How far ahead can we forecast?
Evidence from cross-country surveys

Gultekin Isiklar and Kajal Lahiri
Department of Economics
University at Albany – SUNY
Albany, NY 12222. USA

Corresponding Address: K. Lahiri, Department of Economics, University at Albany – SUNY, Albany, NY 12222. USA. Tel.: (518) 442 4758; Fax: (518) 442 4736.

Email addresses: klahiri@albany.edu (K. Lahiri); gisiklar@gmail.com (G. Isiklar).

How far ahead can we forecast?
Evidence from cross-country surveys

Abstract:

Using monthly GDP forecasts from *Consensus Economics Inc.* for 18 developed countries reported over 24 different forecast horizons during 1989-2004, we find that the survey forecasts do not have much value when the horizon goes beyond 18 months. Using two alternative approaches to measure the flow of new information in fixed-target survey forecasts, we found that the biggest improvement in forecasting performance comes when the forecast horizon is around 14 months. The dynamics of information accumulation over forecast horizons can provide both the forecasters and their clients with an important clue in their selection of the timing and frequency in the use of forecasting services. The limits to forecasting that these private market forecasters exhibit are indicative of the current state of macroeconomic foresight.

1. Introduction

How far ahead into the future do macroeconomic forecasts have value and how does the information content of forecasts change over forecast horizon is a question that has been the focus of various studies.[1] Most of these studies, however, provide measures for the information content of *optimal* forecasts over forecast horizons by modeling the actual data generating process. For example, Öller (1985) and Galbraith (2003) provide estimates of the length of the forecast horizon at which the optimal forecasts contain information by assuming that the actual process follows an ARIMA process. Similarly, Oke and Öller (1999) provide estimates by modeling the actual process using VARMA process.

Granger (1996) pointed out that a feature that will provide limits to how far ahead one can forecast is when the (forecastable) signal gets lost in the (unforecastable) noise. In other words, forecasts will not provide any information when the measurement errors start to make the information content of signals negligible compared to noise. In reality, the measurement errors are not only driven by the level of noise attributed to the data generating process but also to other factors. For example, delays in data releases and data revisions, not to mention structural breaks that are only detectable *ex post*, are some of the factors that may affect the information content of real-time forecasts. In these situations, a forecaster will seem to respond to information that is relevant but also to what is not. These factors do not cause problems in *ex post* analysis of historical data but may induce serious deformation in the information content of real-time forecasts.

[1] See for example, Parzen (1982), Öller (1985), De Gooijer and Klein (1992), Diebold and Kilian (1997), Oke and Öller (1999), and Galbraith (2003).

Only a handful of studies have used real-time survey data to estimate the information content of forecasts (*e.g.*, Mills and Pepper, 1999; Vuchelen and Gutierrez, 2005), and no one has examined the dynamics of how the information content of forecasts change over horizons and how new information increases the information value of forecasts. However, understanding the changes in the information content of forecasts over horizons and, for example, the timing of the arrival of the most important information is critical for both the forecasters and their clients. It is well known that many forecasting agencies like the OECD, Blue Chip, etc. produce forecasts several times a year from an initial 24-month ahead forecast. Some knowledge on the dynamics of information accumulation over forecast horizons can provide forecasters with an important parameter in their selection of the timing and the frequency in the use of forecasting service. For the clients of forecasting firms, the information content of forecasts can be an important consideration in their decisions on how to use and when to use these forecasts.

In this study we address these issues using 15 years of monthly private sector forecast data for 18 developed countries reported over 24 different forecast horizons. We study various characteristics of real GDP growth forecasts over forecast horizons and their differences across countries, and propose two measures for the content of new information in forecasts. We find that the flow of new information to year-over-year GDP growth forecasts follows a hump-shaped curve over horizons with a peak point when the forecast horizon is around 14 months.

The remainder of the paper is structured as follows. Section 2 presents the data. Section 3 discusses certain stylized facts on the evolution of forecasts in a cross–country setting, and reports estimates on the flow of new information at various horizons using two alternative approaches. Section 4 concludes.

2. Data

In this study three data sets are used. The main data of the study on real GDP forecasts come from the *Consensus Economics, Inc*. The second and third data sets contain the actual data series but with different vintages. Our historical data for real GDP growth rates (to calculate the 5-year GDP growth averages and to model the actual GDP growth but not to evaluate forecasts) is constructed from the IMF's *International Financial Statistics* (February 2002 edition). Our real time data set for the purpose of forecast evaluation is mainly constructed from the OECD's mid-year *Economic Outlook,* 1990 to 2004, DEA's *Survey of Business* and Bundesbank's *Monthly Economic Reports*. The details of the data sets follow.

Since October 1989, the *Consensus Economics Inc*. has been polling more than 600 forecasters each month and recording their forecasts for principal macroeconomic variables (including GDP growth, inflation, interest rates and exchange rates) for a large number of countries. Forecasts are made for the current year (based on partial information about developments in that year) and for the following year. The number of panelists ranges from 10 to 30 for most of the countries, and for the major industrialized countries the panelists are based in countries they forecast.

We study the consensus forecasts of annual average real GDP growth. Survey respondents make their first forecasts when there are 24 months to the end of the year they are forecasting; that is, they start forecasting GDP growth in January of the previous year, and their last forecast is reported in the beginning of December of the target year.

So for each country and for each target year we have 24 forecasts of varying horizons. Our data set ranges from October 1989 to June 2004. The countries we study are the eighteen industrialized countries for which forecasts are available from *Consensus Economics, Inc.*[2]

There have been several major changes in the definition of forecast variable since the survey started in 1989. For example, while real GNP was being forecast in the first few years for some countries, the real GDP became the forecast variable since early 1990s. For example, this switch occurred in January 1992 for the US and in January 1993 for Germany. In our data sample, the most significant changes were for Germany. While West Germany's real GNP growth was being forecast through December 1992, after January 1993 the forecast variable became real GDP of West Germany. In addition, unified Germany's GDP growth was added to the survey and West Germany's GDP forecast was removed in May 1997.

In order to evaluate the forecast errors correctly, the forecasts should be matched with the actual data being forecast. It is well documented in the literature that data revisions may have an important impact on the perceived performance of the forecasters. Since forecasters cannot possibly be aware of data revisions after they report their forecasts, we use an early revision as the actual value, which is compiled from the mid-year reports of OECD *Economic Outlook* immediately following the target year. But because of the changes in definitions of target variables (*e.g.,* GNP to GDP or West Germany to Unified Germany) some of the data are not available in the June issues of

---

[2] There are only a very few of studies that have used the *Consensus Forecasts* data set. These are Artis and Zhang (1997), Batchelor (2001), Harvey, Leybourne and Newbold (2001), Loungani (2001), Juhn and Loungani (2002), Gallo, Granger and Jeon (2002), and Isiklar, Lahiri and Loungani (2006). However, none of these studies consider the empirical findings analyzed in this paper.

OECD *Economic Outlook*. We collected these missing data from the original sources, *viz.,* May and June issues of BEA's *Survey of Business* and Bundesbank's *Monthly Economic Reports* for the year immediately following the target year.

3. Evolution of Fixed-Target Forecasts over Horizons

Figure 1 presents the reported forecasts and the realized actual values between 1991 and 2002. Each country's forecasts are divided into three separate panels, which are located horizontally in the figures. Plots start when the forecast horizon is 24, which is reported in January of the previous year, and end when the forecast horizon is 0, which gives the actual realization. Gallo *et al.* (2002) presented this type of graphs for 1993-96 and for three major countries: U.S., U.K., and Japan. We can now examine certain stylized characteristics of the forecast evolution in greater depth.

First, note that for the first six months or so (*i.e.,* for horizons 24 to 18 months), the consensus forecasts do not seem to change very much. This empirical observation leads us to believe that over these horizons, forecasters do not receive dependable information to revise their forecasts systematically. There are important exceptions, however. For the target year 1994, forecasts for Belgium, France, Ireland, and Spain were active from the beginning.

Second, except for Ireland and Japan, the initial forecasts for all other countries seem to be starting from a relatively narrow band and then tend to diverge from these initial starting points. For example, for Austria, Belgium, Denmark and several other countries 24-month ahead forecasts are located between 2 percent and 3 percent, and as information is accumulated these forecasts tend to move towards their final destination.

One may conjecture that these initial long-term forecasts are nothing but unconditional means of the processes. While this conjecture seems to hold for most of the forecasts, the initial forecasts of Irish and Japanese GDP growth rates seem to behave differently. For Ireland, the forecasts tend to move upward and for Japan the forecasts tend to move downward as we go from the far left panel (forecasts for 1991 to 1994) to the far right panel (forecasts for 1999 to 2002). This is understandable given Japan's stagnation and Ireland's extraordinary growth during the 90s. This movement of the 24-month ahead forecasts implies that the long-term expectations have been changing for these two countries, and that recent short run forecast errors have affected the longer-run expectations. See Frenkel (1975) who hypothesized such a feedback.[3]

Third, Gallo *et al.* (2002) noted that the consensus forecasts sometimes do not converge to the right target value due to possible copycat behavior by non-dominant forecasters. In our more comprehensive data set, even though we see some indication of such behavior in certain years for some countries, evidence is not pervasive. For Ireland the one-month ahead forecasts underestimated the targets repeatedly, but this can be explained by the exceptional Irish growth during the nineties. As documented by them, and we also find in Figure 1, U.S. growth for 1995 was seriously overestimated even a month before the end of the target. The last US consensus forecast for 1995 was 3.24% whereas the actual growth based on July revision was 2.03%. This again can be explained by the fact that in the U.S., 1995 was a sudden unanticipated growth slowdown year.[4]

[3] Strictly speaking, the 24-ahead forecasts should be considered as medium-term rather than long-term forecasts.

[4] Öller and Barot (2000) have argued that the divergence between the final forecasts and the preliminary announcements can also be due to the fact that the former tends to underestimate the final figures during upturns and overestimate during downturns. In estimating the preliminary figures, statistical agencies use sample of firms from the previous year sample in which freshly established companies could be missing and conversely, bankrupted companies would be included and impute group averages. In support of the above conjecture, we note that even though the US growth based on preliminary data was 2.03%, it was later revised upwards to 2.50%.

All in all, a close look at these graphs reveals certain undeniable regularities on how the fixed-target forecasts evolve over time. We now proceed to examine more rigorously the timing of the arrival of important information when forecasters break away from their initial estimates.

4. Forecast variance and forecast horizons

Forecasts presented in Figure 1 clearly show that the initial 24-month ahead forecasts are reasonably stable over years, and as the forecast horizon decreases, they tend to diverge. In other words, forecast variability increases as the forecast horizon decreases. While variability of forecast errors is usually associated with uncertainty, forecast variability is inversely related to uncertainty. This argument may look counter intuitive but can be easily understood by the following logic. Consider

$$y_t = f_{t,h} + \varepsilon_{t,h}$$

where y_t is the actual GDP growth, $f_{t,h}$ is the h-period ahead forecast and $\varepsilon_{t,h}$ denote the *ex ante* error associated with this forecast. Since rational expectations imply that $Cov(f_{t,h}, \varepsilon_{t,h}) = 0$ we have $Var(y_t) = Var(f_{t,h}) + Var(\varepsilon_{t,h})$, which implies that the variations in forecasts and forecast errors move in opposite directions as the forecast horizon changes (note $Var(y_t)$ does not depend on the forecast horizon).[5] Therefore as the forecast horizon increases the forecast error variability, and therefore the uncertainty, increases, but the forecast variability decreases.

This observation is confirmed more clearly in Figure 2, where we present the sample variances of the forecasts, *i.e.,* $\sum_t (f_{t,h} - \bar{f}_h)^2 / N_t$, over our sample period at each

[5] This point is noted by several studies. See, for example, Muth (1985).

forecast horizon. The last points in the charts, the points when the horizon is zero, give the variances of the actual values over the sample. As is seen from these figures, as the forecast horizon decreases the variance of the forecast steadily increases. Another way of looking at this increasing variability of forecasts is that as forecast horizon increases more information is accumulated and as more information is accumulated in the forecasts the variation of forecasts increases. This information accumulation process can be mimicked using a simple MA model for the data generating process. Suppose that the actual process has a moving average representation of order q so that

$$y_t = \mu + \sum_{k=0}^{q} \theta_k \varepsilon_{t-k} \,. \tag{1}$$

Then the optimal forecast at horizon h will be

$$f_{t,h} \equiv E(y_t \mid I_{t,h}) = \mu + \sum_{k=h}^{q} \theta_k \varepsilon_{t,k} \tag{2}$$

and the variance of the forecast is

$$Var(E(y_t \mid I_{t,h})) = \sigma^2 \sum_{k=h}^{q} \theta_k^2 \,. \tag{3}$$

Similarly the variance of the forecast when forecast horizon is *h-1* is

$$Var(E(y_t \mid I_{t,h-1})) = \sigma^2 \sum_{k=h-1}^{q} \theta_k^2$$

so that

$$Var(f_{t,h-1}) = Var(f_{t,h}) + \theta_{h-1}^2 \sigma^2 \,.$$

So when the forecast horizon is very long, *i.e.* several years, the forecasts tend to converge towards the mean of the process, and as information is accumulated, the forecasts change increasing the forecast variability. It is interesting to note that Mankiw and Shapiro (1986) used the same argument to conclude that U.S. GDP revisions are "news" rather than "noise". If successive revisions incorporate useful information about past GDP growth, then we will expect the successive revised figures to have more variance than the initial announcement.

While the positive slope in the forecast variance graphs is clear in all figures, there are some differences across countries that are worth mentioning. First, for some countries, *e.g.*, Japan and the USA, the positive slope is not very distinct in the longer-run forecasts, especially when the horizon is more than 18 months. As just shown, the forecast variability increases because of the variability of the accumulated shocks, *i.e.* $\theta_k \varepsilon_{t-k}$ s. Therefore, if forecast variability does not change much over several horizons as is the case, *e.g.*, for Japanese forecasts for horizons from 24 to 15, this may mean that the information acquired 15 months ago does not have much impact on the actual value, *i.e.* $|\theta_k|$ is small. Of course, this may also be related to the informational inefficiency of the forecasts. It is possible that even if the information over this period were important, the forecasters do not incorporate the information in their forecasts causing less than optimal variability in the forecasts. This issue will be addressed later on when we present forecast evaluation measures that are based on forecast errors over forecast horizons.

It is interesting to note that for some countries, the variation of actual values is much larger than the variation of one-month ahead forecasts. This is particularly true for Finland, Ireland, Norway, Portugal, and Switzerland.[6] There can be several reasons for this. One-period ahead forecasts can be written as:

$$y_t = f_{t,1} + u_{t,1}$$

[6] Finland has one of the largest GDP growth variances among industrialized countries, see Öller and Barot (2000).

where $u_{t,1}$ is the error associated with one-step ahead forecasts. Suppose that forecasts are efficiently constructed so that $E(y_{t,h} \mid I_{t,h}) = f_{t,h}$ which implies that $Cov(f_{t,1}, u_{t,1}) = 0$. Then, if the variation of the actual values is very large compared to the forecasts then this implies that the variance of $u_{t,1}$ is very large, which means that there is significant information revealed in the last month of the year.[7] But if the forecasts are not efficient and if $Cov(f_{t,1}, u_{t,1}) \neq 0$ then we have $Var(y_t) = Var(f_{t,1}) + Var(u_{t,1}) + 2Cov(f_{t,1}, u_{t,1})$ which implies that the large difference between the variation of actual and the forecasts can be due to both the variation in noise and the inefficiency $Cov(f_{t,1}, u_{t,1})$. So the large difference between the variances may occur even if the actual process is not very noisy.

Finally, it is worth noting the implications of rational expectations and implicit expectations in the graphs. As it is well known, Muth's (1961) rational expectations hypothesis requires that forecasts should be uncorrelated with the forecast errors, which also implies that the variance of the actual process should be larger than the variance of the forecasts, *i.e.* $Var(y_t) > Var(f_{t,h})$. On the other hand, implicit expectations, pioneered by Mills (1957), state that the actual realizations should be uncorrelated with the forecast errors, and the variance of actual realizations should be lower than the variance of forecasts, *i.e.* $Var(y_t) < Var(f_{t,h})$.[8] The evidence in Figure 2 clearly supports the implications of rational expectations since variances of actual realizations are larger than

[7] Large data revisions between December and June may also be responsible for high variation in $u_{t,1}$.

[8] Using this difference between the rational and implicit expectations, Muth (1985) claimed that firm production forecasts are not consistent with rational expectations hypothesis and proposed a hybrid model of expectation formation in which rational and implicit expectations are special cases. Also see Lovell (1986) for a comparison of the rational expectation and the implicit expectations hypotheses. Lahiri and Lee (1979) justify additional variance in forecasts in terms of possible errors in measurement in the survey data in a rational expectations model.

those of forecasts in majority of the cases. However, short-run forecasts of some countries, namely France, Germany, Denmark, Japan, and the UK seem to mildly violate this relation possibly due to measurement errors.

5. Information content of forecasts

Information value of a forecast is related to how accurate the forecasts are. In this section, we will provide statistics such as mean square error (MSE), mean absolute error (MAE) and Theil's *U* statistic along with another statistic recently proposed by Diebold and Kilian (2001). While MSE and MAE depend on the variability of the actual process, Theil's *U* statistic scales the RMSE by the variability of underlying data and has the advantage of being independent of the variance of the actual process. Formally,

$$U_h(y_n) = \sqrt{\sum_{t=1}^{T}\left(y_t - f_{t,h}\right)^2 / \sum_{t=1}^{T}\left(y_t - y_n\right)^2}\ , \qquad (4)$$

which compares the forecast errors with a naive forecast y_n. If U_h is larger than one, the forecast does not beat the naive forecast. An important issue in calculating the U_h is the selection of the naive forecast. While many studies have used the no-change forecast as the naive forecast, in this study, we will use the 5-year rolling GDP growth averages two years before the end of the target year as the benchmark forecast. We also tried forecasts of no change as y_n. One problem in using the lagged actual value as the benchmark in our case is that when the forecast horizon is more than 12 months, the forecasters do not know y_{t-1}. So the benchmark y_{t-1} may be considered unduly stringent. On the other hand, the benchmark y_{t-2} can be considered too lenient because the current year forecasts will have y_{t-1} known. Also, due to data revisions the lagged value may have to be changed depending on one's assumption on the forecasters' knowledge and beliefs about the latest

GDP. Given that the actual GDP growth is stationary and known, rolling averages can be considered to be a more suitable and transparent benchmark.[9]

A measure of predictability due to Diebold and Kilian (2001) is defined as[10] $p(s,k) = \left[1 - E\left(L\left(e_s\right)\right) / E\left(L\left(e_k\right)\right)\right]$, where $E\left(L\left(e_k\right)\right)$ denotes the expected loss in the long-run forecasts and $E\left(L\left(e_s\right)\right)$ denotes the expected loss in the short-run forecasts. If mean squared errors are used as the loss function, we have $p(s,k) = 1 - MSE_S / MSE_k$. Diebold and Kilian (2001) used this measure to compute the predictability of several macro variables using realized data and noted that it would be interesting to use this measure on the forecast survey data. Thus, when *k*-period ahead (*e.g.,* 24 months) survey forecast is used as the naïve forecast, $p(s,k)$ will give the improvement in forecasts as the horizon decreases. To the best of our knowledge, no study has ever used this statistic on survey data.

Table 1 presents MAE, MSE, and *U* statistics for 12-month and 24-month ahead forecasts. As expected, MAE and MSE are uniformly less for 12-month forecasts compared to the 24-month forecasts for all countries. For 24-month ahead forecasts, *U* is less than one for all countries except for Italy and Switzerland, but very marginally. For 12-month ahead forecasts, all the countries have *U* statistics less than one, suggesting that the forecasts have predictive value over the naïve forecast.[11]

---

[9] We thank the two anonymous referees for making this suggestion.

[10] This measure is also related to the *forecast content function* proposed by Galbraith (2003). In Galbraith's *forecast content function*, *MSE* of the unconditional mean forecast replaces MSE_k, which also defines the so-called *skill score* that has been used extensively in other disciplines (see, for instance, Murphy (1988)).

[11] Following Artis and Marcellino (2001) and Diebold and Mariano (1995), we also tested the statistical significance of the MSE and MAE figures against our benchmark. At the 5% significance level, the 24-month forecasts were insignificant for all countries except one, and the 12-months forecasts were significant only for seven countries. Due to the relatively small sample size, these tests might not have adequate power to reject the null of equal forecast accuracy.

Figure 3 presents Theil's $U_h(\bar{y})$ and Diebold and Kilian's $p(h,24)$ for each forecast horizon and country. Notice that large values of Theil's U imply large forecast errors. On the other hand, large values of $p(h,24)$ imply that forecasts improve considerably over the 24-month ahead forecast $f_{t,24}$. The left axes in the figures show $U_h(\bar{y})$ whereas the right axes show the values of $p(h,24)$. To pinpoint the longest horizon at which the forecasts beat the naive forecast, Figure 3 includes a vertical line through the longest horizon at which the estimated $U_h(\bar{y})$ is lower than one. This provides an easy way to compare the countries with each other. For all countries, as expected, the quality of the forecasts increases as the forecast horizon decreases. The graphs also point out certain amount of heterogeneity across countries.

When we look at the performance rankings based on $U_h(\bar{y})$, we again observe that other than Italy and Switzerland, all the country forecasts beat the naïve forecast when the forecast horizon is 24. We also observe that, as expected, for all countries the $U_h(\bar{y})$ decreases gradually as forecast horizon decreases. Even the Switzerland and Italian forecasts beat the naïve forecasts when forecast horizons are 19 and 17 months, respectively.[12]

The Diebold-Kilian measure of predictability $p(h,24)$ shows the improvement in the information content of the forecasts as measured by the decrease in MSE over that of the 24-month ahead forecasts. As shown in Figure 3, the predictive ability of GDP

[12] For the sake of comparison, we also looked at forecast efficiency with no-change as the naïve forecast. The MSE and MAE associated with 5-year rolling average of actual values as forecasts were higher than those with y_{t-1} for all countries, and significantly so for most. Thus, the information requirement of y_{t-1} as the benchmark is very stringent. As expected, we observe that it is much harder to beat the one-year lagged GDP growth as the naïve forecast. Now, for the 24-month-ahead forecasts, U statistic is less than one for only Canada, Denmark, Germany, and the U.S.; the worst performers are Portugal, Ireland, and Netherlands. For 12-month-ahead forecasts, all countries, with the exception of Ireland and Portugal, have U statistics less than one, implying that the forecasts have value over the no-change forecast.

forecasts for some countries (*e.g.*, Canada, Denmark, Finland, France, Japan and USA) does not improve over the 24-month ahead forecasts when the horizon remains relatively long, but for some other countries (*e.g.,* Germany, Ireland, Spain), each additional month increases the information content of the forecasts over the previous month even in longer-run forecasts. For most of the countries, we see that MSE substantially decreases in the short-run forecasts causing $p(h,24)$ to be close to 100% when the forecast horizon is one month. Two exceptions are Norwegian and Irish GDP growth forecasts, where the final values of $p(h,24)$ are less than 80%, and can possibly be explained by relatively inferior preliminary GDP data.

6. Timing of the most valuable information

The slopes of the plots in Figure 3, which can be interpreted as a measure of the improvement in forecast quality over horizons, is found to be somewhat different from country to country. For example, the Norwegian curve does not have a steep slope, which implies that Norwegian forecasts do not improve rapidly with decreasing horizon, but the Japanese curve has a very steep slope implying that the forecast quality increases sharply as new information is acquired.

This last point brings us to another important query: Around what horizon is the most valuable information received or, in other words, at what horizon do the forecasts improve the most? The answers to these questions are related to the slopes of $p(h,24)$ and U_h curves and are addressed in the next sections, where we provide alternative approaches to measure the content of new information at a particular horizon.

The first measure is based on forecast errors. The second measure is based only on the forecast revisions and can be seen as the content of new information as perceived by the forecasters. Following the literature cited above, another measure of forecast improvement will be constructed from the "optimum forecasts" using the time series representation of the actual quarterly GDP growth. We do this for the purpose of comparing forecast behavior with reality.

6.1. New information based on forecast errors

The first difference in the *MSE*$_h$ will provide an estimate for the new information content in forecasts when the horizon is h. From equation (2) an optimal forecast $f_{t,h}$ satisfies

$$\Delta MSE\left(f_{t,h}\right) \equiv MSE\left(f_{t,h+1}\right) - MSE\left(f_{t,h}\right) = \theta_h^2 \sigma^2 , \tag{5}$$

which is equivalent to the information content of the new information in the actual process. Now suppose that $\tilde{f}_{t,h}$ is not an optimal forecast and is generated according to

$$\tilde{f}_{t,h} \equiv E\left(y_t \mid \tilde{I}_{t,h}\right) = \tilde{\mu} + \sum_{k=h}^{q} \tilde{\theta}_k \tilde{\varepsilon}_{t,k} , \tag{6}$$

where q denotes the largest forecast horizon at which the first forecast is reported. It defines the conditional mean of the actual process when the horizon is q, *i.e.,* $\tilde{\mu} = E\left(y_t \mid \tilde{I}_{t-q}\right)$, $\tilde{\varepsilon}_{t,h}$ denotes the 'news' component utilized by the forecaster, and $\tilde{\theta}_h$ denotes the impact of this news component as perceived by the forecaster.

For convenience, let us assume that the forecasters observe the news $\varepsilon_{t,h}$ correctly but their utilization of news is not optimal, so that $\tilde{\theta}_h \neq \theta_h$ and $\tilde{\varepsilon}_{t,h} = \varepsilon_{t,h}$. From equation (6), we see that the forecast errors follow:

$$y_t - \tilde{f}_{t,h} = (\mu - \tilde{\mu}) + \sum_{k=h}^{H}\left(\theta_k - \tilde{\theta}_k\right)\varepsilon_{t,k} + \sum_{k=0}^{h-1}\theta_k \varepsilon_{t,k} , \tag{7}$$

where the first component on the RHS denotes the bias in the forecast, the second component denotes the errors due to inefficiency, and the third component denotes the errors due to unforecastable events after the forecast is reported. Calculating MSE and assuming that sample estimates converge to their population values, we get

$$MSE_h = \left(\mu - \tilde{\mu}\right)^2 + \sum_{k=h}^{H} \left(\theta_k - \tilde{\theta}_k\right)^2 \sigma^2 + \sum_{k=0}^{h-1} \theta_k^2 \sigma^2 \ . \tag{8}$$

Similarly calculating MSE_{h+1} and taking the first difference we find that $\Delta MSE_h \equiv MSE_{h+1} - MSE_h$ is

$$\Delta MSE_h = \theta_h^2 \sigma^2 - \left(\theta_h - \tilde{\theta}_h\right)^2 \sigma^2 , \tag{9}$$

which gives the improvement in the content of the forecasts with the new information. The first element on the RHS represents the maximum improvement in the quality of forecasts if the information is used efficiently, but the second component represents the mistakes in the utilization of the new information. If the usage of the most recent information $\tilde{\theta}_h$ differs from its optimal value θ_h, then the gain from the utilization of new information decreases. In the special case when $\tilde{\theta}_h = \theta_h$, the equation (9) is equivalent to equation (5). In this case, ΔMSE_h will be an estimate for the content of the new information in the actual process $\theta_h^2 \sigma^2$.

6.2. New information based on forecast revisions

While the use of ΔMSE_h provides the improvement in forecasting performance at horizon *h* and therefore gives the information content of the news in terms of forecasting ability, a similar measure can be constructed based solely on forecasts without using the actual data on GDP growth. Notice that, based on equation (2), the optimal forecast revision $r_{t,h} \equiv f_{t,h} - f_{t,h+1}$ is nothing but

$$r_{t,h} = \theta_h \varepsilon_{t,h} \qquad (10)$$

In the sub-optimum case of equation (6), we have the forecast revision process

$$r_{t,h} = \tilde{\theta}_h \varepsilon_{t,h} . \qquad (11)$$

Calculating the mean squared revisions (MSR) and taking the probability limit we get

$$MSR_h = p\lim_T \quad \frac{1}{T}\sum_{t=1}^{T} r_{t,h}^2 = \tilde{\theta}_h^2 \sigma^2 ,$$

which provides a measure for the reaction of the forecasters to news. But since forecasters react to the news based on their perception of the importance of the news, this measure can be seen as the content of the new information as perceived by the forecasters. Note the clear difference between ΔMSE_h and MSR_h. While the first one is driven by the forecast errors, the latter has nothing to do with the actual process. But both of the measures should give the same values if the survey forecasts are optimal.

The difference between MSR_h and ΔMSE_h may provide important behavioral characteristics of the forecasters such as over or under reaction to the news at a specific forecast horizon. MSR_h can be seen as a measure of how forecasters interpret the importance of news at a specific horizon, and ΔMSE_h can be seen as the "prize" they get as a result of revising their forecasts. Suppose that forecasters make large revisions at horizon h^* but the performance of the forecasts do not improve much at that horizon, then one may conjecture that the forecasters react excessively to the news. To see this more clearly, simple algebra yields

$$MSR_h - \Delta MSE_h = 2\left(\tilde{\theta}_h^2 - \theta_h \tilde{\theta}_h\right)\sigma^2,$$

which is positive when $\tilde{\theta}_h^2 > \theta_h \tilde{\theta}_h$, which is the same as the condition $|\tilde{\theta}_h| > |\theta_h|$. But $|\tilde{\theta}_h| > |\theta_h|$ is equivalent to overreaction to the news when the horizon is *h*.

6.3. Empirical comparisons

Before presenting the graphs for ΔMSE_h and MSR_h, let us try to determine their plausible shapes conceptually. As shown earlier, we expect to see forecast variability to increase as new information is accumulated. If the information content in a particular period is much larger than in the previous period, we expect to see a marked increase in forecasting performance in that period.

When the forecast horizon is very short, we expect the impact of new information on forecasts to be small for two reasons. First, the impact of a shock is determined partly by the length of time for the shock to be totally absorbed by the economy. For instance, one would expect the 9/11 terrorist attack to have affected the 2002 U.S. GDP growth more that the 2001 growth. When there is not enough time for the transmission mechanisms to impact output fully, the observed effect will be small. This implies that as horizon gets smaller the impact on GDP growth of a typical shock will be correspondingly smaller. Second, since the forecast variable is yearly real GDP growth, current year forecasts will be highly driven by the quarterly real GDP announcements and data revisions during the year. So as we approach the end of the target year, a lot of information about the target will already be known and it is expected that in the last few months the impact of the information will be very small. Consequently, we expect that the new information update will be small when the forecast horizon is short. Similar to the first reasoning above, when the horizon is long and the target is next year, the total impact is expected to be small since most of the impact will be consumed before the target year even starts. Also, forecasters may be reluctant to adjust forecasts to news immediately due to uncertainty. The uncertainty factor will tend to make the news arrival curve more concentrated towards the right. Nevertheless, these observations suggest that when the horizon is too short or too long forecast revisions due to new information is expected to be small, so that we expect the impact of shocks to peak in the middle horizons.

Figure 4 presents ΔMSE_h (dots) and MSR_h (boxes) values for our sample of 18 countries. As a very rough approximation, we also present fitted quadratic polynomials to the data for each country. The fitted lines for ΔMSE_h and MSR_h are shown in bold and dashed lines, respectively. As it is clear from the figures, both these lines (with an exception of Finland) display the expected curvature over horizons. Even though our use of a simple quadratic functional form can be questioned, the relatively low information gains at the beginning and at the closing horizons are clearly discernable. For most of the countries the peak of the quadratic line is when forecast horizon is close to 12 months and usually when the horizon is between 14 months and 10 months. The exceptions to this statement (when ΔMSE_h is considered) are Finnish and Irish forecasts. Both of these countries have experienced unusual movements in their real GDP growth rates in 1990s.

In majority of the cases, ΔMSE_h and MSR_h look similar. This implies that forecast revisions are mostly consistent with improvement in forecasting performance with decreasing horizons. Although forecast revision based MSR_h does not use any actual value and does not depend on the traditional forecast error measure (as the ΔMSE_h

graphs), peak points of the two measures still mostly match. But it may also be worthwhile to note that when the peak points do not coincide, MSR_h peaks a few periods later than ΔMSE_h. For most countries ΔMSE_h's are larger than MSR_h's at all horizons, which implies that forecast revisions are sticky, and forecasters stagger their reactions to news.[13]

Given the regularity across countries, and in order to estimate the information arrival curve without imposing any functional from, we pooled all the countries and estimated the ΔMSE_h and MSR_h curves non-parametrically, see Hastie and Tibshirani (1986).[14] These are reported in Figure 5 where we now clearly see that the peaks with both the curves come at about 14 months before the end of the target year. The shapes of the two curves when estimated non-parametrically are remarkably similar, even though MSR_h exhibits a lot more adjustment towards the end of the forecasting period than ΔMSE_h. Thus, the 'currency' rather than the 'time remaining' for a unit perceived shock to affect the economy weighs more heavily in forecast revisions. It seems that the imposition of a simple quadratic polynomial on individual countries with outliers might have shifted the peaks a little to the right for some countries.

The finding that for most of the countries, the biggest adjustment of forecasts to news happens when the horizon is close to 14 months on the average and, depending on the country, it can also vary any where between September of the previous year and

[13] Isiklar *et al.* (2006) estimate the extent of stickiness for G7 countries using the same data source, but a different methodology. See Mankiw and Reis (2001) and Sims (2003) for alternative explanations.

[14] These functions were estimated based on the Spline Smoother with 3 degrees of freedom. The minimization problem was solved using Back-fitting and Local Scoring algorithms. Note that in addition to the Spline Smoother, we also used Kernel Smoother and Local Regression procedures. The results were practically the same. Allowing for fixed country effects also did not change the estimated functions. We used PROC GAM in SAS to do the calculations

February of the current year (*i.e.,* horizons 15 to 10 months) begs another interesting question: What is the source of information that is revealed during this time? Clearly official GDP announcements have to be one of the main information sources that will have a large impact on the forecasting performance. The initial GDP figures for the previous year, which is released as early as late January for some countries and in February for the majority of the countries, can provide important information about the current year GDP growth. But what we find is that by October of the previous year, forecasters begin the make major revisions to next year forecasts, and this is done in part by forming firm ideas about current year's GDP growth. Therefore for most of the countries the improvement in the forecasts cannot be attributed to the release of the first quarter GDP. In fact, the first quarter GDP figures for the target year are not released during the first quarter of the current year. The U.S. and U.K. GDP figures for the first quarter are released during the last week of April and in most other countries GDP figures are released in the second half of May. Thus, it is clear that forecasters extract information from relevant monthly indicators such as employment, industrial production, manufacturing index, etc. that are correlated with GDP. In addition, various leading indicators (*e.g.,* stock market index, interest rate spreads, building permits, unemployment insurance claims, etc.) which have predictive power up to a year and which are available more promptly act as valuable information sources.

But we should note that there is a difference in the nature of information provided by monthly indicators during September of the previous year and February of the current year, and the GDP data releases for the previous year. The previous year's GDP figures increase the information content of the forecasts for several reasons. First, information of the previous year's GDP level determines the base of the GDP growth forecast for the following year, which may have substantial effect on the current year forecasts. In addition, if GDP growth has a large serial correlation and forecasters employ extrapolative expectations to capture this, the release of last year's GDP figure may initiate large revisions and increase the forecasting performance. It is also possible that the previous year's GDP growth will have a large impact on the forecasting performance if the forecasters can learn from their previous mistakes and employ an error correction model to revise their expectations.

6.4. Content of new information implied in the actual process

In this section we provide a measure of new information in an "optimal" forecast, which is based on modeling the actual process. The content of the new information in the actual process can be calculated by estimating equation (1) using the actual quarterly real GDP growth data. For example, one may think of fitting an MA model on the real GDP growth series, and then treating the estimated MA coefficients as estimates for θ_h coefficients. As pointed before, this approach is the main idea behind several studies in the calculations of the information content of optimal forecasts, *e.g.,* Öller (1985).

But, in this study the forecasts are what are called "fixed-target" forecasts. So the target variable represented by y_t is not quarterly real GDP growth but annual real GDP growth. This implies that we have to make a transformation on the MA coefficients estimated using the quarterly real GDP growth series to be comparable with the annual real GDP growth forecasts.

Suppose y_t denotes the annual real GDP growth as before and $\tilde{y}_{t,q}$ denotes the annualized quarterly real GDP growth q quarters before the end of the year t. For example, $\tilde{y}_{t,1}$ is the GDP growth rate in the last quarter of year t and $\tilde{y}_{t,4}$ is the GDP growth rate in the first quarter of the year t. Note that in this notation we have $\tilde{y}_{t,k} = \tilde{y}_{t-1,k+4}$.

Then by definition:

$$y_t = \frac{1}{4}\sum_{k=1}^{4} \tilde{y}_{t,k} \,. \tag{12}$$

Now suppose that $\tilde{y}_{t,q}$ has the following $MA(\infty)$ representation:

$$\tilde{y}_{t,q} = \gamma_0 \varepsilon_{t,q} + \gamma_1 \varepsilon_{t,q+1} + \gamma_2 \varepsilon_{t,q+2} + \ldots + \gamma_k \varepsilon_{t,q+8} + \ldots \tag{13}$$

Then substituting this MA process in the equation (12) gives the MA representation for the actual process:

$$y_t = \frac{1}{4}[\gamma_0 \varepsilon_{t,1} + (\gamma_0 + \gamma_1)\varepsilon_{t,2} + (\gamma_0 + \gamma_1 + \gamma_2)\varepsilon_{t,3} + (\gamma_0 + \gamma_1 + \gamma_2 + \gamma_3)\varepsilon_{t,4} + (\gamma_1 + \gamma_2 + \gamma_3 + \gamma_4)\varepsilon_{t,5}$$
$$+ (\gamma_2 + \gamma_3 + \gamma_4 + \gamma_5)\varepsilon_{t,6} + (\gamma_3 + \gamma_4 + \gamma_5 + \gamma_6)\varepsilon_{t,7} + \cdots]$$

More specifically, the MA form can be represented as follows:

$$y_t = \sum_{k=0}^{\infty} \delta_k \varepsilon_{t,k+1} \tag{14}$$

where

$$\delta_k = \sum_{i=\max(k-3,0)}^{k} \gamma_i \,. \tag{15}$$

The intuition behind this representation should be straightforward. While last quarter shocks have only a unique chance of having an impact on the annual GDP growth, third quarter shocks will have the chance two times: contemporaneous effect on the third quarter GDP (*via* γ_0) and then a secondary effect on the last quarter GDP growth (*via* γ_1). Similarly, first quarter shocks will have four impact coefficients. When the horizon is larger than 4 quarters, the shocks will have 4 chances to have an effect on the current year GDP growth. But in this case, there will not be any contemporaneous impact since the effect will be seen on the previous year's GDP growth.

To estimate the γ_k s, we use the seasonally adjusted quarterly real GDP growth and estimate ARMA models for each country. The AIC criterion was used to select the order of AR and MA terms. Then we transformed this ARMA models into a $MA(\infty)$ representation, which gives us γ_k s. This is a 'safe' way to get a reasonable MA representation. Direct MA modeling is an alternative way but the models may not converge under certain conditions. After getting the MA coefficients of the quarterly model, we construct the MA coefficients of the annual model using equation (15) and then calculate the optimal percentage of variation at horizon k as $100 \times \delta_k^2 / \sum_{i=1}^{8} \delta_i^2$.

To be comparable with the survey forecast data we use the 1990-2001 period to estimate ARMA models for each country. Note that the longest horizon we are interested is two years and we use quarterly GDP growth rates to estimate the γ_k s, so we have only eight observations to plot for each country.[15] For the sake of brevity and in order to see a stable "flow of information" curve, Figure 6 presents the percentage shares of 18 countries aggregated and separately only for the UK and the US by horizon. Since we use quarterly data to generate the shares, we plot each quarter's value in the center of the

[15] There were outliers in the data too: Germany on 1991:1 (8.32%), Portugal on 1988:1 (15.6%), and Norway on 1997:2 (6.9%). In addition, the GDP growth rates of Spain behave abnormally during 2000:2 to 2001:1 having growth rates of 3.8, -3.04, 5.3 and -2.2 percent respectively. With these data points, the model failed to converge so we used the data until 2000:1 for Spain. Except for the Spanish case, the results were, however, not affected by the control of the outliers.

quarter. So, for example, the first estimated share of a contemporaneous shock is plotted when the horizon is 2 months. Despite small samples, the results clearly suggest an asymmetric hump-shaped adjustment with a peak at 11 months horizon. With a few exceptions, a similar pattern was found for each of the sample countries.[16]

We do not expect these graphs based on optimal forecasts from the actual data generating process to be exactly the same as those based on survey forecasts. First, because the flow of information curve with optimal forecasts was estimated using quarterly data, the curve will be slightly shifted to the right. Moreover, as we have pointed out before, forecast errors in real life are driven not only by the level of "well behaved" noise attributed to the data generating process based on revised data, but also by numerous other factors. For example, randomness and systematic features in data revisions, structural breaks, model misspecifications, outliers, etc. that are only detectable *ex post*, are some of the factors that may affect the information content of real-time forecasts. These factors do not affect the *ex post* analysis of historical data of the target variable, but may induce significant noise in the information content of real-time survey forecasts. The observed peak with survey expectations arrives a little earlier (Fig. 5) than the peak using optimal forecasts from a fitted time series model (Fig. 6) possibly because survey forecasts freely absorb information from other indicators in forecasting GDP growth that the time series model cannot. This is an important intuition about the value of survey forecasts that we glean by comparing survey forecasts with those from time series models.

[16] These countries are: Austria, Denmark, Japan, Norway, and Portugal. The anomaly for these countries can possibly be explained by the estimated AR coefficients due to small samples and their robustness.

7. Conclusions

In this paper we study the characteristics of monthly GDP growth forecasts for 18 developed countries during 1989-2004. We study how forecasting performance improves as the forecast horizon decreases, and at what horizons forecasts start to become informative. Since there are many forecasting organizations around the world providing forecasts for many macroeconomic variables with horizons up to two years or more, it is useful to explore the value of these forecasts, and thereby understand the limits to how far ahead today's professionals can reasonably forecast. Since the panel of forecasters in *Consensus Economics*, *Inc.* are all private market agents, the limits to forecasting that these specialists exhibit can safely be taken as indicative of the current state of economic foresight. However, the answer from our exhaustive data analysis did not turn out to be a "single-liner". We have found wide diversity in the quality of the forecasts across countries, and the horizons at which forecasts start becoming useful, possibly reflecting the forecast difficulty of the underlying series.

We used Theil's U statistic with the 5-year rolling GDP growth as the benchmark, and another measure of predictability recently suggested by Diebold and Kilian (2001) with the 24-month ahead forecast as the benchmark. For 24-month ahead forecasts, U is less than one for all countries except for Italy and Switzerland, but very marginally. For 12-month ahead forecasts, all the countries have U statistics less than one, implying that the forecasts have predictive value. Using the Diebold-Kilian skill measure and the variance functions, we however found that for majority of the countries the longer-term forecasts for up to 18 months are no better than the initial 24-month ahead forecasts. That is, over these longer horizons, forecasters do not receive dependable information to adjust their forecasts. We also observed a similar pattern when we looked at the horizons at which the survey forecasts beat the naïve no-change forecast. These findings imply that the survey forecasts do not have much value when the horizon goes beyond 18 months.

In this paper we have proposed two alternative approaches to measure the flow content of new information in survey forecasts. The first measure is based on the improvement in actual forecasting performance over horizons, and the second measure is based on forecast revisions that can be considered as a measure of the importance of new information as perceived by forecasters. Whereas the latter can be interpreted as a measure of how forecasters interpret the importance of news in real time, the former is the *ex post* “prize” they get as a result of revising their forecasts. Under rationality and without much unforeseen errors in the sample period, these two approaches should yield similar results. Using nonparametric methods, we found that both the approaches indicate the largest improvement in forecasting performance comes when the forecast horizon is around 14 months.

Acknowledgements:

We thank the participants of the Leipzig conference on the “Future of Forecasting” for many helpful and encouraging comments. In particular, comments from Roy Batchelor, Anthony Davies, Ullrich Heilemann, Fushang Liu, Lars-Erik Öller, Zulkarnain Pulungan, Franz Seitz, Xuguang Sheng, Herman Stekler, Kenneth Wallis and two anonymous referees are gratefully acknowledged. However, we are responsible for any errors and omissions.

References

Artis, M., & Marcellino, M. (2001). Fiscal forecasting: The track record of the IMF, OECD and EC. *Econometrics Journal*, 4, S20-S36.

Artis, M., & Zhang, W. (1997). International business cycles and the ERM: Is there a European business cycle? *International Journal of Finance and Economics*, 2, 1-16.

Batchelor, R. A. (2001). How useful are the forecasts of intergovernmental agencies? The IMF and OECD versus the consensus. *The Review of Economics and Statistics,* 33, 225-35.

De Gooijer, J. G., & Klein, A. (1992). On the cumulated multi-step ahead predictions of vector autoregressive moving average processes. *International Journal of Forecasting,* 7, 501-513.

Diebold, F. X., & Kilian, L. (2001). Measuring predictability: theory and macroeconomic applications. *Journal of Applied Econometrics*, 16, 657-669.

Diebold, F.X., & Mariano, R.S. (1885). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13, 256-263.

Frenkel, J. A. (1975). Inflation and the formation of expectations. *Journal of Monetary Economics,* 1, 403-433.

Galbraith, J. W. (2003). Content horizons for univariate time-series forecasts. *International Journal of Forecasting,* 19, 43-55.

Gallo, G.M., Granger, C. W. J., & Joen, Y. (2002). Copycats and common swings: The impact of the use of the forecasts in information sets. *IMF staff Papers*, 49, 1, 4-21.

Granger, C. W. J. (1996). Can we improve the perceived quality of economic forecasts? *Journal of Applied Econometrics,* 11, 455-473.

Hastie, T., & Tibshirani, R. (1986). Generalized additive models. *Statistical Science,* 1, 297-318.

Harvey, D.I., Leybourne, S.J., & Newbold, P. (2001). Analysis of a panel of UK macroeconomic forecasts. *Econometrics Journal,* 4, 37-55.

Isiklar, G., Lahiri, K., & Loungani, P. (2006). How quickly do forecasters incorporate news? Evidence from cross-country surveys. *Journal of Applied Econometrics*, 21, 713-737.

Juhn, G. & Loungani, P. (2002). Further cross-country evidence on the accuracy of the private sector output forecasts. *IMF Staff Papers*, 49, 49-64.

Lahiri, K., & Lee, J.S. (1979). Tests of rational expectations and the Fisher effect. *Southern Economic Journal*, 46, 413-424.

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.

Lovell, M. C. (1986). Tests of the rational expectations hypothesis. *American Economic Review,* 76, 110-124.

Mankiw, N.G. & M.D. Shapiro. (1986). News or noise? An analysis of GNP revisions. *Survey of Current Business*, 66, 20-25.

Mankiw, N. G. & Reis, R. (2002). Sticky information versus sticky prices: A proposal to replace the new Keynesian Phillips curve. *Quarterly Journal of Economics,* 117, 1295-1328.

Mills, E.S. (1957). The theory of inventory decisions. *Econometrica,* 25, 222-238.

Mills, T. C. & Pepper, G. T. (1999). Assessing the forecasters: an analysis of the forecasting records of the treasury, the London Business School and the National Institute. *International Journal of Forecasting,* 15, 247-257.

Mincer, J.A. (1969). Models of adaptive forecasting. In J. Mincer (Ed.), *Economic forecasts and expectations: Analysis of forecasting behavior and performance*, Vol. 6, NBER, New York, 83-111.

Murphy, A. H. (1988). Skill scores based on the mean square error and their relationships to the correlation coefficient. *Monthly Weather Review*, 116, 2417-2424.

Muth, J. F. (1961). Rational expectations and the theory of price movements. *Econometrica,* 29, 315-335.

Muth, J. F. (1985). Properties of some short-run business forecasts. *Eastern Economic Journal,* 11, 200-210.

Oke, T. & L.-E. Öller (1999). Testing for short memory in a VARMA process. *Journal of Forecasting,* 18, 477-487.

Öller, L.-E. (1985). How far can changes in business activity be forecasted? *International Journal of Forecasting,* 1, 135-141.

Parzen, E. (1982). ARARMA models for time series analysis and forecasting. *Journal of Forecasting,* 1, 67-82.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics,* 50, 665-690.

Vuchelen, J., & Gutierrez, M.-I. (2005). A direct test of the information content of the OECD growth forecasts. *International Journal of Forecasting,* 21, 103-117.

Table 1

MAE, RMSE, and Theil's *U*

	12-month ahead forecasts			24-month ahead forecasts		
Country	MAE	RMSE	Theil's *U*	MAE	RMSE	Theil's *U*
Austria	0.98	1.16	0.69	1.24	1.48	0.94
Belgium	0.99	1.15	0.62	1.28	1.68	0.89
Canada	1.21	1.36	0.58	1.44	1.7	0.77
Denmark	0.72	0.99	0.65	0.96	1.14	0.73
Finland	2.24	2.89	0.59	2.7	3.37	0.68
France	0.79	0.99	0.58	1.15	1.5	0.84
Germany	0.79	1.03	0.37	1.49	1.96	0.67
Ireland	2.35	2.76	0.75	2.98	3.67	0.96
Italy	0.77	0.87	0.62	1.39	1.61	1.12
Japan	1.41	1.58	0.64	1.9	2.3	0.91
Netherlands	0.89	1.06	0.45	1.38	1.72	0.71
Norway	0.92	1.13	0.58	1.14	1.33	0.68
Portugal	0.98	1.31	0.43	1.4	1.89	0.60
Spain	0.61	0.86	0.40	1.18	1.58	0.70
Sweden	0.9	1.13	0.47	1.46	1.84	0.76
Switzerland	1.22	1.45	0.75	1.71	2.04	1.02
UK	0.77	1.02	0.43	1.08	1.62	0.71
USA	0.96	1.09	0.57	1.28	1.59	0.92

U statistic uses 5 year rolling average GDP growth as the naïve forecast.

Fig. 1. Evolution of fixed-target forecasts over horizons

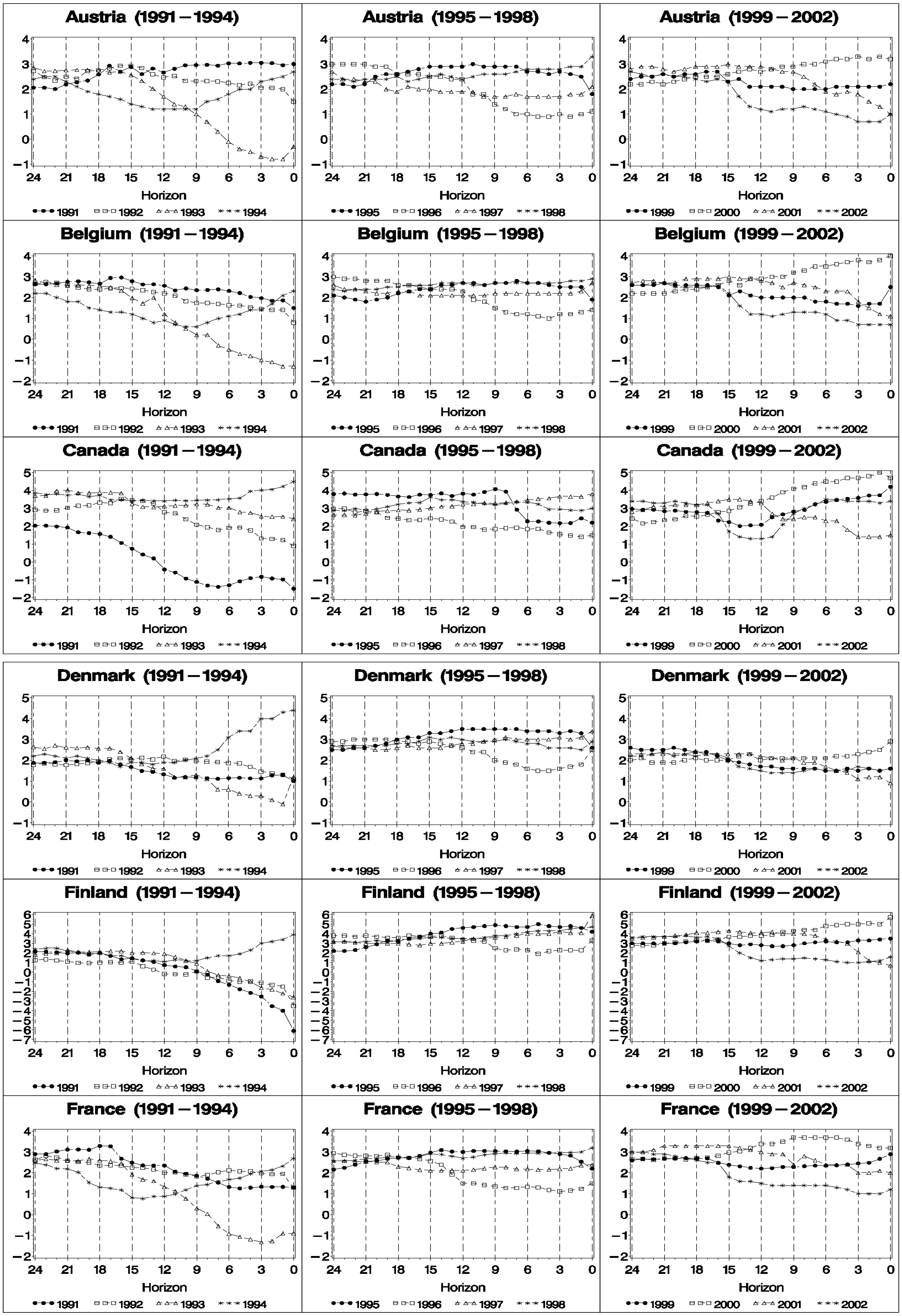
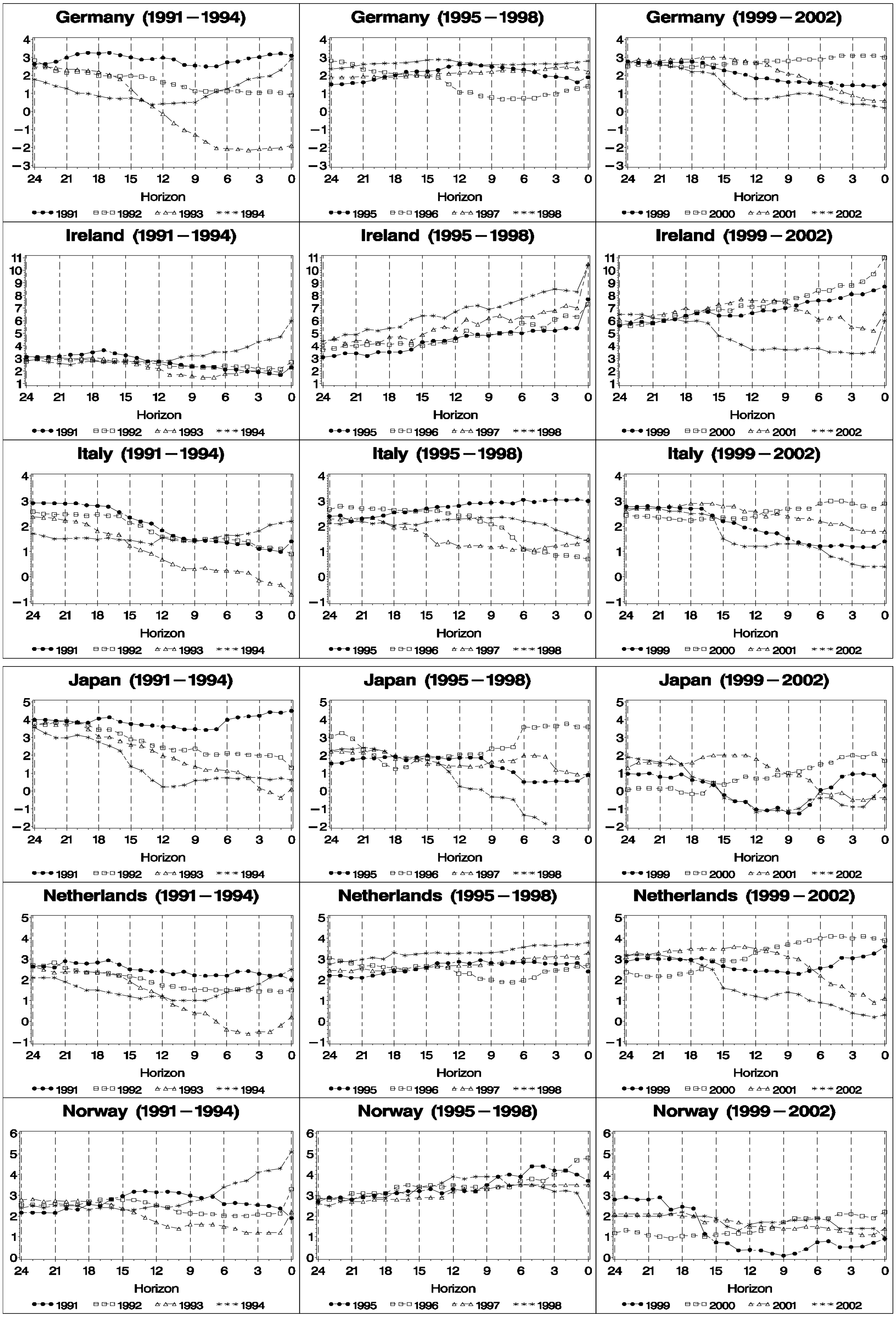
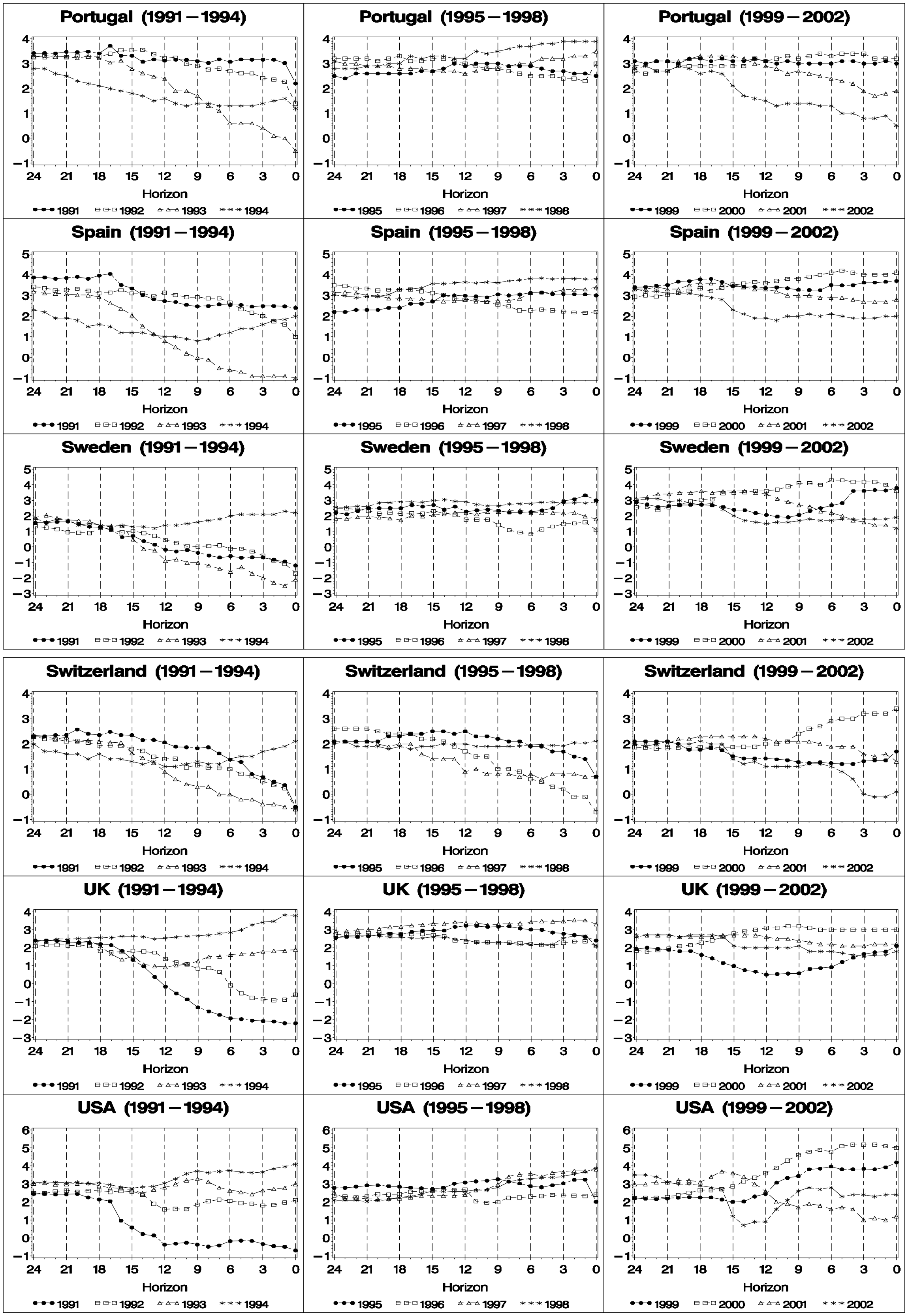


Fig. 1. Cont.



Germany (1991–1994)
Horizon
1991 1992 1993 1994
Germany (1995–1998)
Horizon
1995 1996 1997 1998
Germany (1999–2002)
Horizon
1999 2000 2001 2002
Ireland (1991–1994)
Horizon
1991 1992 1993 1994
Ireland (1995–1998)
Horizon
1995 1996 1997 1998
Ireland (1999–2002)
Horizon
1999 2000 2001 2002
Italy (1991–1994)
Horizon
1991 1992 1993 1994
Italy (1995–1998)
Horizon
1995 1996 1997 1998
Italy (1999–2002)
Horizon
1999 2000 2001 2002
Japan (1991–1994)
Horizon
1991 1992 1993 1994
Japan (1995–1998)
Horizon
1995 1996 1997 1998
Japan (1999–2002)
Horizon
1999 2000 2001 2002
Netherlands (1991–1994)
Horizon
1991 1992 1993 1994
Netherlands (1995–1998)
Horizon
1995 1996 1997 1998
Netherlands (1999–2002)
Horizon
1999 2000 2001 2002
Norway (1991–1994)
Horizon
1991 1992 1993 1994
Norway (1995–1998)
Horizon
1995 1996 1997 1998
Norway (1999–2002)
Horizon
1999 2000 2001 2002

Fig. 1. Cont.



Portugal (1991 – 1994)
Portugal (1995 – 1998)
Portugal (1999 – 2002)
Spain (1991 – 1994)
Spain (1995 – 1998)
Spain (1999 – 2002)
Sweden (1991 – 1994)
Sweden (1995 – 1998)
Sweden (1999 – 2002)
Switzerland (1991 – 1994)
Switzerland (1995 – 1998)
Switzerland (1999 – 2002)
UK (1991 – 1994)
UK (1995 – 1998)
UK (1999 – 2002)
USA (1991 – 1994)
USA (1995 – 1998)
USA (1999 – 2002)
Horizon
1991 1992 1993 1994
1995 1996 1997 1998
1999 2000 2001 2002

Fig. 2. Forecast variance over forecast horizons, 1989:10 – 2004:06

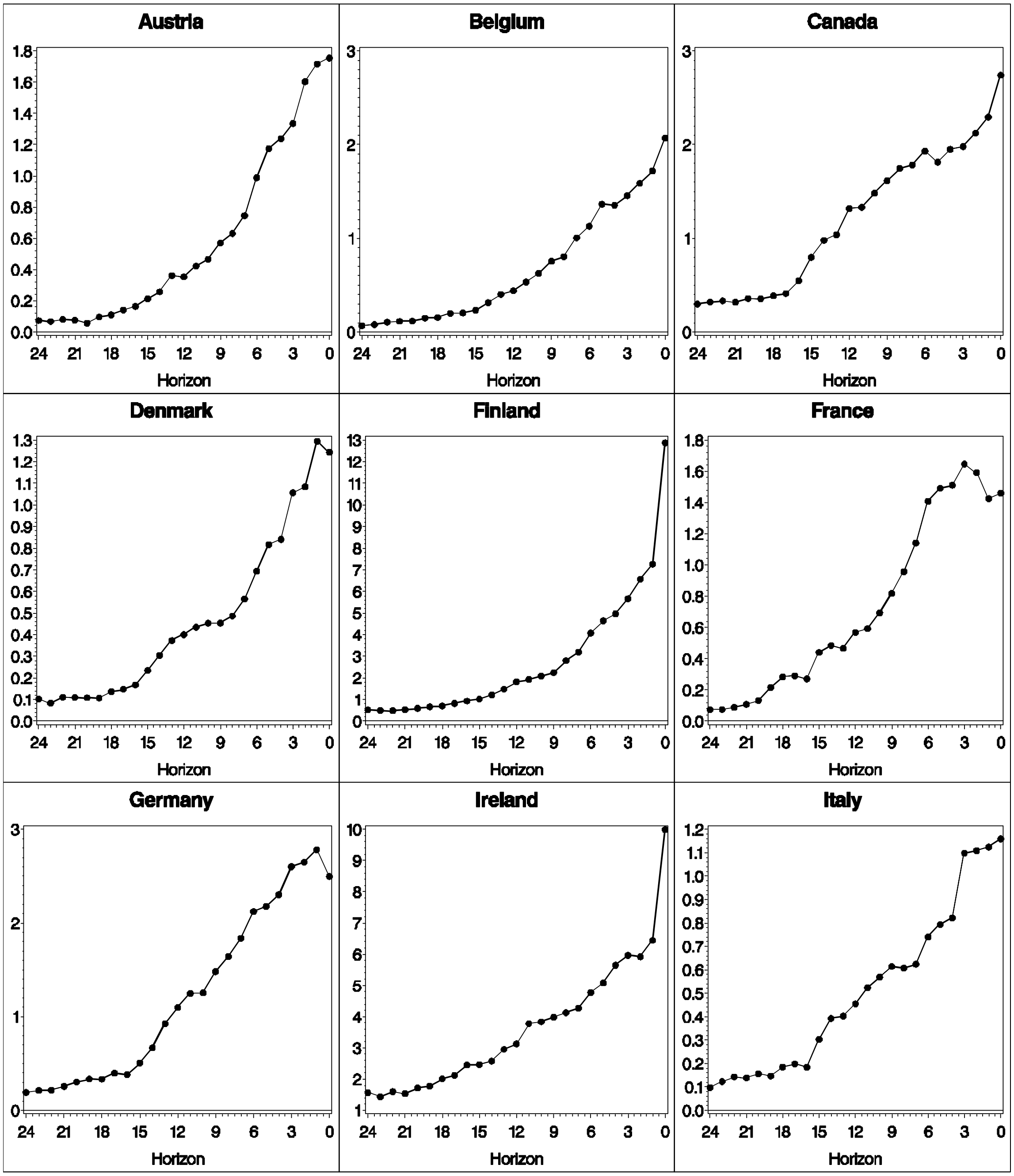
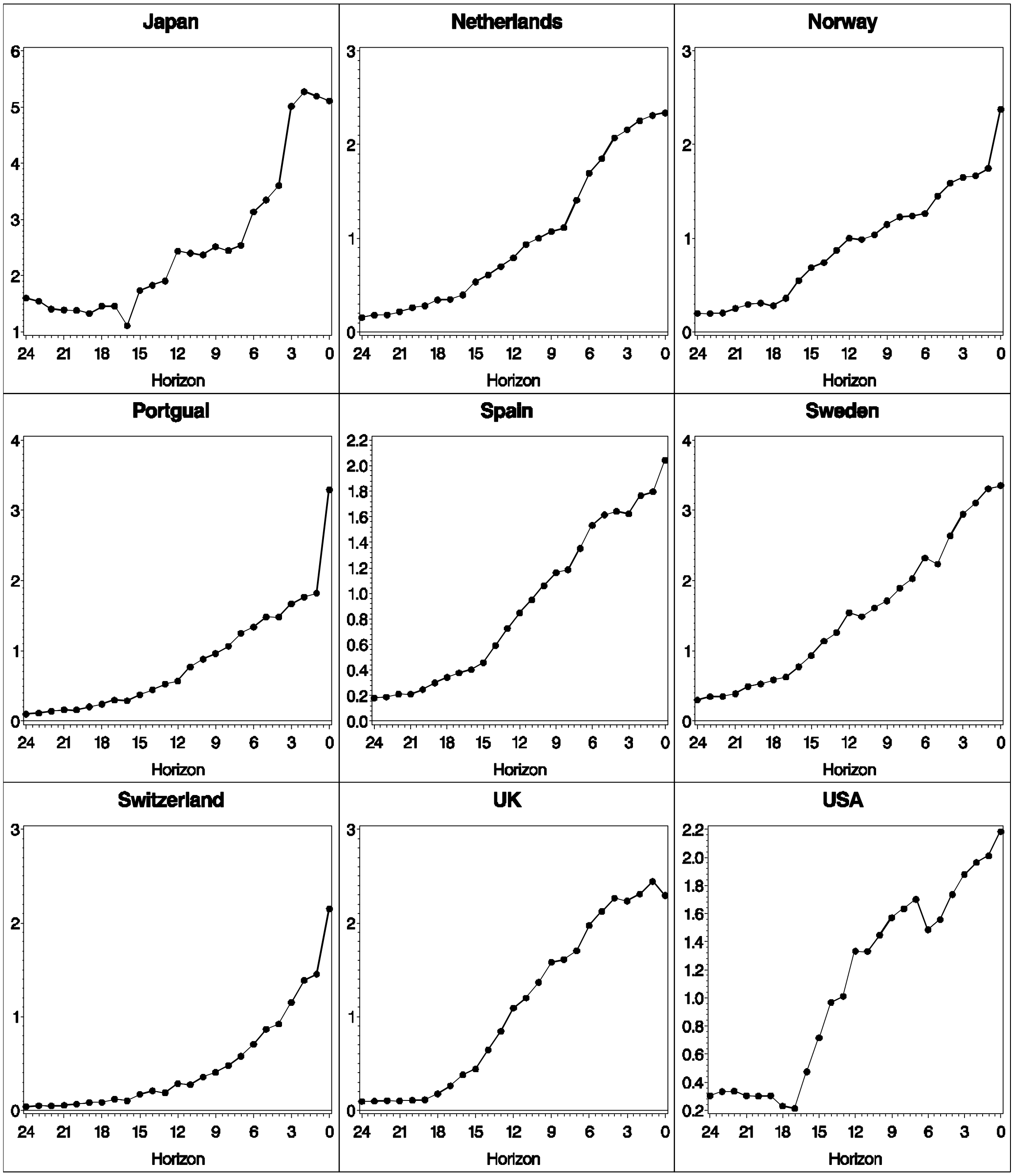


Fig. 2. Cont.



Japan
Netherlands
Norway
Portgual
Spain
Sweden
Switzerland
UK
USA
Horizon

Fig. 3. Information content of forecasts over horizons (*U* statistic uses 5 year rolling average GDP growth as the naïve forecast)

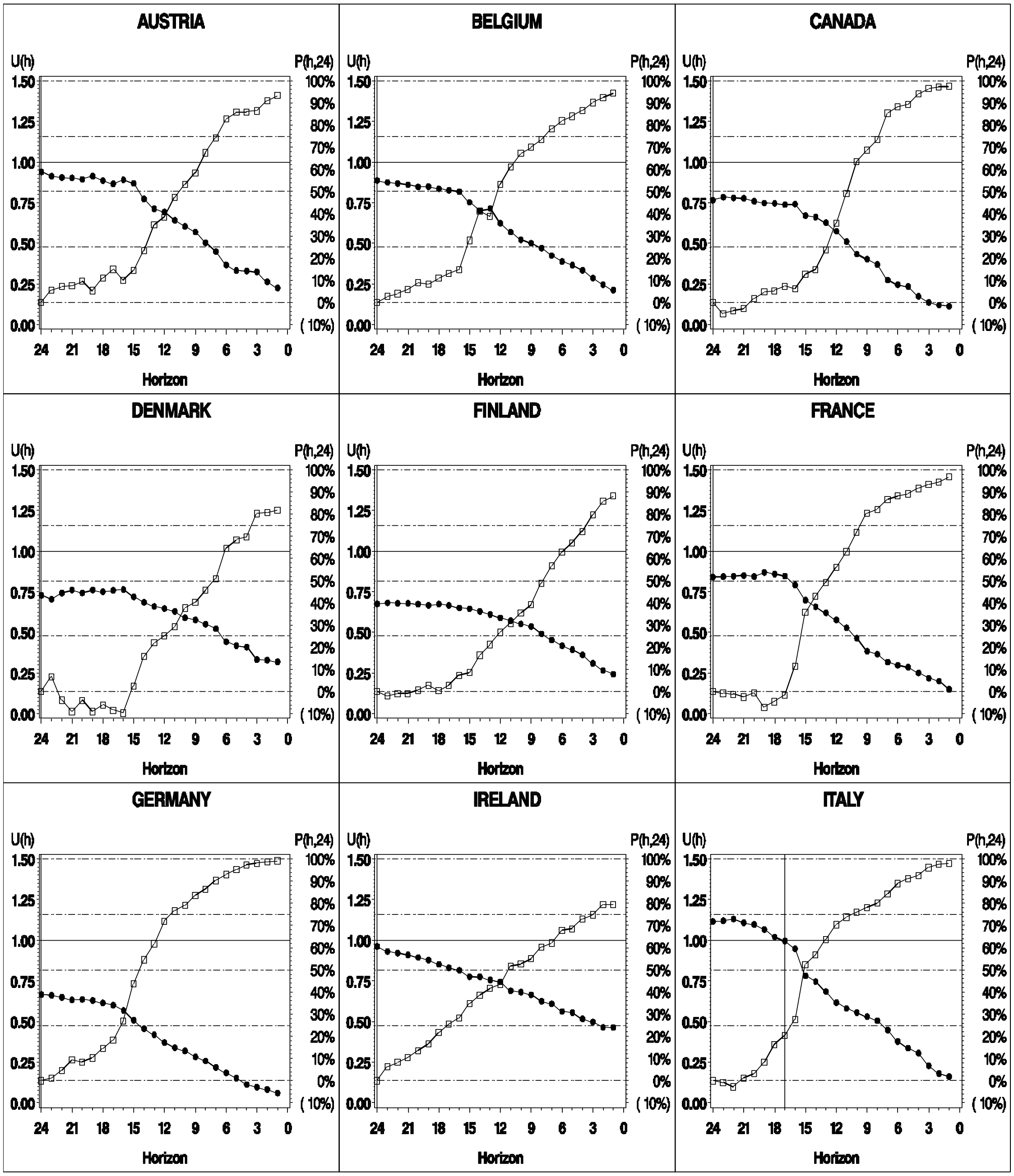


Fig. 3a. Cont.

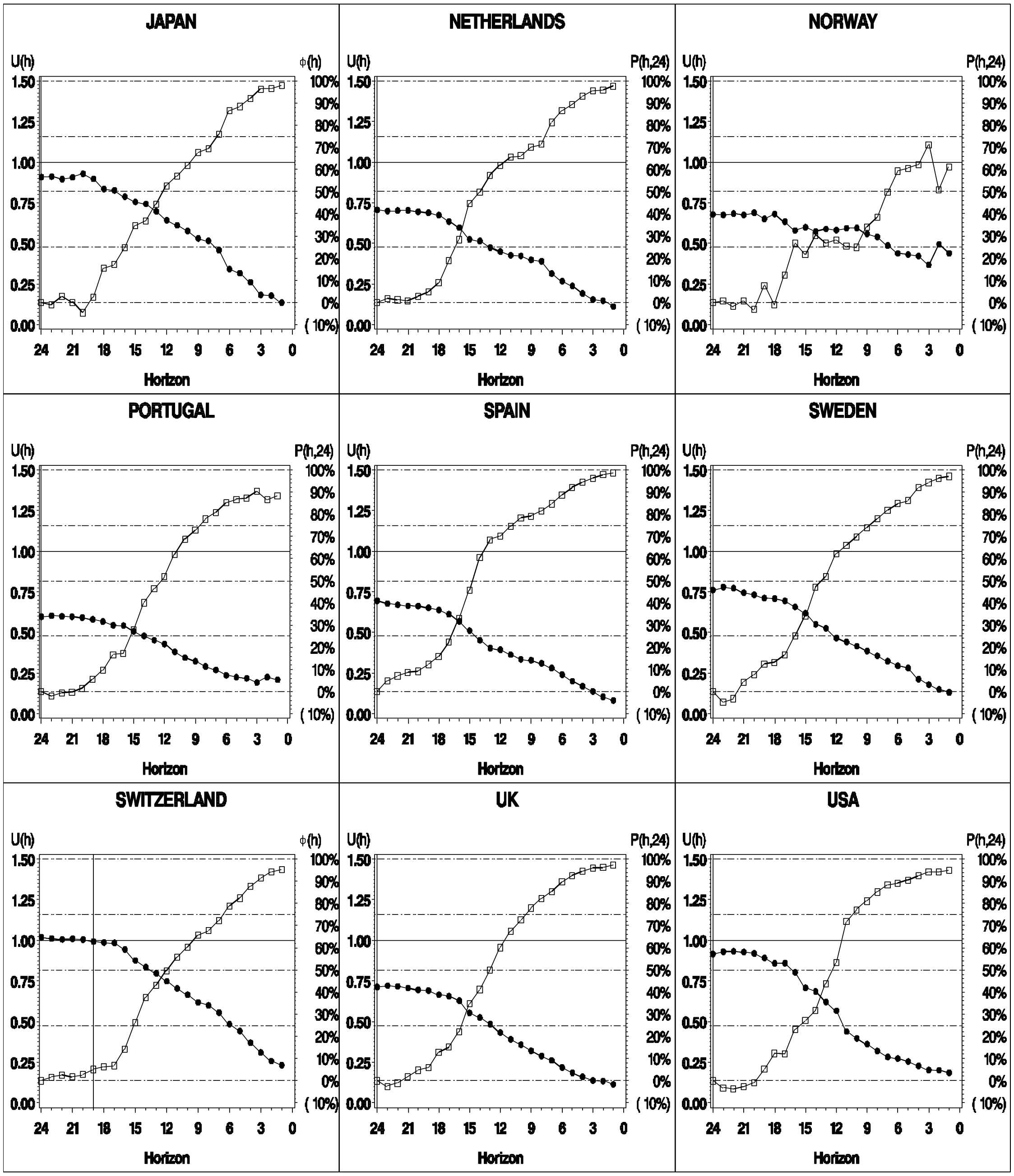

JAPAN
NETHERLANDS
NORWAY
PORTUGAL
SPAIN
SWEDEN
SWITZERLAND
UK
USA
U(h)
φ(h)
P(h,24)
1.50
1.25
1.00
0.75
0.50
0.25
0.00
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
(10%)
24 21 18 15 12 9 6 3 0
Horizon

Fig. 4. Flow of information arrival over horizons, 1989:10 – 2004:06

ΔMSE = *dots*; Mean Square Revision (MSR) = boxes

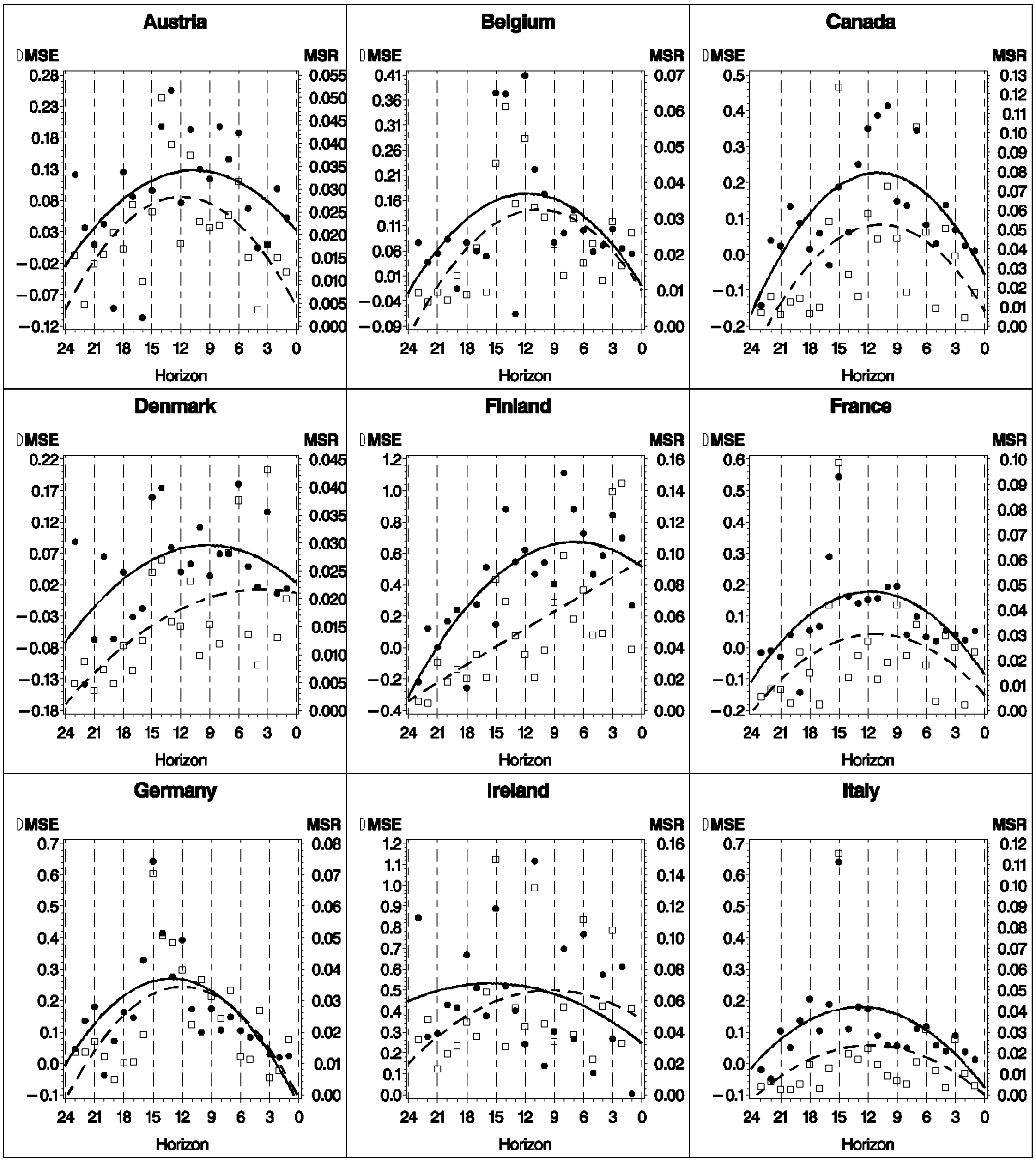


Fig. 4. Cont.

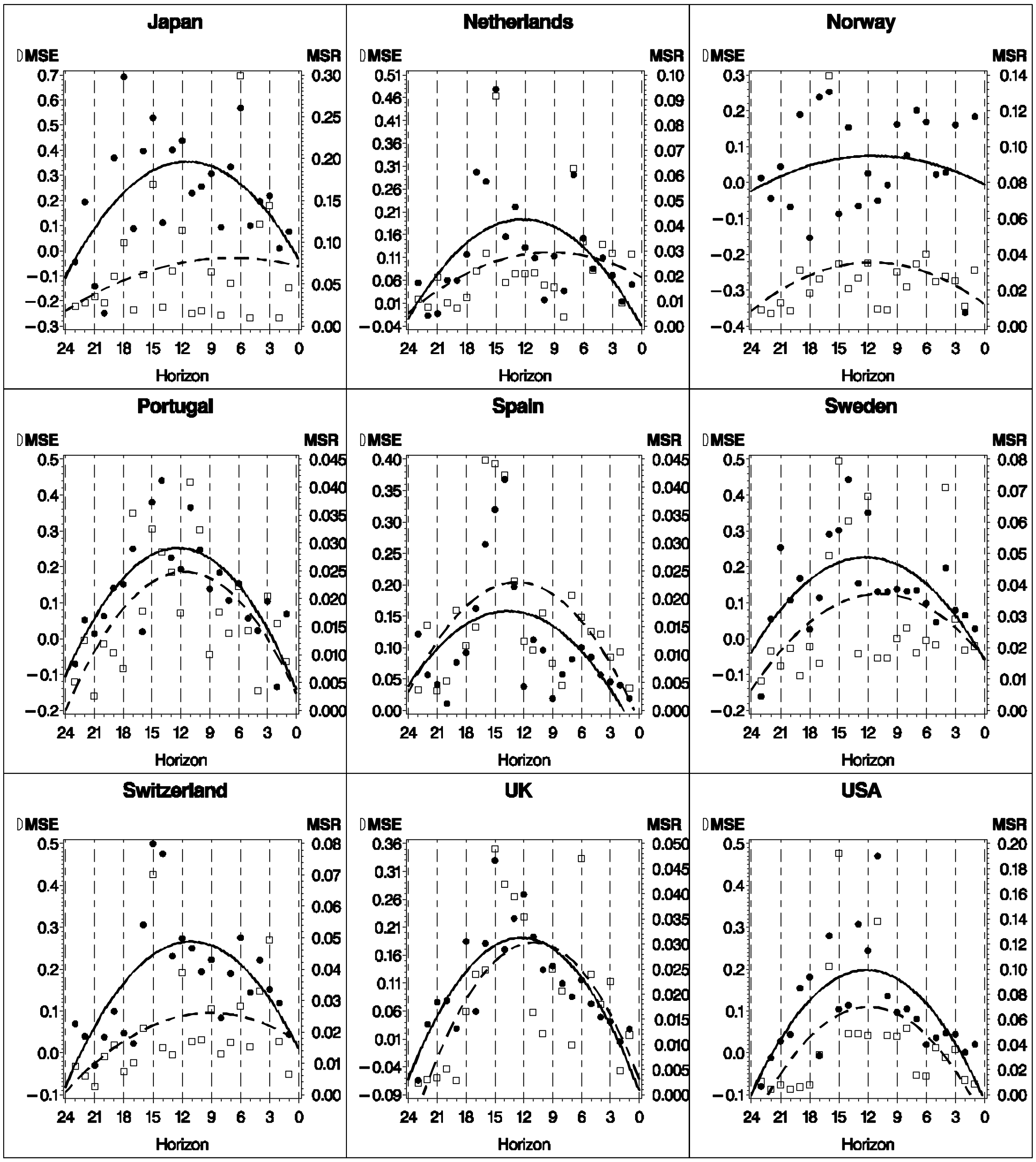

Japan
Netherlands
Norway
Portugal
Spain
Sweden
Switzerland
UK
USA
DMSE
MSR
Horizon

Fig. 5. Nonparametric Information Arrival Curve - All Countries Pooled

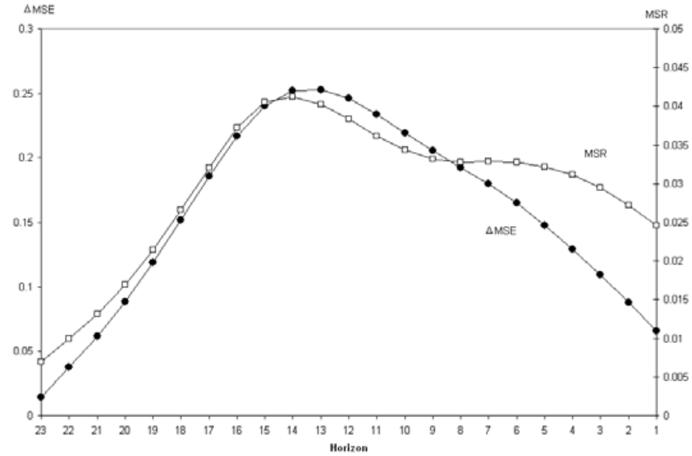


Fig. 6. Flow of information arrival based on ARMA model of GDP growth