Consensus forecasts and inefficient information aggregation

Christopher Crowe*

PRELIMINARY AND INCOMPLETE.

August 25, 2009

Abstract

Consensus forecasts are inefficient, because the weight placed on aggregate new private information relative to the prior reflects the optimal choices of the underlying individual forecasters—whose private signals are noisier individually than in aggregate—and as a result the consensus forecast over-weights the prior (Lim, Kim and Shaw, 2001). This paper assesses the empirical relevance of this result using a cross-country panel of growth forecasts. There is strong evidence that consensus forecasts are in fact inefficient. Perhaps more surprisingly, there is no evidence that this inefficiency is understood and corrected for by forecasters themselves. These results have implications for information aggregation via the price mechanism—offering one explanation for the positive short run serial correlation observed in stock prices and the apparent success of momentum trading strategies, and posing a challenge for the efficient markets hypothesis more generally—and also for the wider literature on information processing (Morris and Shin, 2002; Benabou, 2009).

JEL classification:

Keywords: Consensus Forecasts; Momentum Trading

*International Monetary Fund, Research Department. ccrowe@imf.org. +1 202 623-5303.
This work reflects the views of the author alone and does not reflect the views of the IMF, its Executive Board or Management.

1 Introduction

Consensus forecasts—mean forecasts from a panel of individual forecasters—are inefficient. This result is not new, but it appears not to be well known and its empirical relevance has yet to be tested.[1] This paper's illustrates the empirical relevance of the phenomenon using the well-known and widely used forecasts dataset produced by Consensus Economics, and discusses some potential ramifications.

To understand this paper's empirical strategy, one must first appreciate the intuition behind the theoretical result. Consider a simple signal extraction model in which a group of agents form independent forecasts based on a private signal and a common prior. Each individual forecaster, if producing an honest minimum variance forecast, will update their prior based on their private information about the state of the economy: the noisier a forecaster's private signal, the less weight he or she places on it compared to the prior. The consensus forecast will then weight the prior and the average private signal using the same relative weights (on average) as the private forecasters. But the mean private signal has a lower variance than each of the individual signals (this is the argument for employing the consensus forecast in the first place). In other words, the consensus forecast over-weights the prior, and as a result is inefficient. Notably, it is (negatively) correlated with its own forecast error. This result is a general one and does not rely on strategic forecasting behavior or heterogeneous signal precision.

One can test the empirical relevance of the result by analyzing the relationship between the *forecast errors* associated with the consensus forecast (the difference between the consensus forecast and the actual realization) and the *revisions* to the consensus forecast (the difference between the current consensus forecast and the previous one). In assessing this relationship, one can differentiate between two competing assumptions about how agents use the information in the last published consensus to update their prior. The first potential assumption is that agents update their prior naively, using the most recent previous consensus forecast for the same period as the best signal of the existing aggregate information available. In this case, the inefficiency of the consensus forecast will show up in a significant negative regression coefficient in a simple regression of the consensus forecast error on the most recent forecast revision. Under the second assumption, agents update their prior rationally, taking into account the inefficiency of the last consensus forecast. In this case, the forecast error will be correlated with not only the most recent consensus forecast revision, but with *every previous revision* for the forecast period in question. In a regression of the consensus forecast error on the current and lagged forecast revisions, the coefficients should be significantly different from zero and alternate in sign from negative to positive.

[1] Kim, Lim and Shaw (2001) is the only published reference on the theoretical finding, although an early draft of Ottaviani and Sorensen (2006) also makes reference to it. However, a reading of the literature on forecasting suggests that the finding does not appear to be well-known.

Hence, a significant negative coefficient on the most recent revision is evidence that the predicted inefficiency in the consensus forecasts is quantitatively significant, under both assumptions concerning how forecasters update their prior. By contrast, significant coefficients on lagged forecast revisions—with a pattern of alternating signs—provide evidence in favor of the second assumption, and against the first.

The results of this empirical exercise are unambiguous. There is very strong evidence that the consensus forecasts are indeed inefficient, and that this inefficiency is quantitatively important, particularly at forecast horizons of around 12 months or less. However, there is no evidence that this inefficiency is understood by the forecasters themselves and incorporated into their forecasts. This latter result is, in a sense, unsurprising, given that the theoretical result concerning consensus forecast inefficiency does not appear to be well known. Nevertheless, it is useful to have robust empirical evidence based on a well-specified model of forecaster behavior. Moreover, the second result reinforces the relevance of the first. If agents were able to back out efficient forecasts from the consensus forecasts, then the fact that the consensus forecast is inefficient would be of only theoretical interest. But since even professional forecasters—typically sophisticated observers with some economics training whose analysis may also inform financial market trading strategies—appear to take the consensus at face value, then the inefficiency of the consensus forecast seems likely to be an empirically relevant phenomenon.

These empirical results have a number of interesting implications. Consensus forecasts—the mean forecast from a panel of professional forecasters—are a widely used tool for researchers and policymakers. For instance, inflation-targeting central banks frequently present consensus private sector forecasts alongside their own in-house forecasts as a means of increasing the credibility of their own forecasts and policymaking. Consensus forecasts are also employed in the academic literature as proxies for public knowledge about the future state of the economy (cite) or as a benchmark against which to test individual public or private sector forecasters (Romer and Romer, 2000). However, this paper argues that more caution should be exercised when using consensus forecasts. In particular, the consensus forecast does not typically represent the best (most efficient) estimate of the true state of the world, conditional on the aggregate information set. Its use as a benchmark may therefore flatter the performance of alternative forecasts.

This paper also has implications for the debate over the efficacy of releasing more transparent public information sparked in part by the contribution of Morris and Shin (2002) and, in the policy sphere, by moves towards greater transparency by central banks, notably in the context of inflation targeting (IT) regimes. One can show that—even without Morris and Shin's "Beauty Contest" component to the forecasting process—more transparent public information can increase the forecast error for the consensus forecast, even as the forecast errors of individual forecasts are reduced. Moreover, the parameter restrictions (with respect to the relative noisiness of the public and private information) for this negative effect on consensus forecast accuracy are considerably less restrictive than in the Morris and Shin case for individual forecasts. Whereas, in the Morris and Shin case, the condition for more accurate public information to reduce the accuracy of private forecasts requires that the public signals are no less than one-eighth as accurate as the private signals, the negative effect for *consensus* forecasts holds if the public signals are simply less accurate than the private. While the former condition is indeed unlikely, the latter seems plausible in some cases. More broadly, the lesson that better public information can lead to worse collective judgments is, I believe, a general one with many applications.

A similar phenomenon of over-weighting the prior may manifest itself in the adjustment of asset prices to new information. Asset prices are forward looking, and therefore represent the aggregate response of heterogeneous market participants to diffuse and noisy signals about future returns. Essentially, asset prices are consensus forecasts. Since asset prices may therefore initially underreact to new information, this could then help to explain the observed short-run serial correlation of asset returns and the resulting success of momentum trading strategies (Jegadeesh and Titman, 1993, 2001; Hong and Stein, 1999). Such trading strategies are all the more likely to be successful given the evidence from the forecasts data that professional forecasters (many of whom are also financial market participants) do not make the necessary adjustment to the consensus forecast in practice. Moreover, since efficient forecasts of future prices rely on information on past prices as well as on fundamentals, this model suggests a channel for self-fulfilling beliefs and the emergence of bubbles and crashes.

This paper's results can also be related to other models of information aggregation and social learning. For instance, in Benabou's (2009) model of "groupthink" or collective delusion, agents' anticipatory feelings (deriving utility not just from the realization of an event but the anticipation of its realization) can generate individually-rational delusion that is collectively reinforcing, creating incentives for "groupthink" (collective blindness to reality). This paper demonstrates that a simple signal extraction model can generate a similar phenomenon—the collective judgment embodied in the consensus forecast puts too much emphasis on the group's shared prior over new information—without departing from standard specifications of utility. The model outlined here differs from models of social learning that generate herd behavior (e.g. Banerjee, 1992) in that the inefficiency of the consensus forecast does not stem from forecasts being made public sequentially, but from the communication of posterior beliefs rather than raw signals. Nevertheless, it shares the core feature of these models identified by Benabou (2009): "the key problem is a failure to aggregate private signals and its cure resides in more communication." A subtle insight from this model is that the *form* of communication is central to efficient aggregation. Agents must be induced to communicate not their best guess, but rather their idiosyncratic information. Moreover, the evidence suggests that agents fail to optimally use the aggregate information that does become available.

Finally, our results point to an additional kind of aggregation problem—moving between microeconomic and macroeconomic models. For instance, it may be invalid to use macroeconomic or average data to derive insights into individual behavior: as we demonstrate in this paper, the empirical finding that consensus forecasts are inefficient does not imply that *individual* forecasts are not rational. More significantly, our results suggest that insights based on individual behavior do not necessarily carry over to a macroeconomic (multi-agent) environment. Even when agents' behavior imposes no externality, insights derived from individual rationality (i.e. that expectations are rational) may not hold in aggregate. The economy does not behave "as if" it were populated by a single rational representative agent.[2]

Section 2 provides a formal discussion of the paper's key argument. It goes on to demonstrate that an efficient forecast can be recovered by reducing the overweighting on the prior using a linear combination of the consensus forecast and the prior. The size of the required adjustment is increasing in the relative variance of the individual forecasters' idiosyncratic private information.[3]

Section 3 illustrates the argument using a large dataset of individual monthly forecasts of annual economic growth for a cross-section of countries between 1989 and 2006. It describes the data, then presents evidence of the existence and size of the inefficiency associated with employing consensus forecasts for different countries and time horizons. There is strong evidence of inefficiency for consensus forecasts, much less so for individual forecasts, as predicted by the model.[4] Using the adjustment method outlined above and employing the most recent previous consensus forecast as a proxy for the prior, one can obtain adjusted forecasts with mean squared errors up to 17 percent lower than the raw consensus forecasts, even out of sample. This effect may appear small; however, it would be surprising, given the inherent uncertainty about the future path of the economy, if a simple adjustment technique could eliminate a substantial fraction of the uncertainty. Moreover, since consensus forecasts are generally thought to be at, or near, the frontier in terms of forecast accuracy (see, e.g., Ang et al., 2007, for an application to inflation forecasts), the fact that a simple adjustment technique can reduce forecast errors *at all* is significant. However, there is no evidence that forecasters make the necessary adjustment to the previous month's consensus forecast in updating their prior.

To illustrate an application of the model, section 4 uses our adjusted forecasts to show that private forecasters may have been better informed on average about the Asian economic crisis in 1997-8 than the raw consensus forecasts would suggest. Section 5 provides a fuller discussion of how our model relates to some of the existing literature discussed above.

Section 6 then concludes with a discussion of implications for optimal forecasting methods. Forecasting technologies that decentralize information-gathering

[2] The fact that this aggregation problem shows up with respect to expectations is particularly worrisome given the central role ascribed to rational expecations in most modern macroeconomic models.

[3] This section replicates the key results in Lim, Kim and Shaw (2001) for the case with a continuum rather than a finite number of individual forecasters.

[4] The difference in performance between consensus and individual forecast rationality tests has been noted before (Batchelor and Dua, 1991, for instance, who argue that "rationality tests on the consensus forecasts provide a poor guide to the extent of rationality among individual forecasters").

but centralize the production of a final forecast (as are typically employed in a central bank) may help to alleviate this problem of over-weighting the prior. Forecasting aggregating services (such as the one whose data are employed in this paper) may need to adjust their methodology to provide more efficient consensus forecasts (although, for many countries, the efficiency gain is likely to be relatively small). One method is to employ an ex post adjustment similar to the one used in this paper. An alternative method would be to attempt to elicit individual forecasts that place a higher weight on each forecaster's private signal compared to their best forecast. One means of achieving this in practice would be to provide incentives for strategic forecasting such that forecasters attempt to differentiate their forecasts from those of others (Ottaviani and Sørensen, 2006).

2 Model and Proof

The following model captures the basic features of this discussion. A variable y is distributed iid with mean μ and variance σ_ν^2:

$$y = \mu + \nu \tag{1}$$

A continuum of agents $i \in [0,1]$ each receives a private signal of y, given by x_i:

$$x_i = y + u + \varepsilon_i \tag{2}$$

where u is an aggregate error, iid mean-zero with variance σ_u^2, and ε_i are iid mean-zero idiosyncratic errors with variance σ_ε^2. The information structure, including all variances, are common knowledge.

Agents form forecasts following the usual Bayesian updating:

$$f_i = F_i x_i + (1 - F_i)\,\mu \tag{3}$$

where F_i minimizes the root mean squared forecast error and is therefore given by:

$$F_i = F = \frac{\sigma_\nu^2}{\sigma_\nu^2 + \sigma_u^2 + \sigma_\varepsilon^2} \forall i \tag{4}$$

Proposition 1 *The consensus forecast is inefficient, even as individual forecasts are all efficient.*

Proof. The individual forecasts are uncorrelated with their forecast errors:

$$\begin{aligned} E\left[f_i\left(f_i - y\right)\right] &= E\left[\left(\mu + F\left(u + \nu + \varepsilon_i\right)\right)\left(F\left(u + \varepsilon_i\right) - \left(1 - F\nu\right)\right)\right] \\ &= F\left(F\left(\sigma_u^2 + \sigma_\varepsilon^2\right) - \left(1 - F\right)\sigma_\nu^2\right) \\ &= 0 \end{aligned} \tag{5}$$

Define the consensus forecast as the mean forecast:

$$\overline{f} = E\left[f_i\right] = F\left(y + u\right) + \left(1 - F\right)\mu \tag{6}$$

Then the consensus forecast is correlated with its forecast errors:

$$\begin{aligned} E\left[\overline{f}\left(\overline{f}-y\right)\right] &= E\left[\left(\mu+F\left(u+\nu\right)\right)\left(Fu-\left(1-F\right)\nu\right)\right] \\ &= F\left(F\sigma_u^2-\left(1-F\right)\sigma_\nu^2\right) \\ &= -F^2\sigma_\varepsilon^2 \leq 0 \end{aligned} \tag{7}$$

■

The easiest way to understand this result is to compare the individually-rational relative weight F on the individual signal $(y+u+\varepsilon_i)$ with the weight that would arise from minimizing the root mean squared forecast error associated with the aggregate signal $(y+u)$:

$$F_C = \frac{\sigma_\nu^2}{\sigma_\nu^2+\sigma_u^2} = \left(1+\frac{\sigma_\varepsilon^2}{\sigma_\nu^2+\sigma_u^2}\right)F > F \tag{8}$$

Because the aggregate signal has a lower variance than the individual signals, it should receive a higher weight. Then the optimal centralized forecast is uncorrelated with its forecast errors:

$$\begin{aligned} E\left[f_C\left(f_C-y\right)\right] &= E\left[\left(\mu+F_C\left(u+\nu\right)\right)\left(F_C u-\left(1-F_C\nu\right)\right)\right] \\ &= F_C\left(F_C\sigma_u^2-\left(1-F_C\right)\sigma_\nu^2\right) \\ &= 0 \end{aligned} \tag{9}$$

If the prior μ and the variances of the various shocks are known, and once the consensus forecast is observed, the efficient forecast can be derived as a linear combination of the consensus forecast and the prior:

$$\begin{aligned} f_C &= \overline{f}+\frac{F_C-F}{F}\left(\overline{f}-\mu\right) \\ &= \overline{f}+\frac{\sigma_\varepsilon^2}{\sigma_\nu^2+\sigma_u^2}\left(\overline{f}-\mu\right) \end{aligned} \tag{10}$$

3 Application to Consensus Economics Data

3.1 Data

This study employs forecasts of economic growth for 38 countries over the period 1989-2008 at the level of the individual forecaster, collected by Consensus Economics. Forecasters are surveyed at the start of each month and their responses are published around the 10th of the month. Country coverage is more limited at the start of the sample period; moreover, some countries are surveyed only once every two months. The forecasts are for annual economic growth over the current and next calendar year: the forecast horizon (dated to the end of the forecast year) therefore varies from 0 to 23 months. The number of forecasters per country, forecast period and forecast horizon varies between

5 and 39, with an interquartile range of 14 to 21 and a mean value of around 18. The baseline sample drops outliers from the underlying dataset of individual forecasters, to prevent coding errors or extreme values from distorting the results.[5] The dataset includes 167,802 individual forecasts. Efficiency gains are calculated using a truncated sample of forecasts up to 2006, in order to allow the calculation of gains out of sample. The out of sample dataset includes all forecasts for 2008 (made in 2007 and 2008) and forecasts for 2007 made in 2007 (i.e. with a forecast horizon of less than 12 months).

Consensus forecasts are obtained by taking a simple arithmetic average across all forecasters for the forecast year (τ), time horizon (h) and country (c) in question. The forecast by individual forecaster i is therefore denoted $f^i_{ch\tau}$ while the consensus forecast is denoted $\overline{f}_{ch\tau}$. The consensus dataset is therefore a three-dimensional panel covering up to 38 countries, 20 years and 24 forecast horizons. Forecast errors are calculated with respect to annual GDP growth in the calendar year in question, $y_{c\tau}$ taken from the IMF's World Economic Outlook (WEO) database. We abstract from questions of data revision by using final data.

3.2 Methodology

To test the main proposition of the paper, this section provides evidence on the correlation between forecast errors and forecasts for individual and consensus forecasts. To take this proposition to the data, one must first modify the model to match the information structure of the data. The key difference between the simple model outlined in the previous section and the forecasting environment of the data is that repeated forecasts are made for each forecast period, so that the public information reflected in the prior $\mu_{ch\tau}$ will include previously published consensus forecasts.

The assumed information structure is then as follows: each forecast period, agents receive a signal $x^i_{ch\tau}$ which includes an idiosyncratic component $\varepsilon^i_{ch\tau}$ and a common component $u_{ch\tau}$, as before. To simplify the analysis, all the error terms are assumed iid. From (10), for a given prior underlying the previous forecast, $\mu_{c(h+1)\tau}$ (with variance given by σ^2_{h+1}), the *optimal forecast* going into the current forecasting period (i.e. before individual signals $x^i_{ch\tau}$ are observed), which should form the *prior* for this period, is therefore given by:

$$\mu_{ch\tau} = \overline{f}_{c(h+1)\tau} + \frac{\sigma^2_{\varepsilon}}{\sigma^2_{h+1} + \sigma^2_u} \left(\overline{f}_{c(h+1)\tau} - \mu_{c(h+1)\tau} \right) \forall h = \{0, 1, ..., 22\} \quad (11)$$

where $\mu_{c23\tau} = \mu_{c\tau}$ (i.e. the prior for the forecast at the maximum forecasting horizon is the simple ex ante prior).

To test the empirical relevance of the inefficiency result, I make two alternative assumptions about the prior $\mu_{ch\tau}$ that is adopted in each forecasting period. The first assumption is that forecasters are fully rational and arrive at the prior

[5] Outliers are defined as the bottom and top 0.5 percent of the sample of individual forecast errors. In fact the results are very similar when these outliers are not dropped.

using (11). The second assumption is that forecasters do not undertake the necessary adjustment implied by (11), but naively use the previous consensus forecast ($\mu_{ch\tau} = \overline{f}_{c(h+1)\tau}$), perhaps due to bounded rationality.

The condition for rational individual forecasts in (5) can be tested by testing the following empirical hypothesis:

$$E\left[\left(f^i_{ch\tau} - \mu_{ch\tau}\right)\left(f^i_{ch\tau} - y_{c\tau}\right)\right] = 0 \tag{12}$$

If individual agents are forming rational forecasts in the traditional sense, then their update with respect to the previous period's consensus forecast will be uncorrelated with their forecast error (otherwise, agents could exploit this correlation to arrive at forecasts with lower-variance forecast errors).

Similarly, the assertion in (7) that consensus forecasts are inefficient is tested via the empirical hypothesis:

$$E\left[\left(\overline{f}_{ch\tau} - \mu_{ch\tau}\right)\left(\overline{f}_{ch\tau} - y_{c\tau}\right)\right] \leq 0 \tag{13}$$

To test these hypotheses in a regression setting, one can use (10) to give the difference between the consensus and optimal forecast:

$$\overline{f}_{ch\tau} - f^C_{ch\tau} = -\frac{\sigma^2_\varepsilon}{\sigma^2_h + \sigma^2_u}\left(\overline{f}_{ch\tau} - \mu_{ch\tau}\right) \tag{14}$$

While we do not observe the optimal forecast $f^C_{ch\tau}$, this is by definition equal to the actual value $y_{c\tau}$ plus a residual that is orthogonal to $f^C_{ch\tau}$:

$$f^C_{ch\tau} = y_{c\tau} + e_{cht} \tag{15}$$

One can therefore estimate the following regression in order to derive estimates of the degree of inefficiency in the consensus forecasts and the size of the correction required to give efficient forecasts:

$$\begin{aligned} \overline{f}_{ch\tau} - y_{c\tau} &= a_h - b_h\left(\overline{f}_{ch\tau} - \mu_{ch\tau}\right) + e_{cht} \\ \text{where } E\left[b_h\right] &= \frac{\sigma^2_\varepsilon}{\sigma^2_h + \sigma^2_u}; E\left[a_h\right] = 0 \end{aligned} \tag{16}$$

Hence, assuming that our simple signal extraction model is valid, the slope coefficient b_h, in addition to signaling the degree of inefficiency in the consensus forecasts, provides an estimate of the relative variance of the individual signals.

An estimate of the efficient forecast $f^C_{ch\tau}$ can then be obtained by transforming the consensus forecast using the estimated regression coefficient from (16):[6]

$$\widehat{f}^C_{ch\tau} = \overline{f}_{ch\tau} + b_h\left(\overline{f}_{ch\tau} - \mu_{ch\tau}\right) \tag{17}$$

[6] Since the theory predicts that consensus forecasts are unbiased (as long as the underlying individual forecasts are unbiased), the adjustment does not take into account the estimated bias coefficient a_h.

To provide a metric for the efficiency gain that can be obtained via this transformation, we compare the root mean square error (RMSE) of the raw forecasts $\left(\overline{f}_{ch\tau} - y_{c\tau}\right)$ with the RMSE of the adjusted forecasts:

$$\text{Efficiency Gain } (\%) = 100 \times \frac{RMSE\left(\overline{f}_{ch\tau} - y_{c\tau}\right) - RMSE\left(\widehat{f}^{C}_{ch\tau} - y_{c\tau}\right)}{RMSE\left(\overline{f}_{ch\tau} - y_{c\tau}\right)} \tag{18}$$

This measure can be compared across forecast horizons and countries to ascertain in which environments our adjustment is most useful for improving forecasting efficiency.

Rational Priors

Using (11) iteratively to substitute for $\mu_{ch\tau}$, $\mu_{c(h+1)\tau}$,..., $\mu_{c(22)\tau}$ in (16), one then obtains the following regression specification:[7]

$$\begin{aligned} \overline{f}_{ch\tau} - y_{c\tau} &= a + \sum_{j=h}^{22} \left[B_{jh} \left(\overline{f}_{cj\tau} - \overline{f}_{c(j+1)\tau} \right) \right] + e_{cht} && (19) \\ \text{where } B_{jh} &= \prod_{k=j}^{h} (-b_k) ; E[b_h] = \frac{\sigma^2_{\varepsilon}}{\sigma^2_h + \sigma^2_u} \geq 0 \end{aligned}$$

Note the distinctive pattern of alternating signs predicted for the B_{jh} coefficients:

$$B_{hh} \leq 0, B_{(h+1)h} \geq 0, B_{(h+2)h} \leq 0, B_{(h+3)h} \geq 0, ... \tag{20}$$

While the first inequality is implied by the model with both rational and naive priors, the subsequent inequalities are necessary under the rational priors case. Hence, evidence against this pattern of alternating signs would lead one to reject the hypothesis that priors are rational, in favor of the naive case.

Naive priors

Under the second assumption, the condition for rational individual forecasts in (5) can be tested by modifying (12) to give:

$$E\left[\left(f^i_{ch\tau} - \overline{f}_{c(h+1)\tau}\right)\left(f^i_{ch\tau} - y_{c\tau}\right)\right] = 0 \tag{21}$$

Similarly, (7) is tested by modifying (13):

$$E\left[\left(\overline{f}_{ch\tau} - \overline{f}_{c(h+1)\tau}\right)\left(\overline{f}_{ch\tau} - y_{c\tau}\right)\right] \leq 0 \tag{22}$$

while (16) is modified to give:

$$\begin{aligned} \overline{f}_{ch\tau} - y_{c\tau} &= a_h - b_h \left(\overline{f}_{ch\tau} - \overline{f}_{c(h+1)\tau} \right) + e_{cht} && (23) \\ \text{where } E[b_h] &= \frac{\sigma^2_{\varepsilon}}{\sigma^2_h + \sigma^2_u} ; E[a_h] = 0 \end{aligned}$$

[7] Since the initial prior $\mu_{c\tau}$ underlying the first consensus forecast $\overline{f}_{c23\tau}$ is not observed, we substitute the expected value of $\overline{f}_{c23\tau} - \mu_{c\tau}$, which is simply zero.

3.3 Efficiency Tests

To provide some initial evidence in support of our main argument, Table 1 presents simple bivariate correlations between the most recent period's forecast updates and forecast errors for the individual forecasts and for the consensus, for the full sample as well as for different forecast horizons. In particular, column I provides correlation coefficients for the individual forecasts; column II provides coefficients for the consensus forecasts; finally, given the evidence of some (mostly positive) correlation for the individual forecasts, column III provides coefficients for consensus forecasts based on orthogonalized individual forecast errors.[8]

[Table 1 about here]

There is evidence for some positive correlation for the individual forecasters, suggesting that at the individual level forecasters may over-weight their private information. This is in line with other findings (cite) and is consistent with the kind of strategic forecasting behavior discussed by Ottaviani and Sørensen (2006).[9] However, results for consensus forecasts are very different. Here, there is evidence for negative correlation overall and at every forecast horizon. Estimated negative correlation coefficients are as high as 40 percent in absolute value. The estimated correlation is even stronger when one corrects for the under-weighting of the prior in the underlying individual forecasts by using the consensus derived from orthogonalized individual forecast errors.

We then turn to regression analysis to shed more lights on these results. We first test whether a naive or rational model of prior updating is more appropriate. Table 2 presents two sets of results. The first column presents coefficient estimates for (19), including the current and up to 22 previous forecast updates, based on a pooled sample across all 23 forecast horizons. In this specification, $\left(\overline{f}_{cj\tau} - \overline{f}_{c(j+1)\tau}\right)$ is set equal to zero for $j < h$. The second column summarizes results for 23 sets of by-horizon estimates (in which $\left(\overline{f}_{cj\tau} - \overline{f}_{c(j+1)\tau}\right)$ is excluded for $j < h$). The decimal presented here gives the proportion of regressions in which the relevant coefficient estimate is significant at (at least) the 10 percent level and "correctly" signed (i.e. aligns with (20)), minus the fraction in which it is significant but enters with the "wrong" sign. The evidence supports the view that the consensus forecast is inefficient, and in particular underreacts to new information so that the forecast error is negatively correlated with the most recent update to the consensus forecast. However, there is no evidence for the pattern of significant coefficients with alternating signs on longer-dated updates that is predicted under rational forecasting (20). The coefficients in the pooled regression are as likely to have the "wrong" sign as the correct one, a pattern repeated in the per-forecast horizon regressions. Hence, the naive

[8] Individual forecast errors are orthogonalized by regressing individual forecast errors on the forecast update by forecast horizon and using the OLS residuals. In other words, the orthogonalized individual forecast errors mimic the behavior of forecasts from rational individual forecasters.

[9] Table A1 in the Appendix presents regression results for the individual forecasts.

updating of priors appears the relevant case.[10]

Given the evidence in favor of naive forecaster behavior, the remainder of the results impose the naive case and focus only on the most recent forecast update, $\left(\overline{f}_{ch\tau} - \overline{f}_{c(h+1)\tau}\right)$.[11] Table 3 presents results for regressions of consensus forecast errors on the update to the consensus forecast. The first column reports results for a pooled sample across all forecast horizons, while the remaining columns present results for each of the 23 horizons. There is very significant evidence of underweighting of new information in the consensus forecasts, as predicted. The coefficient is significantly negative in the pooled specification as well as 15 out of the 23 per-horizon specifications—including all horizons of 10 months or less. Overall, three broad ranges are apparent in the data: a significant effect with coefficients generally around -1 for horizons of 0 to 6 months; a significant effect with slightly larger coefficients for horizons of 7 to 13 months, and generally no significant effect for horizons of 14 months and above. Table 4 presents results for consensus forecast errors based on orthogonalized individual forecast errors: results here are marginally stronger.

These results shed some light on the relative variance of idiosyncratic and shared information available to forecasters at different forecasting horizons. The estimated coefficient on the consensus forecast update is generally in the range of -1 to -3, suggesting that the variance of forecasters' idiosyncratic signals, at the shorter forecast horizons where the effect is significant, is at least as high as the variance of the common signals, and potentially up to three times as high. The fact that the coefficient is generally not significantly different from zero at longer forecasting horizons is also instructive. This finding is consistent with forecasters using different quantitative forecasting models, which are more likely to deliver divergent results for near-term projections but to assume a reversion to historical trend at longer horizons.

3.4 Efficiency Gains from adjusted Consensus Forecasts

The next step is to calculate the efficiency gain associated with an appropriate adjustment of the consensus forecasts, as in (18). In all cases, the naive priors case is assumed to be the relevant one. A complicating factor is that the relative variance of private and public signals is likely to differ systematically across countries as well as across time periods, and a good correction technique

[10] In fact, the evidence in favor of the "wrong" sign on the penultimate forecast revision ($update_1$) suggests that forecasters may fail to immediately update their prior using the new consensus forecast—even allowing for the fact that they do not adjust the consensus to correct for its inefficient incorporation of new information. Priors therefore appear to be even more "sticky" than under the naive strategy, perhaps because forecasters recognize that the consensus forecast is inefficient but fail to realize the source of the inefficiency and its potential remedy.

[11] The results in Table 2 include only forecasts produced monthly; however, for some countries and time periods forecasts are produced only every two months. To increase coverage, the remaining results include forecasts produced bi-monthly, as in Table 1; in these cases, the forecast update therefore reflects the change between the current forecast and the forecast produced two months ago.

should allow for this. Allowing b_j to differ across each country and time horizon would risk over-parameterizing the model, reducing the precision of parameter estimates and undermining the legitimacy of the resulting adjusted forecasts (particularly out of sample). However, the systematic pattern noted above across three broad time horizons suggests a relatively parsimonious parameterization, in which the slope coefficient b_j is assumed to differ across each country c and within the three broad time horizons $H = \{0-6, 7-13, 14-22\}$. The following specification is therefore run:

$$\overline{f}_{ch\tau} - y_{c\tau} = a_{c,H} + b_{c,H}\left(\overline{f}_{ch\tau} - \overline{f}_{c(h+1)\tau}\right) + e \tag{24}$$

and corrected forecasts are derived using (17). Tables 5 and 6 report the resulting (in-sample) efficiency gains by country and by forecast horizon, respectively. The most significant efficiency gains are felt at a relatively short horizon, between 3 and 12 months. Overall, efficiency gains average 3.5 percent, or almost 6 percent for horizons of 12 months or less. The maximum gain, at 12 months, is 10 percent. Among countries, efficiency gains are generally largest for emerging markets and smallest for industrial countries. Efficiency gains may be largest for emerging market economies due to the relatively smaller time-series available for these countries, resulting in some over-fitting of the model in sample, as well as the greater growth volatility typically experienced. These opportunities for efficiency gains are not generated by inefficiencies at the level of individual forecasts. When this exercise is replicated using consensus forecasts based on orthogonalized individual forecasts, the efficiency gains are even larger (5.2 percent overall and 9.0 percent for forecast horizons of 12 months or less). This reinforces the evidence from Table 1 that individual forecasts *over-weight* idiosyncratic signals, helping to offset some of the negative effect of the inefficiency created by aggregating the forecasts.

[Tables 5 and 6 about here]

Out of sample efficiency gains are somewhat lower overall, but not markedly so (Tables 7 and 8).[12] Some countries with very good in-sample performance (e.g. Indonesia), register a very poor out of sample performance, suggesting that some caution should be exercised in extrapolating in-sample patterns based on extreme events (such as the Asian crisis). Overall, however, one could have improved near-term forecast accuracy for forecasts made in 2007 and 2008 by more than 5 percent, with respect to the raw consensus forecast, based on the adjustment technique outlined here and data available through end-2006.

4 Case Study: The Asian Financial Crisis

There is broad agreement that the Asian financial crisis came as a genuine surprise to most observers and market participants. For instance, Radelet and Sachs (2000) argue that "one of the most unusual aspects of the Asian crisis is the extent to which it was unpredicted by market participants and market

[12] In fact, gains are significantly higher for some countries and time periods.

analysts." Even following the sharp depreciation of the Thai Baht in early July 1997, economic projections were slow to react. Consensus growth forecasts for 1998 remained—in comparison to the eventual outturn—wildly over-optimistic, reacting only slowly to events.

We present some graphical results relating to this question, from two perspectives. The first perspective is a purely conceptual one: using all the information available and allowing for heterogeneity across countries, can one uncover some optimal forecast, and if so, what does it suggest about the underlying signals that forecasters received? This is akin to an in-sample forecasting exercise. The second perspective is practical: given the information available to forecasters at the time, could they *themselves* have arrived at a better forecast? This is essentially an out-of-sample exercise. One would expect these approaches to yield different answers, because in practice the experience of the crisis in these countries might lead one to put a higher weight on forecast revisions, but agents would not be aware of this at the time.[13]

Figure 1 presents results for the first perspective. It shows forecasts for output growth in 1998 for Indonesia, Korea, Malaysia and Thailand. The solid line shows the consensus forecast, the dashed line the adjusted forecast based on running (24) for the 1989-2006 period and applying the transformation given by (17), and the dotted line gives the actual outturn. The vertical line at July 1997 signals the start of the crisis.

[Figure 1 about here]

To analyze the question from the second perspective we modify both the sample and the specification of (24). Since the sample will be significantly smaller in this out of sample exercise, we do not expect to be able to accurately estimate country-specific b coefficients (data for these countries is only available from December 1994); however, to exploit the information on different forecast horizons we allow b to differ by horizon.[14] To match the information available to forecasters at the time, the sample includes only forecasts to 1996 (i.e., available when the first forecasts for 1998 growth were made in January 1997).[15] The results are shown in Figure 2.

[Figure 2 about here]

Figure 1 demonstrates that forecast updates appear to have contained significant information that is discarded in the information aggregation process. Forecasters started to downgrade their forecasts for Thailand in the first half of 1997, for Indonesia and Malaysia in the second half of the year and for Korea

[13] Purely in data terms, the significant forecast errors during the crisis combined with constant downward revisions of forecasts would lead one to estimate particularly large b coefficients in (23), particularly when one allows these coefficients to be country-specific.

[14] The sample in this case does not exclude outliers, since the Asian crisis accounts for a large share of the significant individual forecast errors that were dropped in the cropped sample.

[15] Only preliminary, unrevised, estimates will be available at the end of each year, so we may be overstating the amount of information actually available to forecasters. However, since we ignore information available during 1997 and 1998 that may have given the forecasters better estimates of the adjustment coefficient b, we may also be understating the information available for some periods. We expect these effects to counteract each other and to not be very important in any case.

by the start of 1998; however, the reaction to new information was slow. As a result, our “optimal” adjusted growth forecasts for 1998 were significantly closer to the actual outturn by the end of 1997. The estimated efficiency gain varies from 4.7 percent in Korea to 20.5 percent in Thailand. Gains for forecasts made during 1998 itself are even more dramatic, ranging from 37.8 percent in Korea to 60.6 percent in Indonesia. Even using only information available to forecasters at the time, more accurate adjusted forecasts were feasible. Figure 2 shows adjusted forecast errors below consensus forecast errors for most of the second half of 1997 and 1998. Efficiency gains vary from 2.2 percent in Korea to 5.0 percent in Thailand; efficiency gains for forecasts made during 1998 are again larger, ranging from 5.7 percent in Indonesia to 8.9 percent in Korea.

This evidence suggests that, while forecasters may not have predicted the start of the crisis, they may have been better informed about its growth impact as it developed than a simple reading of the consensus forecasts would imply. Moreover, they could have used our adjustment technique to combine information in consensus forecast revisions with the consensus forecasts themselves to arrive at better forecasts in real time. Had the forecasters known that a larger adjustment coefficient b was necessary than the historical data implied, then they could have arrived at even better forecasts (closer to those achieved in the in-sample exercise). This has implications for improving forecast performance during other crisis events and severe recessions, including the current one.

5 Discussion

5.1 Morris and Shin (2002): A Reassessment

Morris and Shin (2002) use a global games framework to model Keynes’s (1936) “beauty contest.” Keynes argues that agents’ attempts to second guess the consensus forecast can damage the information-revelation role of market prices. Morris and Shin extend this logic to show that more transparent (lower variance) public signals could lead to higher forecast errors on the part of market participants. In their model, public information is over-weighted relative to agents’ private signals because it is more helpful for second-guessing other agents and hence for aligning agent’s own forecast more closely with the consensus. More accurate public signals can exacerbate this over-weighting problem, potentially increasing the volatility of agents’ individual forecasts.

Morris and Shin’s arguments have been criticized, notably for the ad-hoc nature of the beauty contest element and because the parameter values necessary for their argument on the negative effect of public information provision to hold are unrealistic (Svensson, 2006). In particular, the public information must be at least eight times noisier than the private information, which seems unlikely (in the context of making central bank economic forecasts public) in light of evidence on the apparent superiority of central bank forecasts over their private sector counterparts (Romer and Romer, 2000).

As we have already argued, this apparent superiority may be exaggerated by comparing central bank forecasts with inefficient consensus forecasts rather than an efficient aggregator of the private sector's information set. However, the logic of our model suggests that, even without the beauty contest element, the provision of more accurate public information could harm the accuracy of the consensus forecast, even if it has an unambiguously positive effect on the accuracy of individual forecasts. Given the attention paid to the consensus forecast and its potential role for uninformed agents in the economy in forming their expectations, this could argue against greater transparency.

To formalize this argument, we first illustrate the effect of changing the accuracy of public information in the context of our simple signal extraction model; we then analyze the effect in Morris and Shin's model that includes the beauty contest element. In both cases we model more accurate public information as a reduction in the variance of the prior, σ_ν^2.

In the context of our signal extraction model, the mean squared consensus forecast error is given by:

$$E\left[\left(\overline{f}-y\right)^2\right] = F^2\sigma_u^2 + (1-F)^2\,\sigma_\nu^2 \tag{25}$$

Differentiating equation (25) with respect to σ_ν^2 yields:

$$\frac{\partial E\left[\left(\overline{f}-y\right)^2\right]}{\partial \sigma_\nu^2} = \frac{\sigma_u^2+\sigma_\varepsilon^2}{\left(\sigma_\nu^2+\sigma_u^2+\sigma_\varepsilon^2\right)^3}\left(\left(\sigma_u^2+\sigma_\varepsilon^2\right)^2 - \sigma_\nu^2\left(\sigma_\varepsilon^2-\sigma_u^2\right)\right) \tag{26}$$

which is positive if and only if:

$$\left(\sigma_\varepsilon^2-\sigma_u^2\right)\sigma_\nu^2 < \left(\sigma_u^2+\sigma_\varepsilon^2\right)^2 \tag{27}$$

Hence, more transparent public information can increase the forecast error for the consensus forecast if the public information is a relatively noisy signal and the common component of the private signal (u) is relatively unimportant compared to the idiosyncratic component (ε).

In the special case where the private signals are purely idiosyncratic ($\sigma_u^2 = 0$), the condition (27) collapses to $\sigma_\nu^2 < \sigma_\varepsilon^2$: this implies that greater transparency can increase the *consensus* forecast error if the public information is initially more noisy than agents' own signals. This condition seems likely to hold in some situations. Moreover, the negative effect on the accuracy of the consensus forecast does not require the "beauty contest" element, whereas greater transparency always improves the accuracy of individual forecasts without the beauty contest.

However, one can include the beauty contest element, as in Morris and Shin (2002), so that forecasters now weight the public and private signals to minimize a loss function that is a weighted sum of the mean squared forecast error and the mean squared deviation of individual forecasts from the consensus. The weight on the beauty contest element is given by r:

$$L = (1-r)\,E\left[(f-y)^2\right] + rE\left[\left(f-\overline{f}\right)^2\right] \tag{28}$$

As shown by Morris and Shin, agents now over-weight the public signal:

$$F_b = \frac{(1-r)\,\sigma_\nu^2}{\sigma_\varepsilon^2 + (1-r)\left(\sigma_u^2 + \sigma_\nu^2\right)} \tag{29}$$

and differentiating (25) with respect to σ_ν^2 yields:

$$\frac{\partial E\left[\left(\bar{f} - y\right)^2\right]}{\partial \sigma_\nu^2} = \frac{(1-r)\,\sigma_u^2 + \sigma_\varepsilon^2}{\left((1-r)\left(\sigma_\nu^2 + \sigma_u^2\right) + \sigma_\varepsilon^2\right)^3}\left(\left((1-r)\,\sigma_u^2 + \sigma_\varepsilon^2\right)^2 - (1-r)\,\sigma_\nu^2\left(\sigma_\varepsilon^2 - (1-r)\,\sigma_u^2\right)\right) \tag{30}$$

which is positive if and only if:

$$(1-r)\,\sigma_\nu^2\left(\sigma_\varepsilon^2 - (1-r)\,\sigma_u^2\right) < \left((1-r)\,\sigma_u^2 + \sigma_\varepsilon^2\right)^2 \tag{31}$$

For the special case considered by Morris and Shin ($\sigma_u^2 = 0$), this condition collapses to $(1-r)\,\sigma_\nu^2 < \sigma_\varepsilon^2$. By contrast, the condition for more accurate public information making *individual* private forecasts more accurate is $(2r-1)\,(1-r)\,\sigma_\nu^2 < \sigma_\varepsilon^2$.

Comparing these two expressions yields a number of insights. First, while the condition for better public information to reduce the mean squared error of the *individual* forecasts always holds when the beauty contest is not present, it can fail to hold for consensus forecasts even without the beauty contest. This is because consensus forecasts overweight the public information even without the beauty contest element, so that the mechanism by which more public information can be harmful is also present without the beauty contest. In fact, while the beauty contest has to be relatively important ($r > 0.5$) for the harmful effect of transparency to (potentially) be felt for the individual forecasts, the harmful effect is most likely to hold for *consensus* forecasts when the beauty contest is relatively *unimportant* (r is low). The rationale for this is that the beauty contest is not the cause of the consensus forecast overweighting the prior: if one were to set σ_ε^2 to zero in (29) to eliminate the cause of overweighting the prior in the simple signal extraction model then the $(1-r)$ terms drop out and one is left with the expression for the efficient forecast weight F_C, regardless of the beauty contest. Hence, introducing the beauty contest reduces the impact of the mechanism by which the prior is overweighted, diluting the impact and making it less likely that an increase in transparency is harmful. Finally, since $(2r-1) \leq 1$, the parameter space over which more accurate public information can be harmful is strictly larger for consensus forecasts than for individual forecasts.

5.2 Asset Prices as Inefficient Information Aggregators

[preliminary—either shorten to brief section noting the potential application to asset prices and noting the implications, or a longer section with a more developed model]

So far we have focused on consensus forecasts as a particular aggregator of private information. For economic theorists the price mechanism has traditionally had a key role as a means of efficiently collating and communicating the diffuse economic signals of atomistic market participants. Asset prices in particular reflect forward-looking expectations, since decisions to buy or sell assets for a given price reflect agents' solution to their individual signal-extraction problem with respect to the asset's expected return. Since equilibrium asset prices therefore reflect the consensus or average solution to the signal extraction problem, they may also be inefficient. The key point is that prices do not reflect raw information, but information filtered through economic agents. Even if these agents are individually rational, the collective signal revealed through the price mechanism is less informative than their underlying collective information set.

Consider an asset whose uncertain payoff in period 2 is given by y. To simplify the argument, assume that agents—distributed along a continuum with unit measure—are risk averse and do not discount the future. Moreover, they are restricted to either buy or sell one unit of the asset in each period. The equilibrium price is the value at which the number of buyers equals the number of sellers (so that the market clears). Uncertainty is resolved in period 2, so that $p_2 = y$. Agents come into period 1 with a private signal of y, x_i and a prior μ given by the previous market clearing price, p_0.

The characteristics of the signals and the prior are the same as in section II. In addition, we assume that the distribution of ε_i is symmetric. Then it follows that the market price in period 1 reflects the median forecast of y in period 1 (i.e. the solution to the median agent's signal extraction problem):

$$p_1 = p_0 + F\left(u + \nu\right) \tag{32}$$

Note that, given our symmetry assumption, this is equal to the mean or consensus forecast of y.

Then it follows from the results in section 2 that market prices tend to be inefficient, particularly when there is more pronounced heterogeneity in expectations (i.e. uncertainty) among market participants. Comparing our forecasting model with this simple asset pricing model, the consensus forecast update from the former $\left(\overline{f} - \mu\right)$ is equal to the capital gain on the asset between period 0 and period 1 in the latter $(p_1 - p_0)$. Similarly, the consensus forecast error $\left(\overline{f} - y\right)$ is equal to (minus) the capital gain between period 1 and period 2, $-\left(p_2 - p_1\right)$. Hence, our previous results imply that returns over the two periods will be positively correlated. Our model could therefore account for the short-run positive correlation of stock market returns that has been widely discussed in the finance literature (Jegadeesh and Titman, 1993, 2001; Hong and Stein, 1999). The intuition is similar to that of Hong and Stein's (1999) behavioral model, although the mechanism is much simpler.

Of course, once agents observe p_1, and assuming that they know the relative variances of the signals and that no new information becomes available, then prices will instantly readjust to a new efficient equilibrium price level and there will be no opportunities for arbitrage. However, in reality it is unlikely that agents will be able to easily distinguish between movements in market prices that reflect changes in fundamentals and movements that reflect responses to previous changes, particularly since the relevant variances are likely to be significantly time-varying and the timing and distribution of signals about fundamentals across diverse market participants are not likely to be well known. Well-informed agents may be able to exploit opportunities for arbitrage implied by this correlation in returns (the evidence alluded to earlier suggests that even relatively naive momentum-trading strategies earn small positive returns). Meanwhile, a key implication is that asset prices become reflexive, responding to themselves as well as to fundamentals. This introduces a potent role for self-fulfilling beliefs and hence the emergence of asset price bubbles and crashes.

Finally, note that the model can also be modified to generate situations in which market prices *overreact* to new information. Suppose the prior is not common to all agents, but rather has an aggregate and an idiosyncratic component, whereas the new information is common across agents. This would correspond to a situation in which there is considerable uncertainty over the true state of the world (e.g. model uncertainty) but new information (e.g. macro forecasts or a political shock) is revealed with little room for individual interpretation. Then it is the *new* aggregate information rather than the prior which would be over-weighted in p_1, and as a result prices will initially overshoot rather than only partially adjust to the new information.

5.3 "Groupthink"

The insights from this paper can also be related to recent work on "groupthink" or collective delusion by Benabou (2009). Benabou's model is based on a modification to a standard utility function to incorporate anticipatory feelings (i.e. deriving utility from expected as well as actual consumption). With this modification, Benabou is able to generate externalities from individual behavior that give rise to equilibria with self-fulfilling collective delusion. Agents rationally choose to ignore pertinent information, giving rise to collective judgments that differ systematically from reality.

The phenomenon of groupthink involves a number of features that distinguish it from the model in this paper, notably the introduction of anticipatory feelings that lead to rational self-denial (and ultimately mutually-reinforcing denial within groups) and the absence of a major role for idiosyncratic private information. Benabou's model is much richer than that of the current paper; nevertheless, there are some important complimentarities between the two approaches. The collective failure to rationally process the available aggregate information—leading to the discarding of useful information—is the most obvious common feature to both models. Another commonality is the role of common group beliefs. In the "Groupthink" model agents choose to ignore information at odds with the group prior, leading to a less efficient aggregation of private signals. In the simple signal extraction model in this paper the common prior is also overweighted collectively, although this behavior is individually rational.[16]

Moreover, as discussed in relation to Morris and Shin's paper, a more accurate (lower variance) prior can lead to a less accurate collective forecast in our model. This result can provide insights into some of the groupthink phenomena discussed in Appendix A to Benabou's paper. For instance, banks' increasing use of sophisticated risk-management tools and quantitative models for pricing assets may have genuinely provided a more accurate assessment of expected returns that led to a lower-variance prior among managers, leading to a (positive) direct effect on the accuracy of their consensus forecast. However, this more accurate signal then caused the managers, receiving diffuse but ultimately correlated signals of potential downside risk, to rationally put greater weight on the group prior, collectively underweighting their noisy but (in aggregate) informative private signals, and therefore reducing consensus forecast accuracy. The condition for this underweighting effect to dominate the direct effect—that the group prior be no less noisy than the managers' own idiosyncratic signals—seems plausible, given that the models turned out to perform rather poorly out of sample.[17] In other words, the claims prior to the crisis that these new techniques were reducing risk, and the emerging consensus post-crisis that their use reduced the ability to foresee and forestall emerging risks, may both be accurate assessments. The first relates to the group prior; the second to the group posterior.

6 Conclusions

[to follow]

[References to follow]

[16] Benabou's mechanism is also individually rational conditional on the inclusion of anticipatory feelings in the utility function.

[17] Explanations for the crisis have tended to emphasize that the models performed worse than expected (i.e. their signal to noise ratio turned out to be lower than believed). However, while this may indeed have been the case, it is not necessary as an explanation for the crisis.

Tables and Figures

Table 1. Correlation Coefficients

Horizon	I	II	III
0	0.03	-0.13	-0.15
1	0.10	-0.19	-0.24
2	0.05	-0.21	-0.24
3	-0.01	-0.28	-0.28
4	0.06	-0.24	-0.27
5	0.02	-0.36	-0.37
6	0.01	-0.28	-0.28
7	0.03	-0.40	-0.41
8	0.03	-0.22	-0.24
9	0.06	-0.32	-0.35
10	0.06	-0.25	-0.28
11	0.14	-0.07	-0.18
12	0.02	-0.37	-0.37
13	0.14	-0.18	-0.23
14	0.15	-0.10	-0.17
15	0.21	0.03	-0.08
16	0.20	-0.08	-0.16
17	0.21	-0.03	-0.09
18	0.23	0.08	-0.01
19	0.19	-0.11	-0.16
20	0.24	-0.03	-0.11
21	0.22	-0.05	-0.11
22	0.24	-0.05	-0.11
Pooled	0.13	-0.15	-0.19

Correlation between forecast errors and forecast updates
Column I: results for individual forecasts
Column II: results for consensus forecast
Column III: results for adjusted consensus forecasts

Table 2. Rational versus Naïve Prior Updating

Model	Pooled	Signs[1]
Update0	-0.997***	0.57
	(0.259)	
Update1	-0.527***	-0.14
	(0.193)	
Update2	-0.254**	0.10
	(0.108)	
Update3	-0.078	0.15
	(0.071)	
Update4	-0.019	-0.16
	(0.064)	
Update5	0.004	0.17
	(0.077)	
Update6	0.111*	-0.18
	(0.059)	
Update7	0.181***	0.13
	(0.059)	
Update8	0.132**	-0.07
	(0.065)	
Update9	0.105*	0.14
	(0.055)	
Update10	0.120*	-0.08
	(0.064)	
Update11	0.068	0.08
	(0.051)	
Update12	0.268***	-0.27
	(0.071)	
Update13	0.266**	0.20
	(0.115)	
Update14	0.139	-0.33
	(0.112)	
Update15	0.142	0.25
	(0.111)	
Update16	0.043	0.00
	(0.147)	
Update17	-0.08	-0.17
	(0.162)	
Update18	-0.156	0.20
	(0.152)	
Update19	-0.106	0.00
	(0.159)	
Update20	-0.265	0.00
	(0.356)	
Update21	0.09	0.00
	(0.277)	
Update22	0.252	0.00
	(0.345)	
Constant	-0.283***	
	(0.101)	
Observations	9,050	
R^2	0.035	

For the pooled specification, SEs are clustered by country and year.
[1] Proportion of "correct" signs minus proportion of "incorrect" signs;
from 23 per-forecast horizon regressions.
For the by-forecast horizon results, SEs are clustered by country.
Significance level denoted by *** (1 percent); ** (5 percent); * (10 percent).

Table 3. Efficiency Tests: Consensus Forecasts

Horizon	Pooled	0	1	2	3	4	5	6
Update	-1.044***	-0.605***	-1.145***	-1.150***	-1.057***	-1.521***	-1.963***	-1.218***
	(0.218)	(0.176)	(0.299)	(0.177)	(0.205)	(0.270)	(0.380)	(0.438)
Constant	-0.381***	-0.430**	-0.514***	-0.512**	-0.476***	-0.459**	-0.493***	-0.422**
	(0.096)	(0.203)	(0.139)	(0.206)	(0.141)	(0.194)	(0.153)	(0.197)
Observations	10,330	475	495	468	489	466	489	459
R^2	0.023	0.016	0.035	0.046	0.08	0.06	0.129	0.076
Horizon	7	8	9	10	11	12	13	14
Update	-3.424***	-1.136***	-2.112***	-1.580***	-0.262	-3.176***	-2.245**	-0.702*
	(0.565)	(0.274)	(0.768)	(0.423)	(0.452)	(0.630)	(0.948)	(0.399)
Constant	-0.469***	-0.379*	-0.571***	-0.390**	-0.556***	-0.344	-0.588***	-0.272
	(0.160)	(0.202)	(0.174)	(0.188)	(0.180)	(0.207)	(0.174)	(0.221)
Observations	475	453	467	441	467	436	457	428
R^2	0.157	0.05	0.1	0.063	0.005	0.134	0.034	0.01
Horizon	15	16	17	18	19	20	21	22
Update	0.193	-0.768	-0.358	0.789	-2.188*	-0.332	-0.635	-0.821
	(0.587)	(0.928)	(0.748)	(0.865)	(1.271)	(0.935)	(1.141)	(1.858)
Constant	-0.383*	-0.1	-0.342*	0.055	-0.313	0.017	-0.401*	-0.16
	(0.193)	(0.189)	(0.198)	(0.201)	(0.218)	(0.199)	(0.227)	(0.204)
Observations	449	424	448	418	434	412	427	353
R^2	0.001	0.007	0.001	0.006	0.013	0.001	0.002	0.003

Standard Errors clustered by country (country/year for pooled specification)
Significance Level denoted by *** (1 percent); ** (5 percent); * (10 percent).

Table 4. Efficiency Tests: Adjusted Consensus Forecasts

Horizon	Pooled	0	1	2	3	4	5	6
Update	-1.355***	-0.708***	-1.497***	-1.323***	-1.023***	-1.717***	-2.023***	-1.258***
	(0.207)	(0.176)	(0.299)	(0.177)	(0.205)	(0.270)	(0.380)	(0.438)
Constant	-0.095	-0.06	-0.058	-0.114	-0.042	-0.134	-0.059	-0.114
	(0.095)	(0.203)	(0.139)	(0.206)	(0.141)	(0.194)	(0.153)	(0.197)
Observations	10,330	475	495	468	489	466	489	459
R^2	0.038	0.022	0.058	0.06	0.076	0.075	0.136	0.08
Horizon	7	8	9	10	11	12	13	14
Update	-3.551***	-1.249***	-2.331***	-1.789***	-0.661	-3.247***	-2.828***	-1.250***
	(0.565)	(0.274)	(0.768)	(0.423)	(0.452)	(0.630)	(0.948)	(0.399)
Constant	-0.025	-0.096	-0.091	-0.156	-0.142	-0.204	-0.183	-0.188
	(0.160)	(0.202)	(0.174)	(0.188)	(0.180)	(0.207)	(0.174)	(0.221)
Observations	475	453	467	441	467	436	457	428
R^2	0.167	0.06	0.12	0.079	0.032	0.139	0.053	0.03
Horizon	15	16	17	18	19	20	21	22
Update	-0.584	-1.543	-1.243	-0.104	-3.014**	-1.248	-1.516	-1.756
	(0.587)	(0.928)	(0.748)	(0.865)	(1.271)	(0.935)	(1.141)	(1.858)
Constant	-0.086	-0.166	-0.09	-0.023	-0.035	-0.026	-0.061	-0.069
	(0.193)	(0.189)	(0.198)	(0.201)	(0.218)	(0.199)	(0.227)	(0.204)
Observations	449	424	448	418	434	412	427	353
R^2	0.006	0.026	0.008	0	0.024	0.013	0.012	0.012

Standard Errors clustered by country (country/year for pooled specification)
Significance Level denoted by *** (1 percent); ** (5 percent); * (10 percent).

Table 5. In-Sample Efficiency Gains, by Country

Country	Efficiency Gain (Percent)
Argentina	3.1
Australia	1.8
Brazil	0.8
Canada	1.3
Chile	3.9
China P.R.: Hong Kong	4.7
China P.R.: Mainland	3.0
Colombia	0.3
Czech Republic	6.1
France	3.0
Germany	1.6
Hungary	2.7
India	1.0
Indonesia	7.7
Italy	2.9
Japan	1.3
Korea Republic of	10.3
Malaysia	13.0
Mexico	3.1
Netherlands	1.0
New Zealand	0.5
Norway	0.3
Peru	5.3
Poland	5.7
Romania	1.3
Russian Federation	5.4
Singapore	4.6
Slovak Republic	3.0
Spain	-0.8
Sweden	2.7
Switzerland	2.1
Taiwan Province of China	6.1
Thailand	9.6
Turkey	2.3
Ukraine	2.9
United Kingdom	0.3
United States	0.1
Venezuela	1.7

Author's calcuations, as in text
Forecasts made for years to 2006.

Table 6. In-Sample Efficiency Gains, by Forecast Horizon

Horizon (months)	Efficiency Gain (Percent)
0	0.0
1	1.9
2	1.6
3	5.9
4	3.2
5	7.1
6	6.6
7	8.1
8	4.7
9	7.2
10	4.6
11	7.7
12	10.1
13	5.1
14	2.3
15	2.9
16	1.5
17	0.4
18	0.4
19	-0.4
20	0.3
21	2.6
22	1.0
≤12	5.9
All	3.5

Author's calcuations, as in text

Forecasts made for years to 2006.

Table 7. Out of Sample Efficiency Gains, by Country

Country	Efficiency Gain (Percent)
Argentina	2.9
Australia	5.7
Brazil	2.8
Canada	2.2
Chile	-0.9
China P.R.: Hong Kong	4.6
China P.R.: Mainland	4.9
Colombia	5.0
Czech Republic	-43.1
France	3.7
Germany	-6.7
Hungary	-0.1
India	1.9
Indonesia	-59.5
Italy	3.3
Japan	1.6
Korea Republic of	5.8
Malaysia	9.2
Mexico	3.4
Netherlands	-5.9
New Zealand	2.0
Norway	2.6
Peru	12.0
Poland	6.2
Romania	-4.9
Russian Federation	4.5
Singapore	8.3
Slovak Republic	5.9
Spain	2.2
Sweden	5.4
Switzerland	4.1
Taiwan Province of China	4.7
Thailand	-9.1
Turkey	3.4
Ukraine	1.9
United Kingdom	2.0
United States	0.8
Venezuela	-25.5

Author's calcuations, as in text

Forecasts made from Jan 2007 onwards, for 2007/08.

Table 8. Out of Sample Efficiency Gains, by Forecast Horizon

Horizon (months)	Efficiency Gain (Percent)
0	6.2
1	5.8
2	6.9
3	0.3
4	-0.1
5	5.3
6	-2.3
7	-1.3
8	16.5
9	3.0
10	15.5
11	5.9
12	9.2
13	6.1
14	-1.2
15	-0.9
16	-0.6
17	0.2
18	1.1
19	-1.7
20	2.7
21	-4.0
22	0.7
≤12	5.1
All	3.0

Author's calcuations, as in text

Forecasts made from Jan 2007 onwards, for 2007/08.

Table A1. Efficiency Tests: Individual Forecasts

Horizon	Pooled	0	1	2	3	4	5	6
Update	0.455***	0.103	0.351***	0.173*	-0.034	0.196*	0.06	0.04
	(0.028)	(0.129)	(0.125)	(0.091)	(0.103)	(0.118)	(0.142)	(0.105)
Constant	-0.285***	-0.370***	-0.456***	-0.398***	-0.434***	-0.325***	-0.433***	-0.308***
	(0.023)	(0.060)	(0.071)	(0.068)	(0.077)	(0.074)	(0.088)	(0.083)
Observations	167,802	7,751	8,223	7,625	7,959	7,503	8,030	7,589
R^2	0.016	0.001	0.009	0.003	0	0.004	0	0
Horizon	7	8	9	10	11	12	13	14
Update	0.127	0.113	0.219*	0.209	0.399***	0.071	0.583***	0.548***
	(0.136)	(0.103)	(0.130)	(0.208)	(0.109)	(0.125)	(0.125)	(0.097)
Constant	-0.444***	-0.283***	-0.480***	-0.235**	-0.414***	-0.14	-0.405***	-0.084
	(0.102)	(0.098)	(0.109)	(0.107)	(0.116)	(0.119)	(0.121)	(0.132)
Observations	7,943	7,493	7,751	7,221	7,706	7,099	7,535	6,936
R^2	0.001	0.001	0.004	0.004	0.021	0	0.019	0.021
Horizon	15	16	17	18	19	20	21	22
Update	0.777***	0.775***	0.885***	0.893***	0.826***	0.915***	0.881***	0.935***
	(0.130)	(0.108)	(0.107)	(0.075)	(0.056)	(0.093)	(0.080)	(0.085)
Constant	-0.297**	0.067	-0.252*	0.078	-0.278**	0.043	-0.340**	-0.09
	(0.129)	(0.131)	(0.132)	(0.133)	(0.136)	(0.132)	(0.133)	(0.131)
Observations	7,201	6,720	7,226	6,722	7,001	6,504	6,672	5,392
R^2	0.045	0.04	0.044	0.051	0.035	0.056	0.049	0.058

Standard Errors clustered by country, year and forecast horizon.
Significance Level denoted by *** (1 percent); ** (5 percent); * (10 percent).

Figure 1.

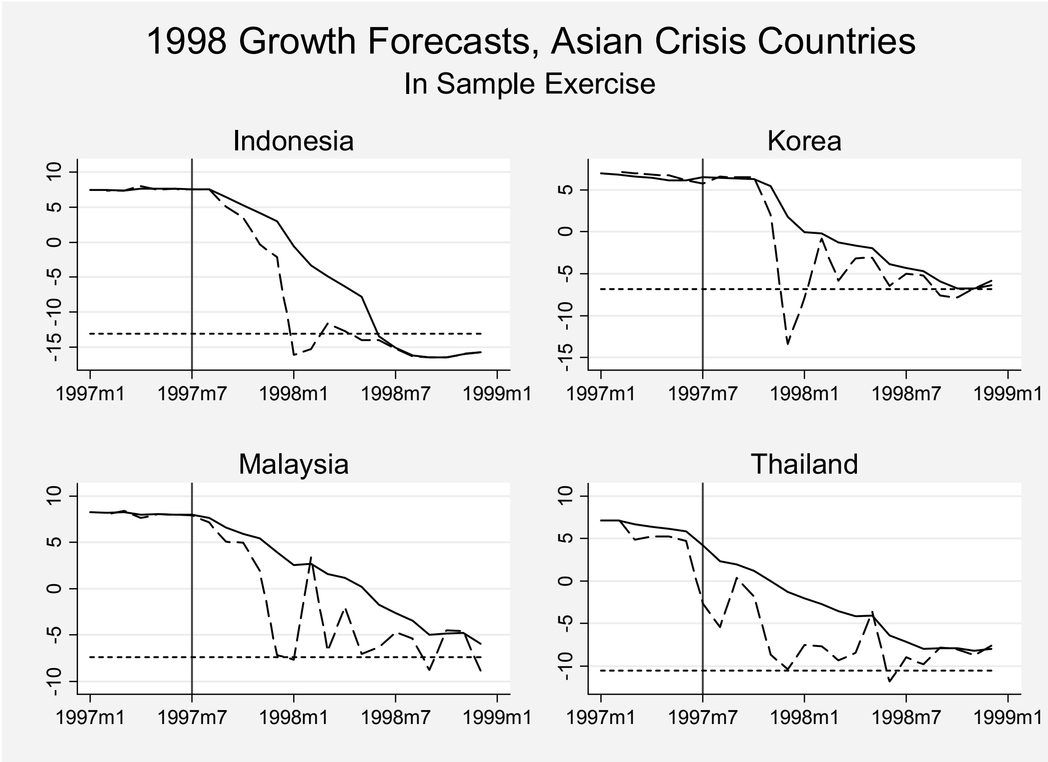

Figure 2.