_i|y]$. A forecast close to the consensus then indicates a small absolute value of $\varepsilon_i$ and is therefore good news about the forecaster's ability.

## 7. Conclusion

In this paper we have formulated and contrasted two distinct theories of strategic forecasting within the normal model, widely used in information economics and empirical studies. In the process we have gained some novel insights on the forces driving informed agents to deviate from honest reporting of their conditional expectations. Mis-reporting results from the effect on individuals' payoff of the subtle interaction of the private information available to each individual with the public prior information available to the market and commonly shared by all agents.

Our first theory posits that competition for best accuracy takes place with *pre-specified rules*. Since the forecasters share the same public information, competition is highest when the state turns out to be equal to the prior mean. At the posterior expectation a small deviation away from the prior mean results in a first-order gain due to reduced competition and a second-order loss due to lower probability of winning. Equilibrium forecasts in a winner-take-all contest are then excessively differentiated in comparison with the corresponding conditional expectations. Note that the evaluation in a forecasting contest is ex post optimal when the market can only observe the accuracy ranking. But in reality the market has more information to evaluate the quality of forecasts.

Our second theory posits that the market has all the information contained in the forecasts and the realization of the state and uses it to *ex post optimally* evaluate the forecasters. We have assumed that better informed forecasters observe signals on average closer to the state and that forecasters who are reputed to be better informed have a higher payoff. We have shown that forecasters wish to appear to have received a signal equal to the posterior expectation of the state conditional on the signal actually received. In the presence of public information, the observed signal is necessarily different from the posterior expectation. If the market naively believes that forecasters are honest, forecasters then shade their forecasts toward the prior mean. If the market is fully aware of the forecasters' strategic incentives, equilibrium forecasts are imprecise but not shaded.[56]

[56] A possible criticism of our approach is that in reality competition among forecasters combines elements

In both the contest and the reputational theory the incentive to deviate from honesty is driven by the property that private signals are unimodally centered around the state and the fact that public information is available at the moment of forecasting. In the absence of public information, honest forecasting is an equilibrium in both models. In the empirical literature, Zitzewitz (2001a) takes this public information into account and finds significant and strong exaggeration in the forecasts of I/B/E/S analysts. This finding is in line with the equilibrium of our forecasting contest, but inconsistent with the deviation or the equilibrium of our reputational cheap talk model.[57]

Our results raise questions about the interpretation and use of professional forecasts to test the predictions of theories on how agents' decisions depend on expectations, as done in financial, international, and macro economics.[58] Further development of the theory of strategic forecasting should lead to methods to adjust for the induced biases and thus improve the interpretation of the results of econometric studies that use professional forecasts as proxies of market expectations.[59]

While interesting phenomena emerge in simple and plausible models without the need to depart from the rationality paradigm, there is some experimental evidence on deviations from Bayesian rationality. For example, according to Kahneman and Tversky's (1973) representativeness bias, forecasters often disregard prior information when making intuitive predictions. Experimental subjects put excessive weight on their signal and are overconfident in their predictions, similarly to what happens in equilibrium of our forecasting contest.[60] More work needs to be done on building and testing behaviorally plausible models of forecasting.[61]

In order to test the different theories, it might be useful to compare non-anonymous

of both theories. For example, *Institutional Investors* ranks analysts based on the opinions of large institutional investors. See Ottaviani and Sørensen (2003) for results on mixed incentives and relative reputational concerns.

[57] As argued by Zitzewitz (2001b), the observed exaggeration is also consistent with an alternative version of the reputational signaling model in which forecasters are privately informed about their ability and are evaluated according to an econometric technique.

[58] See e.g., Frankel and Froot (1987) for a study of exchange rate expectations using surveys of professional forecasters.

[59] See e.g., Romer and Romer (2000) and Prati and Sbracia (2002).

[60] Our reputational theory only requires that the forecaster *believes* that the market *perceives* the underlying talents of forecasters to be heterogenous. In an early behavioral model of financial advice, Denton (1985) assumes that investors listen to financial advisers who have no real information. In some cases, this might well be the case. For example, Hartzmark (1991) found that futures forecasters depend on luck rather than forecasting ability. Zitzewitz (2001a) instead finds that security analysts differ greatly in performance, justifying the assumption that they are fundamentally heterogeneous.

[61] See De Bondt and Thaler (1990) for arguments in favor of behavioral biases. It is worth remarking that a behavioral bias can be reinforced by strategic incentives. Ehrbeck and Waldmann (1996) run an early horse race between economic and psychological explanations for biases in forecasting.

with anonymous forecasting surveys. The existence of anonymous surveys (starting in 1946 with the Livingston Survey) presupposes a belief that forecasters might not honestly report their best predictions if anonymity were not preserved. While the identity of the forecasters belonging to the panel is typically available, anonymous surveys do not reveal which individual made which forecasts. A possible rationale for preserving anonymity is that it can (but need not) guarantee honest forecasting.[62]

In this paper we have adopted a positive approach, leaving unanswered a number of normative questions: How can forecasting surveys be improved to counteract the gaming incentives? How do different mechanisms affect the forecasters' incentives for information acquisition?[63]

[62] As reported by Croushore (1993): "This anonymity is designed to encourage people to provide their best forecasts, without fearing the consequences of making forecasts errors. In this way, an economist can feel comfortable in forecasting what she really believes will happen [...]. The negative side of providing anonymity, of course, is that forecasters can't claim credit for particularly good forecast performance, nor can they be held accountable for particularly bad forecasts. Some economists feel that without accountability, forecasters may make less accurate predictions because there are fewer consequences to making poor forecasts." A problem with the hypothesis of honest forecasting in perfectly anonymous surveys is that we do not have a theory to predict behavior in this situation. There is no clear objective. Possibly, the forecasters are concerned that the survey editor discovers something about their behavior, partly undermining the survey anonymity. In fact, as we were told by Croushore and Lamont, forecasters often seem to submit to the anonymous surveys the same forecasts they have already prepared for public (i.e. non-anonymous) release.

[63] Combining the insights of the statistical literature on scoring rules with the economic theory of regulation, Osband (1989) studies the optimal provision of forecasting incentives in the presence of costly information acquisition. See also Prendergast (1993).

## Appendix

**Proof of Proposition 1.** Recall that $x|s_i$ is normal with mean $(\tau s_i+\nu\mu)/(\tau+\nu)$ and precision $\tau+\nu$. Suppose that all opponents use the linear strategy $c(s) = As + (1-A)\mu$, with $A \in (0,1]$. Then $c|x$ is normal with $E[c|x] = Ax + (1-A)\mu$ and $V[c|x] = A^2/\tau$. Disregarding an irrelevant constant term, we find

$$\log(\gamma(c|c)) = -\frac{\tau(c-(Ac+(1-A)\mu))^2}{2A^2} = -\frac{\tau(1-A)^2(c-\mu)^2}{2A^2},$$

a concave quadratic function of $c$ with peak at $\mu$. The forecaster maximizes $\log(q_i(c_i|s_i)) - \log(\gamma(c_i|c_i))$, the difference of two concave quadratic functions. The objective function is concave when the first concave term prevails, i.e. for $\tau+\nu > \tau(1-A)^2/A^2$.

When $\tau+\nu > \tau(1-A)^2/A^2$, the forecaster has a unique best reply $c_i = Bs_i + (1-B)\mu$, with $B = \tau/\left(\tau+\nu-\tau(1-A)^2/A^2\right) \in [\tau/(\tau+\nu), +\infty)$. When instead $\tau+\nu < \tau(1-A)^2/A^2$, there is no best response because the incentive to move away from $\mu$ is so strong that forecaster $i$ wishes to go to the extremes $\pm\infty$. In the knife-edge case $\tau+\nu = \tau(1-A)^2/A^2$ the objective function is linear — whenever $s_i \neq \mu$ there is again no best reply, as the forecaster wishes to go to one of the extremes.

In particular, in the honest case $A = \tau/(\tau+\nu)$, it is optimal to reply with $B = \tau^2/(\tau^2+\tau\nu-\nu^2)$, provided that $\nu/\tau < \left(1+\sqrt{5}\right)/2$. We conclude that the best reply against truthtelling by all the opponents is to exaggerate. □

**Proof of Proposition 2.** $C = 0$ is not compatible with a symmetric equilibrium since in this case the opponents' forecasts are all equal to $c = \mu$ so that all replies other than $\mu$ yield forecaster $i$ higher payoff. Assume then that the forecasters use linear strategies of the form $c(s) = Cs + (1-C)\mu$ with $C \in (0,1]$. As shown in the proof of Proposition 1, forecaster $i$ has as best reply a linear strategy with weight $\tau/\left(\tau+\nu-\tau(1-C)^2/C^2\right)$ on the signal, provided that $\tau+\nu > \tau(1-C)^2/C^2$. The fixed-point condition for a symmetric Nash equilibrium is that this linear strategy be equal to the one posited, or $(1-C)\tau = C^2\nu$. Insert the values $C = 0, \tau/(\tau+\nu), 1$ in this quadratic equation to conclude that it possesses only one positive solution, and that this solution is in $(\tau/(\tau+\nu), 1)$. The second-order condition for the forecaster's optimization requires $\tau+\nu > \tau(1-C)^2/C^2$. Using $(1-C)\tau = C^2\nu$, this condition reduces to $\tau > -\nu C$, clearly satisfied by the positive solution for $C$. Finally, the solution of the quadratic equation is $C = \left(\sqrt{\tau^2+4\nu\tau}-\tau\right)/2\nu$. □

**Proof of Proposition 3.** Observing $m = h(s)$ and $x$, the evaluator infers the realized signal $\hat{s} = h^{-1}(m)$ and error $\hat{\varepsilon} = \hat{s} - x$. The updated reputation is then $p(t|m,x) = \tilde{g}(\hat{s}|x,t)\,p(t)/g(\hat{s}|x)$. This posterior reputation satisfies two intuitive properties due to

the assumptions on $\tilde{g}$. First, the posterior reputation depends on $m$ and $x$ only through the absolute size of the error $|\hat{\varepsilon}|$. Second, a small realized absolute error is good news about the forecaster's talent: for any $t < t'$, the likelihood ratio $p(t|m,x)/p(t'|m,x) = (\tilde{g}(\hat{s}|x,t)/\tilde{g}(\hat{s}|x,t'))(p(t)/p(t'))$ is increasing in $|\hat{\varepsilon}|$. These two properties imply (see Milgrom (1981)) that $W(m|x)$ is a strictly decreasing function of the inferred absolute error $|\hat{\varepsilon}|$.

Consider now the best response of a forecaster with signal $s$. The posterior distribution on $x$ is normal with mean $h(s)$ and variance $1/(\nu+\tau)$. The inferred forecast error $\hat{\varepsilon} = h^{-1}(m) - x$ is then normally distributed with mean $h^{-1}(m) - h(s)$ and variance $1/(\nu+\tau)$. The best reply maximizes the expected value of $W$, or equivalently minimizes a symmetric loss function of the error $h^{-1}(m) - x$. The forecaster then chooses $m$ such that the error has mean zero, by setting $h^{-1}(m) = h(s)$. □

**Lemma 1** *If a p.d.f. $\hat{g}(.)$ satisfies the property*

$$\hat{g}(t'\varepsilon)\,\hat{g}(t\varepsilon') < \hat{g}(t\varepsilon)\,\hat{g}(t'\varepsilon') \text{ for } \varepsilon' > \varepsilon \geq 0 \text{ and } t' > t, \tag{7.1}$$

*its counter-cumulative distribution satisfies it as well:*

$$\left[1-\hat{G}(t'\varepsilon)\right]\left[1-\hat{G}(t\varepsilon')\right] < \left[1-\hat{G}(t\varepsilon)\right]\left[1-\hat{G}(t'\varepsilon')\right] \text{ for } \varepsilon' > \varepsilon \geq 0 \text{ and } t' > t. \tag{7.2}$$

**Proof.** Integrating (7.1) for $\varepsilon'' > \varepsilon'$, we get

$$t'\hat{g}(t'\varepsilon)\left[1-\hat{G}(t\varepsilon')\right] < t\hat{g}(t\varepsilon)\left[1-\hat{G}(t'\varepsilon')\right] \tag{7.3}$$

for $\varepsilon' > \varepsilon$. Notice that the left-hand side and the right-hand side of (7.2) are equal for $\varepsilon' = \varepsilon$, and that by (7.3) we know that the derivative of the left-hand side is larger than the derivative of the right-hand side of (7.2). We conclude that (7.2) holds. □

**Proof of Proposition 5.** To support this equilibrium, we also need to specify the evaluator's beliefs following out-of-equilibrium messages. When receiving any message different from $m_L$ and $m_H$, the evaluator assumes that the forecaster possessed a signal below the threshold, thereby resulting in the same posterior reputation as message $m_L$. These beliefs satisfy the requirements of a perfect Bayesian equilibrium.

Assume that the evaluator conjectures a binary strategy with threshold signal $\check{s}$. We find $\varphi(m_H|x,t) = \int_{\check{s}}^{\infty} \tilde{g}(s|x,t)\,ds = \int_{\check{s}}^{\infty} t\hat{g}(t|s-x|)/2ds$. This reduces to $(1-\hat{G}(t|\check{s}-x|))/2$ when $\check{s} > x$ and $(1+\hat{G}(t|\check{s}-x|))/2$ when $\check{s} < x$, where $\hat{G}$ is the c.d.f. corresponding to the p.d.f. $\hat{g}$. Therefore, $\varphi(m_H|x,t) = 1 - \varphi(m_H|2\check{s}-x,t) = \varphi(m_L|2\check{s}-x,t)$. From (5.1), this implies the symmetry property $W(m_H|x) = W(m_L|2\check{s}-x)$.

When $\check{s} > x$, it follows from Lemma 1 that message $m_H$ is worse news about the talent than the observation that $s \geq x$. Thus we have $W(m_H|x) < W(m_H|\check{s})$ for $x < \check{s}$. Symmetrically, $W(m_H|x) > W(m_H|\check{s})$ for $x > \check{s}$. These inequalities and symmetry imply that $W(m_H|x) > W(m_L|x)$ for $x > \check{s}$.

We now show that when $\check{s} = \mu$, the forecaster does not wish to deviate from the putative equilibrium strategy. By symmetry, it suffices to assume that $s \geq \mu$ and check that $U(m_H|s) \geq U(m_L|s)$. Using (5.2) and symmetry of $W$, $U(m_H|s) - U(m_L|s)$ is

$$\int_{-\infty}^{\infty} (W(m_H|x) - W(m_L|x))\, q(x|s)\, dx = \int_{\mu}^{\infty} (W(m_H|x) - W(m_L|x))\, (q(x|s) - q(2\mu - x|s))\, dx.$$

Since $q(x|s)$ is the p.d.f. of the symmetric normal distribution with a mean weakly above $\mu$, we have $q(x|s) \geq q(2\mu - x|s)$ when $x \geq \mu$. We already had $W(m_H|x) > W(m_L|x)$ when $x > \mu$, so the integrand of the last integral is everywhere non-negative, implying that the integral is non-negative, i.e. that $U(m_H|s) \geq U(m_L|s)$ as desired.

Finally, we show that when $\check{s} \neq \mu$, the forecaster wishes to deviate from the binary strategy. Without loss of generality, focus on the case $\check{s} > \mu$. We show that there exists a signal $s > \check{s}$ such that $U(m_H|s) < U(m_L|s)$. As above, $U(m_H|s) - U(m_L|s) = \int_{\check{s}}^{\infty} (W(m_H|x) - W(m_L|x))\, (q(x|s) - q(2\check{s} - x|s))\, dx$. We have again $W(m_H|x) > W(m_H|\check{s})$ for $x > \check{s}$. At $x = \check{s}$ we have $q(x|s) = q(2\check{s} - x|s)$. By properties of the normal p.d.f., we obtain $q(x|s) < q(2\check{s} - x|s)$ for $x > \check{s}$ provided $E[x|s] = (\tau s + \nu\mu) / (\tau + \nu) < \check{s}$, which is certainly true for $s$ slightly greater than $\check{s}$. □

**Proof of Proposition 6.** By applying the well-known result that $E[y|y > 0] = \sigma\sqrt{2/\pi}$ for a normal variable $y \sim N(0, \sigma^2)$ (cf. Johnson and Kotz (1970)), we see that $m_H = E[x|s \geq \mu]$ is equal to $E[E[x|s]|s \geq \mu] = E[\frac{\tau s + \nu\mu}{\tau + \nu}|s \geq \mu] = \mu + \frac{\tau}{\tau+\nu} E[s - \mu|s > \mu] = \mu + \sqrt{2\tau / (\pi\nu(\tau + \nu))}$. By symmetry, we have $m_L = \mu - \sqrt{2\tau / (\pi\nu(\tau + \nu))}$. Given $x$, the chance of the forecast $r$ taking the high value $m_H$ is $1 - \Phi(\sqrt{\tau}(\mu - x))$ where $\Phi$ is the c.d.f. of the standard normal distribution. The binomial distribution of $r$ has mean $E[r|x] = \mu + (1 - 2\Phi(\sqrt{\tau}(\mu - x)))\sqrt{2\tau / (\pi\nu(\tau + \nu))}$ and variance $V[r|x] = 4(1 - \Phi(\sqrt{\tau}(\mu - x)))\, \Phi(\sqrt{\tau}(\mu - x))\, 2\tau / (\pi\nu(\tau + \nu))$. For any $x$, we have $\Phi(\sqrt{\tau}(\mu - x)) \in [0, 1]$ so $4(1 - \Phi(\sqrt{\tau}(\mu - x)))\, \Phi(\sqrt{\tau}(\mu - x)) \leq 1$, where the bound is tight being achieved for $x = \mu$. We conclude that $V[r|x] \leq 2\tau / (\pi\nu(\tau + \nu))$. Finally, insert $\rho = \tau/\nu$. □

## References


Avery, Christopher and Judith Chevalier, "Herding over the Career," *Economics Letters*, 1999, 63(3), 327–333.

Banerjee, Abhijit V., "A Simple Model of Herd Behavior," *Quarterly Journal of Economics*, 1992, 107(3), 797–817.

Batchelor, Roy and Pami Dua, "Conservatism and Consensus-seeking among Economic Forecasters," *Journal of Forecasting*, 1992, 11, 169–181.

Bates, John M. and Clive Granger, "The Combination of Forecasts," *Operations Research Quarterly*, 1969, 20(4), 451–468.

Bernheim, B. Douglas, "A Theory of Conformity," *Journal of Political Economy*, 1994, 102(5), 841–877.

Bikhchandani Sushil, David Hirshleifer and Ivo Welch, "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades," *Journal of Political Economy*, 1992, 100(5), 992–1026.

Bhattacharya, Sudipto and Paul Pfleiderer, "Delegated Portfolio Management," *Journal of Economic Theory*, 1985, 36(1), 1–25.

Bunn, Derek W., "A Bayesian Approach to the Linear Combinations of Forecasts," *Operations Research Quarterly*, 1975, 26(2), 325–329.

Chevalier, Judith and Glenn Ellison, "Career Concerns of Mutual Fund Managers," *Quarterly Journal of Economics*, 1999, 114(2), 389–432.

Crawford, Vincent and Joel Sobel, "Strategic Information Transmission," *Econometrica*, 1982, 50(6), 1431–1452.

Croushore, Dean, "Introducing: The Survey of Professional Forecasters," Federal Reserve Bank of Philadelphia Business Review, 1993, November/December, 3–13.

Croushore, Dean, "The Livingston Survey: Still Useful After All These Years," Federal Reserve Bank of Philadelphia Business Review, 1997, March/April, 15–26.

Dawid, A. Philip, "Probability Forecasting", *Encyclopedia of Statistical Sciences: Volume 7*, 1986, edited by S. Kotz, N. L. Johnson and C. B. Read, New York: Wiley-Interscience, 210–218.

De Bondt, Werner F. M. and Richard H. Thaler, "Do Security Analysts Overreact?" *American Economic Review*, 1990, 80(2), 52–57.

de Finetti, Bruno, "Methods for Discriminating Levels of Partial Knowledge Concerning a Test Item", *British Journal of Mathematical and Statistical Psychology*, 1965, 18-I, 87–123. Reprinted in *Probability, Induction, and Statistics: The Art of Guessing*, Chapter 4, London: John Wiley & Sons, 1972, 25–63.

DeGroot, Morris, *Optimal Statistical Decisions*, 1970, New York: McGraw Hill.

Denton, Frank T., "The Effect of Professional Advice on the Stability of a Speculative Market", *Journal of Political Economy*, October 1985, 93(5), 977–93.

Ehrbeck, Tilman and Robert Waldmann, "Why are Professional Forecasters Biased? Agency versus Behavioral Explanations," *Quarterly Journal of Economics*, 1996, 111(1), 21–40.

Frankel, Jeffrey A. and Kenneth A. Froot, "Using Survey Data to Test Standard Propositions Regarding Exchange Rate Expectations," *American Economic Review*, 1987, 77(1), 133–153.

Geanakoplos, John, David Pearce and Ennio Stacchetti, "Psychological Games and Sequential Rationality," *Games and Economic Behavior*, 1989, 1(1), 69–79.

Graham, John, "Herding Among Investment Newsletters: Theory and Evidence," *Journal of Finance*, 1999, 54(1), 231–268.

Granger, Clive W. J., "Prediction with a Generalized Cost of Error Function," *Operational Research Quarterly*, 1969, 20(2), 199–207.

Granger, Clive W. J., "Outline of Forecast Theory Using Generalized Cost Functions," *Spanish Economic Review*, 1999, 1(2), 161–173.

Gul, Faruk and Russell Lundholm, "Endogenous Timing and the Clustering of Agents' Decisions," *Journal of Political Economy*, 1995, 103(5), 1039–1066.

Hartzmark, Michael L., "Luck versus Forecast Ability: Determinants of Trader Performance in Futures Markets, *Journal of Business*, 1991, 64(1), 49–74.

Holmström, Bengt, "Managerial Incentive Problems: A Dynamic Perspective," in *Essays in Economics and Management in Honor of Lars Wahlbeck*, Helsinki: Swedish School of Economics, 1982. Reprinted in *Review of Economic Studies*, 1999, 66(1), 169–182.

Hong, Harrison, Jeffrey D. Kubik and Amit Solomon, "Security Analysts' Career Concerns and Herding of Earning Forecasts," *Rand Journal of Economics*, 2000, 31(1), 121–144.

Hong, Harrison and Jeffrey D. Kubik, "Analyzing the Analysts: Career Concerns and Biased Earnings Forecasts," *Journal of Finance*, 2003, 58(1), 313–351.

Hotelling, Harold, "Stability in Competition," *Economic Journal*, 1929, 39(153), 41–57.

International Monetary Fund, *World Economic Outlook: Fiscal Policy and Macroeconomic Stability*, May 2001.

Johnson, Norman L. and Samuel Kotz, *Distributions in Statistics: Continuous Univariate Distributions - 1*, New York: Wiley, 1970.

Kahneman, Daniel and Amos Tversky, "On the Psychology of Prediction," *Psychological Review*, 1973, 80(4), 237–251.

Kandel, Eugene and Neil D. Pearson, "Differential Interpretation of Public Signals and Trade in Speculative Markets," *Journal of Political Economy*, 1995, 103(4), 831–872.

Kandel, Eugene and Ben Zion Zilberfarb, "Differential Interpretation of Information in Inflation Forecasts," *Review of Economics and Statistics*, 1999, 81(2), 217–226.

Keane, Michael P. and David E. Runkle, "Testing the Rationality of Price Forecasts: New Evidence from Panel Data," *American Economic Review*, 1990, 80(4), 714–735.

Keane, Michael P. and David E. Runkle, "Are Financial Analysts' Forecasts of Corporate Profits Rational?," *Journal of Political Economy*, 1998, 106(4), 768–805.

Kim, Oliver, Steve C. Lim and Kenneth W. Shaw, "The Use of Forecast Revision in Reducing Built-in Biases in Mean Analyst Forecasts," November 1998, University of Maryland and Long Island University mimeo.

Kutsoati, Edward and Dan Bernhardt, "Can Relative Performance Compensation Explain Analysts' Forecasts of Earnings?" August 2000, Tufts University and University of Illinois mimeo.

Lamont, Owen, "Macroeconomic Forecasts and Microeconomic Forecasters," *Journal of Economic Behavior and Organization*, 2002, 48(3), 265–280.

Laster, David, Paul Bennett and In Sun Geoum, "Rational Bias in Macroeconomic Forecasts," *Quarterly Journal of Economics*, 1999, 114(1), 293–318.

Lehmann, Erich L., "Ordered Families of Distributions," *Annals of Mathematical Statistics*, 1955, 26(3), 399–419.

Lehmann, Erich L., "Comparing Location Experiments," *Annals of Statistics*, 1988, 16(2), 521–533.

Michaely Roni and Kent L. Womack, "Conflict of Interest and Credibility of Underwriter Analyst Recommendations," *Review of Financial Studies*, 1999, 12(4), 654–686.

Milgrom, Paul, "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics*, 1981, 12, 380–391.

Morgan, John and Philip Stocken, "An Analysis of Stock Recommendations," *Rand Journal of Economics*, 2003, 34(1), 183–203.

Morris, Stephen and Hyun Shin, "The Social Value of Public Information," *American Economic Review*, 2002, 92, 1521–1534.

Muth, John F., "Rational Expectations and the Theory of Stock Price Movements," *Econometrica*, 1961, 29(3), 315–335.

Osband, Kent, "Optimal Forecasting Incentives," *Journal of Political Economy*, 1989, 97(5), 1091–1112.

Osborne, Martin and Carolyn Pitchik, "The Nature of Equilibrium in a Location Model", *International Economic Review*, 1986, 27(1), 223–237.

Ottaviani, Marco and Peter Sørensen, "Herd Behavior and Investment: Comment," *American Economic Review*, 2000, 90(3), 695–704.

Ottaviani, Marco and Peter N. Sørensen, "Forecasting and Rank-Order Contests," 2002, London Business School and University of Copenhagen mimeo.

Ottaviani, Marco and Peter N. Sørensen, "Reputational Cheap Talk," 2003, London Business School and University of Copenhagen mimeo.

Prati, Alessandro and Massimo Sbracia, "Currency Crises and Uncertainty about Fundamentals," 2002, IMF and Bank of Italy mimeo.

Prendergast, Canice, "A Theory of "Yes Men"," *American Economic Review*, 1993, 83(4), 757–770.

Prendergast, Canice and Lars Stole, "Impetuous Youngsters and Jaded Oldtimers: Acquiring a Reputation for Learning," *Journal of Political Economy*, 1996, 104(6), 1105–1134.

Romer, Christina D. and David H. Romer, "Federal Reserve Information and the Behavior of Interest Rates," *American Economic Review*, 2000, 90(3), 429–457.

Savage, Leonard, "Elicitation of Personal Probabilities and Expectation," *Journal of the American Statistical Association*, 1971, 66(336), 783–801.

Scharfstein, David and Jeremy Stein, "Herd Behavior and Investment," *American Economic Review*, 1990, 80(3), 465–479.

Stark, Tom, "Macroeconomic Forecasts and Microeconomic Forecasters in the Survey of Professional Forecasters," 1997, Federal Reserve Bank of Philadelphia Working Paper 97-10.

Steele, J. Michael and James Zidek, "Optimal Strategies for Second Guessers," *Journal of the American Statistical Association*, 1980, 75(371), 596–601.

Stekler, Herman O., "Who Forecasts Better?" *Journal of Business and Economics Statistics*, 1987, 5(1), 155–158.

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

Welch, Ivo, "Herding among Security Analysts," *Journal of Financial Economics*, 2000, 58(3), 369–396.

Zarnowitz, Victor and Luis A. Lambros, "Consensus and Uncertainty in Economic Predictions," *Journal of Political Economy*, 1987, 95(3), 591–621.

Zellner, Arnold, "Biased Predictors, Rationality and the Evaluation of Forecasts," *Economics Letters*, 1986, 21, 45–48.

Zitzewitz, Eric, "Measuring Herding and Exaggeration by Equity Analysts," 2001a, Stanford GSB mimeo.

Zitzewitz, Eric, "Opinion-Producing Agents: Career Concerns and Exaggeration," 2001b, Stanford GSB mimeo.

Zwiebel, Jeffrey, "Corporate Conservatism and Relative Compensation," *Journal of Political Economy*, 1995, 103(1), 1–25.

## Omissible Proofs

**Proof of Proposition 7.** Assume that the opponents use a linear strategy $\hat{m}(s) = Cs + (1-C)\mu$ where $C \in (0,1]$. The hypothetical observation of $x = c_i$ and of signal $s_i$ gives two independent sources of information about $y$. Updating normal beliefs as usual, we find $y|c_i, s_i \sim N\left(\left(\nu\mu + \tau_0 c_i + \tau s_i\right)/\left(\nu + \tau_0 + \tau\right), 1/\left(\nu + \tau_0 + \tau\right)\right)$. Conditionally on $x = c_i$ and $s_i$, the message $\hat{m}(s_j) = Cy + C\varepsilon_j + (1-C)\mu$ is then normally distributed with mean $C\left(\nu\mu + \tau_0 c_i + \tau s_i\right)/\left(\nu + \tau_0 + \tau\right) + (1-C)\mu$ and variance $C^2\left(\nu + \tau_0 + 2\tau\right)/\left(\left(\nu + \tau_0 + \tau\right)\tau\right)$. This gives the mass of correct opponent guesses,

$$\gamma(c_i|c_i, s_i) = \sqrt{\frac{(\nu+\tau_0+\tau)\tau}{(\nu+\tau_0+2\tau)C^2 2\pi}} \exp\left(-\frac{(\nu+(1-C)\tau_0+\tau)^2\tau}{2(\nu+\tau_0+2\tau)(\nu+\tau_0+\tau)C^2}\left(c_i - \frac{(\nu+(1-C)(\tau_0+\tau))\mu + C\tau s_i}{\nu+(1-C)\tau_0+\tau}\right)^2\right).$$

Clearly, $\gamma(c_i|c_i, s_i)$ is centered between $\mu$ and $s_i$. Nevertheless, when $C < 1$ this center remains closer to $\mu$ than the honest estimate $h(s_i)$ since the weight on $s_i$ is smaller: $C\tau/\left(\nu + (1-C)\tau_0 + \tau\right) < \tau/\left(\tau + \nu\right)$. Provided there exists a best response, this response is therefore biased away from $\mu$, by the same logic as before.

Recall that $x|s_i \sim N\left(\left(\nu\mu + \tau s_i\right)/\left(\nu + \tau\right), \left(\nu + \tau_0 + \tau\right)/\left(\left(\nu + \tau\right)\tau_0\right)\right)$. This objective function $\log\left(U_i(c_i|s_i)\right) = \log\left(q_i\left(c_i|s_i\right)\right) - \log\left(\gamma(c_i|c_i, s_i)\right)$ is quadratic in the choice variable $c_i$. The first order condition characterizing the unique maximizer is

$$\frac{(\nu+(1-C)\tau_0+\tau)^2\tau}{(\nu+\tau_0+2\tau)C^2}\left(c_i - \frac{(\nu+(1-C)(\tau_0+\tau))\mu + C\tau s_i}{\nu+(1-C)\tau_0+\tau}\right) = \tau_0\left(\nu+\tau\right)\left(c_i - \frac{\nu\mu+\tau s_i}{\nu+\tau}\right).$$

Gathering terms, this can be rewritten as $c_i = Ks_i + (1-K)\mu$. The equilibrium fixed point condition requires that the weight on $s_i$ should be $C$, i.e.

$$\begin{aligned} & \tau_0\left(\nu+\tau\right)\left(\nu+\tau_0+2\tau\right)C^2 - \left(\nu+(1-C)\tau_0+\tau\right)^2\tau \\ = \; & \tau\left(\tau_0\left(\nu+\tau_0+2\tau\right)C - \left(\nu+(1-C)\tau_0+\tau\right)\tau\right). \end{aligned}$$

This quadratic equation in $C$ can be easily solved. The total coefficient on $C^2$ on the left hand side is positive. At $C = 0$, the right hand side exceeds the left hand side. At $C = 1$ the opposite is true. The unique solution $C \in (0,1)$ defines an equilibrium, if it satisfies the second order condition. The second order condition requires positivity of the left hand side, or equivalently of the right hand side, i.e. $C > \tau\left(\nu+\tau_0+\tau\right)/\left(\tau_0\left(\nu+\tau_0+3\tau\right)\right)$. This condition can be checked by inserting $\tau\left(\nu+\tau_0+\tau\right)/\left(\tau_0\left(\nu+\tau_0+3\tau\right)\right)$ for $C$ in the fixed point equation and verifying that the right hand side exceeds the left hand side. This criterion for equilibrium existence is then

$$\begin{aligned} & \left(\nu+\tau\right)\left(\nu+\tau_0+2\tau\right)\frac{\tau^2(\nu+\tau_0+\tau)^2}{\tau_0(\nu+\tau_0+3\tau)^2} - \left(\frac{(\nu+\tau_0+\tau)(\nu+\tau_0+2\tau)}{\nu+\tau_0+3\tau}\right)^2\tau \\ < \; & \tau\left(\left(\nu+\tau_0+2\tau\right)\frac{\tau(\nu+\tau_0+\tau)}{(\nu+\tau_0+3\tau)} - \left(\frac{(\nu+\tau_0+\tau)(\nu+\tau_0+2\tau)}{\nu+\tau_0+3\tau}\right)\tau\right). \end{aligned}$$

For small $\tau_0$ this fails since $\left(\nu+\tau\right)^3\left(\nu+2\tau\right)\tau^2/\left(\nu+3\tau\right)^2 > 0$. For large $\tau_0$ it holds since the coefficient on $\tau_0^2$ is $-\tau < 0$. □

**Proof of Proposition 8.** First, assume that all opponents $j \neq i$ use a fully separating strategy. Besides the forecast of forecaster $i$, the evaluator observes $n$ independent signals about $y$, namely every $s_j$ where $j \neq i$ and $x$. From the well-known updating of beliefs on a normal state, this is equivalent to the observation of just one more precise signal about $y$. So, without loss of generality, we can imagine that $x$ itself contains all the evaluator's external information on $y$. Since $x = y + \varepsilon_0$ where $\varepsilon_0$ is independent of $\varepsilon_i$ and $t_i$, $y$ is a sufficient statistic for $x$ and the law of iterated expectations gives $p_i(t_i|m_i, x) = E[p_i(t_i|m_i, y)|x]$. Observe then that

$$\begin{aligned} U_i(m_i|s_i) &= \int_{-\infty}^{\infty} \left[ \int_0^{\infty} u(t) p_i(t_i|m_i, x)\, dt \right] q_i(x|s_i)\, dx \\ &= \int_{-\infty}^{\infty} \left[ \int_0^{\infty} u(t) \int_{-\infty}^{\infty} p_i(t_i|m_i, y) q(y|x)\, dy\, dt \right] q_i(x|s_i)\, dx \\ &= \int_{-\infty}^{\infty} \left[ \int_0^{\infty} u(t) p_i(t_i|m_i, y)\, dt \right] \left[ \int_{-\infty}^{\infty} q(y|x) q_i(x|s_i)\, dx \right] dy. \end{aligned}$$

This resembles the original expression for $U_i(m_i|s_i)$, except that the old $q_i(y|s_i)$ has been replaced by the average $\int_{-\infty}^{\infty} q(y|x) q_i(x|s_i)\, dx$ of the evaluator's beliefs. Both $q(y|x)$ and $q_i(x|s_i)$ are normal p.d.f.s, and their product can be rewritten as $A_0 \exp(-A_1(x - A_2(y, s_i))^2 - A_3(y - A_4(s_i))^2)$ where $A_0, A_1, A_3$ are constants not depending on $x, y, s_i$, the constant $A_2$ depends on $y, s_i$ only, and $A_4(s_i) = [\tau_0 \tau_i s_i + (\nu + \tau_0 + \tau_i)\nu\mu]/[\tau_0 \tau_i + (\nu + \tau_0 + \tau_i)\nu]$. We then find $\int_{-\infty}^{\infty} q(y|x) q_i(x|s_i)\, dx = A_5 \exp\left(-A_3(y - A_4(s_i))^2\right)$ where $A_5$ is independent of $y, s_i$ since the normalizing constant of a normal p.d.f. does not involve the mean. We conclude that the forecaster's objective function is of the same form as previously, where $\int_{-\infty}^{\infty} q(y|x) q_i(x|s_i)\, dx$ is a normal p.d.f. with mean $A_4(s_i)$ strictly between $\mu$ and $s_i$. As in Proposition 3, forecaster $i$ will deviate from a fully separating strategy $m_i$ by issuing $m_i(A_4(s_i)) \neq m_i(s_i)$ for any $s_i \neq \mu$.

Second, assume that all the opponents apply the binary strategy with threshold $\mu$. We use again that $U_i(m_i|s_i) = E[W_i(m_i|y) E[q(y|x, m_{-i})|s_i]]$ as derived above, where $W_i(m_i|y)$ is precisely the same as in the proof of Proposition 5. That proof carries over to this new situation, once we verify that $E[q(y|x, m_{-i})|s_i] \geq E[q(2\mu - y|x, m_{-i})|s_i]$ when $y \geq \mu$ and $s_i \geq \mu$. Using symmetry of $q(y|x, m_{-i})$, we have

$$\begin{aligned} & E[q(y|x, m_{-i})|s_i] - E[q(2\mu - y|x, m_{-i})|s_i] \\ &= \int_{-\infty}^{\infty} \sum_{M_{-i}} \left[ (q(y|x, m_{-i}) - q(2\mu - y|x, m_{-i}))\, q_i(x, m_{-i}|s_i) \right] dx \\ &= \int_{\mu}^{\infty} \sum_{M_{-i}} \left[ (q(y|x, m_{-i}) - q(2\mu - y|x, m_{-i})) \left( q_i(x, m_{-i}|s_i) - q_i\left(2\mu - x, m_{-i}^c|s_i\right) \right) \right] dx \end{aligned}$$

where $M_{-i} = \{m_L, m_H\}^{n-1}$ and $m_{-i}^c$ denotes the vector of messages opposite to $m_{-i}$, i.e. in which every $m_H$ is replaced by $m_L$ and vice versa. First, when $y \geq \mu$ and $x \geq \mu$, for any $m_{-i}$, $q(y|x, m_{-i}) \geq q(2\mu - y|x, m_{-i})$ since $x$ is closer to $y$ than to $2\mu - y$. This implies that $\int_{\mu}^{\infty} \sum_{M_{-i}} [q(y|x, m_{-i}) - q(2\mu - y|x, m_{-i})] dx \geq 0$. Second,

$\int_{\mu}^{\infty} \sum_{M_{-i}} \left[ q_i \left( x, m_{-i} | s_i \right) - q_i \left( 2\mu - x, m_{-i}^c | s_i \right) \right] dx = \Pr(x \geq \mu | s_i) - \Pr\left( x \leq \mu | s_i \right) \geq 0$ since $s_i \geq \mu$. Finally, the result that $E\left[ q(y|x, m_{-i}) | s_i \right] \geq E[q(2\mu - y|x, m_{-i}) | s_i]$ follows from the positive correlation of $q(y|x, m_{-i}) - q(2\mu - y|x, m_{-i})$ with $q_i(x, m_{-i}|s_i) - q_i(2\mu - x, m_{-i}^c|s_i)$. This is due to the fact that when $s_i, x, y \geq \mu$, every opponent's message $m_H$ is believed more frequent than $m_L$, and every opponent's message $m_H$ gives greater $q(y|x, m_{-i})$ and smaller $q(2\mu - y|x, m_{-i})$ than message $m_L$. $\square$