# Are there any reliable leading indicators for US Inflation and GDP Growth?*

Anindya Banerjee¶
Department of Economics
European University Institute
Via della Piazzuola, 43
50133, Firenze
anindya.banerjee@iue.it

Massimiliano Marcellino
IEP-Bocconi University, IGIER and CEPR

Via Salasco, 5
20136, Milano
massimiliano.marcellino@uni-bocconi.it

February 2005

Abstract

In this paper we evaluate the relative merits of three alternative approaches to information extraction from a large data set for forecasting, namely, the use of an automated model selection procedure, the adoption of a factor model, and of single-indicator-based forecast pooling. The comparison is conducted using a large set of indicators for forecasting US inflation and GDP growth. We also compare our large set of leading indicators with purely autoregressive models, using an evaluation procedure that is particularly relevant for policy making. The evaluation is conducted both ex post and in a pseudo real-time context, for several forecast horizons, and using both recursive and rolling estimation. The results indicate a preference for simple forecasting tools, with a good relative performance of pure autoregressive models, and substantial instability in the characteristics of the leading indicators. A pseudo real-time analysis provides useful indications for the selection of the best leading indicator, in particular for GDP growth.

*JEL Classification:* C53, E37, C50

*Keywords:* leading indicator, factor model, model selection, GDP growth, inflation

---

¶ Corresponding author

* The authors thank the Editor Mike Clements, an anonymous Associate Editor, an anonymous referee of the journal, and David Hendry for their comments. We have also benefited from the views of participants in seminars at the Banque de France, Oxford, Universidad Carlos III – Madrid, the European Central Bank, and the Fundación Economía Aragonesa – Zaragoza. We are grateful to Igor Masten for his valuable assistance with the data and estimation in this paper, to Charles Steindel for providing the data in Cecchetti *et al.* (2000), to Anirvan Banerji, Director of Research at the Economic Cycle Research Institute (ECRI), for his insightful comments on the use of indexes such as the JOC-ECRI index and the Future Inflation Gauge (FIG), and to ECRI for providing their historic data for FIG free of charge. The financial generosity of the EUI under research grants EFN 402 and I 216 and of the Italian MIUR is also acknowledged.

## 1. Introduction

Methods for forecasting inflation and output growth have been the subject of intensive research in econometrics. Recent papers include the use of univariate leading indicator models in forecasting US inflation (Cecchetti, Chu and Steindel (2000)), factor models for forecasting US inflation (Stock and Watson (1999a)) and automated procedures for forecasting GDP growth using systems of leading indicators (Camba-Mendez, Kapetanios, Smith and Weale (2001)).

The question of the choice of indicators and models is particularly significant, given the easy availability of ever-increasingly large data sets. We focus on single-equation methods for forecasting US inflation and GDP growth using leading indicators and factor models, where an important feature of our analysis is a consideration of variable selection in a large-dimensional data set.

We take as our starting point the paper by Cecchetti *et al.* who show that models for forecasting inflation using indicators taken individually are often out-performed by simple autoregressions (where inflation forecasts are based only on past values of inflation). These indicators are broadly classified into price, financial and real variables. The Cecchetti *et al.* analysis is based upon looking at the forecasting performance of models where each of the indicators is incorporated individually into autoregressions of inflation to determine if the accuracy of the inflation forecast is improved. The models operate on the basis of fixed lag lengths for all regressions and a particular root-mean-squared-error criterion is used to judge forecasting accuracy (denoted RMSE-*h* and described in detail below).

We generalise the Cecchetti *et al*. analysis in three essential ways. Firstly, we allow for the selection of the 'best' leading indicator and the appropriate lag length using a model selection procedure developed by Hendry and Krolzig (1999) that relies on the joint application of information criteria, significance testing on the parameters, and residual based tests for correct model specification. The procedure is implemented with their software *PcGets*. Especially within the context of large-dimensional data sets, we are saved a considerable amount of effort by not having to input the indicators individually but by allowing the programme to choose the best fitting model. We consider this to be a fundamental and essential step in handling large data sets for forecasting and think of *PcGets* as one such method of model selection.

Secondly, we also consider groups of indicators, where grouping is based either on economic considerations or on the forecasting performance of the single indicators. In this context, where the starting model can be very large, the use of *PcGets* as a model specification device is even more important.

Thirdly, we consider pooling the single indicator forecasts. Pooled forecasts have been shown to perform very well for macroeconomic variables (see *e.g.* Stock and Watson (1999b) and Hendry and Clements (2004) for recent assessments). In our context, forecast pooling can also be considered as an alternative tool for information extraction from large datasets, where the information is combined not in sample but directly out of sample.

The final contribution of our paper is to re-assess the usefulness of factor models in forecasting inflation and GDP growth, a method that has a long tradition in macroeconomic forecasting and has been recently re-proposed by Stock and Watson (2002) and others. Factor models extract and summarise information by the use of principal components and are seen, within the framework of our research here, as an alternative to *PcGets* as methods of dealing with large-dimensional data sets. In contrast with the latter approach where variable selection is an important part of the analysis, factor analysis extracts the main driving factors from the entire data set and the factors themselves can usually not be given natural or self-evident economic interpretations.

The forecast comparison is conducted using both an ex-post and a pseudo ex-ante approach. In the ex-post evaluation, future values of the exogenous regressors are assumed known, and the grouping of the leading indicators is based on the overall (average over all the periods) forecasting performance of the single indicators. This kind of evaluation is relevant in a policy making environment where there is interest in detecting the component of the forecast error due to the assumptions on the development of the exogeneous variables, see e.g. Keereman (1999, 2003). Moreover, treating the future values of the indicators as known provides the maximum advantage *against* the autoregressive models but, as we will see, in many cases it is not enough to beat them.

In the ex-ante framework, no future information is used, future values of the regressors are forecast, and the choice of the indicators is based on their past forecasting records. This provides an indication for the construction of feasible leading indicator forecasts.

It is also worth pointing out that we follow Cecchetti *et al.* in the choice of a loss function that is not standard in the academic literature on forecasting but likely reflects the current practice in policy making institutions. Specifically, for each estimation period we compute the square root of the average squared forecast errors one to eight steps ahead: *i.e.*, we average over the forecast horizons instead of over time, since we argue that the use of the standard average of the fixed horizon root mean squared errors over a reasonably long period of time can be misleading by hiding many interesting characteristics of the indicators. In particular, some indicators can outperform autoregressive models on average but forecast very poorly in some periods. This has serious consequences if the forecasts are used in a policy-making environment. The fact that the indicators should be changed from period to period, depending on the likelihood of particular economic shocks over the forecasting period, does not emerge under the choice of a loss function which averages over time. The robustness of leading indicators is therefore overemphasised. More details are provided below.

We should stress that in this paper we focus on the role of leading indicators for forecasting GDP growth and inflation rather than their turning points, as it is sometimes considered in the literature. A comprehensive evaluation of the leading properties of the indicators for turning point forecasts, along the lines of our exercise for forecasting growth rates, would be very interesting, see e.g. Clements and Galvao (2005) for an assessment of the forecasting performance of the Conference Board Composite Leading Index and its components. Yet, the theory underlying the use of factor models, automatic model selection procedures, or pooling for forecasting binary variables such as an expansion/recession indicator using very large information sets is still rather underdeveloped. We therefore leave this important topic for future research. However, several of the available results on the usefulness of a limited number of leading indicators for forecasting turning points seem to be in line with the instability we find for predicting growth rates, see e.g. Marcellino (2004) for a recent overview.

Our paper has six sections following this introduction. Sections 2, 3 and 4 deal with some relevant methodological aspects. Section 2 describes in some more detail the Cecchetti *et al*. paper and the results contained therein. Section 3 outlines the model selection procedure we adopt while Section 4 provides a brief introduction to the factor models developed by Stock and Watson *inter alia*. Section 5 describes the data, with further details given in the data appendix. Section 6 presents the main empirical results and comparisons. Section 7 concludes.

## 2. The unreliability of inflation indicators (Cecchetti *et al.)*

Cecchetti *et al*. use a 19-variable data set, grouped into three broad classes (commodity prices, financial indicators and indicators of the status of the real economy) over the time span

1975:1 to 1998:4, to provide forecasts for one- to eight-step-ahead quarterly inflation in the consumer price index (*CPI*) of the US. [1]

The estimated model takes the fixed form

$$INFL_t = \alpha + \sum_{i=1}^{4} \beta_i INFL_{t-i} + \delta IND_{t-1} + \varepsilon_t \ ,$$

where $INFL$ denotes quarterly CPI inflation and $IND_{t-1}$ is the lagged value of the particular indicator variable chosen. The object of the exercise is to determine the value of the addition of the leading indicator in terms of improving the accuracy of the inflation forecast. The model is estimated first from 1975:1 to 1984:4 to provide forecasts for the eight quarters up to 1986:4. The estimation sample is next augmented by one year (*i.e.* until 1985:4) and the model is re-estimated to forecast inflation for the 1986-87 period. This exercise of augmentation is continued recursively until the estimation sample extends to 1996:4 and forecasts are provided for the 1997-98 period.

Whenever out-of-known-sample values of the $INFL$ variable are required to generate forecasts, the forecast value (from the previous regressions) is used. Thus, for example, to generate an inflation forecast for 1985:2 where the estimation sample is until the end of 1984:4, requires a value for inflation for 1985:1 which must be assumed to be unknown. In the Cecchetti *et al.* methodology, the inflation forecast for 1985:1 generated by the model estimated until 1984:4 is used to forecast inflation for 1985:2. The inflation forecasts generated for 1985:1 *and* 1985:2 are used to forecast inflation for 1985:3 and so on. In contrast, unknown values of the lagged leading indicator variable are replaced by the actual lagged value of the indicator in the forecast period. The comparison is thus made ex-post, since future value(s) of the leading indicator(s) are assumed to be known. This provides the maximum advantage for the leading indicators against the autoregressive model, in the sense that the indicators can only perform worse in a real-time context where their future values have also to be predicted.

For each estimation period, Cecchetti *et al.* compute their eight-step-ahead RMSE statistic as the square root of the average squared forecast errors one to eight steps ahead. As mentioned in the introduction, this procedure differs from the standard practice of taking averages over the whole forecasting period of the forecast errors computed for a fixed horizon. We refer to this loss function as the RMSE-*h* criterion to distinguish it from the standard RMSE.

[1] The complete list is given in Table 1 (page 2) of Cecchetti *et al*.

The main advantage of this method of evaluation is that it is closer to the practice of forecast evaluation by policy makers and practitioners, where the same model is used to forecast at different horizons and the interest is in the periodic evaluation of the model (and possibly in its periodic re-specification). Another important advantage of this procedure is that it is robust to structural changes over the forecast sample, in the sense that if an indicator performs well only over a sub-sample, this information will emerge and not remain hidden by a good average performance. The drawback is that since the series of the computed RMSE-*h* statistics is short and its elements are highly correlated, it is not possible to provide a reliable test for a significant difference in forecasting performance. Even bootstrapping would provide large standard errors around the relative RMSEs-*h*. Nevertheless, we think that the information provided by the point RMSEs-*h* is useful and clear enough to rank the competing forecasting models.

The best leading indicator is defined as the variable (for each sample) that provides the largest reduction in RMSE-*h* compared to a fourth order autoregression of inflation on its lagged values. An important finding reported by Cecchetti *et al.* is that the majority of the indicator-based forecasts are outperformed by the autoregression benchmark. Ten of the nineteen indicators underperform the autoregression in more than 50% of the estimation periods, and no single indicator consistently improves on autoregressive projections. Among the variables that work well are (a) the growth in the *Journal of Commerce – Economic Cycle Research Institute* price index for industrial materials (*JOC-ECRI Index*), (b) M2 growth, (c) growth in average hourly earnings and (d) the number of weekly hours worked. The use of all four of these variables is however problematic, the first because of ex-post changes in the composition of the index and the latter three because of their close relationship with inflation (given that the actual future value of the indicator variable is used for forecasting, its interrelationship with the variable being forecast is somewhat problematic). The exchange rate level and the growth of M1 consistently lead to higher RMSEs-*h* than AR models, as do interest rate variables, the unemployment rate, the monetary base, the employment to population ratio and the National Association of Purchasing Management (NAPM) composite index.

We were able to replicate all these results, and also considered the properties of one- to four-step-ahead forecasts. In this case the first model is estimated from 1975:1 to 1984:4 to provide forecasts for the four quarters up to 1985:4, but the results remain essentially unchanged.

## 3. Using an automated model-selection procedure

An important shortcoming of the above approach is that the lag length is fixed and not subject to testing, while combinations of leading indicators are never taken into consideration. Taking account of both these shortcomings may well lead to greater efficiency in the use of the leading indicator approach. We propose four different routes for doing so. First, we use an automated model-selection procedure to provide the best specified single indicator model of variable lag lengths for each sample. Second, we discuss possible criteria for constructing combinations of leading indicators and evaluate whether these lead to gains in RMSE-*h*. Third, we consider factor methods as a way of summarising efficiently the information contained in a large set of indicators. Finally, we evaluate whether pooling the information not in sample but out of sample matters.

The automated model selection procedure we consider was developed by Hendry and Krolzig (1999), Krolzig and Hendry (2001) and Hoover and Perez (1999), and is implemented with the software *PcGets*. The starting point for the algorithm is a general unrestricted model (GUM) containing all variables likely (or specified) to be relevant, including the maximum lag length of the independent and dependent variables. For example, in reconsidering the Hendry and Ericsson (1991) model of narrow money demand in the UK, Hendry and Krolzig (2001) specify the GUM as a regression of $\Delta(m-p)_t$ on $(m-p-x)_{t-1}$, up to one lag each of $\Delta p_t$ and $r_t$ and up to four lags each of $\Delta(m-p)_{t-1}, \Delta^2 p_t$ and $\Delta r_t$. Denoting logarithms of data in lower case, $m$ is M1, $x$ is real total final expenditure in 1985 prices, $p$ is its deflator and $r$ is the opportunity cost of holding money given by the 3-month local-authority interest rate minus the sight deposit rate).

The algorithm starts from a 'pre-search' simplification by applying tests for variable deletion, following which the GUM is simplified. This step uses a loose significance level such as 10%, to delete highly non-significant regressors. The procedure is refined at the second stage, where many alternative further reductions of the GUM are considered, using both *t* and *F* tests and information criteria as reduction (or deletion of variables) criteria. Diagnostic tests ensure that the models chosen as valid simplifications/reductions are congruent representations of the data. The third stage is the encompassing step (see *e.g.* Mizon and Richard (1986)) where all valid reduced models from the second step are collected, and encompassing tests are used to evaluate the relative merits of these competing models. Only models that are not encompassed are retained. If more than one model survives the third stage, their union forms the new general model and the algorithm recommences.

This process continues until the set of non-encompassed models reduces to one or the union is repeated. In the case of the Hendry and Ericsson GUM, only one model survives the selection process and gives the original Hendry and Ericsson specification. Notice that selection criteria like AIC and BIC are used in the specification search but these are employed in conjunction with statistical tests for the significance of the variables and the congruence of the models as a statistical representation of the data.

For our purposes, when we focus attention only on single indicators (in order to generalize Cecchetti *et al*. directly), the lag length of the autoregression of inflation or GDP growth on its past is left specified only up to a maximum in the GUM, as is the lag length of the indicator variable. *PcGets* then provides the most parsimonious model that is used for forecast comparisons. When more than one indicator is contemplated, we need only to extend the set of independent variables, specify a maximum lag length and let *PcGets* do the rest. Since this is a regression-based approach, only a limited number of indicators can be considered in order not to exhaust degrees of freedom. In what follows, we select the indicators to be included in the GUM based either on economic criteria (real, nominal, financial variables) or on their forecasting performance as single indicators.

## 4. Factor Models

Dynamic factor-models have recently been successfully applied to forecasting US, UK and Euro-area macroeconomic variables (Stock and Watson (2001a), Artis, Banerjee and Marcellino (2004) and Marcellino, Stock and Watson (2003) respectively). This technique can be viewed as a particularly efficient means of extracting information from a large number of data series, so that instead of a single indicator variable or a group of indicator variables, we may contemplate the use of the most important factors extracted from the whole data set for forecasting. Here we briefly introduce the representation and estimation theory for the dynamic factor model.

Let $X_t$ be the $N$-macroeconomic variables to be modelled, observed for $t=1,...,T$. $X_t$ admits an approximate linear dynamic factor representation with $\bar{r}$ common factors, $f_t$, if:

$$X_{it} = \lambda_i(L) f_t + e_{it} \tag{1}$$

for $i=1,...,N$, where $e_{it}$ is an idiosyncratic disturbance with limited cross-sectional and temporal dependence, and $\lambda_i(L)$ are lag polynomials in non-negative powers of $L$; see for example Geweke (1977), Sargent and Sims (1977), Forni, Hallin, Lippi, and Reichlin (2000) and, in particular, Stock and Watson (2002). If $\lambda_i(L)$ have finite orders of at most $q$, equation (1) can be rewritten as,

$$X_t = \Lambda F_t + e_t \tag{2}$$

where $F_t = (f_t^{'},\ldots, f_{t-q}^{'})'$ is $r\times 1$, where $r\leq (q+1)\bar{r}$ , and the *i-th* row of $\Lambda$ in (2) is $(\lambda_{i0},\ldots,\lambda_{iq})$ .

The factors provide a summary of the information in the data set, and can therefore be expected to be useful for forecasting. From a more structural point of view, the factors can be considered as the driving forces of the economy. In both cases, it is extremely important to have accurate estimators of the factors.

Stock and Watson (2002) show that, under some technical assumptions (restrictions on moments and stationarity conditions), the column space spanned by the dynamic factors $f_t$ can be estimated consistently by the principal components of the variables in $X_t$. A condition that is worth mentioning for the latter result to hold is that the number of factors included in the estimated model has to be equal or larger than the true number. Normally two or three factors are sufficient to explain a large proportion of the variability of all the time series. We use up to six factors in what follows, since some of the lower-ordered factors may be relevant for forecasting the variable of interest even though their explanatory power for the set of indicators is only marginal.

It should be stressed that the principal-component-based estimator is consistent for the space spanned by the factors, *not* for the factors themselves. This follows from the lack of identification of the factors, since the representation in equation (2) is identical to

$$X_t = \Lambda P^{-1} P F_t + e_t = \Theta G_t + e_t , \tag{3}$$

where $P$ is any square matrix of full rank $r$ and $G_t$ is an alternative set of $r$ factors. While this lack of identification is problematic when interpreting the factors in a structural way, it is unproblematic for forecasting, since the factors $F$ and $G$ are equivalent summaries of the information in $X$.

Finally, it is worth noting that, under additional mild restrictions on the model, the principal component based estimator remains consistent even in the presence of changes in the factor loadings, *i.e.* $\Lambda = \Lambda_t$ . In particular, Stock and Watson (2002) allow either for a few abrupt changes, or for a smooth evolution as modelled by a multivariate random walk for $\Lambda_t$ .

## 5. The Data

Cecchetti *et al.* (2000) group 19 inflation indicators in three main groups: commodity prices, financial indicators and indicators of real economic activity (like capacity utilization rate and unemployment rate), to which they add also average hourly earnings. Commodity prices include specific prices for oil, industrial materials, precious metals and indexes for groups of similar goods. The group of financial indicators contains exchange rates, different monetary aggregates, interest rates and term premia.

For the data in this paper we use a slightly more detailed categorisation that is more in line with the one used by Stock and Watson (2001). As in Cecchetti *et al.* (2000) we use quarterly data with the sample starting in 1975:1, but the end of the sample has been extended to 2001:4. The main difference with the Cecchetti dataset is that we do not use commodity prices because of problems with data availability for these series (although we include fuel and electricity prices which are the most important categories). Our primary data source is the OECD Main Economic Indicators database and the data are seasonally adjusted at the same source. Altogether we use 64 indicators and 74 GDP growth indicators as given in Tables 3 and 9. Note that some variables are used as an indicator in both levels and in growth rates, to check whether a certain variable can perform well as an indicator (for some sub-periods) in levels even though we would expect it to perform better if suitably transformed *i.e.* in growth rates. This provides yet another check of the reliability of the forecasting technology. Inflation and GDP growth are both treated as stationary, but we also repeat the evaluation for first differenced inflation.

The group of output indicators is the largest and is composed mainly of data for different indexes of industrial production, plus aggregate demand components in the casr of GDP growth. Capacity utilization rate also falls into this group. Other variables fall into the groups of employment and working hours, retail, manufacturing and trade sale data, housing, stock prices, exchange rates, interest rates, money and credit aggregates, price indexes, labour costs. Finally, there is a miscellaneous group that contains balance of payments data, euro area HICP inflation and GDP growth, the ECRI Future Inflation Gauge (FIG), the Conference Board composite leading index, a set of diffusion indexes constructed by the Institute of Supply Management, and the Chicago Fed National Activity Index. These composite indexes are included in the comparison since they already contain a summary of a large amount of information. Yet, most of them are constructed to predict turning points rather than growth rates, so that their performance should be interpreted with care in our context.

## 6. Empirical analysis

In this section we discuss the results of the forecast comparison exercise for inflation and GDP growth. The first subsection summarizes the findings for the ex-post analysis. The second subsection focuses on the ex-ante evaluation. The final subsection summarizes the main conclusions. Additional details are available upon request.

### 6.1 Ex-post analysis

In order to benchmark our subsequent analysis for inflation, Table 1 reports the results derived from the use of the same set of variables and time span as Cecchetti *et al.* but using *PcGets* to automate the selection of the best indicator and lag length. The GUM consists of inflation on its own lags and lags of a single indicator variable (with a maximum of 6 lags both for the dependent variable and for the indicator. This is compared with a pure autoregression (with lags determined by *PcGets*). The eight-step-ahead RMSEs-*h*, computed as in Section 2, are used for evaluation.

We have 19 indicators and 13 evaluation periods. In 9 out of the 13 periods the autoregression does better than at least 50% of the models with an indicator. Although the best performing indicator is always better than a pure autoregression, no indicator consistently out-performs the autoregression or is best more than twice. This depends on the different types of shocks hitting the economy at different periods. Moreover, some indicators do much worse than the autoregression, but are not deleted from the sample. This may be a reflection of the result emphasized by Clements and Hendry (1999) that models that work within sample may have very poor forecasting properties.

The RMSEs-*h*, both from the autoregressions and the leading indicator augmented models show a tendency to change and (on the whole) decline over time. This feature may be attributed to the slow down in the rate of inflation over time and emphasises the virtue of not relying on averages of fixed horizon forecast errors.

*PcGets* provides lower RMSEs-*h* for the autoregressive model than Cecchetti *et al*. in 7 out of 13 periods. This increases to 9 out of 13 periods when looking at the best indicator. Moreover, Cecchetti *et al.* and *PcGets* give different best performing indicators, so that allowing for lag selection matters. This provides justification for using a selection rule instead of using a fixed number of lags in the estimating (and forecasting) models.

Repeating the Cecchetti *et al.* analysis with our larger and longer data set[2] we find that in 10 out of the 16 evaluation periods the autoregression does better than at least 50% of the models with an indicator. The best indicator remains better than the autoregression, but there is a lot of variation in the best indicator and only 7 out of the 64 indicators we consider do better than the autoregression more than half of the time. The set of the best indicators includes reasonable variables from an economic point of view such as the growth of industrial production (aggregate, non-durable), wholesale sales, the 10 year interest rate and the growth in hourly wage earnings. The growth in food and energy prices also work well, but this result should be interpreted with care since at this stage of the analysis future values of these variables are treated as known, and they are important components of inflation. Care should be also exerted in the interpretation of the intermediate performance of diffusion indexes, for example FIG can beat the autoregression in only 5 out of 16 evaluation periods, while CFNAI does so in 8 out of 16 periods. As mentioned in Section 5, these indexes are constructed for predicting turning points while here we are focusing on forecasting growth rates.

In the case of GDP growth we find that the autoregression does better than at least 50% of the models with an indicator in 13 of the 16 evaluation periods, more often than in the case of inflation. The best indicator, however, remains better than the autoregression although it changes over time, with a component of industrial production performing well recently, which is particularly interesting in light of the 2001 US slowdown. Only 2 out of the 74 indicators do better than an autoregression at least half the time, namely an index of growth in import prices and a diffusion index for manufacturing inventories. Other good indicators for GDP growth include an index of growth in export prices, and components of industrial production and gross domestic product. All three of CFNAI, the composite leading index produced by ECRI (ECRIuswlim) and the CLI produced by the Conference Board outperform the AR model in only 6 out of 16 evaluation periods. As in the case of the Inflation Gauge, this finding should be interpreted with care since these indexes aim at anticipating cyclical turning points rather than forecasting GDP growth.

[2] The detailed tables on which the results reported on this page are based are not included in order to save space and are available from us upon request.

We now turn to an evalution of the three approaches for forecasting in the presence of a large information set, namely, using (six) factors extracted from the large dataset, grouping the variables into subclasses supplemented with automatic model specification,[3] and pooling the single indicator based forecasts. The results are summarized in Tables 2 and 3, where we report the RMSE-*h* of the different forecasts relative to the AR benchmark for, respecitively, inflation and GDP growth.

Factor forecasts (group 2) are on the whole disappointing. They outperform the other multivariate methods in only 2 out of 10 evaluation periods for inflation and 3 for GDP growth, but even in these periods they are usually beaten by the best single indicator model. Adding the three best factors to the best single indicators (those outperforming the AR benchmark in at least 50% of the evaluation periods) is helpful for inflation but not for GDP growth. For the former, the relative RMSE-*h* decreases with respect to the case without factors in 7 out of 10 periods, for the latter in only 2 out of 10 periods, as may be seen by comparing groups 1 and 3 in Tables 2 and 3.

These findings are in disagreement with the good performance for the US of factor models compared with AR reported by Stock and Watson (2002). However, direct comparison with Stock and Watson on the basis of Table 2 is inappropriate for several reasons. First we make use of both a different evaluation criterion (as described in Section 2 above), and a different estimation method (static versus dynamic). Secondly, quarterly instead of monthly data are used here, and thirdly Stock and Watson use more variables to extract the factors.

To evaluate whether grouping the indicators by their economic category helps, we divide them into subsets of real variables, price variables and financial variables. Broadly speaking, the 'Real 1' group contains growth rates of measures of industrial production plus the unemployment rate. 'Real 2' contains measures of turnover, consumers' confidence and the capacity utilization rate. 'Real 3', used only for GDP growth, is related to imports, exports and the balance of payments. 'Financial' includes interest rates, spreads, money growth, interest rates and measures of exchange rates. 'Prices' contains fuel and energy prices, earnings and measures of unit labour costs.[4] It may be seen from Table 2 that an economic based grouping beats either the factors or the group of best single indicators (*i.e.*

[3] *PcGets* is used for model selection amd the indicators are included conservatively in the final version of the GUM, based on a 1% significance level criterion (include if statistic rejects at 1%). No major changes are noted when a more liberal inclusion rule is adopted (include if statistic rejects at 5%)

[4] Some indicators are included both in levels and in growth rates to allow *PcGets* to select the best transformation.

group 1) in only 3 out of 10 evaluation periods for inflation. For GDP growth, this occurs in 5 out of 10 periods for GDP growth, see Table 3.

We can now consider the role of forecast pooling as an alternative tool for forecasting inflation and GDP growth when a large information set is available. Starting with the pioneering work of Bates and Granger (1969) it is well known that a combination of forecasts may perform better than each of the single constituent forecasts. As discussed by Hendry and Clements (2004), possible reasons for this finding may be model misspecification and parameter non-constancy that are attenuated by weighting.

The pooling weights should in principle depend on the entire covariance matrix of the forecasts to minimize the RMSE. Since this is too complicated in our framework with many forecasts, we consider two simple procedures that have performed well, for example, in Stock and Watson (1999b). First, a simple average of all the single indicator forecasts, and second the median of the forecasts. The latter could be more robust since some indicators produce forecasts with high RMSEs-*h*.

The results are again reported in Tables 2 and 3 for, respectively, inflation and GDP growth, under the columns P-Mean and P-Med. As expected, the median performs systematically better than the average of the single forecasts. Yet, the median forecasts are better better than the AR in only 2 out of 10 cases for inflation, never for GDP growth. The periods when pooling works better are those when a large fraction of the indicators outperforms the autoregression. Similarly, the performance of pooling improves substantially when some of the worst performing indicators are not included in the set of single forecasts under consideration for pooling. Yet, even in this case, a single (though time-varying) indicator typically beats the pooled forecast.

Overall, the main message that emerges from this analysis is that single indicator forecasts in general outperform more sophisticated multivariate methods, when using the RMSE-*h* loss function for evaluation. The gains from simplicity are larger than those from the use of a larger information set. Hence, we return to considering in a more realistic context the performance of single indicator forecasts.

### 6.2 Ex-ante analysis

The results so far have been obtained assuming that future values of the leading indicators are known, which provides the most favourable environment for the use of such indicators (in the sense that if indicators do not perform well here they can be expected not to do so in real time). We have seen above that single indicators are generally preferable to using combinations of such variables and out-perform autoregressions. In this section we evaluate whether this latter finding still holds in a pseudo-real-time framework.

Our method of ex-ante evaluation can best be described by an example. Say that we are in the last quarter of 1992. Then we can use 1990:4 for estimation and produce forecasts for 1991:1 until 1992:4 and compute the RMSE-*h* for each indicator (at this step actual values of the indicators for 1991-92 are used since they are in the information set). The indicator that provides the lowest RMSE-*h* can then be used to forecast from 1993:1 until 1994:4, where the estimation sample is extended until the last available observation *i.e.* 1992:4. Moreover, since values of the indicator variable(s) over the year 1993:1 until 1994:4 are not known in 1992:4, autoregressive models are used to forecast them. This procedure is repeated for each year.

The procedure above is implemented both for $h$=8 and $h$=4, and the results for inflation are reported in Table 4. When $h$=8 the feasible indicator forecasts are better than autoregressions in 5 out of 9 periods, while when $h$=4 the performance of the feasible indicators deteriorates, and the autoregression is out-performed only in 2 out of 11 periods.

For GDP growth, Table 5 shows that when $h$=8 the feasible indicator forecasts are better than autoregressions in 7 out of 10 periods, while for $h$=4 the autoregression is out-performed in 10 out of 11 cases. This represents a noteworthy improvement with respect to the equivalent results for forecasting inflation, and provides a very useful indication for real time forecasting of GDP growth.

To conclude, to assess the robustness of the findings reported so far, we have considered the role of differencing inflation (and nominal indicators) rather than using levels, and the consequences of rolling rather than recursive estimation. Both features do not seem to play a major role for the evaluation of the leading indicators; detailed results are available upon request.

### 6.3 Summary

Five main conclusions can be drawn for forecasting inflation from the results in this section. First, ex post, autoregressions are beaten by univariate leading indicator models, but the best indicator changes over time, reflecting the varying nature of the major sources of inflation. Second, grouping either according to economic categories or to the performance of the single indicators, complemented by the automatic model selection procedure implemented with *PcGets*, is better than using factor models, but in general the RMSE-*h* is higher than when using single indicators. Third, the results are robust to the degree of differencing, the use of rolling estimation and choice of forecasting horizon. Fourth, the median pooled estimator performs better than the average, but for it to beat the AR a careful selection of the single forecasts to be pooled is required. Finally, in a pseudo ex ante context, the indicators can hardly beat the autoregressions more than 50% of the time, which provides support for the AR model as a robust forecasting device for inflation when using our loss function.

As for inflation, five main conclusions can be drawn for forecasting GDP growth. First, ex post, univariate leading indicator models are better than autoregressions, but the best indicator changes over time and there are fewer indicators with a satisfactory performance than for inflation. Second, grouping either the factors or the indicators according to their univariate performance, complemented by the automatic model selection procedure implemented with *PcGets*, is often better than the autoregression, but worse than the single indicators (as for inflation). Third, the results are robust to the use of rolling estimation and choice of forecasting horizon. Fourth, forecast pooling works only in a fraction of cases, smaller than for inflation. Finally, and more importantly, ex ante the indicators can beat the autoregressions more than 80% of the time.

## 8. Conclusions

The first contribution of this paper is the empirical comparison of three alternative approaches to information extraction from a large data set for forecasting, namely, the use of an automated model selection procedure, the adoption of a factor model to summarize the available information, and single indicator based forecast pooling. Both for inflation and GDP growth it turns out that all methods are systematically beaten by single indicator models.

The second main contribution is the comparison of a large set of leading indicators with purely autoregressive models, using an evaluation procedure that is particularly relevant for policy making. Ex post, *i.e.* assuming that future values of the indicators are known, they systematically outperform autoregressive models. But, even in this unrealistic context that biases the comparison against the autoregression, the best indicator changes continuously over time, and most indicators generate higher RMSE-*h* than the autoregression in at least 50% of the evaluation periods.

Finally, in an ex-ante context, we have developed a feasible procedure that allows the construction of indicator based forecasts that outperform the autoregressions in about 50% of the cases for inflation and 80% for GDP growth.

## References


Artis, M., Banerjee, A., and M. Marcellino. 'Factor Forecasts for the UK', *Journal of Forecasting*, (2004), forthcoming.

Bates, J.M. and C.W.J. Granger. 'The Combination of Forecasts', *Operations Research Quarterly* 20, (1969), 415-468.

Camba-Mendez, G., Kapetanios, G., Smith, R.J., and M.R. Weale. 'An Automatic Leading Indicator of Economic Activity: Forecasting GDP Growth for European Countries', *Econometrics Journal* 4, (2001), S56-S90.

Cecchetti, S.G., Chu, R.S., and C. Steindel. 'The Unreliability of Inflation Indicators.', *Current Issues in Economics and Finance* 4/6, Federal Reserve Bank of New York, 2000.

Clements, M.P., and D.F. Hendry. *Forecasting Non-stationary Economic Time Series*. Cambridge, Massachusetts: MIT Press, 1999.

Clements, M.P. and Galvao, A. (2005), 'Combining predictors and combining information in modelling: Forecasting US recession probabilities and output growth', in *Nonlinear Time Series Analysis of Business Cycles*, C. Milas, P. Rothman and D. van Dijk (eds.), Elsevier.

Geweke, J. 'The Dynamic Factor Analysis of Economic Time Series', In *Latent Variables in Socio-Economic Models*, edited by Aigner, D.J. and A.S. Goldberger, Ch. 19. Amsterdam: North-Holland, 1977.

Granger, C.W.J. and P. Newbold. *Forecasting Economic Time Series*, 2nd edn., New York: Academic Press, 1986.

Forni, M., Hallin, M., Lippi, M. and L. Reichlin. 'The Generalized Dynamic Factor Model: Identification and Estimation', *Review of Economics and Statistics* 11, (2000), 369-379.

Hendry, D.F., and M.P. Clements. ‘Pooling of Forecasts’, *Econometrics Journal* 7, (2004), 1-31.

Hendry, D.F., and N.R. Ericsson. ‘Modeling the demand for narrow money in the United Kingdom and the United States’, *European Economic Review* 35 (1991), 833-886.

Hendry, D.F., and H.M. Krolzig. ‘Improving on `Data mining reconsidered by Hoover, K.D. and S.J. Perez’, *Econometrics Journal* 2, (1999), 202—219.

Keereman, F. (1999), ‘The track record of the Commission forecasts’, European Commission Economic Papers, 137.

Keereman, F. (2003), ‘External assumptions, the international environment and the track record of the Commission forecasts’, European Commission Economic Papers, 189.

Krolzig, H.M. and D.F. Hendry. ‘Computer automation of general-to-specific model selection procedures’, *Journal of Economic Dynamics and Control*, 25, (2001), 831-866.

Marcellino, M., Stock, J.H. and M.W. Watson. ‘Macroeconomic Forecasting in the Euro Area: Country Specific Versus Euro Wide Information’, *European Economic Review*, 47, (2003), 1-18.

Marcellino, M. (2004). ‘Leading indicators: What have we learned?’, in preparation for Elliott, G., Granger, C.W.J. and A. Timmermann. “Handbook of Economic Forecasting”, Elsevier – North Holland.

Mizon, G.E., and J.F. Richard. ‘The encompassing principle and its application to non-nested hypothesis tests’, *Econometrica* 54 (1986), 657-678.

Sargent, T.J., and C.A. Sims. (1997). ‘Business Cycle Modelling without Pretending to have Too Much a-priori Economic Theory’, In *New Methods in Business Cycle Research*, edited by C.A. Sims, Minneapolis: Federal Reserve Bank of Minneapolis.

Stock, J.H. and M.W. Watson (1993). 'A procedure for predicting recessions with leading indicators: econometric issues and recent experience', in: Stock, J. H., and M. W. Watson, eds., Business Cycles, Indicators, and Forecasting, The University of Chicago Press (Chicago): 95-153.

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

Stock, J.H. and M.W. Watson. (1999b) 'A comparison of linear and non-linear univariate models for forecasting macroeconomic time series', in Engle, R. and White, H.. (eds), *Cointegration, causality, and forecasting: A festschrift in honor of Clive W.J. Granger*, Oxford: Oxford University Press, 1-44.

Stock, J.H. and M.W. Watson. 'Forecasting Output and Inflation: The Role of Asset Prices', NBER Working Paper 8180 (2001).

Stock, J.H. and M.W. Watson. 'Macroeconomic Forecasting Using Diffusion Indexes', *Journal of Business and Economic Statistics*, 20, (2002) 147-62

# Data Appendix

## List of variables and transformations used

| Variable | Trans | Description |
|---|---|---|

## Output variables

| | | |
|---|---|---|
| gdp | DLV | Gross Domestic Product – Total (BN $, 1996 prices, S.A.) |
| ip | LV, DLV | Industrial production – total (1995=100, S.A.) |
| ipc | LV, DLV | Industrial production – consumer goods (1995=100, S.A.) |
| ipcd | LV, DLV | Industrial production – durable consumer goods (1995=100, S.A.) |
| ipcnd | LV, DLV | Industrial production – non-durable consumer goods (1995=100, S.A.) |
| ipint | LV, DLV | Industrial production – intermediate goods (1995=100, S.A.) |
| ipi | LV, DLV | Industrial production – investment goods (1995=100, S.A.) |
| ipman | LV, DLV | Industrial production – manufacturing (1995=100, S.A.) |
| ipcons | LV, DLV | Industrial production – construction (1995=100, S.A.) |
| cap | LV | Capacity utilization rate (%, S.A.) |
| gdpc | DLV | GDP-private consumption (1996 prices BN$, S.A.) |
| gdpgov | DLV | GDP-government consumption (1996 prices BN$, S.A.) |
| gdpcons | DLV | GDP-construction (1996 prices BN$, S.A.) |
| gdpi | DLV | GDP-fixed capital formation (1996 prices BN$, S.A.) |

## Employment and hours

| | | |
|---|---|---|
| lurat | LV | unemployment rate (% of civilian labor force, S.A.) |
| lhman | LV | weekly hours worked – manufacturing (hours, S.A.) |

## Retail, manufacturing and trade sales

| | | |
|---|---|---|
| rtvaltot | LV, DLV | retail sales – total (MN$, S.A. |
| rtvaldur | LV, DLV | retail sales – durables (MN$, S.A.) |
| whval | LV, DLV | wholesale sales – total (MN$, S.A.) |
| cars | LV | passeneger car registrations (000 number, S.A.) |

## Housing

| | | |
|---|---|---|
| ccost | LV, DLV | cost of construction, residential 1995=100 |

## Stock prices

| | | |
|---|---|---|
| fs | LV, DLV | NYSE share prices (1995=100) |

## Exchange rates

| | | |
|---|---|---|
| ereff | LV, DLV | US real effective exchange rate (1995=100) |
| eneff | LV, DLV | US nominal effective exchange rate (1995=100) |

## Interest rates and spreads

| | | |
|---|---|---|
| fy10gov | LV | government composite bonds (>10 years, % p.a.) |
| fcod | LV | certificates of deposits (3 month, % p.a.) |

| | | |
|---|---|---|
| ff | LV | federal funds rate (% p.a.) |
| spread3 | LV | fcod – ff |
| spread10 | LV | fy10gov – ff |

## Money and credit quantity aggregates

| | | |
|---|---|---|
| m1 | DLV | monetary aggregate M1 (BN$, S.A.) |
| m2 | DLV | monetary aggregate M2 (BN$, S.A.) |
| m3 | DLV | monetary aggregate M3 (BN$, S.A.) |
| lebank | DLV | bank lending (BN$, S.A.) |

## Price indexes

| | | |
|---|---|---|
| cpi | DLV | (1995=100, S.A.) |
| cpf | DLV | food (1995=100, S.A.) |
| cpidur | DLV | durables (1995=100, S.A.) |
| cpifuel | LV, DLV | fuel and electricity (1995=100, S.A.) |
| pexp | DLV | Export prices (1995=100, S.A.) |
| pimp | DLV | Import prices (1995=100, S.A.) |

## Wages

| | | |
|---|---|---|
| wheman | DLV | hourly earnings – manufacturing (1995=100) |
| whetot | DLV | hourly earnings – total private (1995=100, S.A.) |
| wc | DLV | unit labor cost – manufacturing (1995=100, S.A.) |

## Miscellaneous

| | | |
|---|---|---|
| fbopnet | LV | current account balance (BN$, S.A.) |
| fgoodsnet | LV | net trade (BN$, S.A.) |
| fimp | LV, DLV | Imports (BN$, S.A.) |
| fexp | LV, DLV | Exports (BN$, S.A.) |
| conf | LV | Consumer sentiment (1995=100, S.A.) |
| ECRIuswlim | DLV | ECRI US leading indicator |
| FIG | LV | ECRI US Future Inflation Gauge |
| CLI | DLV | Conference Board Composite Leading Index |
| ISMpDI | LV | Institute for Supply Management prices diffusion index, manufacturing |
| ISMempDI | LV | Institute for Supply Management employment diffusion index, manufacturing |
| ISMinvDI | LV | Institute for Supply Management inventories diffusion index, manufacturing |
| ISMordDI | LV | Institute for Supply Management new orders diffusion index, manufacturing |
| ISMPMIDI | LV | Institute for Supply Management PMI diffusion index, manufacturing |
| ISMprodDI | LV | Institute for Supply Management production diffusion index, manufacturing |
| CFNAI | LV | Chicago FED National Activity Index |
| EUgdp | DLV | Euro area real GDP |
| EUinfl | LV | Euro area HICP inflation |

Transformations used: LV – levels, DLV – annual growth rate (variables denoted by ‘g’ in the endings of the names)
S.A. indicates that the data have been seasonally adjusted at source (OECD Main Economic Indicators).

# Tables

### Table 1: Reproduction of Table 2 in Cecchetti *et al.* (2000) using their data and *PcGets*

| | Number of Indicators That Performed | | RMSE-*h* | | | *PcGets* deletes |
|---|---|---|---|---|---|---|
| **Estimation period** | **Better Than AR** | **Worse Than AR** | **Autoreg.** | **Best Indicator** | **Worst Indicator** | |
| 75:1 84:4 | 3 | 16 | 1.76 | 1.70 (3-m yield - $r^{ff}$) | 5.36 (M1) | M2 |
| 75:1 85:4 | 2 | 17 | 2.25 | 2.17 (NAPM com. index) | 9.7 (Exchange rate, growth) | Price of oil |
| 75:1 86:4 | 14 | 5 | 3.67 | 1.15 (Federal funds rate [$r^{ff}$]) | 3.86 (10-y bond rate - $r^{ff}$) | Weekly hours index, growth |
| 75:1 87:4 | 9 | 10 | 1.74 | 0.81 (Exchange rate, growth) | 5.89 (M1) | M2 |
| 75:1 88:4 | 12 | 7 | 2.32 | 1.55 (JOC index, level) | 5.17 (M1) | - |
| 75:1 89:4 | 7 | 12 | 2.28 | 1.73 (Cap. util. rate) | 4.21 (Employment/pop ratio) | M2, Hourly earnings, growth |
| 75:1 90:4 | 11 | 8 | 2.86 | 1.17 (JOC index, growth) | 3.99 (Exchange rate, growth) | M2 |
| 75:1 91:4 | 4 | 15 | 0.54 | 0.38 (Weekly hours index, growth) | 3.79 (Mh) | - |
| 75:1 92:4 | 1 | 18 | 0.64 | 0.49 (Price of oil) | 4.20 (Mh) | - |
| 75:1 93:4 | 3 | 16 | 0.73 | 0.52 (Weekly hours index, growth) | 4.10 (JOC index, level) | - |
| 75:1 94:4 | 8 | 11 | 0.83 | 0.56 (Unemployment rate) | 2.74 (10-y bond rate - $r^{ff}$) | - |
| 75:1 95:4 | 4 | 15 | 0.88 | 0.74 (Exchange rate, growth) | 2.89 (Price of oil) | - |
| 75:1 96:4 | 17 | 2 | 2.67 | 0.52 (M1) | 2.77 (Cap. util. rate) | - |

The complete list and definitions of the indicator-variables are given in Table 1 (page 2) of Cecchetti *et al*.

### Table 2a: Performance of groups of variables in forecasting inflation up to eight quarters ahead (conservative strategy) – RMSE-*h* relative to benchmark AR

| Estimation period | Real 1 | Real 2 | Financial | Prices | Group 1 | Group 2 | Group 3 | P-Mean | P-Med | Best single |
|---|---|---|---|---|---|---|---|---|---|---|
| 75:1 90:4 | 0,73 | 0,92 | 1,50 | **0,43** | 0,74 | 0,69 | 0,82 | 2,67 | 2,54 | 0,49 |
| 75:1 91:4 | **0,82** | 3,29 | 4,82 | 1,46 | 2,54 | 2,55 | 2,55 | 1,13 | 0,94 | 0,88 |
| 75:1 92:4 | 1,56 | 3,86 | 2,89 | 1,71 | 1,41 | 2,18 | 4,05 | 2,02 | 1,32 | **0,94** |
| 75:1 93:4 | 1,66 | 84,01 | 5,97 | 1,36 | 1,96 | 1,01 | 1,67 | 2,39 | 1,19 | **0,84** |
| 75:1 94:4 | 2,30 | 0,85 | 2,64 | 0,91 | 1,27 | 2,30 | 0,95 | 1,42 | 0,96 | **0,56** |
| 75:1 95:4 | 3,35 | 1,17 | 1,68 | 1,09 | 1,28 | 2,04 | 0,97 | 1,56 | 1,35 | **0,81** |
| 75:1 96:4 | 2,88 | 1,31 | 6,33 | 1,61 | 1,37 | 0,85 | 1,16 | 2,12 | 1,95 | **0,58** |
| 75:1 97:4 | 0,77 | 0,44 | 1,39 | 0,97 | 0,35 | 0,51 | 0,33 | 1,30 | 1,19 | **0,29** |
| 75:1 98:4 | 1,19 | 0,57 | 1,39 | 0,73 | 0,32 | 0,44 | **0,23** | 1,16 | 1,03 | 0,26 |
| 75:1 99:4 | | 1,02 | 2,43 | 1,79 | 0,68 | 2,25 | **0,51** | 1,39 | 1,23 | 0,63 |

**BOLD** indicates the lowest RMSE-*h* for the estimation period.

### Table 2b: Groupings of variables in Table 2a

| **Real 1** | **Real 2** | **Financial** | **Prices** | **Group 1** | **Group 2** | **Group 3** | **P-Mean** | **P-Med** |
|---|---|---|---|---|---|---|---|---|
| Ipg<br>Ipcg<br>Ipcdg<br>Ipcndg<br>Ipintg<br>Ipig<br>ipmang<br>ipconsg<br>Lurat | cap<br>conf<br>lhman<br>rtvaltotg<br>rtvaldurg<br>whvalg<br>cars | fs<br>ff<br>spread10<br>spread3<br>m1g<br>m2g<br>m3g<br>lebankg<br>ereff<br>ereffg<br>eneff<br>eneffg | cpifuel<br>cpfg<br>cidurg<br>cpifuelg<br>whemang<br>whetotg<br>wcg<br>ccost<br>ccostg | ipcndg<br>whval<br>whvalg<br>cpifuelg<br>whetotg<br>fy10gov<br>cpfg | 6 US factors | ipcndg<br>whval<br>whvalg<br>cpifuelg<br>whetotg<br>fy10gov<br>cpfg<br>US F1<br>US F3<br>US F6 | Mean of the single indicator based forecasts | Median of the single indicator based forecasts |

See the Data Appendix for the definition of each variable

**Table 3a: Performance of groups of variables in forecasting GDP growth up to eight quarters ahead (conservative strategy) – RMSE-*h* relative to benchmark AR**

| Estimation period | Real 1 | Real 2 | Real 3 | Financial | Prices | Gr 1 | Gr 2 | Gr 3 | P-Mean | P-Med | Best single |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 75:1 90:4 | 1,86 | 2,69 | 1,43 | * | * | 2,02 | 1,80 | 4,63 | 2,23 | 2,08 | **0,72** |
| 75:1 91:4 | 1,58 | 2,31 | 1,35 | * | * | 1,36 | 1,30 | 1,81 | 2,28 | 2,11 | **0,93** |
| 75:1 92:4 | 1,27 | 1,22 | 1,50 | 1,40 | * | 1,23 | 1,05 | 1,20 | 2,94 | 2,27 | **0,87** |
| 75:1 93:4 | 1,39 | 20,03 | 1,20 | 1,11 | 1,03 | 1,39 | 1,67 | 1,46 | 2,35 | 1,73 | **0,95** |
| 75:1 94:4 | 1,22 | 1,23 | 1,27 | 1,00 | 1,62 | 1,37 | 1,53 | 1,46 | 2,43 | 2,06 | **0,87** |
| 75:1 95:4 | 1,12 | 1,04 | 1,28 | 0,83 | 2,07 | 1,38 | 1,40 | 1,43 | 2,16 | 2,05 | **0,73** |
| 75:1 96:4 | 1,25 | 1,13 | 0,90 | 1,48 | 1,92 | 0,95 | 1,08 | 1,29 | 2,15 | 2,10 | **0,84** |
| 75:1 97:4 | 1,01 | 1,08 | 1,07 | 0,86 | 2,30 | 0,88 | **0,64** | 1,00 | 2,66 | 2,60 | 0,72 |
| 75:1 98:4 | 1,09 | 4,71 | 1,05 | 1,01 | 1,28 | **0,78** | 0,98 | 0,82 | 2,33 | 2,35 | 0,81 |
| 75:1 99:4 | 0,92 | 0,00 | 0,88 | 2,11 | | 0,62 | 0,80 | 0,53 | 2,39 | 2,53 | **0,41** |

**BOLD** indicates the lowest RMSE-*h* for the estimation period.
* Model contained only a constant. Pure AR terms also deleted by *PcGets*.

**Table 3b: Groupings of variables in Table 3a**

| Real 1 | Real 2 | Real 3 | Financial | Prices | Group 1 | Group 2 | Group 3 | P-Mean | P-Med |
|---|---|---|---|---|---|---|---|---|---|
| Ipg<br>Ipcg<br>ipcdg<br>ipcndg<br>ipintg<br>Ipig<br>ipmang<br>ipconsg<br>Lurat | cap<br>conf<br>lhman<br>rtvaltotg<br>rtvaldurg<br>whvalg<br>cars | fbopnet<br>fgoodsnet<br>fimp<br>fexp<br>fimpg<br>fexpg | Fs<br>Ff<br>spread10<br>spread3<br>m1g<br>m2g<br>m3g<br>Lebankg<br>Ereff<br>Ereffg<br>Eneff<br>Eneffg | cpifuel<br>cpfg<br>cpidurg<br>cpifuelg<br>whemang<br>whetotg<br>wcg<br>ccost<br>ccostg | ipcndg<br>Ipconsg<br>fexpg<br>pimpg<br>ISMinvDI | 6 US factors | ipcndg<br>Ipconsg<br>fexpg<br>pimpg<br>ISMinvDI<br>US FI<br>US F2<br>US F3 | Mean of the single indicator based forecasts | Median of the single indicator based forecasts |

See the Data Appendix for the definition of each variable

### Table 4: Ex-ante performance of indicators in forecasting inflation

| | RMSE-*h*=8 | | RMSE-*h*=4 | |
|---|---|---|---|---|
| **Point in Time** | **$AR_{rec}$** | **$IND_{rec}$** | **$AR_{rec}$** | **$IND_{rec}$** |
| 90:4 | 2.90 | 5.66 | 3.37 | **3.36** |
| 91:4 | 0.56 | **0.51** | 0.48 | 1.38 |
| 92:4 | 0.66 | 1.32 | 0.60 | **0.55** |
| 93:4 | 0.70 | **0.59** | 0.71 | 1.74 |
| 94:4 | 0.96 | **0.58** | 0.50 | 0.98 |
| 95:4 | 1.02 | 1.10 | 1.28 | 2.87 |
| 96:4 | 1.09 | 1.95 | 1.15 | 1.89 |
| 97:4 | 1.76 | **1.06** | 0.62 | 1.10 |
| 98:4 | 1.76 | **0.66** | 1.18 | 1.76 |
| 99:4 | | | 0.81 | 0.82 |
| 00:4 | | | 0.30 | 1.66 |

Column 2 contains the RMSEs-*h* of one- to eight-step-ahead forecasts with pure AR models obtained with recursive estimation. Column 3 contains the RMSEs-*h* of forecasts produced with the best feasible indicator using recursive estimation. Analogous numbers are reported in columns 4 and 5 for *h*=4.

**BOLD** indicates that the corresponding RMSE-*h* is smaller than the RMSE-*h* of the pure AR model.

### Table 5: Ex-ante performance of indicators in forecasting GDP growth

| | RMSE-*h*=8 | | RMSE-*h*=4 | |
|---|---|---|---|---|
| **Point in time** | **$AR_{rec}$** | **$IND_{rec}$** | **$AR_{rec}$** | **$IND_{rec}$** |
| 90:4 | 1.55 | 2.90 | 1.82 | 3.17 |
| 91:4 | 1.91 | **0.56** | 1.36 | **0.99** |
| 92:4 | 2.22 | **0.79** | 2.66 | **2.35** |
| 93:4 | 1.51 | 1.53 | 1.60 | **0.71** |
| 94:4 | 1.88 | **1.61** | 1.71 | **0.92** |
| 95:4 | 1.87 | **1.19** | 2.06 | **0.58** |
| 96:4 | 2.00 | 2.85 | 1.47 | **1.31** |
| 97:4 | 2.60 | **1.78** | 2.42 | **2.31** |
| 98:4 | 2.35 | **2.28** | 2.82 | **0.52** |
| 99:4 | 2.60 | **2.13** | 2.04 | **0.76** |
| 00:4 | | | 2.91 | **2.10** |

Column 2 contains the RMSEs-*h* of one- to eight-step-ahead forecasts with pure AR models obtained with recursive estimation. Column 3 contains the RMSEs-*h* of forecasts produced with the best feasible indicator using recursive estimation. Analogous numbers are reported in columns 4 and 5 for *h*=4.

**BOLD** indicates indicates that the corresponding RMSE-*h* is smaller than the RMSE-*h* of the pure AR model.