Macroeconomic Forecasting with Real-Time Data: An Empirical Comparison

Christiaan Heij[*], Dick van Dijk, Patrick J.F. Groenen

Econometric Institute
Erasmus University Rotterdam

Econometric Institute Report EI2009-27

Abstract

Macroeconomic forecasting is not an easy task, in particular if future growth rates are forecasted in real time. This paper compares various methods to predict the growth rate of US Industrial Production (IP) and of the Composite Coincident Index (CCI) of the Conference Board, over the coming month, quarter, and half year. It turns out that future IP growth rates can be forecasted in real time from ten leading indicators, by means of the Composite Leading Index (CLI) or, even somewhat better, by principal components regression. This amends earlier negative findings for IP by Diebold and Rudebusch. For CCI, on the other hand, simple autoregressive models do not provide significantly less accurate forecasts than single-equation and bivariate vector autoregressive models with the CLI. This amends some of the more positive results for CCI recently reported by the Conference Board. Not surprisingly, all forecast methods improve considerably if 'ex post' data are used, after possible data updates and revisions.

Keywords: vintage data, leading indicators, forecast evaluation, recessions, industrial production, composite coincident index

[*]Corresponding author, Econometric Institute, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands, email: heij@ese.eur.nl, fax: +31-10-4089162, tel: +31-10-4081269.

1 INTRODUCTION

When making decisions, it helps to know as much as possible of the future, and forecasts are intended to serve this purpose. Of course, to support decisions now, such forecasts can use only current and past data. This is known as 'real-time' forecasting, as the forecasts are based on data that are available at the real decision moment. This obvious requirement poses some restrictions in macroeconomic forecasting, as the final values of many key economic variables, like monthly production and income, become available only after some delay. It takes some time to collect the relevant information, but nowadays this causes a delay in the US of less than a month. Data revisions pose a much more serious problem, as these revisions may be substantial and they may still take place after several months and even years after initial releases. Therefore, many macroeconomic data relating to recent months are only provisional, and it is precisely these data that are of crucial importance to forecast the near future. This type of data uncertainty provides an extra challenge in real-time forecasting.

Diebold and Rudebush (1991) are among the first who studied real-time forecasting. Their main findings, based on monthly vintage data for the seventies and eighties of the last century, are the following. In order to forecast monthly Industrial Production (IP), the Composite Leading Index (CLI) of the Conference Board helps substantially if 'ex post' data are used. However, the usefulness of CLI evaporates completely in real-time forecasting, and in many cases a simple autoregressive (AR) model involving only lags of IP performs best. This finding illustrates that practically relevant real-time results may differ widely from results obtained within unrealistic scenario's employing final release data.

The main motivation of this paper is to reconsider the questions posed by Diebold and Rudebush now, two decades later, with IP and CLI data over the last twenty years. A related question concerns the best way to extract the predictive information available from leading indicators. The CLI is a much used index in forecasting, but it is of interest to consider alternative index methods, such as principal components.

Real-time forecasting is also discussed, for instance, by Swanson, Ghysels and Callan (1999), Croushore and Stark (2001), Pesaran and Timmermann (2005), and Clements and Galvao (2006), see also the survey paper of Croushore (2006). Recently, the Conference Board has reported positive results for real-time forecasts of the Composite Coincident Index (CCI), as the CLI provides significant forecast gains over the AR benchmark model, see McGuckin, Ozyildirim, and Zarnowitz (2007). These and most other real-time forecast studies are concerned with the US, but some recent applications deal with the euro area, for instance, Golinelli and Parigi (2008) and Ozyildirim, Schaitkin, and Zarnowitz (2009).

This paper reconsiders the (negative) real-time forecast findings of Diebold and Rudebush (1991) for IP and the (positive) findings of McGuckin, Ozyildirim, and Zarnowitz (2007) for CCI. Our results for vintage data of the most recent two decades do actually turn out to be quite the opposite of what was found in earlier studies. That is, CLI provides significant gains in IP forecasts for one, three, and six months ahead. However, CLI does not provide any forecast gain in CCI forecasts over the same three horizons.

Our main results are the following. Real-time forecasts of IP for one, three, and six months ahead, are all significantly improved by incorporating the information of leading indicators. This result holds true if the leading indicators are summarized by CLI, as in Diebold and Rudebush, with gains in mean squared error (MSE) of 10-15%. Even better results, with MSE gains of 15-20%, are obtained by employing principal component regressions (PCR). This turns the negative result of Diebold and Rudebush for the period 1968-1988 into a positive result for the period 1989-2009.

On the other hand, it turns out to be hard to forecast CCI in real time. CLI does worse than AR in one, three, and six months ahead forecasting, and although PCR does improve about 10% on AR, this gain is significant only for forecasts one month ahead and not for three or six months ahead. The positive findings for CLI reported in McGuckin, Ozyildirim, and Zarnowitz are possibly due to the fact that their analysis is not completely real-time. First, the majority of their forecast models employ (in real time unavailable) future values of CCI and CLI. [1] Second, their forecast models employ

[1] To be precise, at the end of their Section 4.2, they report that their equations (3)-(5) are considered only for $j \geq p$, that is, where the forecasted variable is the growth rate over months $t-j$ till t and the values of the explanatory variables run till month $t-p$; for $j > p$, that is, in the majority of cases, the real-time restriction is violated, as at time $t-j$, the values for months $t-j+1$ to $t-p$ are not yet available.

(in real time unavailable) final vintage lagged CCI values (as AR terms), instead of real-time values.

The structure of the rest of this paper is as follows. Section 2 discusses the data set, with vintage data of coincident and leading indicators. Eight methods participate in the forecast competition, as described in Section 3. The real-time forecast results for IP and CCI are in Section 4, and Section 5 shows results for various other data specifications, including the forecast results obtained for 'ex post' (final release) data. Real-time recession forecasts are compared in Section 6, and Section 7 concludes.

2 DATA

2.1 Coincident and Leading Indicators

The data consist of monthly observations of a set of US macroeconomic variables over the period January 1959 till April 2009 (604 observations).[2] The data set contains ten leading indicators and four coincident indicators (including industrial production, IP). The Conference Board constructs two composite indexes from these indicators, the composite leading index (CLI) and the composite coincident index (CCI).

<< TABLE 1 to be inserted somewhere over here. >>

Table 1 provides a data overview. The ten leading indicators are average weekly hours in manufacturing, average weekly initial claims for unemployment insurance, manufacturers' new orders for consumer goods and materials, manufacturers' new orders for non-defense capital goods, building permits for new private housing units, vendor performance slower deliveries diffusion index, M2 money supply, the S&P 500 stock price index, the spread between the 10-year Treasury bond rate and the Federal Funds rate, and the University of Michigan index of consumer expectations. The Business Cycle Indicators Handbook of the Conference Board (2001) provides further background on these leading indicator variables.

For most variables, data become available after a delay of one month, whereas this delay is two months for three of the leading indicators. For vintages up to November 2000, the release delay of published data is two months. Since December 2000, the

[2]Source: The Conference Board, June 2009. We thank Ataman Ozyildirim of The Conference Board for making these data available to us.

three indicators with a delay of two months are extrapolated by one month, by means of an autoregressive model of order two, as suggested in McGuckin, Ozyildirim, and Zarnowitz (2001, 2007). In this way, the release delay is reduced to one month. This means, for instance, that one-month-ahead out-of-sample forecasts for month t are made with the vintage data published in month $t+1$ until November 2000, and with the vintage of month t from December 2000 onwards, as the vintage data run till month $t-1$ in both cases.

All variables are transformed to stationarity. The employed data transformations are indicated in Table 1, and they are the same as in Stock and Watson (2002a, 2005). After this transformation, all leading indicators and the CLI are standardized by Z-scores. Several of the leading indicators contain one or more breaks, in particular, changes of measurement scale, but all these breaks are removed by the data transformations. Some of the resulting series of leading indicators contain some outliers, and these are not adjusted. The predicted variables, IP and CCI, are measured as (annualized) growth rates, without standardization and without outlier adjustment. Let Y_t be the level of IP or CCI, then the transformed series $y_t = 1200\text{dln}(Y_t) = 1200(\ln(Y_t) - \ln(Y_{t-1}))$ is the annualized one-month growth percentage. The target variable in forecasting is the annualized future h-month growth percentage, that is defined by $y_t^h = (1200/h)(\ln(Y_{t+h}) - \ln(Y_t)) = (1/h)\sum_{j=1}^{h} y_{t+h}$.

The predictor set to forecast IP and CCI is restricted to the ten leading indicators in Table 1. This may seem a somewhat restricted data set, as an abundance of macroeconomic information is available nowadays. However, the restriction to the ten indicators can be motivated in several ways. First, any real-time analysis requires use of vintage data, and these are available from the Conference Board for all variables in Table 1. For larger data sets, as in Stock and Watson (2002a, 2002b, 2005), vintage data are not available yet. Second, leading indicators provide a strong type of predictor variables, as is discussed, for instance, in Marcellino (2006) and Banerjee and Marcellino (2006). Third, more data do not always lead to better forecasts, see Boivin and Ng (2006). This holds in particular in real-time forecasting, as will be seen later in this paper. If data are updated over time, advanced methods may well suffer more than simple methods from the preliminary nature of initial releases. As discussed before, Diebold and Rudebush (1991) showed that simple AR models often provided more accurate forecasts of IP than models with the CLI. In Section 4, it turns out that AR models perform better than (univariate and bivariate) models with the CLI. More in general, the additional uncertainty due to data revisions requires the use of relatively simple models.

2.2 Vintages

Real-time data with data revisions have a so-called vintage structure. This means that, in each month and for each variable, the current value is published together with all the, possibly revised, past values. The resulting data matrix, with rows for each observation month and columns for each vintage (publication month), has a characteristic triangular structure, see, for instance, Diebold and Rudebush (1991). Our data set consists of monthly vintages, from January 1989 till May 2009, with data from January 1959 onwards. For most variables, with release delay 2 for early vintages and 1 for later vintages, the data of the first vintage run till October 1988 and those of the last vintage run till April 2009. This means that there are in total 245 values for these variables for October 1988, one for each vintage, as opposed to only a single value for April 2009, namely from the vintage of May 2009.

The target variable in forecasting is the actual (as opposed to the initially reported) growth rate of IP and CCI. The final vintage provides the most reliable information on this actual historical growth rate, so that the final vintage values are taken as the target values in prediction. This is also in line with the approach taken by Diebold and Rudebush (1991) and McGuckin, Ozyildirim, and Zarnowitz (2007). It is therefore of special interest to compare the final vintage values with those reported in real time (first release), and we also compare this with the second release value, that is, after the first update. The three values (first, second, and final release) are compared in Figure 1. Data revisions have a larger effect for shorter forecast horizons. Final vintage values may differ considerably from first release (real-time) values. For CCI, the largest revisions are mostly upward and less downward. This finding holds for all horizons, but most clearly so for growth rates over three and six months. The large upward revisions occur mostly for initial releases with large negative growth rates (recessions). For IP, upward and downward revisions occur roughly equally often, and large negative growth rates are much less revised as compared to CCI. The largest upward revisions, especially for CCI, occur during the 1991 recession. The revisions are smaller and tend to be downward during the recession in 2001, and also during the recent recession since 2008.

<< **FIGURE 1 to be inserted somewhere over here.** >>

In Figure 2, the growth rate (as annualized percentage) of IP and CCI over the last month, quarter, and half year of 1989 and 2007 is displayed, as computed from the various vintages. The revisions are sometimes quite dramatic and occur also after long periods (ten years and more). Perhaps this is not completely unexpected for monthly growth rates, as these are highly volatile, but also the quarterly and half year growth rates are subject to major revisions. For instance, the growth rate of CCI over the last quarter of 1989 is revised from 1.0% to about 2.0%, and the growth rate of IP over the last half year of 2007 is revised from 0.8% to 2.2%. Figure 3 provides similar information, now for the first six and the final vintage values for the (annualized percentage) growth rate over the last month, quarter, and half year of each year from 1988 till 2008. For some years, for instance in 1990, initial releases are far from final vintage values, and most of the largest revisions (of more than 5%) are upward. For years with recessions (early 1990, mid 2001, and from 2008 onwards), revisions are upward in 1990, small in 2001, and downward in 2008. It should be noted that the final value for 2008 may still change in later vintages.

<< **FIGURES 2 and 3 to be inserted somewhere over here.** >>

3 FORECAST METHODS

3.1 Real-time, recursive, out-of-sample forecasting

The real-time forecast performance of alternative methods is evaluated by comparing their simulated out-of-sample forecasts. This means that, for each forecast origin T, the employed data set consists of the vintage of month T. This vintage contains observations up to and including month $T-2$ prior to December 2000, and up to and including month $T-1$ from December 2000 onwards. First, this vintage is used to estimate forecast models for the h-month growth rate of IP or CCI, for forecast horizons

h of one, three, and six months. (We also considered $h = 12$ and $h = 24$, but these results are not reported here. For $h = 12$, the results are qualitatively similar to those for $h = 6$, whereas for $h = 24$ the considered real-time forecast methods are no longer able to provide any significant gains over the AR benchmark.) Second, the obtained models generate growth rate forecasts over the future h months, that is, for the period running from $T - 1$ till $T - 2 + h$ (prior to December 2000) or from T till $T - 1 + h$ (from December 2000 onwards). As the models are estimated with data that are all prior to the forecast period, the forecasts are out-of-sample.

This forecasting procedure is repeated for each forecast origin T, starting with the first vintage in January 1989 (to forecast growth rates for the h months after October 1988) and ending h vintages before May 2009 (to forecast growth rates for the h-month period ending in April 2009, which is the most recent value available for the final vintage of May 2009). The number of forecasts for horizon h is equal to $245 - h$. As the forecast evaluation period 1989-2009 may contain structural breaks, the models are estimated with a moving window of ten years, that is, with the most recent 120 observations of the employed vintage. The accuracy of the h-month growth rate forecasts will be evaluated by means of the mean squared forecast error (MSE).

This recursive, simulated out-of-sample forecast evaluation method is the same as in a series of forecast studies of Stock and Watson (1999, 2002a, 2002b, 2006). The only (but important) distinction is that those studies use final vintage data, whereas our forecasts are based on real-time data.

3.2 AR and CLI

As before, let Y_t be the level of IP or CCI and let

$$y_t = 1200\Big(\ln(Y_t) - \ln(Y_{t-1})\Big)$$

be the annualized one-month growth percentage. The target variable in forecasting is the annualized future h-month growth percentage,

$$y_t^h = \frac{1200}{h}\Big(\ln(Y_{t+h}) - \ln(Y_t)\Big) = \frac{1}{h}\sum_{j=1}^{h} y_{t+j}.$$

Forecasts of this variable are denoted by $\widehat{y}_t^h$. Autoregressive models, relating y_t^h to y_t and its lags, are taken as a benchmark. The AR forecast model with r lags is given by

$$\widehat{y}_t^h = \alpha + \sum_{j=0}^{r} \beta_j y_{t-j}.$$

The main question studied in Diebold and Rudebush (1991) and McGuckin, Ozyildirim, and Zarnowitz (2007), is whether or not the composite leading index (CLI) provides significant additional forecast power. Our CLI forecast model with q lags of CLI is defined by

$$\widehat{y}_t^h = \alpha + \sum_{j=0}^{r} \beta_j y_{t-j} + \sum_{j=0}^{q} \gamma_j \mathrm{CLI}_{t-j}.$$

Here CLI_t denotes the one-month (annualized percentage) growth rate of CLI. As discussed before, the target variable y_t^h is the 'true' growth rate of IP (or CCI), and this is taken to be the one computed from the final vintage. However, to estimate the forecast model, only real-time data of IP (or CCI) and CLI are used. In order for real-time CLI to have forecast power for (final release) IP (or CCI), it should be correlated with IP (or CCI) both in real time (for purposes of estimation) and in final form (for purposes of forecasting). Table 2 provides information on various real-time and final vintage correlations. Table 2.a shows that the correlation between first and final vintage values of one-month growth rates of CCI is .55 (as compared to .85 for IP). The correlations for horizons of three and six months are larger, but especially for CCI the effects of data revisions remain rather large (with correlations of .76 and .83). These results suggest that, for real-time forecasting, forecasting at short horizons ($h = 1$) may be difficult, and that it may be harder to forecast CCI than IP. Table 2.b shows the correlations of real-time CLI with IP and CCI, both for real-time and for final vintage data of IP and CCI. The 'Average' correlations in Table 2.b are relevant for estimation, and the 'Final' ones for forecasting. In most cases (except for $h = 1$ at lag 0), both the 'Average' and the 'Final' correlations are larger for IP than for CCI. All these correlations are quite modest, with values mostly between 0.2 and 0.4. Table 2.b shows also the autocorrelations of CLI, which are relevant for VAR forecast models where CLI is itself forecasted from the past (see Section 3.4). The real-time autocorrelations, relevant in estimation, are small, only about 20% for lags 1 and 2 and negligible for larger lags. This indicates that VAR models with CLI may not add much forecast power to univariate models.

<< TABLE 2 to be inserted somewhere over here. >>

The above type of model specification, relating h-month growth rates in IP (or CCI) to one-month growth rates (and their lags) of IP (or CCI) and one-month growth rates of CLI, is similar to the methodology of Stock and Watson (1999, 2002a, 2002b). In their terminology, the CLI model is called the DI-AR-Lag model, as the forecasts are based on the 'diffusion index' CLI and its lags and on autoregressive terms corresponding to current and lagged values of the one-month growth rate. To apply this model, specific values for the lag orders q and r should be chosen. The results in Stock and Watson (2002a, 2006) show that the Bayes Information Criterion (BIC) works rather well in this respect. For all methods, BIC is therefore used to determine the lag orders q and r at each forecast origin, based on a moving estimation window consisting of the past ten years of observations. The considered set of forecast models have CLI lag $q \leq 3$ and autoregressive lag $r \leq 3$. Models without autoregressive terms (DI-Lag, in the terminology of Stock and Watson) are also considered.

The chosen set-up differs in several respects from that of some previous studies. The main differences with Diebold and Rudebush (1991) are the following: (i) They consider solely the one-month growth rate, which is forecasted h periods ahead. (ii) They consider a fixed set of models, without selecting a model. (iii) They use final vintage (instead of real-time) AR terms y_{t-j} in their forecast models. The main differences with McGuckin, Ozyildirim, and Zarnowitz (2007) are: (i) As mentioned in the introduction, most of their forecast models use future data, that is, models for y_t^h containing right-hand-side explanatory variables at times beyond the current time t. (ii) They use final vintage (instead of real-time) AR terms y_{t-j} in their forecast models. (iii) They use h-month growth rates of IP (or CCI) and of CLI as explanatory variables in their forecast models, instead of one-month growth rates, with the disadvantage that h-month growth rates show considerably more autocorrelation than one-month growth rates. (iv) They report the average MSE over a set of fixed forecast models, instead of the MSE of models selected by BIC. In our opinion, the set-up of Stock and Watson is a rather natural one, and we prefer to follow their methodology.

3.3 OLS and PCR

The CLI of the Conference Board is constructed from the ten underlying leading indicators shown in Table 1. Although much effort is invested in composing this index in the best suitable way, it is of interest to consider alternative ways to incorporate the information in the ten leading indicators. A simple method is to include all leading indicators (L_i, $i = 1, \ldots, 10$) as separate predictors. The corresponding forecast model, without any lags and without autoregressive terms, is

$$\widehat{y}_t^h = \alpha + \sum_{i=1}^{10} \gamma_i \mathrm{L}_{i,t}.$$

The number of observations in estimation is 120. We also considered this model with up to 3 AR lags and up to 3 lags of all leading indicators, selected by BIC, but this does not provide any forecast improvements over the above model. The method above will be denoted by 'OLS', as the effect of the leading indicators is estimated simply by regressing the growth rate on all leading indicators. An alternative method is to condense the information in the ten leading indicators in terms of one or a few principal components (PC_i, $i = 1, \ldots, p$), with corresponding forecast model

$$\widehat{y}_t^h = \alpha + \sum_{j=0}^{r} \beta_j y_{t-j} + \sum_{i=1}^{p} \sum_{j=0}^{q} \gamma_{ij} \mathrm{PC}_{i,t-j}.$$

This is called Principal Components Regression (PCR). We consider models with $p \leq 3$ components, and the actual values of p, q and r are selected by BIC. This method is applied in Stock and Watson (1999, 2002a, 2002b, 2006) to condense the information in large sets of (far more than one hundred) predictor variables, but the same methodology can be applied for the relatively small set of ten leading indicators. We refer to Anderson (1984) and Jolliffe (2002) for further background on principal components. The principal components can be seen as an alternative method to construct a composite index or diffusion index, and we refer to Marcellino (2006) for a more comprehensive overview of index construction methods.

3.4 VAR

The forecast models discussed before are all single-equation methods, where the IP or CCI growth rate is related to a set of predictor variables and their lags. It may help in multi-step-ahead forecasting to use relatively short lags in the equation and to forecast the required future values of the predictor variables. This is achieved by vector autoregressive (VAR) models. We will consider bivariate VAR models, with variables IP (or CCI) and CLI, as an alternative to the single-equation CLI model. The VAR methodology can in principle also be applied for the OLS and PCR methods, but this will not be pursued here. (It turns out that 11-dimensional VAR models in terms of the ten leading indicators and IP or CCI do not perform well, which is possibly due to the large number of parameters.) In total, we consider four types of VAR model, that is, with a 'direct' or 'iterated' approach and explaining the past or future h-month growth rate (see Marcellino, Stock, and Watson (2006) and Marcellino (2006) for a discussion of direct and iterated methods). More precisely, let the current time (forecast origin) be denoted by t. Then the most recent available h-month growth rate of IP or CCI is that over the period from $t-h+1$ till t, that is, y_{t-h}^h. The purpose is to forecast the coming h-month growth rate, y_t^h. Let x_t be the one-month CLI growth rate from $t-1$ till t and let y_t be the one-month growth rate of IP or CCI from $t-1$ till t (note that $y_{t-1}^1 = y_t$). The CLI method of Section 3.2 amounts to the regression of y_t^h on y_t, x_t, and their lags. The four VAR methods are defined as follows (all models include constant terms, which are not denoted explicitly).

- VARm- ('iterated past'): Regress (y_t, x_t) on lags of (y_t, x_t). Then, (y_s, x_s) are recursively forecasted one month ahead, for $s = t+1$ till $s = t+h$, and $\widehat{y}_t^h = (1/h)\sum_{j=1}^h \widehat{y}_{t+j}$.

- VARm+ ('iterated future'): Regress (y_t, x_{t-h+1}) on lags of (y_t, x_{t-h+1}). The (iteratively computed) forecast equation for y_s, for $s = t+1$ till $s = t+h$, involves (lags of) x_{s-h}, and these are all observed (non-extrapolated). As before, $\widehat{y}_t^h = (1/h)\sum_{j=1}^h \widehat{y}_{t+j}$.

- VARh- ('direct past'): Similar to VARm-, but with h-month growth y_{t-h}^h instead of 1-month growth y_t. That is, regress (y_{t-h}^h, x_t) on lags of (y_{t-h}^h, x_t). Then, (y_{s-h}^h, x_s) are recursively forecasted for $s = t+1$ till $s = t+h$, and this gives $\widehat{y}_t^h$.

- VARh+: ('direct future'): Similar to VARm+, but with y_{t-h}^h instead of y_t. That is, regress (y_{t-h}^h, x_{t-h+1}) on lags of (y_{t-h}^h, x_{t-h+1}). Then, y_{s-h}^h is recursively forecasted for $s = t + 1$ till $s = t + h$. Similar to VARm+, $\widehat{y}_t^h$ depends on observed (non-extrapolated) values of x_t and its lags.

For $h = 1$, all four VAR methods are identical. The CLI method of Section 3.2 can be seen somewhat as a mixture of these methods, as the dependent variable is the h-month growth rate ('direct') but the explanatory variables are one-month growth rates. The CLI forecast equation is a mixture of that of VARm+ (with monthly predictors) and VARh+ (with h-month dependent variable). In our forecast comparison, we consider VAR models with at most 3 lags, selected by BIC.

4 REAL-TIME FORECAST RESULTS

4.1 Full period results

The eight methods discussed in Sections 3.2-3.4 are applied in real-time forecasting of IP and CCI over horizons of one, three, and six months. Table 3 provides a summary of outcomes, in terms of the mean squared forecast error (MSE) and the bias and variance of the forecast errors. As an additional benchmark, we consider the simple forecast obtained by taking the sample mean over the most recent past 120 observations (denoted by 'M' in Table 3). The MSE's for M and AR are in absolute terms, whereas the MSE's of the other eight methods are taken relative to the MSE of the AR benchmark method.

For IP, PCR provides the lowest MSE's for all three horizons. It has the smallest forecast error variance of all methods, compensating a bias that is sometimes slightly larger than that of competing methods. The MSE gains of PCR as compared to (the best of) M and AR range from 12% for $h = 1$ to 20% for $h = 6$. CLI performs slightly less well, with MSE gains ranging from 7% to 15%. For $h = 3$ and $h = 6$, OLS does equally well as CLI. In comparison with AR, the simple M method is the better one for $h = 1$ (as a larger bias is more than compensated by a smaller variance), it is about equally well as AR for $h = 6$, and for $h = 3$ it is worse. The VAR models show some mixed results. VARh- is best for $h = 3$, with VARh+ close by, but VARm+ is best for $h = 6$. The forecast bias is consistently negative for all methods, except for OLS, so that most methods tend to overestimate the future growth of IP. The models selected by BIC have relatively few lags. For AR, the average lag length ranges from 0.2 (for $h = 1$) to 1.1 (for $h = 3$). For CLI, AR terms are sometimes not included and the average AR lag is close to 0 for all horizons, and the average CLI lag is about 1. For PCR, the far majority of selected models exclude all AR terms, and the average number of principal components is about 2 with an average lag of about 0.5 to 0.9. All four VAR models always have selected lag 1.

The results for CCI differ in some important respects from those for IP. Again, PCR is the best method for all horizons, but the forecast gains of about 10% are smaller than for IP. CLI is not able to improve on the AR benchmark, and this holds also true for OLS and (for $h = 3$ and $h = 6$) for all four VAR methods. For CLI, the average AR lag ranges from 1.5 to 2, and the average CLI lag ranges from 0.3 to 0.6. PCR mostly excludes all AR terms, and the average number of components is about 1.8 with average lag below 0.5.

<< **TABLE 3 to be inserted somewhere over here.** >>

The significance of the forecast gains for IP and CCI is evaluated by the Diebold-Mariano t-test with HAC standard errors, see Diebold and Mariano (1995) and Newey and West (1987). As the forecast model structure is selected by BIC at each forecast origin, the models employed by the various methods are not nested. The t-values are reported in Table 6.a.1 and 6.b.1, in the rows labeled 'X all real, target final'. For the VAR models, the results are shown only for the 'iterated past' (VARm-) and 'direct past' (VARh-) methods, as these methods perform overall somewhat better than VARm+ and VARh+. Table 6 contains many more test outcomes that will be discussed later, see Section 5. Table 6.a.1 indicates that the forecast gains of PCR and CLI for IP are significant (at 5% level) for all three horizons, and VARm- and VARh- obtain significant gains only for $h = 1$. Further, although PCR performs better than CLI, this difference is not significant for any horizon, and the same holds true for the differences in forecast quality of CLI and PCR as compared to OLS. The results in Table 6.b.1 for CCI show that none of the methods provide significant gains for $h = 3$ and $h = 6$. The only significant result (at 5% level) is that, for $h = 1$, PCR beats all its competitors, that is, AR, CLI, and OLS.

Summarizing, considerable real-time forecast gains over AR can be obtained for IP, with PCR as best method and CLI as second-best. For CCI, however, no significant real-time gains over AR can be obtained. These results for the last two decades amend earlier (negative) results for IP of Diebold and Rudebush (1991) for the seventies and eighties of the last century and (positive) results for the CCI in McGuckin, Ozyildirim, and Zarnowitz (2007). For IP, the improved real-time forecast results for the period 1989-2009 as compared to the period 1968-1988 may be partly due to a structural reduction since 1989 in the variance of the revisions of monthly IP vintage data, as reported by Swanson and Van Dijk (2006).

As concerns the choice between iterated and direct VAR forecasts, as noted in Marcellino, Stock, and Watson (2006), this is an empirical question. For IP, the direct method VARh- is best for $h = 3$, whereas the iterated method VARm+ is best for $h = 6$. For CCI, the direct method VARh- is best for $h = 3$ and $h = 6$, but the differences with the indirect method VARm- are small.

4.2 Results in sub-periods

In the foregoing, alternative methods were evaluated by the forecast performance over the full evaluation period from 1989 till early 2009. Table 4 reports the MSE's of the methods over four sub-periods of five years, and also for the most recent twelve forecasts falling mostly in the most recent recession period. The most important message is that the optimal method differs per sub-period, but that PCR and CLI perform reasonably well (as compared to AR) in all cases. The MSE's of AR indicate much variation in forecast error variance. It becomes increasingly hard to forecast IP, whereas the most difficult periods for CCI are 1989-1993 and 2004-2008. For both IP and CCI, the last twelve forecasts are by far the hardest of all. In this period with strongly negative growth, OLS performs by far the best, with gains of about 30% for IP and 50-70% for CCI. The good performance of OLS in this turbulent period may be due to its simplicity, as it is the only method that does not employ any lagged values. VAR methods turn up to be optimal over several sub-periods, but among these four methods there is no clear winner. For horizons $h = 3$ and $h = 6$, VARm- is best in 3 cases, VARm+ in 2, VARh- in 4, and VARh+ never. In most of what follows, we will consider only two of the four VAR methods, that is, VARm- and VARh-

<< TABLE 4 to be inserted somewhere over here. >>

5 ALTERNATIVE DATA SPECIFICATIONS

5.1 Forecast of real-time growth rates

In the foregoing, the target variable to be forecasted is the 'real' growth rate, that is, the 'true' macroeconomic developments as evidenced after possible data revisions. This motivated the use of final vintage values for the target variable. However, as is also noted in Golinelli and Parigi (2008), it is also of interest to forecast the real-time growth rates prior to revisions, as financial markets and policy makers may react strongly to initial releases of macroeconomic data. This holds true in particular for one-month ahead forecasts, as the effect of data revisions may gradually smooth out for longer forecast horizons. Initial releases affect short run reactions of economic agents, whereas the final vintage growth rates are more relevant as indicators of the actual state of the economy.

The results of real-time forecasts of the real-time (initial release) growth rates of IP and CCI are shown in Table 5 (in the rows labeled 'X all real, target real'). For ease of comparison, the table contains also the MSE's of Table 3 for forecasting final vintage growth rates (in the rows labeled 'X all real, target final'). The table contains further results that are discussed in Section 5.2.

As compared to the final vintage growth rates, the MSE of AR for real-time growth rates becomes somewhat smaller, up to 20%, except for CCI at horizons of three and six months, where the MSE increases by up to 20%. For IP, PCR performs best in one-month-ahead forecasting, with an MSE gain of 12% over the AR benchmark. For forecasts of the IP growth rate over the coming three or six months, PCR and CLI perform equally well, with gains of slightly over 10%. OLS performs best for horizon $h = 3$, but worse for $h = 1$ and $h = 6$. Of the VAR methods, VARh- is overall the best. For CCI, the forecast gains are smaller. PCR performs best, with gains of 8-17%, whereas CLI performs less well, with gains of 0-9%. Of the VAR methods, VARm- and VARh- perform relatively best.

<< TABLE 5 to be inserted somewhere over here. >>

The differences in forecast quality of the various methods is evaluated by the Diebold-Mariano t-values in Table 6.a.1 and 6.b.1 (the rows labeled 'X all real, target final'). For IP, the forecast gains of PCR are significant (at 5% level) for $h = 1$ and $h = 3$, bur not for $h = 6$. The gains of CLI are only significant at the 10% level, for $h = 3$ and $h = 6$. For CCI, the gain is significant (at 5% level) only for VARh- for $h = 6$ (and PCR gains significantly on CLI for $h = 1$, but not on AR).

5.2 Use of final vintage data

If real-time aspects of the data are neglected, we can study the forecast performance of the various methods for final vintage values, that is, the values reported in the most recent available vintage of May 2009. One option is to take final release data for all variables. It is particularly interesting to compare these 'ex post' results with those of real-time, 'ex ante' data, see also Swanson (1996). The results are given in Table 5, 6.a.1, and 6.b.1 (rows labeled 'X all final, target final'). Another option, followed by Diebold and Rudebush (1991) and by McGuckin, Ozyildirim, and Zarnowitz (2007), is to take final release data for all variables except for the ten leading indicators, which are taken in real time, as well as CLI and the principal components derived from the leading indicators. In this way, attention is focused solely on the effect of data revisions in the leading indicators. The results of this case are given in Table 5, 6.a.1, and 6.b.1 (rows labeled '(C)LI real, IP (or CCI) final, target final').

For IP, the outcomes are as follows. If all data are taken from the final vintage, CLI, PCR, and OLS perform about equally well, with gains of about 10-15%. For $h = 3$ and $h = 6$, VARh- performs best of all methods, with VARh+ close by. For CCI, OLS performs best for $h = 1$ and $h = 3$ with gains of about 10%, and VARh- is best for $h = 6$ with a gain of 14%. CLI does reasonably well for $h = 3$ and $h = 6$. Tables 6.a.1 and 6.b.1 indicate that only few gains are significant (at 5% level). That is, for IP the gain is significant for $h = 1$ for CLI, PCR, and VAR, for $h = 3$ for CLI and VARh-, and for $h = 6$ for CLI and PCR, and for CCI this holds true only for CLI for $h = 3$ and for VARh- for $h = 6$. The results for the case where only the leading indicators are taken in real time are roughly similar. The only difference is that the MSE gains are consistently somewhat smaller, and the only gains that are significant (at the 5% level) are, for IP, for CLI, PCR, and VAR for $h = 1$, and for CCI, for $h = 6$ and VARh-.

Tables 6.a.2 and 6.b.2 compare three options to forecast the final growth rates of IP and CCI, that is, with all predictor variables in real time (as in Section 4), with all data taken from the final vintage, and with the leading indicators measured in real time but the other predictors (lags of IP and CCI) in final release form. It might be anticipated that is is hardest to realize forecast gains in real time, and easiest when all data are in final release form. This is indeed the case, at least, to a certain extent. The results for IP in Table 6.a.2 show that the differences are rarely significant for $h = 1$. For $h = 3$ and $h = 6$, it is indeed harder for the AR benchmark method to forecast in real time than with some or all predictors in final form. This is also true for VARm- and VARh-, to a lesser extent for CLI, but not for OLS and PCR. The case with only the leading indicators in real time is not harder than with all data in final form, with a single exception for CLI at $h = 6$. The results for CCI in Table 6.b.2 are somewhat stronger. Now, for all horizons it is easier for AR, CLI, PCR, OLS, and VARh-, to forecast with final data than with all or some of the variables measured in real time. Again, the case with only the leading indicators in real time is not harder than with all data in final form, except for VAR at $h = 1$ and for CLI and PCR at $h = 3$.

<< **TABLE 6 to be inserted somewhere over here.** >>

In a sense, 'ex ante' (real-time, first vintage) and 'ex post' (final vintage) forecasts are two extreme cases. Intermediate cases are obtained for intermediate vintages, for instance, by using second-release or later vintage values that become available after one or more months (so that these scenarios can not be realized in practice). Table 5 shows that the ranking of forecast methods in real time differs from that for final vintage data. Figure 4 shows the MSE gains over 1989-2008 of CLI, PCR, and OLS, over 1989-2008, if the data used in prediction are taken from vintages that are k years ahead of the forecast origin (for $k = 0, \ldots 20$, so that $k = 0$ corresponds to real-time forecasting and $k = 20$ to final vintage forecasting). For all three forecast horizons and for both IP and CCI, PCR performs best in real time. If future vintages could be used, the relative performance of PCR deteriorates while that of CLI improves in most cases. For CCI, the CLI and OLS methods do not improve on AR in real time, and this is only achieved for vintages of about five years or more in the future.

<< **FIGURE 4 to be inserted somewhere over here.** >>

6 RECESSION FORECASTS

6.1 Method

The aim of the methods considered before is to forecast future growth rates for the coming one, three, or six months. These forecasts can also be used in a simple way to provide forecasts of oncoming expansions and recessions, and this comes closer to the original objective of leading indicators as envisaged by Mitchell and Burns (1938). We use the common rule of thumb to define a recession as the occurrence of two subsequent quarters of negative growth in the CCI, as measured in the final vintage release. An expansion corresponds to two subsequent quarters with positive growth, and the regime is mixed if the growth is negative in one of the two quarters and positive in the other one. Over the considered period from 1989 till early 2009, 28 times there is a recession in the next six months, 172 times an expansion, and 39 times a mixed regime. The well-known NBER recession indicator reports 28 recessions and 211 expansions.

At each forecast origin T, a probability forecast for the occurrence of a recession during the next six months is obtained from the real-time forecasts of the three and six month growth rates of the CCI, $\widehat{y}_T^3$ and $\widehat{y}_T^6$, as follows. As the growth rate forecasts are annualized, the predicted (non-annualized) growth rates over the coming two quarters are given by $\widehat{y}_{Q1,T} = \frac{1}{4}\widehat{y}_T^3$ and $\widehat{y}_{Q2,T} = \frac{1}{2}\widehat{y}_T^6 - \frac{1}{4}\widehat{y}_T^3$. The real-time recession probability forecast, $\widehat{p}_T$, is obtained by estimating the probability that both these growth rates are negative. To compute this probability, it is assumed that the two growth rates are jointly normally distributed, with means $\widehat{y}_{Q1,T}$ and $\widehat{y}_{Q2,T}$ and with a covariance matrix estimated from the past ten years of observations of the (current release) quarterly growth rates.

Two statistical tools, QPS and ROC, will be used to evaluate the real-time recession forecast power of the various methods. Let the 'true' (empirical or NBER) recession indicator R_t be defined by $R_t = 1$ for a recession in the coming six months and $R_t = 0$ for an expansion. Then the Quadratic Probability Score (QPS) is defined as

$$\text{QPS} \;=\; \frac{1}{N}\sum_{t=1}^{N}(R_t - \widehat{p}_t)^2.$$

This statistic was proposed by Brier (1950) (for multinomial outcomes; for binomial outcomes, as in the current setting, the original formula of Brier is obtained by multiplying the above formula by a factor 2, but this factor is usually omitted). For the empirical recession indicator, the average is taken over the $N = 200$ expansion and recession periods, leaving out the 39 months with mixed regimes. For the NBER indicator, the average is taken over all $N = 239$ forecast periods.

Recession probability forecasts can be transformed into recession signals by imposing a threshold value, say τ, and forecasting a recession if $\widehat{p}_T > \tau$. The Receiver Operating Characteristic (ROC) depicts the true positive rate (the proportion of true recession signals in all recession periods) against the false positive rate (the proportion of false recession signals in all expansion periods), as the threshold τ ranges from 0 to 1. The ROC value is defined as the area under the ROC curve, which runs in the unit square from (0,0) (for $\tau = 1$) to (1,1) (for $\tau = 0$). Purely random predictions are expected to produce an ROC curve running along the main diagonal, with ROC value 0.5, and successful predictions have an ROC value (sufficiently much) above 0.5, with a maximum of 1.

6.2 Results

Table 7.a shows results for recession probabilities predicted in real time by means of six methods. The average predicted probability of an expansion is larger that 50% for all methods in nearly all situations, also during recessions. OLS gives the best forecasts during recessions, with an average predicted recession probability of 46%, but this method provides also the most false recession signals during expansions and mixed regimes. The differences between AR, CLI, PCR, and VARh- are very minor, and VARm- is the worst method for recession forecasting.

The more detailed results for official NBER recession periods in Table 7.b confirm this general picture. As the task is to provide real-time, 'ex ante' predictions of recessions, the five months prior to the start of a recession are seen as an early recession stage (as it will start within the forecast horizon of six months), and the last five recession months form a late stage (as the recession will end within the forecast horizon). Clearly, the recession of 2001 is missed by all methods, and the same holds true for the early stage of the 2008 recession. The recession of 1990-1991 is predicted with relatively more success, and all methods do about equally well (if the biased high recession signals of OLS are left out). The same holds true for the middle period of the 2008 recession, that was still lasting for the most recent data available in the database, that is, the May 2009 vintage with data running till April 2009 (and final forecast origin October 2008). Even in this long recession period, methods other than OLS do not provide strong recession signals, with predicted recession probabilities of about 25% for CLI, PCR, and VARh-.

Table 7.c reports QPS and ROC statistics of the alternative methods. The QPS results for NBER recessions are very similar to those for the empirically defined recessions. OLS has the lowest QPS of all methods in recessions, especially in the most recent recession period, but at the cost of highest values in expansions. VARm-, on the contrary, has the lowest QPS of all methods in expansions, at the cost of highest values in recessions. This is because OLS is biased to give too many recession signals, and VARm- to give too few. The best ROC value is obtained by VARh-, and AR, CLI, PCR, and OLS give nearly similar results. However, VARm- does not provide acceptable results, as its ROC value of about 0.5 is no better than that of random predictions. The ROC values are larger for the empirical recession indicator than for the NBER recessions. This may be due to the fact that NBER recessions are defined partly in terms of future data, whereas the empirical indicator is defined in terms of (final vintage) growth rates, which are related to the (real-time) target variable fitted by the considered methods. Still, the real-time recession forecast results are rather modest when measured by QPS and ROC. An important reason is that the forecasts are made in real time, whereas the target is to predict 'true', final vintage recessions and expansions. As was seen before, 'ex ante' values may differ quite considerably from the 'ex post' values, causing difficulties in the accurate prediction of recessions in real time.

<< TABLE 7 to be inserted somewhere over here. >>

7 CONCLUSIONS

A realistic comparison of the forecast power of alternative methods demands the use of realistic data, that is, data that are available at the time the forecast of the future is made. This crucial condition is met in real-time forecasting, where every forecast is based on the (most recent) data vintage that is available at the forecast moment.

This paper compares the real-time forecast power of a set of methods to predict the future macroeconomic developments in the US, for the coming month, quarter, and half year, using information in the set of ten leading indicators of the Conference Board. As compared to an autoregressive benchmark, the most notable forecast gains are obtained in forecasting Industrial Production (IP) by means of principal components of the leading indicators. The achieved reduction in mean squared error of monthly forecasts from 1989 till early 2009 is 12%, 17%, and 20%, for growth rate forecasts over respectively the coming month, quarter, and half year. The Composite Leading Index of the Conference Board performs somewhat less well, but the reductions of respectively 8%, 14%, and 15%, are still significant.

The results for the Composite Coincident Index (CCI) are less positive. Principal components achieve a reduction in mean squared error of about 10% for all three forecast horizons, but the reduction is significant only for a horizon of one month. The Composite Leading Index is not found to add any forecast power to the autoregressive benchmark. These results are confirmed by a CCI-based recession forecast study, as the various methods provide at best modest improvements over the autoregressive benchmark in real-time forecasting of future recessions.

The above findings amend previously reported results on real-time forecasting of IP and CCI. Two of the causes why IP is easier to predict than CCI in real time are the following. First, the objective is to forecast the 'real' growth rates, and final vintage data (incorporating data revisions) are taken as the most reliable information on these growth rates. However, the forecasts are based on real-time data that do not yet incorporate future data revisions. For IP, these revisions are less substantial than for CCI. The correlation between the real-time and final data of IP is 85%, 90%, and 91% for the growth rate over respectively a month, a quarter, and a half year. For CCI, these correlations are much lower, respectively 55%, 76%, and 83%. Even if the forecast task is changed to either forecasting the real-time growth rate or forecasting with final vintage data, the results for CCI do not improve much. Even though the forecast errors reduce considerably in forecasting with final vintage data, by far the most of this reduction is already captured by the autoregressive benchmark and none of the other methods is able to consistently achieve further forecast gains. Second, one may ask what the leading indicators actually do lead. In order to be useful in real-time forecasting of final vintage growth rates of IP and CCI, the (real-time) leading indicators should be correlated well with IP and CCI, both in real time (to get significant estimation results) and with the final vintage values of IP and CCI (to get significant forecast improvements). The correlations of the real-time leading indicators tend to be somewhat larger for IP than for CCI. For instance, at lag one and for growth rates over three or six months, the correlation of CLI with real-time IP is .37 and that with final vintage IP is .41-.45, whereas these values for CCI are respectively .22-.27 and .33-.36. Although these correlations are quite modest and the same holds true for lags larger than one, the leading indicators lead IP somewhat more strongly than the CCI.

The methods compared in this paper are all relatively simple. An interesting question for future research is whether more advanced methods can help in real-time forecasting. However, all methods have to cope with the fact that data revisions of macroeconomic variables are substantial, so that the methods should be sufficiently robust in this respect. For example, the best growth rate forecasts of IP and CCI over the recent recession period since 2008 are obtained by simply regressing the growth rates on the set of ten leading indicators. This method provides forecast gains that are much larger than the ones obtained by the Composite Leading Index or principal components, as the gains differ by a factor of up to more than 2. This illustrates that, in difficult periods, it may be best to use simple models.

References

Anderson, T.W. (1984). *An Introduction to Multivariate Statistical Analysis*, 2-nd ed., New York: John Wiley.

Banerjee, A. & Marcellino, M. (2006). Are there any reliable leading indicators for US inflation and GDP growth? *International Journal of Forecasting*, 22, 137-151.

Boivin, J. & Ng, S. (2006). Are more data always better for factor analysis?. *Journal of Econometrics*, 132, 169-194.

Brier, G.W. (1950). Verification of forecasts expressed in terms of probability. *Monthly Weather Review*, 78, 1-3.

Clements, M.P., & Galvo, A.B. (2006). Combining predictors and combining information in modelling: Forecasting US recession probabilities and output growth. In: C. Milas, P.A. Rothman & D. van Dijk (Eds.), *Nonlinear time series analysis of business cycles.* Amsterdam: Elsevier, pp. 5573.

Conference Board, The (2001). The Business Cycle Indicators Handbook.

Croushore, D. (2006). Forecasting with real-time macroeconomic data. In: G. Elliott, C.W.J. Granger & A. Timmermann (Eds.), *Handbook of Economic Forecasting, Volume 1.* Amsterdam: North-Holland, pp. 961-982.

Croushore, D., & Stark, T. (2001). A real-time data set for macroeconomists. *Journal of Econometrics*, 105, 111-130.

Diebold, F.X. & Mariano, R.S. (1995). Comparing Predictive Accuracy. *Journal of Business and Economic Statistics*, 13, 253-263.

Diebold, F.X. & Rudebusch, G.D. (1991). Forecasting output with the composite leading index: A real-time analysis. *Journal of the American Statistical Association*, 86, 603-610.

Golinelli, R., & Parigi, G. (2008). Real-time squared: A real-time data set for real-time GDP forecasting. *International Journal of Forecasting*, 24, 368-385.

Jolliffe, I.T. (2002). *Principal Component Analysis*, 2-nd ed., Berlin: Springer.

Marcellino, M. (2006). Leading indicators. In: G. Elliott, C.W.J. Granger & A. Timmermann (Eds.), *Handbook of Economic Forecasting, Volume 1.* Amsterdam: North-Holland, pp. 879-960.

Marcellino, M., Stock, J.H., & Watson, M.W. (2006). A comparison of direct and iterated AR methods for forecasting macroeconomic series h-steps ahead. *Journal of Econometrics*, 135, 499-526.

McGuckin, R.H., Ozyildirim, A., & Zarnowitz, V. (2001). Making the composite index of leading economic indicators more timely. In: *Business Cycle Indicators Handbook*. New York: The Conference Board, pp. 37-45.

McGuckin, R.H., Ozyildirim, A., & Zarnowitz, V. (2007). A more timely and useful index of leading indicators. *Journal of Business and Economic Statistics*, 25, 110-120.

Mitchell, W.C., & Burns, A.F. (1938). Statistical Indicators of Cyclical Revivals. *NBER Bulletin* 69, New York.

Newey, W.K., & West, K.D. (1987). A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, 55, 703-708.

Ozyildirim, A., Schaitkin, B., & Zarnowitz, V. (2009). Business cycles in the euro area defined with coincident economic indicators and predicted with leading economic indicators. *Working paper*, The Conference Board.

Pesaran, M.H., & Timmermann, A. (2005). Real-time econometrics. *Econometric Theory*, 21, 212-231.

Stock, J.H. & Watson, M.W. (1999). Forecasting Inflation. *Journal of Monetary Economics*, 44, 293-335.

Stock, J.H. & Watson, M.W. (2002a). "Macroeconomic Forecasting Using Diffusion Indexes. *Journal of Business and Economic Statistics*, 20, 147-162.

Stock, J.H. & Watson, M.W. (2002b). Forecasting Using Principal Components from a Large Number of Predictors. *Journal of the American Statistical Association*, 97, 1167-1179.

Stock, J.H. & Watson, M.W. (2005). An Empirical Comparison of Methods for Forecasting Using Many Predictors. *Working Paper.*

Stock, J.H. & Watson, M.W. (2006). Forecasting with Many Predictors. In: G. Elliott, C.W.J. Granger & A. Timmermann (Eds.), *Handbook of Economic Forecasting, Volume 1.* Amsterdam: North-Holland, pp. 515-554.

Swanson, N.R. (1996). Forecasting using first-available versus fully revised economic time-series data. *Studies in Nonlinear Dynamics and Econometrics*, 1, 47-64.

Swanson, N.R., Ghysels, E. & Callan, M. (1999). A Multivariate Time Series Analysis of the Data Revision Process for Industrial Production and the Composite Leading Indicator. In: R.F. Engle & H. White (Eds.), *Cointegration, Causality and Forecasting: A Festschrift in Honor of Clive W.J. Granger*. Oxford: Oxford University Press, pp. 45-75.

Swanson, N.R. & Van Dijk, D. (2006). Are Reporting Agencies Getting It Right? Data Rationality and Business Cycle Asymmetry, *Journal of Business and Economic Statistics*, 24, 24-42.

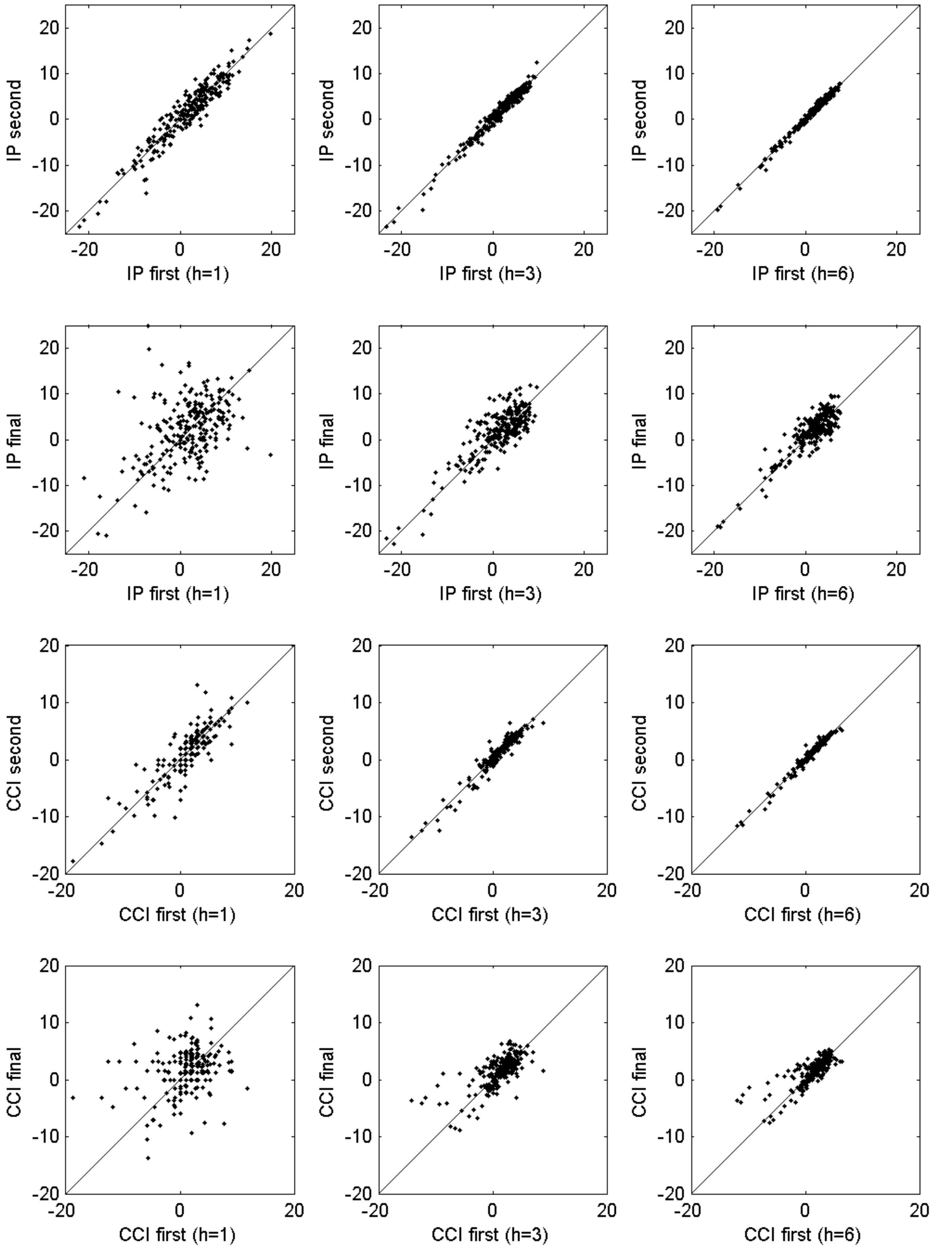

Figure 1: Annualized growth rates of IP (first two rows) and CCI (last two rows) over periods of one month (left), three months (middle), and six months (right), scatter diagrams of second and final release against first (real time) release

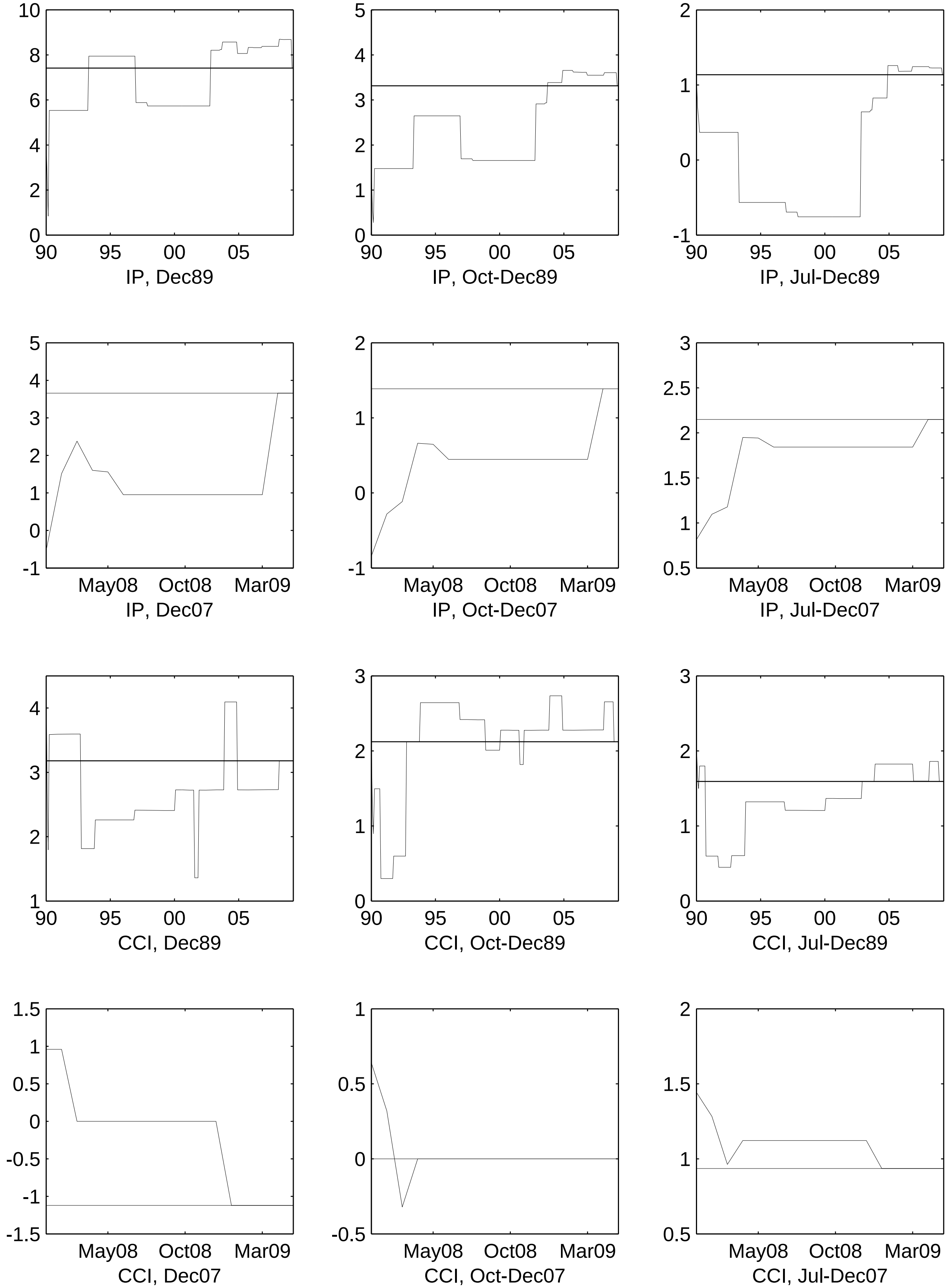

Figure 2: Growth rates over final h months (h=1, 3, 6, from left to right) of 1989 and 2007, for IP (first two rows) and CCI (last two rows), as derived from first to final releases (Feb90-May09 for 1989, and Jan08-May09 for 2007)

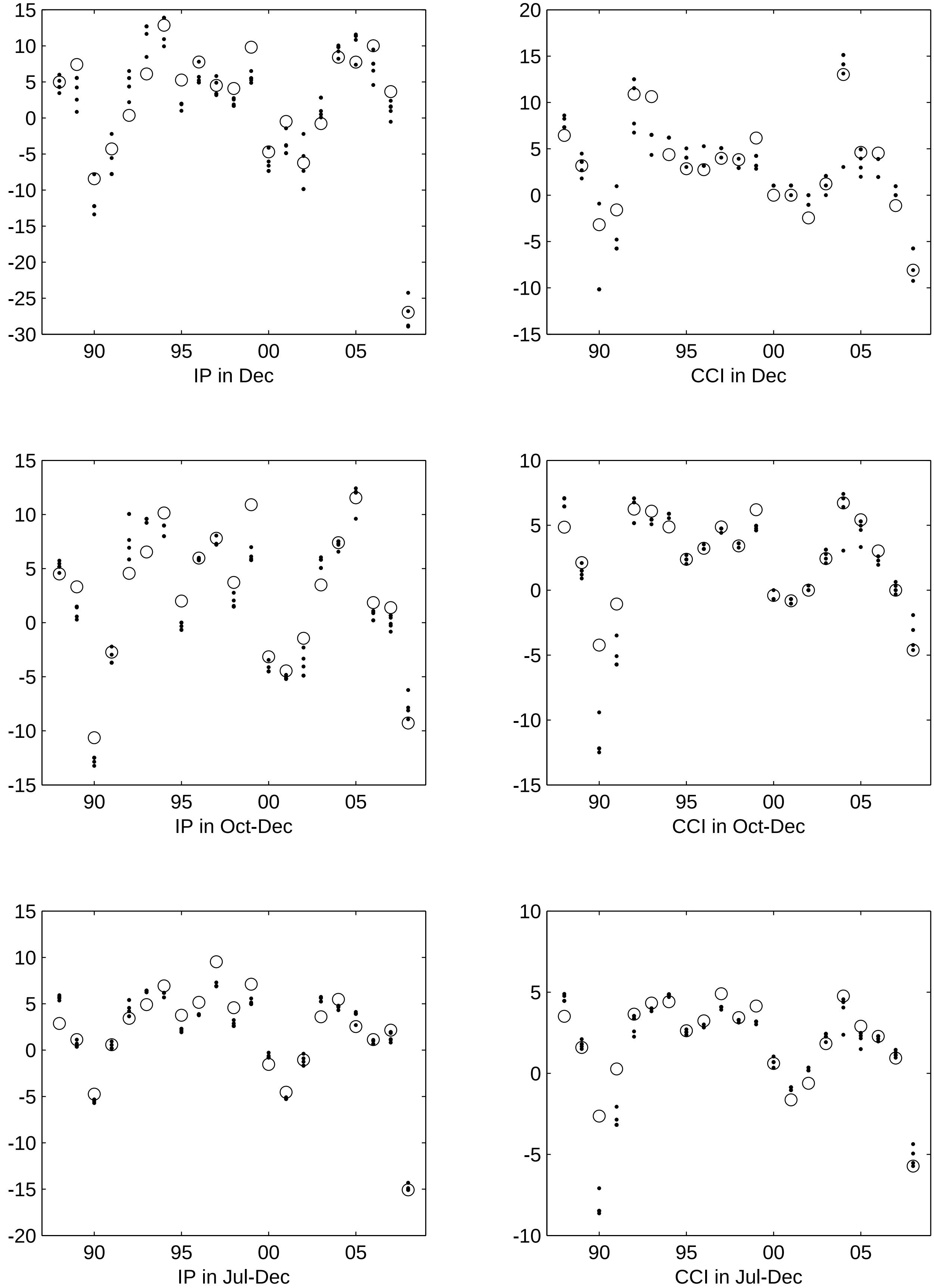

Figure 3: Growth rates over final h months (h=1, 3, 6, from left to right) of each year 1988-2008, for IP (left) and CCI (right), as derived from the first six releases (dots) and from the final release (circles)

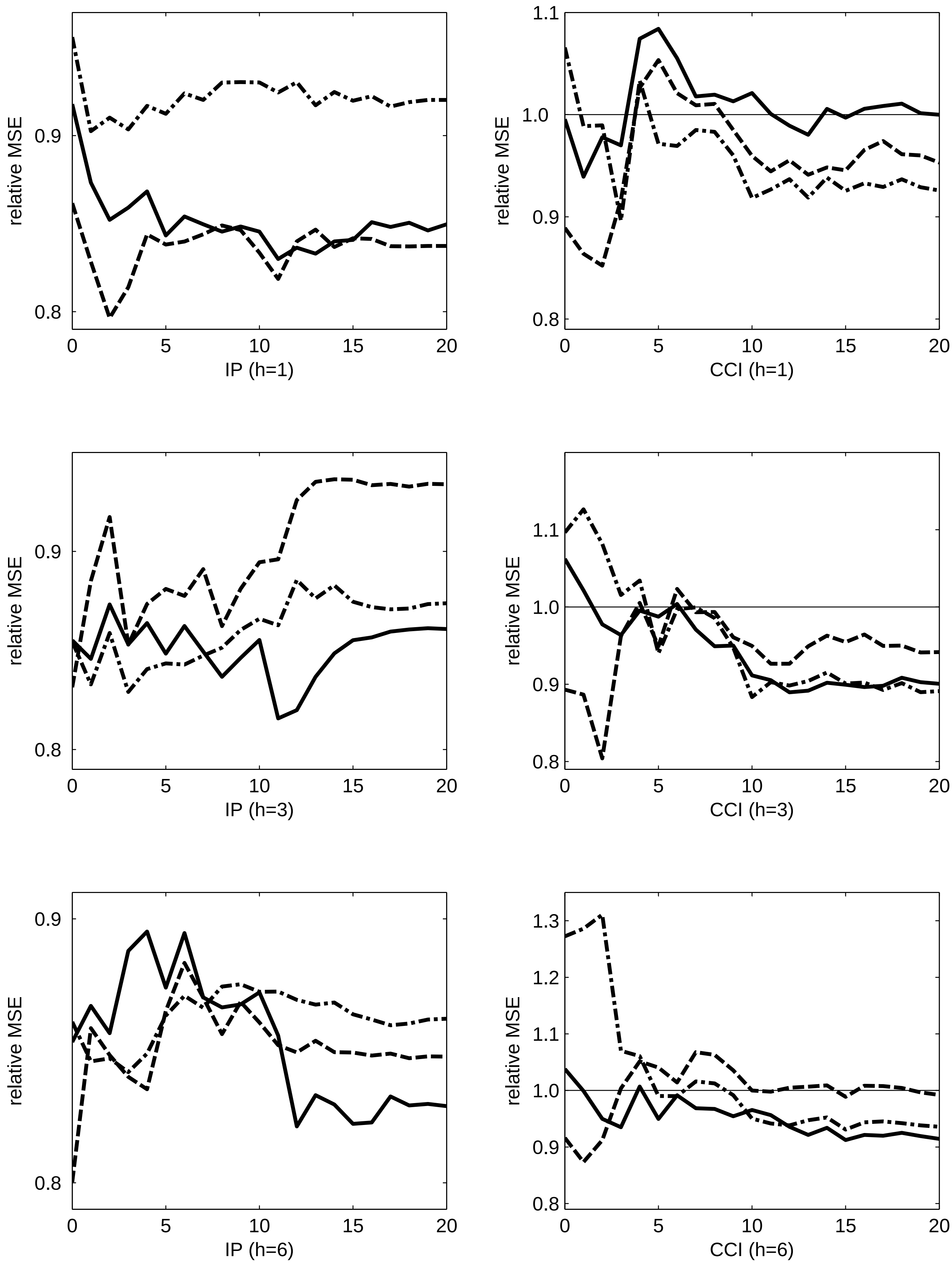

Figure 4: Relative MSE (as compared to the AR benchmark) of growth rate forecasts of IP (left) and CCI (right), for the coming month (top), quarter (middle), and half year (bottom), for vintage shifts from 0 to 20 years, for methods CLI (solid line), OLS (dash-dotted), and PCR (dashed)

Table 1: Vintage data: data period Jan59-Apr09, vintages Jan89-May09

Variable	ConfBrd	DataStream	Delay	Trafo
Leading Indicators				
Average weekly hours, manufacturing (hours per week, SA)	AOM001	USHKIM..O	2, 1	Z(lv)
Average weekly initial claims for unemployment insurance (thousands per week, SA)	AOM005	USUNINSCE	2, 1	Z(dln)
Manufacturers' new orders, consumer goods and materials (millions, chained 96$, SA)	AOM008	USCNORCGD	2, 2	Z(dln)
Manufacturers' new orders, nondefense capital goods (millions, chained 96$, SA)	AOM027	USNOIDN.D	2, 2	Z(dln)
Building permits for new private housing units (thousands, SA)	AOM029	USHOUSATE	2, 1	Z(ln)
Vendor performance, slower deliveries diffusion index (percent, SA)	AOM032	USVENDOR	2, 1	Z(lv)
Money supply, M2 (billions $, SA)	AOM106	USM2....D	2, 2	Z(dln)
S&P 500 common stock index (index: 1941-43=10, NSA)	UOM019	US500STK	2, 1	Z(dln)
Index of consumer expectations, University of Michigan (Jan1966=100, NSA)	UOM083	USUMCONEH	2, 1	Z(dlv)
Interest rate spread, 10-year Treasury bonds less federal funds (percent, NSA)	UOM129	USYSTNFF	2, 1	Z(lv)
Composite Leading Index (CLI)	GOM910	USCYLEAD	2, 1	Z(dln)
Coincident Indicators				
Employees on nonagricultural payrolls (thousands, SA)	AOM041	USEMPNAGE	2, 1	dln
Industrial production (IP) (index: 1992=100, SA)	AOM047	USINPRODG	2, 1	dln
Personal income less transfer payments (billions, chained 96$, SA)	AOM051	USPILESTD	2, 2	dln
Manufacturing and trade sales (millions, chained 96$, SA)	AOM057	USBSSALED	3, 3	dln
Composite Coincident Index (CCI)	GOM920	USCOININ	2, 1	dln

Notes

* The column 'ConfBrd' shows the variable codes used by the Conference Board.
* For most variables, the delay of the first release is 2 months up to November 2000 and 1 month afterwards; for four variables, the delay is 2 months throughout, and for one variable, the delay is three months throughout.
* All variables are transformed to stationarity in the same way as in Stock and Watson (2005); 'Z' indicates z-scores (standardization per vintage), 'd' denotes first-differencing, 'ln' is the natural logarithm, and 'lv' stands for 'leave as is'.
* The coincident indicators are transformed to annualized monthly growth percentages (1200*dln).
* Breaks between vintages are removed by the data transformations; outliers are not removed.
* For all variables, 245 vintages are available (Jan89-May09), mostly with 604 observations (Jan59-Apr09); the number of observations is 603 for four variables with delay 2, and 602 for one variable with delay 3.

Table 2: Correlations of CLI, IP, and CCI (data period Jan59-Apr09, vintages Jan89-May09)

Table 2.a: Correlations of IP and CCI for first, second, and final release

horizon	1		3		6	
release	2	final	2	final	2	final
IP						
1	0.95	0.85	0.99	0.90	0.99	0.91
2		0.88		0.91		0.92
CCI						
1	0.86	0.55	0.96	0.76	0.99	0.83
2		0.72		0.82		0.85

Table 2.b: Real-time correlations of (lags of) CLI with CLI, IP, and CCI

lag of CLI	0	1	2	3	4	5
CLI						
Average	1	0.20	0.24	0.08	0.08	-0.01
Final	1	0.28	0.28	0.25	0.21	0.29
IP						
Average						
h = 1	0.29	0.17	0.36	0.23	0.22	0.12
h = 3	0.41	0.37	0.39	0.27	0.26	0.19
h = 6	0.40	0.37	0.35	0.27	0.25	0.23
Final						
h = 1	0.24	0.34	0.37	0.17	0.30	0.36
h = 3	0.45	0.41	0.40	0.38	0.45	0.41
h = 6	0.46	0.45	0.42	0.37	0.36	0.35
CCI						
Average						
h = 1	0.40	0.10	0.26	0.18	0.19	0.14
h = 3	0.35	0.22	0.28	0.25	0.25	0.23
h = 6	0.34	0.27	0.29	0.26	0.25	0.26
Final						
h = 1	0.32	0.35	0.19	0.25	0.27	0.41
h = 3	0.37	0.33	0.30	0.39	0.39	0.41
h = 6	0.41	0.36	0.36	0.34	0.33	0.34

Notes

* In Table 2.a, IP (CCI) at horizon h is the annualized percentage growth rate of IP (CCI) over the past h months. For each month (Nov88-Mar09), three values are compared: the first release (vintage Jan89-Apr09), the second release (vintage Feb89-May09), and the final release (vintage May09). The table shows the correlations between the resulting three time series (of length 244).
* Table 2.b shows correlations of (lags of) CLI with CLI, IP, and CCI, for growth rates of IP and CCI over horizons of h =1, h = 3, and h=6 months. Two correlations are shown, 'Average' for the average over all vintages (Jan89-May09) and 'Final' for the final vintage (May09).

Table 3: Error statistics of real time forecasts of IP and CCI

Method	M	AR	CLI	VAR m-	VAR m+	VAR h-	VAR h+	OLS	PCR	Gain %
IP										
h = 1 (var = 61.91)										
(Relative) MSE	***61.49***	63.00	0.92	0.92	0.92	0.92	0.92	0.96	**0.86**	12
mean error	-1.09	-0.57	-0.57	-0.61	-0.61	-0.61	-0.61	0.60	-0.61	
variance error	60.55	62.92	57.74	57.97	57.97	57.97	57.97	60.11	54.13	
h = 3 (var = 29.65)										
(Relative) MSE	29.72	***24.89***	0.86	1.08	1.02	0.89	0.93	0.85	**0.83**	17
mean error	-1.06	-0.47	-0.43	-0.86	-0.81	-0.34	-0.42	0.51	-0.68	
variance error	28.71	24.77	21.19	26.21	24.76	22.24	23.15	21.10	20.31	
h = 6 (var = 21.54)										
(Relative) MSE	22.17	***21.82***	0.85	0.95	0.86	0.89	0.95	0.86	**0.80**	20
mean error	-0.96	-0.64	-0.56	-0.86	-0.74	-0.38	-0.31	0.72	-0.54	
variance error	21.33	21.50	18.38	20.10	18.27	19.31	20.70	18.34	17.24	
CCI										
h = 1 (var = 13.19)										
(Relative) MSE	***13.28***	13.79	1.00	0.93	0.93	0.93	0.93	1.06	**0.90**	7
mean error	-0.63	0.00	0.12	-0.25	-0.25	-0.25	-0.25	0.64	-0.20	
variance error	12.95	13.84	13.77	12.84	12.84	12.84	12.84	14.25	12.36	
h = 3 (var = 6.92)										
(Relative) MSE	7.09	***6.38***	1.06	0.98	1.05	0.96	1.09	1.08	**0.89**	11
mean error	-0.63	-0.04	0.02	-0.44	-0.41	-0.06	-0.06	0.56	-0.23	
variance error	6.72	6.40	6.80	6.11	6.53	6.14	6.98	6.62	5.65	
h = 6 (var = 5.32)										
(Relative) MSE	5.59	***5.09***	1.04	0.99	0.99	0.97	1.33	1.27	**0.93**	7
mean error	-0.59	-0.15	-0.03	-0.48	-0.41	-0.10	-0.07	0.64	-0.31	
variance error	5.27	5.09	5.31	4.85	4.89	4.97	6.81	6.10	4.64	

Notes

* Target variable is the growth rate of IP (top half) or CCI (bottom half) as reported in the final vintage of May09. The annualized percentage growth rates of IP and CCI are forecasted for the coming period of h months (h=1, 3, or 6), using real-time data at forecast origins from Jan89 to May09-h. The first forecast uses vintage Jan89 (with data up to Nov88) to forecast the growth over the h months following Nov 88, and the final forecast uses vintage May09-h (data up to Apr09-h) to forecast the growth over the h months ending in Apr09. The number of forecasts for horizon h is equal to 245-h, and 'var' in the first column denotes the variance of the target variable over this forecast period.
* Methods M and OLS use no lags (q = 0, r = -1). For the seven other methods, the number of (V)AR lags, the number of factor lags, and the number of factors (for PCR) are all at most 3, and the actual number of lags and factors is selected for each forecast origin by BIC. All forecast models are re-estimated and selected at each forecast origin, using the past 10 years of data up to the forecast origin (rolling estimation window with 120 monthly observations).
* At each forecast origin, the three leading indicators with delay 2 are extrapolated for one month by an AR(2) model estimated from past data (similar to the procedure of the Conference Board).
* Error statistics are reported for nine methods. The MSE of M (past mean) and AR report the mean squared prediction error, the RMSE of the other methods are MSE's divided by the MSE of the AR benchmark. 'Gain %' shows the percentage gain of the best performing method (RMSE in bold) compared to AR, or compared to M if this has lower MSE than AR (indicated in bold italic).

Table 4: Mean squared errors of real time forecasts of IP and CCI, for various sub-periods

Method	AR	CLI	VAR m-	VAR m+	VAR h-	VAR h+	OLS	PCR	Gain %
IP									
h = 1 (var = 61.91)									
89-09 (# = 245)	63.00	0.92	0.92	0.92	0.92	0.92	0.96	**0.86**	14
89-93 (# = 60)	41.25	**0.79**	0.91	0.91	0.91	0.91	1.27	0.89	21
94-98 (# = 60)	48.54	0.91	1.00	1.00	1.00	1.00	0.91	**0.85**	15
99-03 (# = 60)	41.25	0.96	**0.80**	**0.80**	**0.80**	**0.80**	0.96	0.90	20
04-08 (# = 60)	108.27	0.92	**0.88**	**0.88**	**0.88**	**0.88**	0.88	0.95	12
LAST 12 (# = 12)	419.11	0.90	0.98	0.98	0.98	0.98	**0.71**	0.75	29
h = 3 (var = 29.65)									
89-09 (# = 242)	24.89	0.86	1.08	1.02	0.89	0.93	0.85	**0.83**	17
89-93 (# = 60)	19.88	0.69	0.73	**0.65**	0.91	0.96	1.05	0.80	35
94-98 (# = 60)	16.61	0.75	0.90	0.84	0.91	0.94	0.92	**0.72**	28
99-03 (# = 60)	25.72	0.86	0.79	0.74	**0.59**	0.60	0.80	0.74	41
04-08 (# = 60)	36.42	0.94	1.30	1.24	1.00	1.06	**0.72**	1.00	28
LAST 12 (# = 12)	128.10	0.94	1.94	1.82	1.12	1.17	**0.63**	0.94	37
h = 6 (var = 21.54)									
89-09 (# = 239)	21.82	0.85	0.95	0.86	0.89	0.95	0.86	**0.80**	20
89-93 (# = 60)	14.71	0.81	**0.73**	0.76	1.18	1.61	1.19	0.80	27
94-98 (# = 60)	11.53	0.77	0.94	0.85	0.80	0.82	0.92	**0.73**	27
99-03 (# = 60)	24.06	0.82	0.75	**0.70**	0.79	0.82	0.76	0.86	30
04-08 (# = 59)	37.24	0.92	1.18	1.01	0.86	0.81	**0.79**	0.86	21
LAST 12 (# = 12)	154.94	0.89	1.30	1.09	0.84	0.78	**0.67**	0.86	33
CCI									
h = 1 (var = 13.19)									
89-09 (# = 245)	13.79	1.00	0.93	0.93	0.93	0.93	1.06	**0.90**	10
89-93 (# = 60)	19.05	1.06	**0.74**	**0.74**	**0.74**	**0.74**	1.40	0.80	26
94-98 (# = 60)	9.85	**0.91**	1.00	1.00	1.00	1.00	1.01	0.96	9
99-03 (# = 60)	7.68	1.04	1.18	1.18	1.18	1.18	1.19	**1.04**	-4
04-08 (# = 60)	17.15	0.89	0.89	0.89	0.89	0.89	**0.70**	0.97	30
LAST 12 (# = 12)	44.11	0.84	1.19	1.19	1.19	1.19	**0.52**	0.73	48
h = 3 (var = 6.92)									
89-09 (# = 242)	6.38	1.06	0.98	1.05	0.96	1.09	1.08	**0.89**	11
89-93 (# = 60)	10.96	1.25	**0.44**	0.62	0.97	1.21	1.47	0.77	56
94-98 (# = 60)	2.58	0.84	1.11	1.04	0.89	0.98	0.99	**0.83**	17
99-03 (# = 60)	4.59	0.94	1.31	1.29	**0.79**	0.84	1.19	1.10	21
04-08 (# = 60)	7.00	0.93	1.36	1.34	1.00	1.06	**0.50**	0.97	50
LAST 12 (# = 12)	22.31	0.87	2.11	2.05	1.18	1.22	**0.31**	0.88	69
h = 6 (var = 5.32)									
89-09 (# = 239)	5.09	1.04	0.99	0.99	0.97	1.33	1.27	**0.93**	7
89-93 (# = 60)	7.60	1.28	**0.45**	0.60	1.25	2.08	2.03	0.90	55
94-98 (# = 60)	1.99	0.85	1.10	1.02	**0.73**	0.81	1.12	0.95	27
99-03 (# = 60)	4.40	0.93	1.25	1.22	**0.75**	0.88	1.04	1.11	25
04-08 (# = 59)	6.40	0.89	1.43	1.30	0.88	0.92	**0.57**	0.82	43
LAST 12 (# = 12)	24.99	0.81	1.63	1.44	0.88	0.91	**0.32**	0.73	68

Notes

* Target variable is the final vintage growth rate (as reported in final vintage of May09) of IP (top half) and CCI (bottom half). The annualized percentage growth rates are forecasted for the coming h = 1, 3, or 6 months, using real-time data at forecast origins from Jan89 to May09-h.
* The table values for AR report the mean squared prediction error. The values of the other methods are MSE's divided by the MSE of the AR benchmark. 'Gain %' shows the percentage gain of the best performing method (RMSE in bold) compared to AR.
* See Table 3 for further details on the estimation and selection of forecast models.

Table 5: Mean squared errors of forecasts of IP and CCI for various data specifications

Data	AR	CLI	VAR m-	VAR m+	VAR h-	VAR h+	OLS	PCR	Gain %
IP									
h = 1 (var = 61.91)									
X all real, target final	63.00	0.92	0.92	0.92	0.92	0.92	0.96	**0.86**	14
X all real, target real	50.09	0.95	0.91	0.91	0.91	0.91	0.94	**0.88**	12
(C)LI real, IP final, target final	63.42	0.88	0.91	0.91	0.91	0.91	0.92	**0.84**	16
X all final, target final	63.42	0.85	0.89	0.89	0.89	0.89	0.92	**0.84**	16
h = 3 (var = 29.65)									
X all real, target final	24.89	0.86	1.08	1.02	0.89	0.93	0.85	**0.83**	17
X all real, target real	23.25	0.89	1.12	1.06	0.88	0.92	**0.88**	0.89	12
(C)LI real, IP final, target final	21.69	0.90	1.20	1.13	**0.86**	0.89	0.87	0.93	14
X all final, target final	21.69	0.86	1.18	1.06	**0.85**	0.88	0.87	0.93	15
h = 6 (var = 21.54)									
X all real, target final	21.82	0.85	0.95	0.86	0.89	0.95	0.86	**0.80**	20
X all real, target real	20.72	**0.88**	0.97	0.88	0.89	0.96	0.95	0.88	12
(C)LI real, IP final, target final	20.01	0.91	1.00	0.90	**0.79**	0.84	0.85	0.86	21
X all final, target final	20.01	0.84	0.99	0.85	**0.78**	0.84	0.85	0.86	22
CCI									
h = 1 (var = 13.19)									
X all real, target final	13.79	1.00	0.93	0.93	0.93	0.93	1.06	**0.90**	10
X all real, target real	11.78	1.00	1.03	1.03	1.03	1.03	1.10	**0.92**	8
(C)LI real, CCI final, target final	11.91	0.98	1.01	1.01	1.01	1.01	**0.92**	0.96	8
X all final, target final	11.91	1.00	0.99	0.99	0.99	0.99	**0.93**	0.95	7
h = 3 (var = 6.92)									
X all real, target final	6.38	1.06	0.98	1.05	0.96	1.09	1.08	**0.89**	11
X all real, target real	6.87	0.91	1.11	1.05	0.87	1.01	**0.82**	0.83	18
(C)LI real, CCI final, target final	4.44	0.96	1.36	1.34	0.94	0.99	**0.89**	0.99	11
X all final, target final	4.44	0.90	1.34	1.29	0.92	0.97	**0.89**	0.94	11
h = 6 (var = 5.32)									
X all real, target final	5.09	1.04	0.99	0.99	0.97	1.33	1.27	**0.93**	7
X all real, target real	6.80	0.98	1.04	0.99	0.90	1.17	**0.84**	0.88	16
(C)LI real, CCI final, target final	3.57	0.96	1.38	1.29	**0.87**	1.01	0.97	1.02	13
X all final, target final	3.57	0.91	1.37	1.25	**0.86**	1.04	0.94	0.99	14

Notes

* The table values for AR report the mean squared prediction error. The values of the other methods are MSE's divided by the MSE of the AR benchmark. 'Gain %' shows the percentage gain of the best performing method (RMSE in bold) compared to AR.
* Target variable is the IP or CCI growth rate, either the final value (as reported in the final, May09 vintage) or the first release value as reported in real time.
* Explanatory variables X (lags of IP or CCI and (lags of) CLI, leading indicators, and principal components) are either all real time (based on the most recent vintage available in real time), or with IP or CCI taken from the final vintage but with CLI, leading indicators, and principal components in real time, or all final vintage (values of May09 vintage).
* See Table 3 for further details on the estimation and selection of forecast models.

Table 6.a: Diebold-Mariano tests on IP forecast accuracy for different methods and data specifications

Table 6.a.1: Comparison of methods for IP

Data	AR vs CLI	AR vs VARm-	AR vs VARh-	AR vs OLS	AR vs PCR	OLS vs CLI	OLS vs PCR	PCR vs CLI
h = 1								
X all real, target final	1.86**	1.96**	1.96**	0.75	2.45***	0.20	1.00	-0.92
X all real, target real	1.00	2.17**	2.17**	0.65	2.00**	-0.12	0.87	-0.97
(C)LI real, IP final, target final	2.27**	1.76**	1.76**	0.91	2.59***	0.35	1.30*	-0.85
X all final, target final	2.83***	2.20**	2.20**	0.89	2.32**	0.83	1.36*	-0.18
h = 3								
X all real, target final	2.32**	-0.38	1.23	1.51*	2.77***	-0.22	-0.08	-0.22
X all real, target real	1.43*	-0.55	1.14	0.93	1.71**	-0.07	-0.08	0.03
(C)LI real, IP final, target final	1.57*	-0.80	1.57*	0.97	0.83	-0.18	-0.59	0.90
X all final, target final	1.72**	-0.76	1.78**	1.06	0.91	0.14	-0.53	1.01
h = 6								
X all real, target final	2.13**	0.44	0.89	0.89	2.22**	0.13	0.33	-0.44
X all real, target real	1.61*	0.26	0.74	0.29	1.24	0.62	0.59	0.15
(C)LI real, IP final, target final	1.52*	0.01	1.50*	0.76	1.33*	-0.30	-0.34	0.13
X all final, target final	1.95**	0.07	1.48*	0.90	1.89**	0.08	-0.06	0.22

Table 6.a.2: Comparison of data specifications for IP

Comparison	AR	CLI	VARm-	VARh-	OLS	PCR
h = 1						
X all real vs only (C)LI real (target final)	-0.14	0.89	0.45	0.45	0.53	1.52*
only (C)LI real vs X all final (target final)		1.29*	1.42*	1.42*	-0.24	-0.41
X all real vs X all final (target final)	-0.14	1.74**	1.29*	1.29*	0.25	0.76
h = 3						
X all real vs only (C)LI real (target final)	2.98***	1.85**	1.96**	3.50***	1.59*	0.52
only (C)LI real vs X all final (target final)		0.83	1.12	0.98	0.27	0.45
X all real vs X all final (target final)	2.98***	2.02**	1.98**	3.31***	1.51*	0.66
h = 6						
X all real vs only (C)LI real (target final)	2.00**	0.46	1.86**	2.86***	1.10	-0.37
only (C)LI real vs X all final (target final)		1.70**	0.95	0.21	0.81	1.57*
X all real vs X all final (target final)	2.00**	1.50*	2.18**	2.74***	1.37*	1.13

Notes

* Table shows t-values of Diebold-Mariano test with HAC standard errors; *** (**, *) denotes significantly better forecast accuracy at the 1% (5%, 10%) significance level (according to the t-distribution).
* X denotes the set of explanatory variables, that is, lags of IP and (lags of) CLI and the leading indicators.
* Tests compare two sequences of squared forecast errors, obtained by different methods applied on real time vintage data to forecast final vintage IP growth rates (Table 6.a.1) or obtained by each of six methods for different data specifications (Table 6.a.2).
* Table 6.a.1 tests whether (i) each of five methods (CLI, VARm-, VARh-, OLS, and PCR) is more accurate than the AR benchmark, (ii) CLI and PCR are more accurate than OLS, and (iii) CLI is more accurate than PCR.
* Table 6.a.2 tests, separately for each of six methods to forecast final vintage values of IP, whether (i) if (C)LI (i.e., CLI and the leading indicators) are taken in real time, then it is harder to forecast with lags of real time IP than with lags of final vintage IP, (ii) if lags of IP are taken from the final vintage, then it is harder to forecast with real time (C)LI than with final vintage data of (C)LI, and (iii) it is harder to forecast with all explanatory variables (lags of IP and (C)LI) taken in real time than with all data taken from the final vintage.
* See Table 3 for further details on the estimation and selection of forecast models, and see Table 5 for a description of the various data specifications.

Table 6.b: Diebold-Mariano tests on CCI forecast accuracy for different methods and data specifications

Table 6.b.1: Comparison of methods for CCI

Data	AR vs CLI	AR vs VARm-	AR vs VARh-	AR vs OLS	AR vs PCR	OLS vs CLI	OLS vs PCR	PCR vs CLI
h = 1								
X all real, target final	0.09	1.29*	1.29*	-0.62	1.82**	0.68	1.77**	-1.61**
X all real, target real	-0.07	-0.34	-0.34	-0.94	1.09	0.95	1.64**	-0.84
(C)LI real, CCI final, target final	0.66	-0.21	-0.21	1.01	0.94	-0.71	-0.53	-0.37
X all final, target final	0.00	0.15	0.15	1.04	1.05	-0.98	-0.38	-0.92
h = 3								
X all real, target final	-0.71	0.08	0.53	-0.54	0.73	0.15	1.07	-1.27
X all real, target real	0.68	-0.63	1.38*	0.94	1.34*	-0.62	-0.09	-0.67
(C)LI real, CCI final, target final	0.94	-1.40	0.83	0.90	0.09	-0.60	-0.97	0.53
X all final, target final	1.77**	-1.39	1.14	0.90	0.78	-0.09	-0.52	0.50
h = 6								
X all real, target final	-0.50	0.03	0.22	-0.94	0.44	1.04	1.14	-0.68
X all real, target real	0.28	-0.37	1.84**	0.53	0.70	-0.55	-0.24	-0.65
(C)LI real, CCI final, target final	0.84	-1.29	1.73**	0.12	-0.19	0.07	-0.28	0.55
X all final, target final	1.24	-1.29	1.82**	0.29	0.06	0.13	-0.37	0.62

Table 6.b.2: Comparison of data specifications for CCI

Comparison	AR	CLI	VARm-	VARh-	OLS	PCR
h = 1						
X all real vs only (C)LI real (target final)	2.43***	2.86***	2.11**	2.11**	4.15***	1.45*
only (C)LI real vs X all final (target final)		-0.99	1.88**	1.88**	-0.07	0.48
X all real vs X all final (target final)	2.43***	2.29**	2.41***	2.41***	3.85***	1.49*
h = 3						
X all real vs only (C)LI real (target final)	2.06**	2.74***	0.80	2.56***	2.90***	2.32**
only (C)LI real vs X all final (target final)		1.96**	1.16	1.57*	-0.16	1.72**
X all real vs X all final (target final)	2.06**	2.94***	0.97	2.61***	2.83***	2.51***
h = 6						
X all real vs only (C)LI real (target final)	2.20**	2.40***	0.63	1.75**	2.07**	1.90**
only (C)LI real vs X all final (target final)		1.28	1.05	1.07	1.05	1.08
X all real vs X all final (target final)	2.20**	2.54***	0.73	1.83**	2.09**	2.12**

Notes

* This table is similar to Table 6.a, now with CCI instead of IP.
* X denotes the set of explanatory variables, that is, lags of CCI and (lags of) CLI and the leading indicators.
* See Table 6.a for further explanation of the table set-up.

Table 7: Real-time predicted probabilities of expansions and recessions

Table 7.a: Probabilities of expansion, recession, and mixed periods

Method	#	AR	CLI	VAR m-	VAR h-	OLS	PCR
Prob(Expansion)							
All	239	0.60	0.58	0.65	0.58	0.51	0.61
Real expansion	172	**0.62**	**0.61**	**0.65**	**0.62**	**0.55**	**0.65**
Real recession	28	0.50	0.44	0.64	0.42	0.31	0.43
Real mixed	39	0.57	0.57	0.66	0.54	0.43	0.58
Prob(Recession)							
All	239	0.15	0.17	0.11	0.15	0.25	0.15
Real expansion	172	0.13	0.14	0.11	0.13	0.20	0.12
Real recession	28	**0.21**	**0.26**	**0.12**	**0.26**	**0.46**	**0.25**
Real mixed	39	0.18	0.19	0.12	0.19	0.32	0.19
Prob(Mixed)							
All	239	0.25	0.25	0.24	0.26	0.24	0.24
Real expansion	172	0.24	0.25	0.24	0.25	0.25	0.23
Real recession	28	0.29	0.30	0.24	0.32	0.23	0.31
Real mixed	39	**0.25**	**0.24**	**0.22**	**0.26**	**0.25**	**0.23**

Table 7.b: Recession probabilities at early, middle, and late stages of NBER recessions

Method	#	AR	CLI	VAR m-	VAR h-	OLS	PCR
Jul90-Feb91	8						
E: Feb90-Jun90	5	0.19	0.21	0.18	0.25	0.49	0.25
M: Jul90-Sep90	3	0.23	0.26	0.22	0.26	0.62	0.26
L: Oct90-Feb91	5	0.47	0.64	0.30	0.61	0.89	0.59
Mar01-Oct01	8						
E: Oct00-Feb01	5	0.09	0.12	0.07	0.11	0.17	0.07
M: Mar01-Apr01	2	0.11	0.12	0.08	0.22	0.13	0.08
L: May01-Sep01	5	0.13	0.13	0.09	0.27	0.06	0.11
Dec07-Apr09	17						
E: Jun07-Nov07	5	0.09	0.11	0.07	0.09	0.17	0.07
M: Dec07-Oct08	11	0.24	0.31	0.09	0.27	0.60	0.32
Average							
E	15	0.13	0.15	0.11	0.15	0.28	0.14
M	16	0.20	0.26	0.09	0.24	0.48	0.25
L	10	0.30	0.39	0.20	0.44	0.48	0.35

Table 7.c: QPS and ROC statistics for empirical and NBER recessions

Method	#	AR	CLI	VAR m-	VAR h-	OLS	PCR
QPS							
Empirical							
Expan. & Reces.	200	0.118	0.114	0.123	0.108	0.125	**0.110**
Expansion	172	0.033	0.039	**0.015**	0.034	0.083	0.030
Recession	28	0.642	0.578	0.783	0.562	**0.383**	0.595
Last recession	13	0.552	0.453	0.817	0.510	**0.176**	0.448
NBER							
All	239	0.113	0.111	0.109	0.109	0.139	**0.108**
Expansion	211	0.038	0.047	**0.017**	0.043	0.104	0.040
Recession	28	0.676	0.598	0.806	0.601	**0.403**	0.619
Last recession	13	0.552	0.453	0.817	0.510	**0.176**	0.448
ROC							
Empirical	239	0.735	0.749	0.523	**0.808**	0.726	0.727
NBER	239	0.571	0.619	0.426	**0.673**	0.639	0.620

Notes

* Table 7.a shows probabilities of expansions, recessions, and mixed periods, estimated in real time by six methods. Expansions, recessions, and mixed periods are empirically defined in terms of the (final release) CCI growth rates, that in the coming two quarters should be respectively both positive, both negative, and one positive and the other negative. Bold values indicate probabilities for the actually prevailing regime.

* Table 7.b shows estimated recession probabilities for NBER recessions, in early stages (from 5 to 1 month before start), late stages (5 to 1 month before end), and middle stages (all 6 coming month in recession period). 'Average' is the average over the three NBER recessions.

* Table 8.c shows quadratic probability scores (QPS) and receiver operating characteristics (ROC) of estimated probabilities, both for the empirically defined recessions and for NBER recessions. Bold values show best performing methods (smallest QPS or largest ROC) .

* See Table 3 for further details on the estimation and selection of forecast models. The text contains additional information on the definition of recessions and on computation methods.