Ifo Working Papers

Forecaster Behaviour and Bias in Macroeconomic Forecasts

Roy Batchelor

Ifo Working Paper No. 39

January 2007

Forecaster Behaviour and Bias in Macroeconomic Forecasts*

Abstract

This paper documents the presence of systematic bias in the real GDP and inflation forecasts of private sector forecasters in the G7 economies in the years 1990–2005. The data come from the monthly Consensus Economics forecasting service, and bias is measured and tested for significance using parametric fixed effect panel regressions and nonparametric tests on accuracy ranks. We examine patterns across countries and forecasters to establish whether the bias reflects the inefficient use of information, or whether it reflects a rational response to financial, reputational and other incentives operating for forecasters. In several G7 countries – Japan, Italy, Germany and France – there is evidence of a change in the trend growth rate. In these circumstances, standard tests for rationality are inappropriate, and a bias towards optimism in the consensus forecast is inevitable as rational forecasters learn about the new trend. In all countries there is evidence that individual forecasters converge on the consensus forecast too slowly. However, the persistent optimism of some forecasters, and the persistent pessimism of others, is not consistent with the predictions of models of "rational bias" that have become popular in the finance and economics literature.

JEL Code: E27, E37.
Keywords: Macroeconomic forecasts, judgemental forecasting, bias.

Roy Batchelor
Sir John Cass Business School
City of London
106 Bunhill Row
London EC1Y 8TZ
United Kingdom
Phone: +44(0)207/040-8733
r.a.batchelor@city.ac.uk

* The author is indebted to Kajal Lahiri, Herman Stekler and Ken Wallis, and participants in the Conference on *The Future of Economic Forecasting* at the University of Leipzig and the Seminar of the Ifo Institute, Munich, for helpful comments on earlier drafts.

1. Introduction

Why would a forecaster publish persistently biased forecasts? There are three possible explanations. One is that the forecaster lacks the skill to utilize information efficiently, and fails to learn from past forecast errors. A second is that the forecaster has the skill to utilize information "rationally", but insufficient data to learn about which changes in the target variable are permanent and which are transitory. The third possibility is that the forecaster has skill and sufficient data, but willfully introduces "rational bias" in response to financial or reputational incentives to make an optimistic or pessimistic forecast. Our aim in this paper is to review explanations for bias in forecasts, and particularly forecasts of the macro-economy. These theories are then evaluated for consistency with stylized facts about the performance of private sector forecasts of real GDP growth and inflation in the G7 economies in the years 1990–2005.

The background to the study is a series of papers written in the early 1990s documenting a number of regularities in the behaviour of a panel of US economic forecasters contributing to the Blue Chip consensus forecasting service in the US in the 1980s. Batchelor and Dua (1990a) surveyed their forecasting methods. A key finding was that most forecasters placed at least as much weight on judgment as on formal econometric modelling. Forecasters exercising a balance of modelling and judgment were generally more accurate than the small number of model-only and judgment-only forecasters. We did not explore exactly how judgment was exercised. However, Batchelor and Dua (1990b) found that individual forecasters tended to consistently adopt either an initially optimistic or an initially pessimistic view of the economy 1-2 years ahead. In the context of the US economy, this is unlikely to be due to bias in model forecasts, but more likely reflects judgmental manipulation. Batchelor and Dua (1992a) further showed that most forecasters were conservative in making revisions to predictions of growth, unemployment, inflation and interest rates for a fixed target year. Forecasters underutilised recent economic information, and did not shrink their individual forecasts towards the previously published average or "consensus" forecast sufficiently. This sluggish reaction to news meant that initial optimism or pessimism rankings also persisted at shorter horizons. In spite of this, over a run of years there was little to choose between forecasters in terms of accuracy, and forecasts were generally unbiased, since large variations in the target variables meant that sometimes the optimists proved to be closer to the truth, and sometimes the pessimists were.

One interpretation of these findings is that forecasters were producing technically irrational forecasts because of behavioural biases in the information processing – specifically anchoring and conservatism – as documented by Kahneman and Tversky (1973). Batchelor and Dua (1990b) suggested an alternative explanation which is consistent with rational behaviour in a market where forecasters compete to sell their services. The biases to optimism and pessimism were attempts by forecasters to "brand" their products, and the failure to converge towards the consensus reflected this product differentiation.

Since then, many papers have recognised that forecasters operate in markets, and are subject to incentives of various kinds that might bias their predictions even if information is processed rationally. These typically lead to models where the utility (to the forecaster) of the forecast is either an asymmetric function of the forecast error, or a function of additional factors such as optimism, consistency, or newsworthiness. A few of these papers have looked at macroeconomic forecasters, but most focus on analysts' forecasts of company earnings and stock prices, where there are rich time series and cross-sectional databases, and some very obvious reasons for analysts to make biased predictions. There are some similarities in the environments in which macroeconomists and stock market analysts work, but there are significant differences too. Section 2 of the paper below sets out a framework for analyzing forecast bias, and reviews the finance and economics literature on rational bias.

Section 3 introduces more recent data on real GDP and inflation forecasts. These are taken from the *Consensus Economics* forecasting service, which has published macroeconomic forecasts since 1989 of various private sector bodies – banks, corporations, industry associations, consultancy firms, research institutes, and universities – operating in all G7 economies. This gives us the opportunity to update the earlier US studies, and extend them to other major economies. The experience of the G7 countries in the 1990s has been challenging for forecasters, with the US, Canada and the UK experiencing healthy growth, France and Italy experiencing some slowdown, Japan trapped in a state of almost permanent recession, Germany subject to the shock of unification, and inflation everywhere falling to very low and sometimes negative levels.

Bias in these forecasts is measured and tested for significance using parametric fixed effect panel regressions, and nonparametric tests of accuracy ranks. Two results stand out. First, we replicate the Batchelor and Dua (1990b) finding that individual forecasters appear to cultivate reputations for relative optimism or pessimism in real GDP forecasts in all countries, and consequently converge too slowly on the consensus forecast. This is not consistent with the predictions of models of "rational bias" that have become popular in the finance and economics literature, which predict excessive "herding" and excessive dispersion of forecasts, respectively, but not consistent bias at the individual level. Second, in some countries we also find evidence of bias toward optimism even in the average "consensus" forecast. However, this occurs in Japan, Italy, Germany and France, and to a lesser extent Canada, where in all cases there is evidence of a downward shift in the trend growth rate. In these circumstances standard tests for rationality are inappropriate, and a bias towards optimism in the consensus forecast does not reflect "behavioural" manipulation, but is inevitable as rational forecasters gradually learn about the new trend.

2. Rational Bias

To frame our discussion of the sources of bias, and to motivate the later empirical tests, consider the following scenario. In month t a forecaster i publishes a forecast of the value of some key variable y, such as real GDP growth. The true value of y will be revealed at target date T, so the horizon of the forecast can be regarded as $h = T\text{-}t$. The forecast published in month t we denote as $f_{it,T}$. If the initial forecast is made in month τ, the maximum horizon is $T - \tau$, and the initial forecast is denoted $f_{i\tau,T}$.

The information set Ω_{it} of the forecaster at t can be partitioned into three components:

- the forecast made by forecaster i last month, $f_{it\text{-}1,T}$
- the consensus (average) forecast of all forecasters last month, $f_{t\text{-}1,T}$
- other news arriving during month t, summarised into a point forecast $g_{it,T}$

If y is normally distributed, the statistical expected value $f^*_{it,T}$ will minimise the expected mean squared error $E(y_T - f_{it,T})^2$. For the sake of simplicity, we assume that this is a weighted average of the three inputs:

$$f^*_{it,T} = E\{y_T \mid \Omega_{it}\} = \alpha^*_i f_{it-1,T} + \beta^*_i f_{t-1,T} + \gamma^*_i g_{it,T} \tag{1}$$

The initial forecast is made without reference to past forecasts, but it is helpful to think of it as a weighted average of a forecast based on cumulative relevant information up to τ, say $G_{i\tau,T}$, and a conjectured initial consensus forecast $F_{\tau,T}$, formed on the basis of the public information in $\Omega_{i\tau}$.

In the terminology of Muth (1960), these forecasts are "rational expectations" of the target variable. Provided that the underlying process driving the target variable is stable, characteristics of rational expectations are unbiasedness, orthogonality of the forecast errors with respect to components of the information set, and serially uncorrelated forecast revisions. These form the basis of many empirical tests of "rationality" in published forecasts.

The empirical studies of forecaster behaviour reviewed below suggest that in practice forecasters do have incentives to publish forecasts that minimise the expected squared error, but also may be consistently biased to optimism or pessimism, either overweight or underweight the consensus (and by implication underweight news), and minimise forecast revisions.

This implies that the forecaster faces a utility maximisation problem of the form

$$\max U_{it,T} = - E(y_T - f_{it,T})^2 + a(f_{it,T} - f_{it-1,T})^2 + b(f_{it,T} - f_{t-1,T})^2 \tag{2}$$

The first term is the squared error, the minimand assumed by conventional theory, and programmed into most econometric models. The second term is the change in the individual forecast from one month to the next. Here we expect weight a on the forecast revision to be negative, since there are benefits to "conservatism". The third term is the gap between the individual forecast and the latest consensus forecast. The weight b on this term could be either positive or negative depending on whether forecasters benefit from "product differentiation", or from "herding" towards the consensus.

First order conditions for a maximum imply that the forecaster will publish a forecast of

$$f_{it,T} = \alpha_i f_{it-1,T} + \beta_i f_{t-1,T} + \gamma_i g_{it,T} <> f^*_{it,T}, \tag{3}$$

where $\alpha_i = (\alpha^*_i - a)/(1 - a - b)$, $\beta_i = (\beta^*_i - b)/(1 - a - b)$, and $\gamma_i = \gamma^*_i /(1 - a - b)$. Without loss of generality we can constrain the weights to $(\alpha^*_i + \beta^*_i + \gamma^*_i) = 1$ in (1), so that $(\alpha_i + \beta_i + \gamma_i) = 1$. The analytics of (3) are then straightforward. If forecasters are conservative ($a < 0$) and "herd" ($b < 0$), then $\alpha_i > \alpha^*_i$, $\beta_i > \beta^*_i$, and necessarily $\gamma_i < \gamma^*_i$, so that news is underweighted. If forecasters are conservative ($a < 0$) and try to differentiate their product by underweighting the consensus ($b > 0$), then $\alpha_i > \alpha^*_i$, $\beta_i < \beta^*_i$ and news is overweighted or underweighted ($\gamma_i > \gamma^*_i$ or $< \gamma^*_i$) according as $(a + b) > 0$ or < 0.

Differences between the published forecast and the error minimizing forecast therefore arise from three sources – the initial bias to optimism or pessimism introduced by the forecaster, conservatism in forecast revision, and over- or under-weighting of the consensus forecast. We discuss each in turn.

2.1 Optimism

We use the terms "optimistic" and "pessimistic" as a convenient shorthand to refer to forecasts of the real GDP that prove too high, and forecasts of inflation that prove too low. So long as there is a penalty for forecast revisions, the influence of the initial forecast will be excessive, and any bias to optimism or pessimism in the initial forecast will be propagated forward into the whole series of forecasts for target date T. There are three ways that the initial forecast can be biased.

First, a forecaster whose initial information set by chance implies an optimistic forecast will be optimistic for too long, and vice versa. Over a run of successive monthly forecasts for a single target year this will cause bias. However, since it is unlikely that a forecaster will by chance receive consistently optimistic or pessimistic information year after year, this is unlikely to cause forecasts to appear biased over a long run of years. Bias due to differences in individual information sets will also disappear when averaged over individuals, and so will not induce bias in the consensus forecast.

Second, Batchelor and Dua (1990b) suggest that there are benefits to forecasters of cultivating a reputation as optimists or pessimists.

Forecasting bodies affiliated with governments have several reasons to make biased forecasts, which we could classify as “instrumental”, “indicative”, and “partisan”. The forecast may be used as an instrument to rationalize a particular policy stance. Heinemann (2005) shows that official medium term projections of economic growth in Germany have been persistently optimistic for several decades, allowing governments to make spending plans on the basis of unrealistically high projections of tax receipts. Jonung and Larch (2006) show that over-optimism has also characterized more recent budget projections in Italy and France, though not those of the UK. The forecast may also be used as a tool to stimulate some private sector behaviour. For example, in the era of “indicative planning” (Estrin & Holmes, 1990), governments – notably in France and Japan, and for a short period the UK – consciously promulgated optimistic forecasts in the belief that this would generate private investment, and perhaps even validate the forecast. Official inflation forecasts are generally close to official inflation targets, which are designed in the same way to moderate wage settlements and pricing decisions by the private sector. Finally the forecast may simply be used to put the incumbent party of government in a favourable light. Ulan, Dewald and Bullard (1995), for example, show that US administration forecasts for the US in the 1980s were typically optimistic, while their forecasts for other countries were unbiased.

Private sector forecasters also have incentives to bias their figures towards optimism or pessimism. Some forecasters or users of forecasts may support the incumbent government and its policies, and prefer to cite optimistic forecasts. Others may oppose current policy, and prefer to cite pessimistic forecasts. Forecasters may bend their predictions to make their forecasts more attractive to particular client groups. Trades union may have incentives to seek out forecasters that produce relatively pessimistic inflation forecasts. Business associations may seek out forecasters who make predictions that would support their programs for, say, industrial subsidies, tariff protection or exchange rate devaluation. In addition to pressure from clients, forecasters may have their own reasons to introduce bias. Forecasts are typically produced jointly with comments on economic policy. If a forecaster writes a newsletter or book predicting the death of inflation or an imminent economic collapse he/she is committed to making consistently optimistic or pessimistic statements for some years thereafter, just to maintain credibility, speaking fees and book sales.

This implies that the initial forecast is not a weighted average of the conjectured consensus forecast and an information based forecast, but the conjectured consensus plus or minus some add-factor. Bias from this source will not disappear when averaged over target years. However, it will disappear when averaged over forecasters, since some will necessarily be optimistic and some pessimistic relative to the average, consensus, forecast.

Third, in certain circumstances, it may be optimal for all forecasters to introduce a bias towards optimism. Consider Figure 1, which shows successive series of monthly forecasts of earnings of the companies in the Standard & Poor's 500 stock index for the target years 1990-2006. The earnings forecasts are aggregated from forecasts for the individual companies, each made by several analysts. Like the Consensus Economics forecasts, these predictions start two years before the end of the target year, so analysts make a series of 24 monthly forecasts for each target year. The dotted lines in the figure join the series of forecasts for each target year, and the firm lines show the 12-month ahead forecasts, and the actual outcomes. In all but the most recent years (and possibly 1995 when forecasts were flat), the forecasts show a consistent "walkdown" profile. The initial earnings forecast is between 10% and 50% too optimistic. Then as the months pass, this is gradually revised down towards the realized level. Because the initial optimism does not disappear under aggregation, it cannot be due to product differentiation activity (which would cause some analysts to produce higher and some lower forecasts than the consensus). The generalized bias suggests that there are some incentives operating on *all* analysts that caused them to make excessively optimistic earnings forecasts.

Some commentators have argued that this bias is illusory, the result of a skewed distribution of analyst forecasts which are mostly unbiased but with a few extreme outliers, and it is true that median forecasts are less biased and converge faster than mean forecasts (Brown, 2001; Richardson, Teoh & Wysocki, 2004). Some commentators argue that the bias is due to self-selection by analysts, who for career satisfaction gravitate to sectors and companies that they perceive as having relatively good prospects (McNichols & O'Brien, 1997). However, Butler and Lang (1991) find biases to pessimism as well as optimism among individual analysts. Others have more plausibly suggested that optimism benefits analysts by building relationships with the target companies, and leads to more access to people and information (Francis & Philbrick, 1993; Francis, Hanna & Philbrick, 1997; Lim, 2001). Another possibility is that optimistic forecasts generate trades, since the investment funds that are the main clients of analysts are typically looking for reasons to buy shares in one company rather than another. As is consistent with this idea, forecasts from banks with business relationships with target firms tend to be more optimistic than forecasts from disinterested forecasters (Dechow, Hutton & Sloan, 2000). It also appears that analysts' promotion prospects inside a firm are related to forecast optimism (Hong & Kubik, 2003).

Although it is clearly relevant to financial analysts, this generalised bias to optimism does not apply obviously to economic forecasters. Even for US equity analysts the environment is changing. The collapse of Enron in 2001 and Worldcom in 2002 led to high profile prosecutions of analysts found to be boosting shares, and the Sarbanes-Oxley Act which *inter alia* requires US companies to make accurate and transparent public statements about projected earnings. It may not be fanciful to attribute the excessive moderation in analyst forecasts after 2002 to this new set of incentives, which on the one hand penalises excessive optimism in company announcements, and on the other hand makes a higher quality stream of company data available to all forecasters equally.

2.2 Conservatism

There are three reasons why forecasters might apparently give excessive weight to their own past forecasts. First, the studies collected in Kahneman, Slovic and Tversky (1982) suggest strongly that in making judgments with imperfect information, individuals are subject to biases in the way information is used. Specifically, they tend to be overconfident about the likely accuracy of their own forecasts, and they tend to anchor subsequent forecasts on readily available information, often regardless of its relevance. Taken together, these cognitive failures are liable to lead forecasters to overweight their own past forecasts, and underreact to news, including other forecasters' predictions (Kahneman & Tversky, 1973).

Second, conservatism in forecast revisions might reflect reputation-building through consistency, of a kind discussed earlier. Ehrbeck and Waldeman (1996) argue that forecasters attempt to create credibility by mimicking the characteristics of good forecasters. One characteristic popularly attributed to good forecasters (though not by Ehrbeck & Waldeman) is that they do not have to revise their forecasts as much as bad forecasters. So a bad forecaster can try to look good by being consistent.

Third, if the underlying process driving the target variable is not stable, it is rational for error-minimising forecasters to make serially correlated forecast revisions, and systematic forecast errors, as they learn about changes in the process. A reasonable model for many economic variables is that they consist of a permanent but time-varying component *z*, say, and a transitory component *u*, so that:

$$z_t = z_{t-1} + v_t \qquad v_t \sim N(0, \rho^2) \tag{4}$$

$$y_t = z_t + u_t \qquad u_t \sim N(0, \sigma^2) \tag{5}$$

where $cov(u_t, v_t) = 0$. In this set-up, the news based forecast g_t is the latest observation y_t. As shown by Muth (1960), the error-minimising forecast of y at some future fixed target date T is the single exponential smoothing model:

$$f^*_{t,T} = \lambda y_t + (1-\lambda) f^*_{t-1,T} \tag{6}$$

where $\lambda = \rho^2 / (\rho^2 + \sigma^2)$. That is, forecasts are gradually adjusted in response to new observations on y, with the speed of response λ depending on the relative variance of permanent and transitory changes in y. When a change in the permanent component of the target variable occurs, adjustment in rational forecasts does not occur instantaneously. There is a learning period with a half-life of $1/\lambda$ during which forecasters determine whether the observed change in the target variable is transitory or permanent. Forecast revisions, and errors in these adaptive forecasts, will not be serially uncorrelated, and forecasts will appear biased as a result of the time taken to learn about permanent changes.

2.3 Herding

Forecasters in financial markets overweight the forecasts of other forecasters, leading to the phenomenon of "herding", or excessive concentration of earnings forecasts. Herding is also observed in the behaviour of fund managers, who tend to make portfolio decisions that are similar to those of other managers. Returns to professionally managed investment funds are consequently less dispersed than would be expected by chance. And buying and selling by some managers of, say, investments in technology shares or investments in Thailand, is liable to trigger market bubbles or collapses as others imitate their behaviour. The voluminous academic literature on herding in finance is surveyed in Bickchandani and Sharma (2001).

There are two main theories of how rational herding can arise. The "information cascade" theory posits that forecasts are made sequentially by different agents, and so as each forecast is published it becomes part of the next forecaster's information set. Later forecasts are therefore biased towards the early forecasts, as is the consensus itself. Graham (1999), for example, finds relationships in published forecasts in runs of investment newsletters. The "incentive concavity" theory assumes that the rewards for making an accurate but "bold" (i.e. away from the consensus) forecast are smaller than the penalties for an inaccurate bold forecast. According to Lamont (2002), one prediction of this theory is that less experienced forecasters herd more since their career prospects are at stake. He finds that in a group of US forecasters publishing in *Business Week*, the less experienced do indeed produce fewer extreme predictions. However, Pons-Novell (2003) finds no tendency for inexperienced forecasters in the long-running US Livingston survey to herd on the consensus, and if anything, Stark (1997) finds the opposite result when the same tests are conducted on the well-known FRB Philadelphia *Survey of Professional Forecasters* data set. Applying Lamont's model to forecasts made by economists in Japan, Ashiya and Doi (2001) similarly find no relationship between age and divergence from the consensus. In any case, it is unclear that the incentive to be bold increases with age. In Prendergast and Stole (1996), for example, it is the “impetuous youngsters” who make bold decisions and the “jaded old-timers” who conform.

Whereas herding appears to be the rule in financial forecasting, all research on economic forecasting finds that forecasters underweight the consensus. We have noted that in the model of Ehrbeck and Waldeman (1996), forecasters try to achieve credibility by mimicking the behaviour of good forecasters. The hypothesis that they test is that, because the consensus is known to be a good forecast, bad forecasters will herd on the consensus. However, using data from the Blue Chip Economic Indicators panel they find exactly the opposite, confirming the earlier results of Batchelor and Dua (1992b) on the same survey.

To rationalise these findings, Laster, Bennett and Geoum (1999) develop a model in which there are two users of forecasts: regular users, who know the whole track record of the forecaster, and occasional users who only register the most recent success or failure of the forecaster. To please regular users, forecasters would make predictions that have the minimum expected squared error, since these will have lowest mean square error over a long run of forecasts. To please occasional users, however, forecasters have an incentive to bias each prediction away from the consensus, so as to improve their chances of being the most accurate in a one-off contest. In a similar vein, Ottaviani and Sorensen (2006) develop a model in which the rewards depend on mean square accuracy, but the rewards have to be shared with neighbouring forecasters. The optimal strategy for all forecasters is to move published forecasts a little away from their estimates of the minimum squared error prediction, towards the tail of the distribution of forecasts, and away from the consensus. This reduces the probability of winning the forecast contest, but increases the size of the reward.

Both of these frameworks imply that the dispersion of economic forecasts around the consensus will be larger than expected if each forecasters published the variance minimising prediction (1). However, over a run of forecasts, each forecaster will appear unbiased, sometimes making exaggerated over-predictions and at other times exaggerated under-predictions.

3. Data and empirical tests

The Consensus Forecasts service started publishing forecasts for the G7 economies in September 1989. Each month the service surveys a number of private sector forecasting bodies, normally based in the target country, and publishes tables showing individual forecasts and the arithmetic average or "Consensus" forecast for a range of macroeconomic variables. The Consensus Economics growth and inflation forecasts have been widely used by practitioners and researchers as a benchmark for official forecasts (Artis, 1996; Batchelor, 2001), and to investigate the accuracy of private sector forecasts (Blix et. al., 2001; Loungani, 2001; Isiklar & Lahiri, 2006; Lahiri & Sheng, 2006). Exchange rate forecasts collected by Consensus Economics have also been used to investigate forecaster behaviour; see for example, MacDonald and Marsh (1994, 1996) and more recently Beine, Bénassy-Quéré, and Colas (2003).

Here we focus on their forecasts for real GDP and consumer price inflation. The forecasting institutions are typically international or national banks, business corporations, trade associations, or research institutes. The number of institutions surveyed differs across countries. There are typically 25-30 forecasters making predictions for the US, and maybe 15-20 for Italy. Responses also change from month to month within countries, and forecasters periodically leave or join the service. So the composition of the panel changes a little from month to month, and some forecasters are regular contributors over a number of years, while others appear less frequently. Forecasts are made for the current year and the following year, so if we imagine that the actual variable is revealed in the January following the target year, the forecast horizon starts at 24 months, and shrinks to 1 month. Our data start with the January 1990 forecasts, and end with the December 2004 forecasts, giving predictions at all horizons for the 14 target years 1991 – 2004.

Two issues arise when comparing these forecasts to the actual outcomes for GDP growth and inflation. The first is that both figures – and especially GDP growth – are subject to revision and rebasings. We have looked at the "preliminary" estimates of GDP growth and inflation, which are released early in the year following the target year, and recorded in the February Consensus Forecasts publications. We have also looked at the "final" estimate of growth and inflation published one year later (that is, in February of the second year following the target year). Although there are some differences in error statistics depending on which measure is used, they make no difference to any of the inferences drawn in this study, and most of the tables and figures below use the final estimates. The second issue concerns the target for the GDP in Germany. In surveys up to May 1997, Consensus Forecasts asked respondents for the future growth in both the former West Germany and Germany as a whole. We have followed the pattern of reporting in the electronic Consensus Forecasts database, and used figures for West Germany in predictions made for the years 1990-1996, and all Germany in the years 1997-2004. Again, looking at the consensus figures for all Germany prior to 1997 suggests that this will not affect inferences about forecast errors.

We measure bias using the error $y_T - f_{it,T}$ in the forecast made at time t by forecaster i for variable y at target date T. This can be decomposed into the bias in the consensus forecast $f_{t,T}$ and the deviation in the individual forecast from the consensus, as:

$$y_T - f_{it,T} \equiv (y_T - f_{t,T}) - (f_{it,T} - f_{t,T}) \tag{7}$$

We discuss the properties of bias in the consensus, and bias in individual forecasts around the consensus, in turn.

3.1 Bias in Consensus Forecasts

Table 1 shows the mean and standard deviation of annual growth and inflation rates in the G7 economies, and the average bias in the consensus forecasts at 24-, 18-, 12- and 6-month horizons, where bias is defined as the final outcome less the consensus forecast.

In predicting growth rates, there is an obvious difference between the track record of forecasters in the US, and that of forecasters for the other G7 countries. On average the initial forecasts for US growth have shown little bias, and indeed less bias than forecasts at shorter horizons, though none of the figures is statistically significant. In all the other countries, forecasts have tended to be initially optimistic (negative bias). The bias in forecasts for the UK and Canada are small (0.5% or lower), and are not statistically significant, though they do persist, and in the case of Canada a tendency to overprediction continues well into the target year. The bias in the consensus forecast for Japan is larger, nearly 1%, but is not consistent from year to year, and so is not statistically significant. The initial bias in Japan is also eradicated relatively quickly, and there is little evidence of bias in forecasts at horizons of less than 18 months.

This contrasts with the experience in forecasting the core European Union economies of Germany, France and Italy. In these countries, initial forecasts have tended to be persistently too high, so that the initial biases of around 1% are all statistically significant. As information accumulates, the biases are reduced somewhat, but much more slowly than in Japan. In the case of Italy, a statistically significant bias to optimism persists even in very short term 6-month forecasts. Figure 2 shows the complete track record of the Consensus growth forecasts for Italy. The 24 successive monthly forecasts for each target year, shown by the solid lines, show a walkdown pattern very similar to that exhibited by the financial analysts in Figure 1. In 11 out of the 15 target years, the initial consensus GDP forecast was too high, and was only slowly revised downwards. The pictures are similar if a little less dramatic for Germany and France.

The biases in inflation forecasts are almost a mirror image of the biases in growth forecasts. In Germany and Italy, where the bias to optimism in real GDP forecasts is most pronounced, inflation forecasts have been unbiased at all horizons. In the other countries, the forecasts have initially been biased upwards by about 0.5%. This bias does not persist very long, however, and is close to zero for all countries by the 12 month horizon. Because the biases in inflation forecasts are rather small we focus here on the causes of bias in the real GDP forecasts.

A bias toward optimism can arise from incentives operating on forecasters. However, it is extremely unlikely that economists are subject to the same kind of pressures to produce optimistic forecasts as investment analysts. Also, while forecasters might have reasons to differentiate themselves from the consensus, this is unlikely to be the cause of bias in the consensus itself. Another possibility is that that statistics are distorted by a few outlying observations. As shown by Loungani (2001) and Loungani and Trehan (2002), economists are notoriously bad at forecasting recessions, but are better at timing recoveries. The resulting small number of large overpredictions made in recession years can give an illusion of a bias to optimism in short data series. This does not seem to be a problem here. For example, Figure 2 shows that Italy suffered only one year of negative growth over our data period, and overprediction was not confined to the worst years.

The most likely reason for bias in the consensus forecasts of real growth is that it reflects errors made during the process of learning about a slowdown in the trend growth rates in Japan and continental Europe. In the upper panel of Table 2 we show growth rates by decade for the G7 economies. The US, Canada and the UK show little sign of a slowing in the growth rates between the 1970s, 1980s, 1990s and the early 2000s. In contrast, the growth rate in Japan fell dramatically in the 1990s, and in Germany, France and Italy growth fell in the 1980s, fell further in the 1990s, and further still in the 2000s. The lower panel of Table 2 shows estimates of the parameter λ for the single exponential smoothing model of equation (6), where the target variable y is the change in the natural logarithm of the GDP in each country, and parameter estimates are based on annual data for the years 1960-2004. For the US, Canada and the UK, $\lambda = 0$, implying that changes in growth rates from year to year are transitory rather than permanent. For the other countries, λ is significantly positive, around 0.2 for Germany and Italy and close to 0.5 for France and Japan. In these countries, 20-50% of movements in growth rates represent permanent changes. In the final row of Table 2 we report the prediction from the smoothing model of the underlying growth rate of each G7 economy based on information up to 2004. This is of course just the long term trend growth in the US, Canada and the UK. However, in Germany, France, Italy and Japan, the projected trend is below even the average (low) growth achieved in the 1990s. In the case of Italy, the projected trend is less than 1% per annum.

In these circumstances, it is inevitable that forecasts will be biased, as forecasters in Germany, France, Italy and Japan learn about the new lower trend growth rate. This is consistent with the cross-country patterns of bias in our data. For example, the fall in trend growth in Japan happened earlier than in Europe, following the stock market crash in 1990, and forecasters now seem to have adapted to the new environment. In

Europe, most bias is observed in Germany and Italy, where trend growth has fallen further than in France.

It is also possible that reliance on econometric methods makes it harder for forecasters to adapt, for two reasons. One is the underdevelopment of the supply side. In macroeconomic forecasting models, descriptions of aggregate demand are generally very rich, reflecting a well developed theory of demand, the ready availability of data on components of demand, and many observations of demand shocks in the form of policy changes, and exchange rate movements. However, trend growth is a supply-side phenomenon. Theories of growth are less well articulated. They rely on variables such as labour productivity, financial fragility and social capital, all of which are hard to measure. Oil prices are the only generalised supply side shock for which we have observations which are usable in a time series model, but even in this case it has proved hard to model the substitution effects that have diminished the impact of oil shocks over time. Changes in supply side conditions are therefore hard to account for in a conventional short term econometric model. A second problem is that many econometric models incorporate "error-correction" or "core model" features. These drive forecasts of growth and inflation to some underlying trend. Given that most models are parameterised using methods that give equal weight to all past observations, this trend will reflect some average relationship between aggregate supply and its determinants observed over the past several decades. As Clements and Hendry (2003) point out, if structural changes affecting the trend are occurring, the error correction feature will slow the rate of adaptation relative to simpler models without error correction features.

3.2 Bias in individual forecasts

In addition to bias in the consensus forecasts, there may also be biases at the level of the individual forecaster, in the sense that particular forecasters may tend to make predictions that are persistently higher or lower than the consensus.

O'Brien (1990) tests for significant differences in the accuracy of a panel of investment analysts by running the fixed effects model:

$$|y_T - f_{i,t,T}| = a_0 + \sum_{i=1}^{M} a_i .ID_i + \sum_{T=1}^{N} c_T .YR_T + \varepsilon_{it,T} , \tag{8}$$

where y is the target variable and f the forecast, so that the dependent variable is the absolute forecast error. ID_i is a dummy taking the value 1 for a forecast made by forecaster i in a total panel of M forecasters, and YR_T is a dummy taking the value 1 for forecasts made for the target year T in the total sample of N target years. The coefficients a_i measure "forecaster effects" and the coefficients c_T "year effects" on forecast accuracy. The coefficient a_0 is set to the mean error, so a significant positive value for a_i would indicate a forecaster who was consistently less accurate than average, and vice versa.

An analogous method can be used to test for bias, using

$$y_T - f_{i,t,T} = b_0 + \sum_{i=1}^{M} b_i .ID_i + \sum_{T=1}^{N} d_T .YR_T + \eta_{it,T} , \tag{9}$$

where the dependent variable is simply the forecast error, and the coefficients b_i measure forecaster-effects on the error. With b_0 set to the average bias, a significant positive value for b_i would indicate a forecaster who was consistently more pessimistic than average, and vice versa.

Assuming independent and identically distributed errors $\varepsilon_{it,T}$ $(\eta_{it,T})$, the hypothesis that there is no significant difference in accuracy across forecasters (or a subgroup of forecasters) can be tested by testing for equality across the coefficients b_i for all forecasters (or the subgroup), using a conventional F-test. In the context of our data, three econometric issues arise. First, the errors in (8) and (9) are bound to fall as the forecast horizon shortens, so pooling data across horizons is problematical. We therefore conduct tests separately on each forecast horizon. Second, in the case of forecasts with horizons greater than 12 months, there is liable to be correlation in the residuals due to fact that periods covered by successive forecasts overlap. We have therefore estimated the parameters of (8) and (9) and their standard errors on the assumption that residuals will follow an appropriate moving average process, namely $\varepsilon_{it,T} \sim \text{MA}\{\max(0, T - t - 12)\}$, and similarly for $\eta_{it,T}$.

Third, in addition to these complications, the distribution of $\varepsilon_{it,T}$ is certainly non-normal, and that of $\eta_{it,T}$ possibly non-normal, so the assumptions underlying the use of the F-test are violated. For this reason, Batchelor and Dua (1990b) use the nonparametric Friedman (1937) analysis of variance by ranks to test for significant differences in accuracy and bias. Skillings and Mack (1981) generalise this test to an unbalanced panel as follows. Let $R_{it,T}$ be the rank of forecaster i in the survey at time t made for forecasts with target date T, with N_t forecasters participating in the survey. The normalised rank of forecaster i in this survey is

$$r_{it,T} = \frac{\left\{R_{it,T} - \dfrac{N_t+1}{2}\right\}}{\sqrt{\left\{\dfrac{1+N_t}{12}\right\}}} \tag{10}$$

We write the sum of ranks over a set of surveys with a common horizon $h = T\text{-}t$ as r_{ih}, and the vector of these sums for the first N-1 forecasters as $\boldsymbol{r}_h = [\ r_{1h},\ r_{2h},\ \ldots r_{N\text{-}1h}]$. To adjust for the fact that each forecaster pair is not present in every survey we construct an (N-1) x (N-1) weighting matrix V with diagonal entries = $-m_{ih}$, the number of surveys where forecaster i makes an h-period forecast, and off-diagonal entries M_{ijh}, the number of surveys where both i and j make forecasts. Then $\boldsymbol{r}_h' V^{-1} \boldsymbol{r}_h \sim \chi^2$ (N-1) under the null hypothesis that the ranks are equal across forecasters. In the case of forecasts with horizons above 12 months, the overlapping forecast periods mean that the ranks from successive surveys are not independent. For these cases we report the average values of tests statistics from all non-overlapping subsets of our data.

Ashiya (2006) applies this to testing for differences in accuracy and bias in a large panel of Japanese economic forecasts in the years 1981-2003, a much longer run of data than in previous studies. He finds that the results from the nonparametric method are very similar to those from fixed effects regressions. Both confirm the presence of persistent biases to optimism or pessimism by individual panellists.

We have conducted exhaustive experiments using the 2 hypotheses x 2 tests x 7 countries x 24 horizons x 2 variables. Because the test statistics have a clear pattern, we do not show them all exhaustively. Table 3 sets out results relating to differences in accuracy and differences in optimism, from the fixed effects regression for all countries for 12-month ahead forecasts of growth and inflation. These results are typical of findings at all but the very shortest horizons. The F-tests for equality of coefficients are conducted for "regular" forecasters, who made predictions in at least 40% of the relevant surveys. The second column of Table 3 shows the numbers of regular forecasters in each country, ranging from 14 for Canada and 15 for Italy up to 30 for the US and 38 for the UK.

The hypothesis that these forecasters do not differ in their mean absolute error cannot be rejected at the 5% level except in two quite marginal cases. The test statistics are just significant for the GDP forecasts in France (p-value = .0517), and the inflation forecasts in Canada (p = .0741). Overall there is little evidence that some forecasters are more accurate than others.

The hypothesis that forecasters do not differ with respect to biases toward optimism or pessimism is very strongly rejected for the real growth forecasts in all countries. The hypothesis is also rejected for the inflation forecasts in all countries except the UK, where there is no evidence of individual bias in inflation forecasts.

Table 4 reports results for tests on the 12-month horizon forecasts using the Skillings-Mack (1981) non-parametric test. Although the p-values are somewhat higher than in the case of the F-tests, the inferences to be drawn are exactly the same.

To establish when biases arise, and whether they are eliminated over time, Table 5 shows the F-statistics by selected horizons for one country, the US. Other countries exhibit similar characteristics. There is little evidence of individual differences in accuracy of real growth forecasts even in the 24-month ahead forecast series. However, statistically significant biases toward optimism and pessimism are present at the longest horizons, and, surprisingly, tend to persist even down to very short 3- to 1-month horizons, when the range of forecasts across individuals is small. With respect to inflation forecasts, there is less consistency. In the US, there are significant differences in accuracy and in optimism at most horizons. However, in some surveys the hypothesis of no differences cannot be rejected, and the differences tend to disappear as the horizon shortens.

4. Concluding Comments

This paper has documented the presence of systematic bias in the real GDP forecasts – and to a lesser extent the inflation forecasts – of private sector forecasters in the G7 economies in the years 1990-2005. The data come from the monthly Consensus Economics forecasting service, and bias is measured and tested for significance using parametric fixed effect panel regressions and nonparametric tests of accuracy ranks. The patterns across countries and forecasters give some clue as to whether bias reflects the inefficient use of information, or whether it reflects a rational response to financial, reputational and other incentives operating on forecasters.

In several G7 countries – Japan, Italy, Germany, France – there is evidence of a bias in the consensus forecast for real GDP growth. We do not interpret this as a symptom of irrationality. Nor can it reflect political bias, since the forecasts come from a variety of sources, all private sector, and with no political reason to manipulate their forecasts. Some may have preferences for or against the party in government, but political biases should cancel out in the consensus forecasts. The most compelling explanation for optimism is that it reflects a rational adaptation of forecasts to the falling in the trend growth rate in these countries. In these circumstances, standard tests for rationality are inappropriate, and a bias towards optimism in the consensus forecast is inevitable as rational forecasters learn about the new trend. The fact that the same bias is not observed in the US, Canada and the UK, where trend growth has not fallen, also supports this hypothesis.

This kind of bias is shared by all forecasters. However, we also find persistent individual biases, again mainly in real GDP forecasts. Specifically, forecasters seem to start their forecasting round by adopting a relatively optimistic or pessimistic view of growth. Forecasters who start optimistic one year also start off being optimistic in other years. These biases persist throughout the forecasting cycle, as more information about the target variable, and the forecasts of other forecasters, arrives. Even at short horizons, less than 3 months, rankings by optimism seem to be preserved.

The persistent optimism of some forecasters, and the persistent pessimism of others, is not consistent with the predictions of models of “rational bias” that have become popular in the finance and economics literature. These models predict that forecasters will move their initial forecasts away from the consensus to achieve certain objectives. However, there is nothing in the theory to explain why an individual forecaster would repeatedly seek to publish an initial forecast that is, say, higher than the consensus. The initial forecast in these theories is determined arbitrarily by the news available to each forecaster, and will be randomly above or below the initial consensus. Our explanation is that since the product of a forecaster is a number, the only way to differentiate that product is to make the number high or low. Forecasters thereby cultivate a reputation as optimists or pessimists. The extent to which this is possible without compromising accuracy depends on how noisy the target variable is. The forecasters in the Consensus Forecasts service seem to have judged this well. Or perhaps only forecasters who manage the tradeoff between accuracy and publicity well survive to become regular contributors to Consensus economics. Although there are clear differences among forecasters when ranked by optimism and pessimism, there are no significant differences among their accuracy ranks.

References

Artis, M. J. (1996). *How accurate are the IMF's short-term forecasts? Another examination of the world economic outlook*. IMF Working Paper 96/89.

Ashiya. M. (2006). Forecast accuracy and product differentiation of Japanese institutional forecasters. *International Journal of Forecasting, 22*, 395-401.

Ashiya, M., and Doi, T. (2001). Herd behavior of Japanese economists. *Journal of Economic Behavior and Organisation*, *46*, 343-346.

Batchelor, R. A. (2001). How useful are the forecasts of intergovernmental agencies? The IMF and OECD versus the consensus. *Applied Economics*, *33*, 225-235.

Batchelor, R. A., & Dua, P. (1990a). Forecasting technique, forecaster ideology, and the accuracy of economic forecasts. *International Journal of Forecasting*, *6*, 3-10.

Batchelor, R. A., & Dua, P. (1990b). Product differentiation in the economic forecasting industry. *International Journal of Forecasting*, *6*, 311-316.

Batchelor, R. A., & Dua, P. (1992a). Conservatism and consensus-seeking among economic forecasters. *Journal of Forecasting*, *11*, 169-181.

Batchelor, R. A., & Dua, P. (1992b). Blue Chip rationality tests. *Journal of Money. Credit and Banking, 23*, 692-705.

Beine, M. A., Bénassy-Quéré, A., & Colas, H. (2003). *Imitation Amongst Exchange-Rate Forecasters: Evidence from Survey Data*. THEMA Discussion Paper 2003-39, Universite de Cergy-Pontoise.

Bickchandani, S., & Sharma S. (2001). Herd behaviour in financial markets. *IMF Staff Papers*, *47*, 279-310.

Blix, M., Wadefjord, J., Wieneke, U., Adahl, M., 2001, How good is the forecasting performance of major institutions? *Sverige Riksbank Economic Review*, 3, 38-67.

Brown, L. D. (2001). A temporal analysis of earnings surprises: profits and losses. *Journal of Accounting Research*, *39*, 221-241.

Butler, K. C., & Lang, L. H. P. (1991). The forecast accuracy of individual analysts: evidence of systematic optimism and pessimism. *Journal of Accounting Research*, *29*, 150-156.

Clements, M. P., & Hendry, D. F. (2003). Economic Forecasting: Some Lessons from Recent Research. *Economic Modelling*, *20*, 301-329.

Dechow, P. M., Hutton, A. P., & Sloan, R. G. (2000). The relation between analyst forecasts of long term earnings growth and stock price performance following equity offerings. *Contemporary Accounting Research*, *17*, 1-32.

Ehrbeck, T., & Waldmann, R. (1996). Why are professional forecasters biased? Agency versus behavioral explanations. *Quarterly Journal of Economics*, *111*, 21-40.

Estrin, S., & Holmes, P. (1990). Indicative planning in developed economies. *Journal of Comparative Economics*, *14*, 531-554.

Francis, J., & Philbrick, D. (1993). Analysts' decisions as products of a multi-task environment. *Journal of Accounting Research*, *31*, 216-230.

Francis. J., Hanna, D., & Philbrick, D. (1997). Management communications with securities analysts. *Journal of Accounting and Economics*, *24*, 363-394.

Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association*, *32*, 675-701.

Graham, J. (1999). Herding among investment newsletters: theory and evidence. *Journal of Finance, 54*, 231-268.

Heinemann, F. (2005). *Planning or propaganda? An evaluation of Germany's medium term budgetary planning*. Discussion paper 05-12, ZEW Mannheim.

Hong, H., & Kubik, J. D. (2003). Analyzing the analysts: career concerns and biased earnings forecasts. *Journal of Finance*, *58*, 313-351.

Isiklar, G., & Lahiri, K. (2006). *How far ahead can we forecast? Evidence from cross-country surveys*. Working Paper, Department of Economics, University at Albany – SUNY.

Jonung., L., & Larch, M. (2006). Improving fiscal policy in the EU. The case for independent forecasts. *Economic Policy*,21, 491-534.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, *80*, 237-251.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.) (1982). *Judgement Under Uncertainty: Heuristics and Biases.* Cambridge: Cambridge University Press.

Lahiri, K., & Sheng, X. (2006). *Evolution of forecast disagreement and uncertainty in a heterogeneous agents model*. Working Paper, Department of Economics University at Albany – SUNY.

Lamont, O. (2002). Macroeconomic forecasts and microeconomic forecasters. *Journal of Economic Behavior and Organisation*, *48*, 265-280.

Laster, D., Bennett, P., & Geoum, I. S. (1999). Rational bias in macroeconomic forecasts. *Quarterly Journal of Economics*, *114*, 293-318.

Lim, T. (2001). Rationality and analysts' forecast bias. *Journal of Finance*, *56*, 369-385.

Loungani, P. (2001). How accurate are private sector forecasts? Cross-country evidence from Consensus Forecasts of output growth. *International Journal of Forecasting*, *17*, 419-432.

Loungani, P., & Trehan, B. (2002). *Predicting when the economy will turn*. Federal Reserve Bank of San Francisco, Economic Letter 2002-7.

MacDonald, R., & Marsh, I. (1994). Combining exchange rate forecasts. What is the optimal consensus measure? *Journal of Forecasting*, *13*, 313-332.

MacDonald, R., & Marsh, I. (1996). Foreign exchange forecasters are heterogeneous: confirmation and consequences. *Journal of International Money and Finance*, *15*, 665-685.

McNichols, M., & O'Brien, P. (1997). Self-selection and analyst coverage. *Journal of Accounting Research*, *35*, 167-199.

Muth, J. F. (1960). Optimal Properties of Exponentially Weighted Forecasts. *Journal of the American Statistical Association*, *77*, 299-306.

O'Brien, P. C., 1990, Forecast accuracy of individual analysts in nine industries, *Journal of Accounting Research*, 28, 2, 286-304.

Ottaviani, M. & Sorensen, P. N. (2006). The strategy of professional forecasting. *Journal of Financial Economics, 81*, 441-466.

Pons-Novell, J. (2003). Strategic bias, herding behavior and economic forecasts. *Journal of Forecasting*, *22*, 67-77

Prendergast, C., & Stole, L. (1996). Impetuous youngsters and jaded oldtimers: acquiring a reputation for learning. *Journal of Political Economy*, *104*, 1105-1134.

Richardson, S., Teoh, S., & Wysocki, P. (2004). The walk-down to beatable analyst forecasts: The role of equity issuance and insider trading incentives. *Contemporary Accounting Research*, *21*, 885-924.

Skillings, J. H., & Mack, G. (1981). On the use of a Friedman-type statistic in balanced and unbalanced block designs. *Technometrics*, *23*, 171-177.

Stark, T. (1997). *Macroeconomic forecasts and microeconomic forecasters in the Survey of Professional Forecasters*. Federal Reserve Bank of Philadelphia Working Paper 97-10.

Ulan, M., Dewald, W. G., & Bullard, J. B. (1995). US official forecasts of G-7 economies 1976-90. *Federal Reserve Bank of St. Louis Review*, March/April.

Table 1 Bias in Consensus Forecasts of growth and inflation

	US	Canada	UK	Germany	France	Italy	Japan
Average growth rate	2.67	2.35	1.95	1.73	1.89	1.35	1.59
Standard deviation	1.60	1.74	1.64	1.65	1.20	0.97	2.03
Bias in Consensus							
24-month	0.04	-0.55	-0.38	-0.96	-0.85	-1.24	-0.72
t-stat	*0.11*	*-1.40*	*-1.06*	*-2.53*	*-2.46*	*-4.22*	*-1.25*
18 month	0.01	-0.59	-0.46	-0.77	-0.74	-1.06	-0.37
t-stat	*0.02*	*-1.45*	*-1.15*	*-2.03*	*-1.98*	*-3.84*	*-0.65*
12-month	0.14	-0.30	-0.18	-0.22	-0.31	-0.63	0.24
t-stat	*0.46*	*-0.83*	*-0.63*	*-0.88*	*-1.16*	*-3.33*	*0.61*
6-month	-0.13	-0.17	0.08	0.10	-0.10	-0.23	0.08
t-stat	*-0.72*	*-1.00*	*0.57*	*0.67*	*-0.72*	*-1.87*	*0.32*

	US	Canada	UK	Germany	France	Italy	Japan
Average inflation rate	2.86	2.27	3.25	2.15	3.29	2.69	-0.67
Standard deviation	*0.96*	*1.37*	*2.01*	*1.08*	*1.02*	*2.23*	*1.52*
Bias in Consensus Inflation forecasts							
24-month	-0.05	0.03	0.15	0.02	-0.13	0.20	0.02
t-stat	*-0.32*	*0.26*	*0.47*	*0.11*	*-0.91*	*0.70*	*0.11*
18 month	-0.06	0.02	-0.16	-0.02	-0.08	-0.12	-0.06
t-stat	*-0.54*	*0.12*	*-0.69*	*-0.19*	*-0.59*	*-0.49*	*-0.46*
12-month	-0.03	0.07	0.11	-0.19	-0.19	-0.07	0.04
t-stat	*-0.22*	*0.47*	*0.31*	*-1.50*	*-1.26*	*-0.26*	*0.27*
6-month	0.08	0.12	0.25	-0.17	-0.03	0.21	0.07
t-stat	*0.73*	*1.00*	*1.09*	*-1.23*	*-0.23*	*1.06*	*0.46*

Data source: *Consensus Forecasts* monthly service, 1989-2004. Figures are percent per annum. Figures in italics are t-statistics testing the null hypothesis that the forecast bias is zero.

Table 2 The growth slowdown in Europe and Japan

	US	*Canada*	*UK*	*Germany*	*France*	*Italy*	*Japan*
Average annual growth rates							
1960-70	4.2	5.3	3.0	4.5	5.4	5.7	10.1
1970-80	3.2	4.3	2.0	2.8	3.3	3.6	4.5
1980-90	3.3	2.8	2.6	2.3	2.4	2.3	3.9
1990-00	3.3	2.9	2.4	1.7	1.9	1.6	1.4
2000-05	2.7	2.6	2.4	0.6	1.5	0.5	1.1
Exponential smoothing model parameters and forecasts							
lambda (λ)	0.00	0.20	0.00	0.16	0.47	0.20	0.48
projected trend	3.32	2.96	2.69	1.22	1.61	0.90	1.62

Data source: OECD. Growth figures are percent per annum. Lambda is the parameter of the single exponential smoothing model of Equation (6), $f^*_{t,\,t+1} = \lambda y_t + (1-\lambda) f^*_{t-1,\,t}$ where y_t is growth rate in year t, and $f^*_{t,\,t+1}$ is the forecast of the underlying growth rate made in year t for the year t+1 (and all subsequent years).

Table 3 Tests for differences in accuracy and optimism among individual forecasters, in all G7 countries, for a 12-month horizon: F-tests

Country	**No. of Forecasters**	**Absolute Errors F-statistic**	**Absolute Errors P-value**	**Errors F-statistic**	**Errors P-value**
		Real GDP:			
US	30	0.71	0.8171	3.19	0.0000
CA	14	0.67	0.8003	2.12	0.0125
UK	38	0.87	0.6904	2.35	0.0000
GE	28	0.69	0.8778	3.46	0.0000
FR	22	1.07	0.3797	1.59	0.0517
IT	15	1.28	0.2193	2.01	0.0178
JP	20	1.28	0.1966	1.95	0.0103
		Consumer Price Inflation:			
US	30	1.40	0.1191	2.62	0.0002
CA	14	1.44	0.1392	1.63	0.0741
UK	38	2.22	0.0001	1.91	0.0016
GE	28	1.48	0.0566	3.61	0.0000
FR	22	1.32	0.1592	2.42	0.0007
IT	15	0.89	0.5811	3.28	0.0001
JP	20	0.97	0.5013	1.80	0.0215

F-statistics are tests of the constraints that the coefficients on forecaster effects are equal in the fixed effects regressions (8) and (9), for regular forecasters making forecasts in more than 40% of the relevant *Consensus Forecasts* surveys. P-values below 0.05 indicate statistically significant differences in accuracy or bias at the 5% level.

Table 4 Tests for differences in accuracy and optimism among individual forecasters in all G7 countries, for the 12-month horizon: the Skillings-Mack χ^2 test

Country	*No. of Forecasters*	*Absolute Errors* $\chi 2$ *(N–1)*	*P-value*	*Errors* $\chi 2$ *(N–1)*	*P-value*
		Real GDP:			
US	72	59.49	0.8332	110.53	0.0019
CA	38	29.15	0.8178	57.56	0.0168
UK	74	66.22	0.6999	99.50	0.0213
GE	52	36.65	0.9349	76.53	0.0119
FR	48	47.31	0.4600	61.79	0.0726
IT	43	48.30	0.2335	60.71	0.0308
JP	56	63.71	0.1969	75.53	0.0346
		Consumer Price Inflation:			
US	72	83.47	0.1477	94.06	0.0350
CA	38	43.67	0.2091	47.16	0.1224
UK	74	106.16	0.0068	101.13	0.0163
GE	52	64.39	0.0986	70.66	0.0355
FR	48	54.96	0.1987	64.40	0.0467
IT	43	39.45	0.5835	63.71	0.0169
JP	56	53.71	0.5238	74.52	0.0410

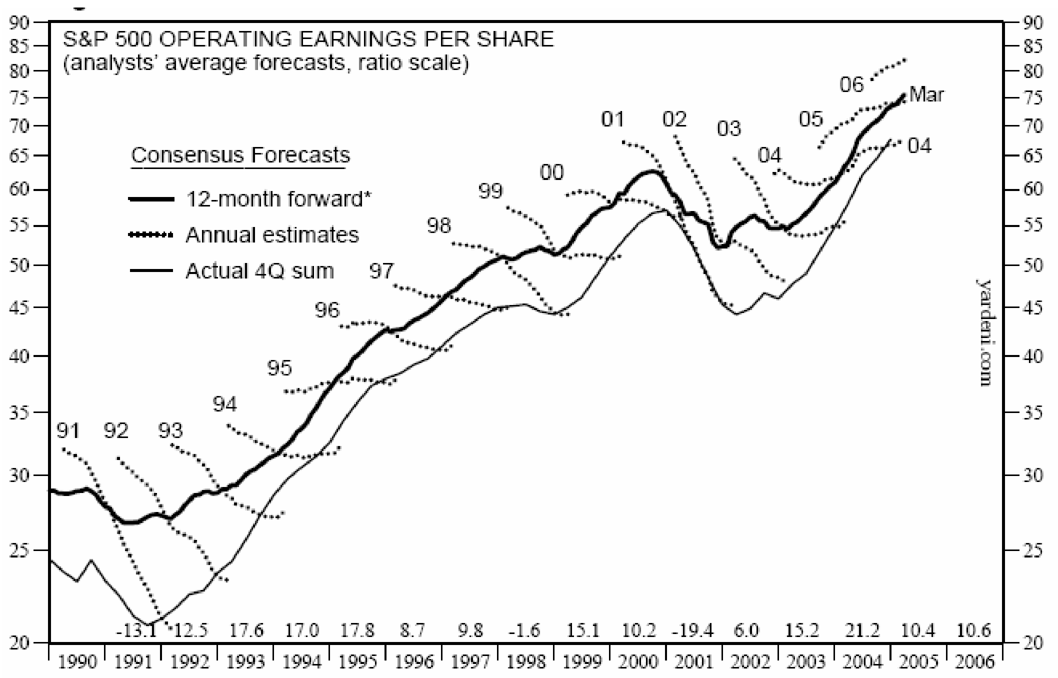
χ^2 statistics are Skillings-Mack test statistics for significant differences in the accuracy and error ranks of forecasters for all forecasters in the relevant *Consensus Forecasts* surveys. P-values below 0.05 indicate statistically significant differences in accuracy or bias at the 5% level.

Table 5 Tests for differences in accuracy and optimism among individual forecasters in the United States, by horizon: F-tests

Horizon (months)	Errors F-statistic	Errors p-value	Absolute Errors F-statistic	Absolute Errors p-value
	Real growth			
24	2.40	0.0010	1.04	0.4122
18	3.94	0.0000	0.51	0.9621
12	3.19	0.0000	0.71	0.8171
6	2.00	0.0074	0.63	0.8913
3	2.68	0.0002	0.72	0.8024
1	0.83	0.6770	1.13	0.3169
	Consumer price inflation			
24	6.14	0.0000	3.29	0.0000
18	3.94	0.0000	2.43	0.0008
12	2.62	0.0002	1.40	0.1191
6	1.24	0.2202	1.48	0.0883
3	2.37	0.0010	2.96	0.0000
1	0.62	0.8986	1.05	0.3993

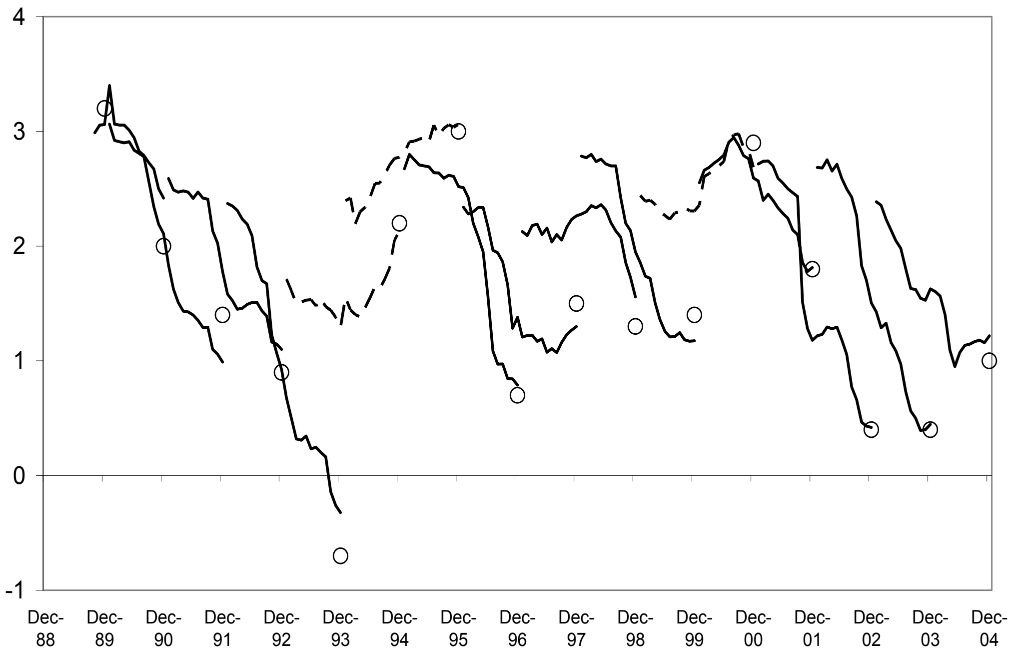
F-statistics are tests of the constraints that the coefficients on forecaster effects are equal in the fixed effects regressions (8) and (9), for regular forecasters making forecasts in more than 40% of the relevant *Consensus Forecasts* surveys. P-values below 0.05 indicate statistically significant differences in accuracy or bias at the 5% level.

Figure 1 Walkdown in Analyst Consensus Forecasts for the S&P500 Company Earnings



The dotted lines connect average monthly analyst forecasts for earnings aggregated over the S&P500 companies for each target year in the period 1990-2006. Each month forecasts are made for the current year and the following year, giving a series of 24 forecasts for each target year. Source: Thomson Financial, Oak Associates (www.yardeni.com).

Figure 2 Walkdown in Consensus forecasts for real GDP growth in Italy



The lines connect the series of monthly *Consensus Forecasts* for real GDP growth in Italy in each of the target years 1989-2004, and circles show the outcome. Each month forecasts are made for the current year and the following year. Source: *Consensus Forecasts*, monthly, 1989-2004.

Ifo Working Papers

No. 38 Sülzle, K., Innovation and Adoption of Electronic Business Technologies, December 2006.

No. 37 Overesch, M. and G. Wamser, German Inbound Investment, Corporate Tax Planning, and Thin-Capitalization Rules – A Difference-in-Differences Approach, December 2006.

No. 36 Kempkes, G. and C. Pohl, The Efficiency of German Universities – Some Evidence from Non-Parametric and Parametric Methods, October 2006.

No. 35 Kuhlmann, A., German Productivity – A Reassessment via the New Ifo Productivity Database, October 2006.

No. 34 Kuhlmann, A., What is the X-Factor in the German Electricity Industry?, September 2006.

No. 33 Temple, J. and L. Wößmann, Dualism and cross-country growth regressions, August 2006.

No. 32 Baumann, F., V. Meier and M. Werding, Transferable Provisions in Individual Health Insurance Contracts, July 2006.

No. 31 Abberger, K., Qualitative Business Surveys in Manufacturing and Industrial Production – What can be Learned from Industry Branch Results?, May 2006.

No. 30 Ruschinski, M., Investigating the Cyclical Properties of World Trade, May 2006.

No. 29 Holzner, Chr., V. Meier and M. Werding, Time Limits in a Two-tier Unemployment Benefit Scheme under Involuntary Unemployment, April 2006.

No. 28 Eggert, W. and A. Haufler, Company Tax Coordination cum Tax Rate Competition in the European Union, April 2006.

No. 27 Lachenmaier, S. and H. Rottmann, Employment Effects of Innovation at the Firm Level, April 2006.

No. 26 Radulescu, D.M. and M. Stimmelmayr, Does Incorporation Matter? Quantifying the Welfare Loss of Non-Uniform Taxation across Sectors, March 2006.

No. 25 Lessmann, Chr., Fiscal Decentralization and Regional Disparity: A Panel Data Approach for OECD Countries, March 2006.

No. 24 Fuchs, Th., Industry Structure and Productivity Growth: Panel Data Evidence for Germany from 1971–2000, December 2005.

No. 23 Holzner, Chr. and A. Launov, Search Equilibrium, Production Parameters and Social Returns to Education: Theory and Estimation, December 2005.

No. 22 Sülzle, K., Stable and Efficient Electronic Business Networks: Key Players and the Dilemma of Peripheral Firms, December 2005.

No. 21 Wohlrabe, K. and M. Fuchs, The European Union's Trade Potential after the Enlargement in 2004, November 2005.

No. 20 Radulescu, D.M. and M. Stimmelmayr, Implementing a Dual Income Tax in Germany: Effects on Investment and Welfare, November 2005.

No. 19 Osterkamp, R. and O. Röhn, Being on Sick Leave – Possible Explanations for Differences of Sick-leave Days Across Countries, November 2005.

No. 18 Kuhlmann, A., Privatization Incentives – A Wage Bargaining Approach, November 2005.

No. 17 Schütz, G. und L. Wößmann, Chancengleichheit im Schulsystem: Internationale deskriptive Evidenz und mögliche Bestimmungsfaktoren, Oktober 2005.

No. 16 Wößmann, L., Ursachen der PISA-Ergebnisse: Untersuchungen auf Basis der internationalen Mikrodaten, August 2005.

No. 15 Flaig, G. and H. Rottmann, Labour Market Institutions and Employment Thresholds. An International Comparison, August 2005.

No. 14 Hülsewig, O., E. Mayer and T. Wollmershäuser, Bank Loan Supply and Monetary Transmission in Germany: An Assessment Based on Matching Impulse Responses, August 2005.

No. 13 Abberger, K., The Use of Qualitative Business Tendency Surveys for Forecasting Business Investing in Germany, June 2005.