Are there any reliable leading indicators

for US Inflation and GDP Growth?[1]

Anindya Banerjee
Department of Economics
European University Institute
Via dei Roccettini, 9
50016 San Domenico di Fiesole (FI)
anindya.banerjee@iue.it

Massimiliano Marcellino
IEP-Bocconi university, IGIER and CEPR

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

31 October 2002

Abstract

In this paper we evaluate the relative merits of two approaches to information extraction from a large data set for forecasting, namely, the use of an automated model selection procedure versus the adoption of a factor model. The comparison is conducted using a large set of indicators for forecasting US inflation and GDP growth. We also compare this 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.

JEL Classification: C53, E37, C50

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

1. Introduction

Methods for forecasting inflation and output growth have been the subject of much 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 (1999)) 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 outperformed 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 on 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 root-mean-squared-error-criterion (RMSE) is used to judge forecasting accuracy.

We generalise the Cecchetti *et al*. analysis in two essential ways. Firstly, we allow for the selection of the 'best' leading indicator and the appropriate lag length to be automated. This automation is undertaken using the *PcGets* algorithm developed by Hendry and Krolzig (1999). 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 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, in addition to *PcGets*, we use the indicator variables in

combination. Both these extensions to the Cecchetti *et al*. methodology afford gains in terms of forecasting performance over autoregressions of inflation. In addition, we update Cecchetti *et al.* until 1999:4 for 8-step-ahead forecasts.

The third contribution of our paper is to re-assess the usefulness of factor models in forecasting inflation. 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 high 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 provides the maximum advantage against the autoregressive models but, as we will see, in many cases this 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

Our paper has seven sections following this introduction. Sections 2, 3 and 4 deal with the methodological aspects. Section 2 describes in some more detail the Cecchetti *et al*. paper and the results contained therein. Section 3 outlines *PcGets* 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. Sections 6 and 7 present the main results and comparisons, for inflation and GDP growth respectively. Some sensitivity analysis is also presented. Section 8 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 1- to 8-step-ahead quarterly inflation in the consumer price index (*CPI*) of the US. [2]

The estimated model takes the fixed form

$$\Delta CPI_t = \alpha + \sum_{i=1}^{4} \beta_i \Delta CPI_{t-i} + \delta IND_{t-1} + \varepsilon_t \,,$$

where IND_{t-1} is the lagged value of the particular indicator variable chosen. The object of the exercise is to determine the value (in terms of improving the accuracy of the inflation forecast or, equivalently, reducing the RMSE) of the addition of the leading indicator. 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. In our evaluation of forecasting methods reported later, we shall also consider the use of rolling samples. The only difference with the use of recursive estimates described in the previous paragraph is the use of a fixed window of ten years, with a recursive updating of the starting date for the estimation period.

Whenever out-of-known-sample values of the *CPI* 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. Instead, 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, in the sense that future value(s) of the leading indicator(s) are assumed to be known.

---

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

For each estimation period, the 8-step RMSE statistic is computed as the square root of the average squared forecast errors one to eight steps ahead. This procedure differs from the standard practice of taking averages over the whole forecasting period of the forecast errors computed for a fixed horizon. 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. The drawback is that since the series of the computed RMSE 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.

The best leading indicator is defined as the variable (for each sample) that provides the largest reduction in RMSE 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 worked well are (a) the growth in the *Journal of Commerce* (*JOC*) price index for industrial materials, (b) M2 growth, (c) growth in average hourly earnings and (d) the number of weekly hours worked. The use of all four variables is however problematic, the first because of the unreliability of the index and the latter three because of their close relationship with inflation. Given that the actual value of the lagged indicator variable is used, its interrelationship with the variable being forecast is somewhat problematic. The exchange rate level and the growth of M1 led consistently to higher RMSEs as do interest rate variables, the unemployment rate, the monetary base, the employment to population ratio and the National Associaton of Purchasing Management (NAPM) composite index.

We can replicate all the results in Cechetti *et al.*. Moreover, we have also considered the properties of 4-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 the somewhat time intensive manner in which the choice of the best leading indicator is made, by looking at the performance of the indicators one at a time. Moreover, the lag length is taken to 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 three different routes to do so. First, we use *PcGets* as 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. Finally, we consider factor methods as a way of summarising efficiently the information contained in a large data set.

PcGets is a computer-automated algorithm focusing on general to specific reductions of models.[3] 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 Δp_t and r_t and up to four lags each of $\Delta(m - p)_{t-1}, \Delta^2 p_t$ and Δr_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.

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. Notice though that 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 (1998), Artis, Banerjee and Marcellino (2001) and Marcellino, Stock and Watson (2000, 2002) 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 data set for forecasting. Here we briefly introduce the representation and estimation theory for the dynamic factor model.

---

[3] See Hendry and Krolzig (1999), Krolzig and Hendry (2001) and Hoover and Perez (1999).

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 (1998). 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 Λ 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 (1998) 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 $T\times T$ covariance matrix of the X's. 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.

It should be stressed that the 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

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

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 (1998) allow either for a few abrupt changes, or for a smooth evolution as modelled by a multivariate random walk for Λ_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 (1998). 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. Our primary data source is the OECD Main Economic Indicators database and the data are seasonally adjusted at the same source. Altogether we use 50 inflation indicators and 60 GDP growth indicators. Note that some variables are used as an indicator in both levels and in growth rates as indicated in the data appendix. We choose to include some variables both in 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 would 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.

The group of output indicators is the largest and it contains mainly data for different indexes of industrial production, plus main aggregate demand components in the case 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 and the miscellaneous group that contains the data from the balance of payments and the consumers confidence indicator.

In their influential paper on forecasting inflation, Stock and Watson (1999) use monthly data for the US, ranging from 1959:1 to 1997:9. The source of data is the DRI - McGraw Hill Basic Economic database (which is not available to us for purposes of updating). The number of variables in their paper is remarkably high (168), since an extensive data set is needed for factor-based forecasting, which is the core of their forecasting exercise. As mentioned above, the categorisation of variables is very comparable to the one used in our paper. The data sets overlap the most for the category "Output variables", where we basically use the same series, but at a different frequency. Stock and Watson's data set is especially richer for categories "employment and hours", which is expected as they use the Phillips curve approach to forecast inflation, and "interest rates", where their data are richer in terms of maturity and include yields of some internationally traded bonds. They also have two additional categories: "inventories and orders", and "consumption (private)".

A more recent paper by Stock and Watson (2001) concentrates on asset prices as forecasting indicators, but is of special interest for our work as it considers forecasting both inflation and GDP growth. In this paper Stock and Watson use quarterly data (ranging from 1959:1 to 1999:4) for 7 OECD countries: Canada, France, Germany, Italy, Japan, the UK and the US. For each country they use 38 series, mostly asset prices, but also selected measures of real activity, monetary aggregates and commodity prices. The basic differences with our data set can be summarised as follows. Due to the focus on asset prices the number of output variables as indicators is reduced but some specific asset prices (house price index, dividend price index, prices of gold and silver) are included. For the same reason they consider also a set of interest rates that is richer in terms of maturity. In addition, while sub-indexes of the

CPI are omitted, they include the GDP deflator, commodity prices and oil prices the latter two being the indicators of external price pressures.

6. Forecasting US Inflation

6.1 Comparison with Cecchetti *et al.* (univariate indicators)

In order to benchmark our subsequent analysis, 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 8-step-ahead RMSEs, 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. The best performing indicator is however always better than a pure autoregression.

No indicator consistently out-performs the autoregression. This can be seen both from the fact that no indicator is best more than twice and that the ranking of the best leading indicator changes over time. In fact, 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 M2 aggregate is deleted in 4 out of 13 periods. The RMSEs, 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 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 such as those incorporated in PcGets instead of using fixed number of lags in the estimating (and forecasting) models.

Table 2 repeats the Cecchetti *et al.* analysis, but uses our larger and longer data set. As mentioned in Section 5, some of the variables in the Cecchetti *et al.* data set had to be omitted because of our inability to update them. However, the only *relevant* variable (in terms of its good performance in Cecchetti *et al.*) not included here is the *JOC* index.

In Table 2 we make use of 50 indicators, as listed in Table 3. There are 16 periods for evaluation. Even with more indicators and a longer sample we see that in 10 out of the 16 periods the autoregression does better than at least 50% of the models with an indicator. The best indicator remains better than the autoregression. There is a lot of variation in the best indicator but a component of industrial production works well in the 1980s. As in Table 1, some indicators do very badly, for example the capacity utilisation rate and the index for IP construction, but are not deleted. Some variables such as the consumer confidence indicator and M2 growth are consistently deleted which is at odds with a specific Cecchetti *et al.* finding noted in Section 2 above but is in broad agreement with the observation that money variables are among the worst indicators.

Table 3 provides a more disaggregate analysis of the leading indicators, by noting the individual performances of each of the indicators. It may be noted that 7 out of the 50 indicators do better than the autoregression more than half of the time. Growth of industrial production (aggregate, durable and non-durable), unemployment rate, growth of retail sales and energy use measures are among the best performing indicators. As noted above, the consumer confidence indicator is the most deleted variable, while lhman (weekly hours worked in manufacturing), fs (share prices), growth of M1 and growth of whemang (growth of hourly manufacturing earnings) also perform poorly.

12

6.2 Comparison with factors and groups of indicators

Table 4 shows the forecasting performance of the first six estimated factors which, taken together, account for about 70% of the variance of the 50 indicators. It turns out that no single factor is better than AR forecasts more than half the time. No single factor produces the best forecast in any time period. Factors 4 and 5 are often deleted, while Factor 6 is not deleted from the forecasting model. This emphasises the importance of factor selection as opposed to looking just at the first few factors, since along with Factors 2 and 3, Factor 6 is in fact one of the better-performing indicators.

Our results for the use of factor models for forecasting inflation are on the whole disappointing. This is in agreement with the findings reported by Stock and Watson (1998), who show that factor models are better for forecasting real variables than prices. However, direct comparison with Stock and Watson on the basis of Table 4 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 and use several factors jointly. We shall return to an evaluation of the last of the three causes in the next paragraph.

Table 5 reports the forecasting performance of groups of variables. Variables are divided into subsets of real variables, price variables and financial variables. Broadly speaking, the Real 1 group contains growth rates of measures of industrial productivity plus the unemployment rate. Real 2 contains measures of turnover, consumers' confidence and the capacity utilization rate. Groups 1 to 4 include some of the best-performing single indicators, based on considering the results reported in Tables 2 and 3. Group 5 contains the six factors analysed separately earlier and Group 6 combines the best indicators with the best factors. Details are contained in Table 5c.

Variables are included conservatively in the final version of the GUM (where *PcGets* is used to make the selection), based on a 1% significance level criterion (include if statistic rejects at 1%). The results of this conservative selection strategy of variables for each subset are reported in Table 5a.

It may be seen that in 8 out of 10 periods some subsets of variables work better than autoregressions.[4] Only in 1 out of 10 periods however is a subset better than the univariate leading indicator and even then the gains are relatively small (less than 5%). Reflective of the poor performance of factors reported above, the group composed of factors (Group 5) is never the best performer, among groups only once.

No major changes are noted when a more liberal inclusion rule is adopted (include if statistic rejects at 5%) as reported in Table 5b. As before, the groups do better than a pure autoregression in 8 out of 10 periods, while the best univariate leading indicator model is out-performed only 2 times out of 10.

6.3 Sensitivity analysis

Table 6 reports the results of rolling estimation for one- to eight-step-ahead forecasts using a 10-year window for estimation. Rolling estimation is sometimes advocated as a means of robustifying the forecasts in the presence of structural changes in the estimation period (see *e.g.* Pesaran and Timmerman (1995)). In 7 out of the 15 evaluation periods of relevance (the first case being exactly the same as the first period for recursive estimation), the mean squared forecasting error of the autoregressive model improves. Most of the gains are for the first part of the sample and the gains occur because dropping early observations does not hurt, since they are the noisy ones. Recursive estimation dominates in the later part of the sample because of more and better-behaved observations.

In 7 out of 15 periods the rolling estimators do better than their recursive counterparts in terms of the number of indicators that perform better than AR. In 6 out 15 periods there is an improvement (over recursive) in the performance of the best indicators. As before, most of the improvements are in the first part of the sample. The best, worst and deleted indicators are different from the recursive case – i.e. the method of estimation matters.

---

[4] There are fewer evaluation periods in Table 5 because of initialisation restrictions and because not all of the variables are available in the most recent periods.

On the basis of these findings, it is not clear whether recursive or rolling methods should be used for forecasting inflation. We will return to this issue in the next sub-section.

Table 7 reports the results for one- to four-step-ahead forecasts using recursive estimation. This is of interest in order to evaluate whether the relatively poor performance of the indicators is attributable to the short or long forecasting horizon. For the shorter forecasting horizon, at least 50% of the indicators do better than the autoregression in 7 out of 17 periods. This is in contrast with the corresponding number of 10 out of 16 periods in Table 2, so that indicators may do better at longer forecasting horizons. Matching the result for $h = 8$, the best univariate indicator is always better than the autoregression, although the best indicators are different from those obtained for $h = 8$. As in the $h = 8$ case, the RMSE is higher in the first part of sample than in the second part.

Table 8 provides the results corresponding to Table 3 for four-step-ahead-forecasts, *i.e.* the ranking of the indicators. We see that fewer indicators are helpful, but when they are helpful they are better than before. In fact 8 out of 50 indicators are better than the autoregression at least 9 out of 17 times, and 18 indicators are better at least 8 out of 17 times. The best performing variables using this criterion are by and large the same as those for the longer horizon.

Finally, Table 9 provides the results of rolling estimates for four-step-ahead forecasts. For 3 out of the 17 periods, at least 50% of the indicators are better than pure autoregressions. In 9 out of the 17 periods, the RMSEs from rolling estimation is lower than those derived from recursive estimation. In 7 out of 17 periods, the best indicator RMSE is better than its recursive counterpart, but in all these cases the gain is less than 10%.

6.4   Forecast pooling

Starting with the pioneering work of Bates and Granger (1969) it is well known that a combination of forecasts may perform betten than each of the single constituent forecasts. The 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 (1999). First, a simple average of all the single indicator forecasts, and second the median of

the forecasts. The latter could be more robust since we have seen that some indicators produce forecasts with high RMSEs.

Table 10 reports the ratios of the RMSEs of the pooled forecasts relative to the AR benchmark. As expected, the median performs systematically better than the average forecasts. For up to eight-step-ahead forecasts, the median forecasts are at least as good as the AR in 7 out of 16 cases for both recursive and rolling estimation. The corresponding figures for up to four-step-ahead forecasts are 8 and 11 out of 17 cases respectively.

Though these results are noteworthy, we see from Table 2 that a single (though time-varying) indicator can beat the AR in each time period. Hence, we return to considering in a more realistic context the performance of single indicator forecasts.

6.5 Ex-ante evaluation

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 for each indicator. The indicator that provides the lowest RMSE 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 for recursive and rolling estimation, and the results are reported in Table 11. For recursive estimation, when $h = 8$ the feasible indicator forecasts are better than autoregressions in 5 out of 9 periods, mostly in recent years. When $h = 4$ the performance of the feasible indicators deteriorates in

line with the ex-post analysis. Now the autoregression is out-performed only in 3 out of 11 periods. When using rolling estimation, $h = 8$ and $h = 4$ provide similar results (autoregression is outperformed 6 out of 10 and 5 out of 11 periods respectively). The relative ranking of recursive and rolling estimation is thus not clear-cut also in this ex-ante framework.

6.6.   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. 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 is higher than when using single indicators. Third, the results are robust to the use of rolling estimation and choice of forecasting horizon, even though the indicators appear to perform better for longer horizons. Fourth, the median pooled estimator performs well in at least 50% of the cases. Finally, in a pseudo ex ante context, the indicators can hardly beat the autoregressions more than 50% of the time.

7. Forecasting GDP growth

7.1   Single indicators for GDP growth

Table 12 provides the results using *PcGets* to specify the model for forecasting GDP growth using univariate indicators (again with a maximum of 6 lags both for the dependent variable and for the indicator). We have the same set of 50 candidate indicators as for inflation, plus a set of 10 GDP components. Details may be consulted in Table 13. In contrast with Table 2, we find that in all the 16 evaluation periods the autoregression does better than at least 50% of the models with an indicator. The best indicator, however, remains better than the autoregression although it changes over time (with a component of industrial production performing well recently). Many more variables are deleted than for the inflation models, but most of the 60are in fact not deleted.

Table 13 provides a disaggregate analysis, where we find that only 3 out of the 60 indicators do better than an autoregression at least half the time (growth in the production of non-durable goods, construction, and export). Broadly speaking, other good indicators for GDP growth include a confidence variable, components of industrial production and gross domestic product, and some foreign exchange variables.

7.2 Factors and groups of indicators

Table 14 shows that no single factor is better than autoregression based forecasts more than half the time. Moreover, no single factor produces the best forecast in any time period, except for one period, while factors 4- 6 are often deleted.

Table 15 considers sets of variables grouped (as in the previous section) into real, financial and price variables, factors, and the best single indicators, details are provided in panel c of the table. Focusing on panel a, where regressors are not deleted from the regression model if significant at the 1%, we find that group 1 (the factors) and group 2 (the best single indicators) produce in general the lowest root mean squared error. The former outperforms the autoregression in 5 out of 9 cases, the latter in 6 out of 9. Similar results are obtained for the factors from the liberal (as opposed to conservative) selection strategy, as reported in panel b. The performance of the other groups deteriorates noticeably. It is worth stressing the better forecasting performance of the factor model for GDP growth than for inflation, in line again with Stock and Watson (1998) who found a better performance for real variables, and the fact that groups 1 and 2 also outperform the single indicators in a few cases (4 out of 9).

7.3 Sensitivity analysis

Table 16 reports the results for rolling estimation, forecasting one to eight steps ahead. In 6 out of the 16 evaluation periods, the mean squared forecasting error of the autoregressive model improves. The best indicator is better than the autoregression, although as before it changes over time. More variables are deleted than for inflation, but most of the 60 are not deleted. The ranking of the indicators remains the same as under recursive estimation.

Table 17 demonstrates the performance of indicators in forecasting GDP growth one to four steps ahead. The results are qualitatively very similar to the $h = 8$ case (see Table 12

for comparison). The autoregression is always better than at least 50% of the models with an indicator. The best indicator model is better than the autoregression, with the indicator changing over time. Industrial production increases in importance, particularly in recent periods.

From the ranking of indicators in providing one to four-step-ahead forecasts, reported in Table 18, only 3 out of 60 indicators are better than the autoregression more than half the time. In particular, ipint (industrial production of intermediate goods) and ipconsg (growth of industrial production in the consumption sector) outperform the autoregression 10 out of 17 and 9 out of 17 times respectively. These two variables also provide the best forecasts 3 and 2 times respectively. The federal funds rate outperforms the autoregression 9 out of 17 times.

Finally, Table 19 gives the results for one to four-step-ahead forecasts using rolling estimation. Focusing on the autoregression, we find that rolling estimation is better than recursive in 6 out of 16 periods, while the best indicator-rolling model is better than the best indicator recursive model in 8 out of 16 periods. These results taken together do not provide a clear ranking of the competing estimation methods, and justify the comparison of rolling versus recursive methods in a pseudo-real time evaluation, which is undertaken below.

7.4 Forecast pooling

The results from average and median forecasts are slightly less encouraging than those reported for inflation, as a likely consequence of the lower average quality of the single indicators. Median forecasts work better than the average of the indicator forecasts, and are at least as good as the AR benchmark in 5 out of 16 cases when h=8 for both recursive and rolling estimation. The figures become 4 and 9 out of 17 respectively when h=4.

7.5 Ex-ante analysis

Table 21 reports the results of the implementation of the ex-ante evaluation method described in Section 6.4. As before, the procedure is implemented both for $h = 8$ and $h = 4$ and for recursive and rolling estimation (columns labelled by rec. and roll.). For both recursive and rolling estimation, when $h = 8$ the feasible indicator forecasts are better than autoregressions in 8 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. Comparing recursive and rolling estimation, the former appears to perform more strongly, in particular for the autoregressive model.

7.6 Summary

As for inflation, four 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, and in a few cases it is also better than the single indicators (while this was never the case for inflation). Third, the results are robust to the use of rolling estimation and choice of forecasting horizon. Finally, and more importantly, ex ante the indicators can beat the autoregressions more than 80% of the times.

8. Conclusions

The first contribution of this paper is the empirical comparison of two alternative approaches to information extraction from a large data set for forecasting, namely, the use of an automated model selection procedure versus the adoption of a factor model. In the case of inflation it turns out that both methods are systematically beaten by single indicator models, with the factor models being the worse of the two. For GDP growth the performance of both methods improves and becomes comparable, though in more than 50% of the cases single indicators are still preferable.

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 than the autoregression in at least 50% of the evaluation periods. In an ex-ante context, we have developed a feasible procedure that allows

20

the construction of indicator based forecasts that outperform the autoregressions in about 50% for inflation and 80% for GDP growth.

Finally, the results we have obtained appear to be robust to the choice of the forecasting horizon and to the use of either recursive or rolling estimation.

References

Artis, Michael, Anindya Banerjee, and Massimiliano Marcellino. "Factor Forecasts for the UK." CEPR Working Paper 3119, 2001.

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

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

Cecchetti, Stephen G., Rita S. Chu, and Charles Steindel. "The Unreliability of Inflation Indicators." Current Issues in Economics and Finance 4/6, Federal Reserve Bank of New York, 2000.

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

Geweke, John. "The Dynamic Factor Anlysis of Economic Time Series." In *Latent Variables in Socio-Economic Models*, edited by Denis J. Aigner and Arthur S. Goldberger, Ch. 19. Amsterdam: North-Holland, 1977.

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

Hendry, David F., and Hans-Martin Krolzig. "Improving on `Data mining reconsidered by K.D. Hoover and S.J. Perez." *Econometrics Journal* 2 (1999), 202—219.

Hendry, David F., and Hans-Martin Krolzig. "Computer automation of general-to-specific model selection procedures." *Journal of Economic Dynamics and Control* 25 (2001), 831-866.

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

Marcellino, Massimiliano, James H. Stock, and Mark W. Watson. "A Dynamic Factor Analysis of the EMU." *mimeo* (2000), Bocconi University.

Marcellino, Massimiliano, James H. Stock and Mark W. Watson. "Macroeconomic Forecasting in the Euro Area: Country Specific Versus Euro Wide Information." *European Economic Review* (2002, forthcoming).

Mizon, Grayham E., and Jean-François Richard. "The encompassing principle and its application to non-nested hypothesis tests." *Econometrica* 54 (1986), 657-678.

Pesaran, Hashem and Alan Timmermann. "Predictability of Stock Returns: Robustness and Economic Significance." *Journal of Finance* 50 (1995), 1201-1228.

Sargent, Thomas J., and Christopher 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 Christopher A. Sims, Minneapolis: Federal Reserve Bank of Minneapolis.

Stock, James H. and Mark W. Watson. "Diffusion Indexes." NBER Working Paper 6702 (1998).

Stock James H. and Mark W. Watson. "Forecasting Inflation." *Journal of Monetary Economics* 44 (1999), 293-335.

Stock James H. and Mark W. Watson. "Forecasting Output and Inflation: The Role of Asset Prices." NBER Working Paper 8180 (2001).

Data Appendix

List of variables and transformations used

Variable	Trans	Description

Output variables

Variable	Trans	Description
Gdp	DLV	GDP – 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

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

Retail, manufacturing and trade sales

Variable	Trans	Description
rtvaltot	LV, DLV	retail sales – total (MN$, S.A.
rtvoldur	LV, DLV	retail sales – durables (MN$, S.A.)
whval	LV, DLV	wholesale sales – total (MN$, S.A.)
cars	LV	passeneger car registrations (000 number, S.A.)

Housing

Variable	Trans	Description
ccost	LV, DLV	cost of construction, residential 1995=100

Stock prices

Variable	Trans	Description
fs	LV	NYSE share prices (1995=100)

Exchange rates

Variable	Trans	Description
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.)

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.)

Transformations used: LV – levels, DLV – annual growth rate
S.A. indicates that the data have been seasonally adjusted at source (OECD Main Economic Indicators).

INFLATION INDICATORS

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

b	Number of Indicators That Performed		Root-Mean-Squared-Error			*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)	-

Table 2: Performance of indicators in forecasting inflation up to eight quarters ahead

Estimation period	Number of Indicators That Performed		Root-Mean-Squared-Error			PcGets Deletes
	Better Than AR	Worse Than AR	Autoreg.	Best Indicator	Worst Indicator	
75:1 84:4	6	44	1.72	1.41 (fy10gov)	6.99 (ipc)	
75:1 85:4	8	42	2.29	2.04 (ipcnd)	7.19 (ipc)	Conf, rtvoldur
75:1 86:4	38	12	3.73	0.91 (ipc-g)	4.88 (ipc)	conf, M2-g, cpidur-g, wheman-g
75:1 87:4	29	21	1.71	0.88 (ipc-g)	5.46 (cap)	conf, lhman, M2-g, wheman-g
75:1 88:4	29	21	2.39	1.96 (cpidur-g)	5.94 (ipint.)	conf, lhman, fs, wheman-g
75:1 89:4	16	34	2.30	1.68 (ipcons)	6.38 (rtvoldur)	conf, fs, M2-g, M3-g, cpifuel, cpifuel-g, wheman-g, ccost
75:1 90:4	26	24	2.90	1.42 (ipi–g)	5.63 (ereff)	conf, lhman, fs, M2-g, M3-g, cpifuel, wheman-g, ccost
75:1 91:4	5	45	0.56	0.49 (ipcnd–g)	3.72 (rtvaltot)	conf, fs, M2-g, ereff, ereff-g, eneff-g, cpidur-g, wheman-g,
75:1 92:4	3	47	0.66	0.62 (cpf–g)	30.05 (cap)	conf, lhman, fs, M2-g, M3-g, ereff, wheman-g
75:1 93:4	2	48	0.70	0.59 (cpf–g)	53.29 (cap)	conf, lhman, rtvaltot, rtvoldur, fs, M2-g, ereff, wheman-g, ccost, ccost-g
75:1 94:4	18	32	0.96	0.54 (rtvaldur-g)	15.17 (cap)	conf, lhman, rtvoldur, fs, M2-g, wheman-g, ccost, ccost-g
75:1 95:4	11	39	1.02	0.83 (m2–g)	5.73 (ipcons)	
75:1 96:4	11	39	1.09	0.63 (m2–g)	6.83 (ipcons)	Cap, conf, fs
75:1 97:4	9	41	1.76	0.51 (fy10gov)	3.72 (cpifuel)	cap
75:1 98:4	46	4	1.76	0.45 (rtvoldur-g)	2.84 (ipcnd)	
75:1 99:4*	27	17	1.22	0.77 (ipi-g)	4.35 (wc–g)	

*6 variables less due to data not available up to 2001q4

Table 3: Ranking the inflation indicators

	Outperforms Autoregression	Underperforms Autoregression	Produces Best Forecast	Produces Worst Forecast	*PcGets* Deletes
			Number of Times the Indicator		
ip	6	10	-	-	-
ipc	4	12	1	-	-
ipcd	4	12	-	-	-
ipcnd	5	11	1	1	-
ipint	4	12	-	1	-
ipi	5	11	1	-	-
ipman	5	11	-	-	-
ipcons	4	12	1	5	-
cap	4	12	-	4	2
conf	2	14	-	-	11
ipg	7	9	-	-	-
ipcg	8	8	1	-	-
ipcdg	6	10	-	-	-
ipcndg	9	7	1	-	-
ipintg	8	8	-	-	-
ipig	6	10	1	-	-
ipmang	7	9	-	-	-
ipconsg	5	11	-	-	-
lurat	8	8	-	-	-
lhman	3	13	-	-	7
rtvaltot	4	12	-	1	1
rtvoldur	5	11	-	1	3
whval	9	7	-	-	-
rtvaltotg	7	9	-	-	-
rtvoldurg	8	8	2	-	-
whvalg	9	7	-	-	-
cars	5	11	-	-	-
fs	2	14	-	-	8
fy10gov	10	6	2	-	-
fcod	6	10	-	-	-
ff	6	10	-	-	-
spread10	6	10	-	-	-
spread3	4	12	-	-	-
M1g	6	10	-	-	-
M2g	4	12	2	-	7
M3g	4	12	-	-	3
lebankg	7	9	-	-	-
ereff	6	10	-	1	3
ereffg	5	11	-	-	1
eneff	8	8	-	-	-
eneffg	5	11	-	-	1
cpifuel	7	9	-	1	2
cpfg	11	5	2	-	-
cpidurg	6	10	1	-	2
cpifuelg	9	7	-	-	1
whemang	4	12	-	-	9
whetotg	9	7	-	-	-
wcg	6	10	-	1	-
ccost	7	9	-	-	4
ccostg	7	9	-	-	2

Table 4: Forecasting performance of US factors for US inflation

	Number of Times the Indicator				
	Outperforms Autoregression	**Underperforms Autoregression**	**Produces Best Forecast**	**Produces Worst Forecast**	*PcGets* **Deletes**
US-F1	4	11	-	-	-
US-F2	6	9	-	-	-
US-F3	6	9	-	-	-
US-F4	3	12	-	-	11
US-F5	1	14	-	-	9
US-F6	6	9	-	-	-

Table 5a: Performance of groups of variables in forecasting inflation up to eight quarters ahead (conservative strategy)

	Root-Mean-Squared-Error											
Estimation period	**Autoreg.**	**Real 1**	**Real 2**	**Financial**	**Prices**	**Group 1**	**Group 2**	**Group 3**	**Group 4**	**Group 5**	**Group 6**	**Best single**
75:1 90:4	2.90	**2.12**	**2.66**	4.35	**1.24**	**1.54**	**1.61**	3.40	**1.59**	**1.60**	**1.69**	1.42
75:1 91:4	0.56	**0.46**	1.84	2.70	0.82	1.16	1.42	1.86	1.71	1.42	2.19	0.49
75:1 92:4	0.66	1.03	2.55	1.91	1.13	1.44	1.48	0.96	1.47	1.26	1.40	0.62
75:1 93:4	0.70	1.16	58.81	4.18	0.95	1.98	1.46	1.37	1.80	1.18	1.62	0.59
75:1 94:4	0.96	2.21	**0.82**	2.53	**0.87**	1.70	**0.93**	1.17	1.00	**0.98**	1.14	0.54
75:1 95:4	1.02	3.42	1.19	1.71	1.11	1.68	1.16	1.44	**1.01**	1.26	1.69	0.83
75:1 96:4	1.09	3.14	1.43	6.90	1.75	1.70	1.16	1.56	**1.05**	**0.95**	**0.96**	0.63
75:1 97:4	1.76	1.36	**0.77**	2.45	1.71	**1.38**	**1.24**	**0.53**	**1.00**	**1.32**	0.69	0.51
75:1 98:4	1.76	**2.10**	**1.01**	2.45	**1.28**	**0.68**	**0.52**	**0.47**	**1.40**	1.92	**1.49**	0.45
75:1 99:4	1.22		1.24	2.96	2.18	1.66	**0.86**	**0.83**	**0.88**			0.77

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

Table 5b: Performance of groups of variables in forecasting inflation up to eight quarters ahead (liberal strategy)

	Root -Mean-Squared-Error											
Estimation period	**Autoreg.**	**Real 1**	**Real 2**	**Financial**	**Prices**	**Group 1**	**Group 2**	**Group 3**	**Group 4**	**Group 5**	**Group 6**	**Best single**
75:1 90:4	2.90	**1.88**	**2.33**	3.40	**2.42**	**1.42**	2.28	3.80	**1.75**	**1.60**	**1.68**	1.42
75:1 91:4	0.56	**0.46**	1.89	1.24	1.92	1.17	1.18	1.53	1.45	1.42	1.83	0.49
75:1 92:4	0.66	1.03	17.86	5.45	5.39	1.23	2.76	2.15	1.20	1.26	1.83	0.62
75:1 93:4	0.70	1.89	19.39	2.01	2.27	2.07	3.09	2.45	3.03	1.18	1.76	0.59
75:1 94:4	0.96	2.21	**0.82**	1.88	**0.75**	1.46	1.57	**0.69**	2.24	**0.98**	1.59	0.54
75:1 95:4	1.02	2.84	1.28	5.53	1.26	2.00	1.16	1.44	**0.94**	1.26	1.85	0.83
75:1 96:4	1.09	5.05	1.69	4.49	1.65	1.40	1.16	**0.95**	**1.05**	**0.95**	1.31	0.63
75:1 97:4	1.76	**1.36**	2.18	**1.02**	**1.52**	**1.42**	**0.56**	**0.51**	**0.66**	**1.32**	1.65	0.51
75:1 98:4	1.76	3.89	**1.63**	3.22	**1.51**	**1.15**	**0.39**	**1.33**	**0.53**	1.92	**1.03**	0.45
75:1 99:4	1.22	5.31	1.24	2.07		1.47	**0.93**	2.01	**1.16**			0.77

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

Table 5c: Groupings of variables in Tables 5a and 5b

Real 1	Real 2	Financial	Prices	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
ipg	cap	fs	Cpifuel	ipc-g	ipc-g	ipcnd-g	ipc-g	US factors	Ipcnd-g
ipcg	conf	ff	Cpfg	ipcnd-g	ipcnd-g	whal	ipcnd-g		Whal
ipcdg	lhman	spread10	Cpidurg	ipint-g	whal-g	whal-g	ipint-g		Whal-g
ipcndg	rtvaltotg	spread3	cpifuelg	lurat	fy10gov	fy10gov	lurat		fy10gov
ipintg	rtvoldurg	M1g	whemang	rtvoldur-g		cpfg	rtvoldur-g		Cpfg
ipig	whvalg	M2g	whetotg	whal			whal		US F2
ipmang	cars	M3g	wcg	whal-g			whal-g		US F3
ipconsg		lebankg	ccost	fy10gov					US F6
lurat		ereff	ccostg	cpf-g					
		ereffg		cpifuel-g					
		eneff		whetot-g					
		eneffg							

Group 1 – 11 variables, all outperform the autoregression at least 50% of times (see Table 2)
Group 2 – variables in this group always (10 subperiods) enter significantly in the final models from Group 1
Group 3 – best 5 variables from Table 2
Group 4 – real variables from Group 1
Group 5 – 6 US factors
Group 6 – best 5 variables from Table 2 + best 3 factors from Table 4

Table 6: Performance of indicators in forecasting inflation eight quarters ahead, 10-year rolling window

Estimation period	Number of Indicators That Performed		Root-Mean-Squared-Error			*PcGets* Deletes
	Better Than AR	Worse Than AR	Autoreg.	Best Indicator	Worst Indicator	
75:1 84:4	6	44	1.72	1.41 (fy10gov)	6.99 (ipcons)	
76:1 85:4	10	40	2.27	2.11 (ipcnd)	7.19 (ereff)	conf, ipc-g, rtvaltot, rtvoldur, m2-g, cpifuel, ccost
77:1 86:4	5	45	1.13	0.75 (ipc)	12.57 (fs)	conf, lhman, m1-g, m2-g
78:1 87:4	15	35	1.43	0.69 (lurat)	10.31 (fs)	conf, rtvaltot, rtvoldur-g, m2-g, m3-g, wheman-g
79:1 88:4	14	36	2.31	1.32 (fcod)	6.31 (ipint-g)	conf, rtvoldur-g, fs, m3-g, wheman-g, whetot-g, wc, ccost, ccost-g
80:1 89:4	14	36	2.61	1.99 (whal-g)	5.42 (rtvoldur)	conf, cars, fs, m1-g, m3-g, ereff, ereff-g, eneff, cpifuel-g, wheman-g, whetot-g, ccost, ccost-g
81:1 90:4	29	21	2.18	0.57 (ff)	6.88 (m3-g)	M1-g, wc-g
82:1 91:4	25	25	1.25	0.50 (cars)	3.48 (ipc)	Ipcons-g, ccost-g
83:1 92:4	17	33	1.44	0.71 (ipint-g)	10.76 (cap)	ipc-g, ipcd-g, ipman-g, ipcons-g, fcod
84:1 93:4	21	29	1.37	0.71 (eneff)	19.67 (cap)	ipc-g, ipcd-g, ipman-g, fs, ccost
85:1 94:4	16	34	0.97	0.49 (rtvoldur-g)	26.71 (cap)	ipcd, ipman, ipcd-g, fs, ccost
86:1 95:4	26	24	1.42	0.83 (wheman-g)	2.74 (ipcons)	ipcd, ipi, ipman, cap, ipcd-g, fs
87:1 96:4	23	27	2.08	0.64 (whal)	4.74 (m3-g)	ipcd, ipi, ipman, cap, ipcd-g, ipi-g, ereff-g, eneff-g, ccost
88:1 97:4	9	41	0.93	0.59 (rtvaltot-g)	3.48 (m2-g)	ip, ipint, ipi, ipma, ipcons, cap, conf, ipi-g, ipcons-g, m1-g, m3-g, ereff, ereff-g, eneff-g, cpidur-g, cpifuel-g, wc-g, ccost-g
89:1 98:4	13	37	1.13	0.32 (fcod)	6.48 (m2-g)	
90:1 99:4*	11	39	1.40	0.87 (ipc-g)	3.79 (ipc)	Cap, whal, fs

*6 variables less due to data not available up to 2001q4

Table 7: Performance of indicators in forecasting inflation up to four quarters ahead

Estimation period	Number of Indicators That Performed		Root-Mean-Squared-Error		
	Better Than AR	Worse Than AR	Autoregression	Best Indicator	Worst Indicator
75:1 84:4	5	45	0.75	0.52 (cpidur-g)	6.42 (eneff)
75:1 85:4	15	35	2.76	1.94 (fy10gov)	12.95 (ereff)
75:1 86:4	43	7	3.92	0.69 (ipc)	6.31 (m3-g)
75:1 87:4	29	21	1.23	0.27 (eneff-g)	8.66 (fs)
75:1 88:4	21	29	1.23	0.87 (rtvoldur-g)	2.91 (spread10)
75:1 89:4	21	29	3.22	1.34 (ereff)	5.25 (rtvoldur)
75:1 90:4	34	16	3.37	1.53 (rtvaltot)	5.28 (ereff)
75:1 91:4	16	34	0.48	0.29 (cap)	4.66 (rtvaltot)
75:1 92:4	6	44	0.60	0.29 (lurat)	2.80 (ccost)
75:1 93:4	5	45	0.71	0.59 (cpifuel-g)	42.34 (cap)
75:1 94:4	2	48	0.50	0.48 (cpf-g)	16.66 (cap)
75:1 95:4	37	13	1.28	0.44 (spread10)	2.20 (ipcons)
75:1 96:4	10	40	1.15	0.64 (cpifuel)	3.76 (ipcons)
75:1 97:4	5	45	0.62	0.34 (ccost)	2.39 (whetot-g)
75:1 98:4	42	8	1.18	0.23 (rtvoldur-g)	2.06 (ipcnd)
75:1 99:4*	30	14	0.81	0.29 (ipc)	1.68 (fs)
75:1 00:4*	25	19	1.65	0.95 (ip-g)	104.24 (fs)

*6 variables less due to data not available up to 2001q4
Note: the list of deleted variables is the same as in Table 2.

Table 8: Ranking the inflation indicators

	Number of Times the Indicator			
	Outperforms Autoregression	**Underperforms Autoregression**	**Produces Best Forecast**	**Produces Worst Forecast**
ip	9	8	-	-
ipc	5	12	2	-
ipcd	6	11	-	-
ipcnd	7	10	-	1
ipint	6	11	-	-
ipi	5	12	-	-
ipman	8	9	-	-
ipcons	4	13	-	2
cap	6	11	1	2
conf	3	14	-	-
ipg	7	10	1	-
ipcg	10	7	-	-
ipcdg	8	9	-	-
ipcndg	10	7	-	-
ipintg	8	9	-	-
ipig	7	10	-	-
ipmang	8	9	-	-
ipconsg	6	11	-	-
lurat	7	10	1	-
lhman	2	15	-	-
rtvaltot	6	11	1	1
rtvoldur	6	11	-	1
whval	9	8	-	-
rtvaltotg	4	13	-	-
rtvoldurg	8	9	2	-
whvalg	8	9	-	-
cars	5	12	-	-
fs	2	15	-	3
fy10gov	9	8	1	-
fcod	7	10	-	-
ff	6	11	-	-
spread10	6	11	1	1
spread3	6	11	-	-
M1g	6	11	-	-
M2g	3	14	-	-
M3g	4	13	-	1
lebankg	8	9	-	-
ereff	5	12	-	-
ereffg	8	9	1	1
eneff	8	9	-	2
eneffg	6	11	1	-
cpifuel	5	12	1	-
cpfg	9	8	1	-
cpidurg	5	12	1	-
cpifuelg	5	12	1	-
whemang	2	15	-	-
whetotg	9	8	-	1
wcg	8	9	-	-
ccost	7	10	1	1
ccostg	9	8	-	-

**Table 9: Performance of Indicators in forecasting inflation up to four quarters ahead
10-year rolling window**

Estimation period	Number of Indicators That Performed		Root-Mean-Squared-Error		
	Better Than AR	Worse Than AR	Autoregression	Best Indicator	Worst Indicator
75:1 84:4	5	45	0.75	0.52 (cpidur-g)	6.42 (eneff)
76:1 85:4	17	33	2.88	2.00 (cpidur-g)	11.22 (ereff)
77:1 86:4	2	48	0.93	0.62 (ipc)	7.19 (fs)
78:1 87:4	16	34	1.04	0.45 (cars)	8.50 (fs)
79:1 88:4	12	38	1.21	0.79 (lurat)	2.81 (ereff)
80:1 89:4	19	31	3.58	1.94 (fy10gov)	5.04 (cpifuel)
81:1 90:4	34	16	2.67	0.65 (whal-g)	5.87 (m3-g)
82:1 91:4	28	22	1.02	0.21 (whetot-g)	2.20 (rtvaltot-g)
83:1 92:4	23	27	1.34	0.50 (eneff)	2.73 (cap)
84:1 93:4	26	24	1.38	0.63 (ipi-g)	18.80 (ccost)
85:1 94:4	24	26	1.13	0.44 (rtvoldur-g)	27.29 (cap)
86:1 95:4	5	45	0.67	0.46 (wc-g)	2.32 (lhman)
87:1 96:4	26	24	1.90	0.61 (whal)	3.29 (conf)
88:1 97:4	6	44	0.60	0.32 (ccost)	2.01 (eneff)
89:1 98:4	8	42	0.84	0.30 (ff)	5.41 (fs)
90:1 99:4*	26	24	0.76	0.33 (ipc)	1.77 (m2-g)
75:1 00:4*	22	28	1.81	1.26 (whal)	414.94 (fs)

*6 variables less due to data not available up to 2001q4
Note: the list of deleted variables is the same as in Table 6.

Table 10: Pooled forecasts of inflation – RMSE relative to benchmark AR

	8q rec		4q rec		8q roll		4q roll	
Period	*Mean*	*Median*	*Mean*	*Median*	*Mean*	*Median*	*Mean*	*Median*
84:4	1,80	1,46	2,47	2,03	1,80	1,46	2,47	2,03
85:4	1,60	1,13	1,41	1,17	1,63	1,10	1,32	1,05
86:4	0,71	0,62	0,63	0,58	2,62	1,92	2,67	2,23
87:4	1,17	0,90	1,14	0,90	1,53	1,13	0,09	0,21
88:4	0,99	0,96	1,14	1,00	1,15	1,05	0,06	0,07
89:4	1,13	1,00	0,98	1,00	1,07	1,00	0,98	1,00
90:4	0,95	0,95	0,87	0,85	1,00	0,91	0,81	0,80
91:4	2,10	1,66	1,76	1,04	1,10	1,00	1,00	0,86
92:4	3,33	2,33	1,90	1,81	1,23	1,07	1,03	1,00
93:4	3,71	1,62	2,82	1,28	1,36	1,00	1,24	0,99
94:4	1,54	1,00	2,52	1,55	1,68	1,07	1,49	1,00
95:4	1,55	1,32	0,87	0,80	1,01	0,99	1,59	1,52
96:4	2,01	1,95	1,56	1,48	0,97	1,00	0,90	0,90
97:4	0,75	0,70	2,03	2,17	1,22	1,00	1,50	1,30
98:4	0,65	0,58	0,75	0,74	1,47	1,47	1,47	1,39
99:4	1,17	1,01	0,94	0,89	1,19	1,08	1,10	0,96
2000:4			2,41	1,04			6,11	1,00

Note to the table: The table reports the RMSE ratios for 1 to 8 (and 1 to 4) step ahead forecasts of inflation using the mean and median of all the single indicator based forecasts, using either recursive (rec) or rolling (10 year window, roll) estimated up to the quarter indicated in the first column.

Table 11: Ex-ante performance of indicators in forecasting inflation

Point in time	RMSE h=8				RMSE h=4			
	AR_{rec}	IND_{rec}	AR_{roll}	IND_{roll}	AR_{rec}	IND_{rec}	AR_{roll}	IND_{roll}
90:4	2.90	5.66	2.18	**2.02**	3.37	**3.36**	2.67	3.32
91:4	0.56	**0.51**	1.25	**0.75**	0.48	**0.48**	1.02	**0.78**
92:4	0.66	1.34	1.44	**1.18**	0.60	**0.55**	1.34	**1.14**
93:4	0.70	0.87	1.37	**0.85**	0.71	1.74	1.38	**0.98**
94:4	0.96	**0.58**	0.97	**0.75**	0.50	0.98	1.13	**0.76**
95:4	1.02	**0.94**	1.42	**1.42**	1.28	**1.22**	0.67	1.18
96:4	1.09	1.95	2.08	**1.30**	1.15	1.89	1.90	**1.25**
97:4	1.76	**1.06**	0.93	1.17	0.62	1.10	0.60	0.61
98:4	1.76	**0.66**	1.13	1.66	1.18	1.76	0.84	1.17
99:4	1.22		1.40	1.54	0.81	0.82	0.76	1.37
00:4					0.30	1.66	1.81	**1.81**

Note to the table: Columns 2 and 4 contain the RMSEs of 1 to 8-step-ahead forecasts with pure AR models obtained with recursive and rolling method respectively. Columns 3 and 5 contain the RMSEs of forecasts produced with the best feasible indicator using recursive and rolling method respectively. Analogous numbers are reported in the right panel of the table for h=4.

GDP GROWTH INDICATORS

Table 12: Performance of indicators in forecasting GDP growth up to eight quarters ahead

Estimation period	Number of Indicators That Performed		Root-Mean-Squared Error			*PcGets* Deletes
	Better Than AR	Worse Than AR	Autoreg.	Best Indicator	Worst Indicator	
75:1 84:4	4	56	1.34	1.22 (3-m COD – ff)	11.09 (eneff)	fbopnet
75:1 85:4	4	56	1.68	1.54 (ccost)	9.82 (eneff)	cpidur-g, fbopnet, fgoodsnet
75:1 86:4	14	46	1.92	1.38 (ccost)	6.38 (fs)	whetot-g, fbopnet, fgoodsnet, fimp, fexp
75:1 87:4	10	50	1.55	1.08 (whal)	8.39 (fs)	eneff-g, cpidur-g, fgoodsnet, fimp, fexp, fexp-g
75:1 88:4	26	34	3.18	1.75 (m1-g)	7.32 (m3-g)	
75:1 89:4	23	37	3.65	1.71 (m1-g)	7.31 (m3-g)	fs, fgoodsnet
75:1 90:4	9	51	1.55	1.11 (cpifuel-g)	4.91 (gdpcons-g)	fgoodsnet
75:1 91:4	3	57	1.91	1.77 (cpifuel-g)	3.75 (rtvaltot-g)	fgoodsnet
75:1 92:4	9	51	2.22	1.94 (rtvoldur-g)	40.08 (cap)	rtvaltot, rtvoldur, cpidur-g, fbopenet, fgoodsnet
75:1 93:4	8	52	1.51	1.43 (lebank-g)	42.78 (cap)	rtvaltot, rtvoldur, rtvaltot-g, cpidur-g, fbopenet, fgoodsnet
75:1 94:4	3	57	1.88	1.64 (fy10gov)	22.74 (cap)	rtvaltot, rtvoldur, rtvaltot-g, cpidur-g, fbopenet, fgoodsnet
75:1 95:4	8	52	1.87	1.36 (ipi– g)	5.33 (m1-g)	cap, rtvaltot, rtvoldur, rtvaltot-g, cpidur-g, fbopenet, fgoodsnet, fimp, fexp
75:1 96:4	11	49	2.00	1.67 (wheman – g)	3.17 (fs)	cap, rtvaltot, rtvoldur, rtvaltot-g, m2-g, eneff-g, cpidur-g, fbopenet, fgoodsnet, fimp, fexp, gdpgov-g
75:1 97:4	19	41	2.60	1.88 (ipint)	3.45 (ipi)	cap, rtvaltot, rtvoldur, rtvaltot-g, m2-g, ereff-g, eneff-g, cpidur-g, wheman-g, ccost, fbopenet, fgoodsnet, fimp, fexp, gdpgov-g
75:1 98:4	21	39	2.35	1.91 (ipcd-g)	3.08 (fy10gov)	cap, rtvaltot, rtvaltot-g, cars, fs, m2-g, ereff-g, eneff-g, cpidur-g, wheman-g, whetot-g, ccost, fbopenet, fgoodsnet, fimp, fexp, gdpgov-g
75:1 99:4*	23	37	2.60	1.07 (ipint)	3.64 (ccost)	cap, rtvaltot, rtvaltot-g, cars, m2-g, ereff-g, eneff-g, cpidur-g, wheman-g, ccost, fbopenet, fgoodsnet, fimp, fexp, gdpgov-g

*6 variables less due to data not available up to 2001q4

Table 13: Ranking the GDP growth indicators (up to 8-quarter forecasts)

	Outperforms Autoregression	Underperforms Autoregression	Produces Best Forecast	Produces Worst Forecast	PcGets Deletes
		Number of Times the Indicator			
infl	5	11	-	-	-
ip	0	16	-	-	-
ipc	5	11	-	-	-
ipcd	2	14	-	-	-
ipcnd	4	12	-	-	-
ipint	7	9	2	-	-
ipi	0	16	-	1	-
ipman	1	15	-	-	-
ipcons	5	11	-	-	-
cap	2	14	-	3	5
conf	7	9	-	-	-
ipg	4	12	-	-	-
ipcg	6	10	-	-	-
ipcdg	4	12	-	-	-
ipcndg	8	8	-	-	-
ipintg	7	9	-	-	-
ipig	4	12	1	-	-
ipmang	2	14	-	-	-
ipconsg	8	8	-	-	-
lurat	3	13	-	-	-
lhman	4	12	-	-	-
rtvaltot	2	14	1	-	8
rtvoldur	3	13	-	-	8
whval	6	10	1	-	-
rtvaltotg	3	13	-	1	7
rtvoldurg	6	10	1	-	-
whvalg	6	10	-	-	-
cars	3	13	-	-	2
fs	0	16	-	3	2
fy10gov	4	12	1	1	-
fcod	4	12	-	-	-
ff	5	11	-	-	-
spread10	5	11	-	-	-
spread3	5	11	1	-	-
m1g	4	12	2	1	-
m2g	2	14	-	-	3
m3g	1	15	-	2	-
lebankg	6	10	1	-	8
ereff	4	12	-	-	-
ereffg	3	13	-	-	3
eneff	2	14	-	-	-
Eneffg	4	12	-	2	5
Cpifuel	2	14	-	-	-
Cpidurg	4	12	-	-	10
Cpifuelg	5	11	2	-	-
Whemang	3	13	1	-	3
Whetotg	6	10	-	-	2
Wcg	4	12	-	-	-
Ccost	4	12	2	1	3

Ccostg	3	13	-	-	-
Fbopnet	3	13	-	-	11
Fgoodsnet	2	14	-	-	14
Fimp	6	10	-	-	7
Fexp	5	11	-	-	7
Fimpg	7	9	-	-	-
Fexpg	8	8	-	-	1
Gdpcg	7	9	-	-	-
Gdpgovg	4	12	-	-	4
Gdpconsg	5	11	-	1	
Gdpig	6	10	-	-	2

Table 14: Forecasting performance of US factors for US GDP growth

	Number of Times the Indicator				
	Outperforms Autoregression	Underperforms Autoregression	Produces Best Forecast	Produces Worst Forecast	Gets Deleted
US-F1	5	10	-	-	-
US-F2	2	12	-	-	-
US-F3	5	10	-	-	-
US-F4	3	12	-	-	7
US-F5	3	12	-	-	4
US-F6	3	12	1 (sample 84)	-	6

Table 15a: Groups of variables (conservative strategy)

	Root-Mean-Squared-Error									
Estimation period	Autoreg.	Real 1	Real 2	Real 3	Financial	Prices	Gr1	Gr2	Gr3	Best single
75:1 90:4	1.55	**2.89**	4.17	**2.21**	*	*	2.65	1.69	3.17	1.11
75:1 91:4	1.91	3.01	4.42	2.57	*	*	**1.32**	2.19	2.63	1.77
75:1 92:4	2.22	2.81	2.71	3.33	3.10	*	**1.88**	**2.12**	2.33	1.94
75:1 93:4	1.51	2.10	30.24	1.81	1.68	1.56	1.67	**1.41**	2.36	1.43
75:1 94:4	1.88	2.30	2.31	2.39	1.88	3.05	**1.27**	**1.77**	2.67	1.64
75:1 95:4	1.87	2.09	1.95	2.40	1.56	3.87	2.00	1.95	2.61	1.36
75:1 96:4	2.00	2.50	2.26	1.80	2.96	3.83	2.43	**1.51**	1.93	1.67
75:1 97:4	2.60	2.63	2.80	2.77	2.23	5.99	**2.31**	**1.59**	2.18	1.88
75:1 98:4	2.35	2.56	11.07	2.46	2.38	3.01	**1.26**	**1.73**	2.51	1.91
75:1 99:4	2.60	**2.40**		**2.29**	5.49					1.07

BOLD indicates that the corresponding RMSE is smaller than the RMSE of pure AR model.
* PcGETS was experiencing problems in these cases. The resulting specific model contained only a constant. It deleted also pure AR terms.

Table 15b: Groups of variables (liberal strategy)

Estimation period	Autoreg.	Real 1	Real 2	Real 3	Financial	Prices	Gr1	Gr2	Gr3	Best single
					Root-Mean-Squared-Error					
75:1 90:4	1.55	3.45	5.35	4.63	**1.98**	3.84	**2.65**	3.54	**2.67**	1.11
75:1 91:4	1.91	3.14	5.49	4.33	2.10	1.96	1.32	2.88	2.13	1.77
75:1 92:4	2.22	2.47	45.59	2.84	3.51	3.66	1.88	3.06	2.38	1.94
75:1 93:4	1.51	2.17	30.24	1.88	3.03	4.52	1.67	2.32	2.36	1.43
75:1 94:4	1.88	2.53	2.34	3.61	2.34	2.62	1.27	2.55	2.67	1.64
75:1 95:4	1.87	2.50	1.61	2.64	6.20	3.08	2.00	2.74	2.61	1.36
75:1 96:4	2.00	2.50	2.02	3.01	2.95	3.83	2.43	2.25	1.93	1.67
75:1 97:4	2.60	2.75	2.69	2.61	3.86	5.25	2.31	2.34	2.20	1.88
75:1 98:4	2.35	2.24	2.35	3.17	2.32	2.90	**1.26**	2.53	2.44	1.91
75:1 99:4	2.60	2.40		4.94	1355.48					1.07

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

Table 15c: Groupings

Real 1	Real 2	Real 3	Financial	Prices	Gr1	Gr2	Gr3
ipg	cap	fbopnet	Fs	Cpifuel	US factors	ipint	ipcndg
ipcg	conf	fgoodsnet	Ff	Cpfg		conf	ipconsg
ipcdg	lhman	fimp	spread10	Cpidurg		ipcndg	fexpg
ipcndg	rtvaltotg	fexp	Spread3	cpifuelg		ipintg	US F1
ipintg	rtvoldurg	fimpg	M1g	whemang		ipconsg	US F3
ipig	whvalg	fexpg	M2g	whetotg		fimpg	
ipmang	cars	gdpcg	M3g	wcg		fexpg	
ipconsg		gdpgovg	Lebankg	ccost		gdpcg	
lurat		gdpconsg	Ereff	ccostg			
		gdpig	Ereffg				
			Eneff				
			Eneffg				

Group 1 – 6 US factors
Group 2 – best 8 variables from Table 2
Group 3 – best 3 variables from Table 2 + best 3 factors

Table 16: Performance of indicators in forecasting GDP growth up to eight quarters ahead, 10-year rolling window

| Estimation period | Number of Indicators That Performed | | Root-Mean-Squared Error | | | *PcGets* Deletes |
	Better Than AR	Worse Than AR	Autoreg.	Best Indicator	Worst Indicator	
75:1 84:4	4	56	1.34	1.22 (spread3)	11.09 (eneff)	fbopnet
76:1 85:4	7	53	1.70	1.56 (ccost-g)	12.64 (ereff)	cpidur-g, fbopnet, fgoodsnet, fecp, fexp-g
77:1 86:4	18	42	2.05	1.37 (ccost)	15.78 (lhman)	eneff, eneff-g, cpidur-g, gdpgov-g
78:1 87:4	11	49	1.55	0.88 (ipcnd-g)	6.89 (fimp)	ereff-g, eneff, eneff-g, cpidur-g
79:1 88:4	26	34	3.12	1.79 (m1-g)	7.05 (m3-g)	
80:1 89:4	27	33	3.73	2.29 (lurat)	5.02 (ff)	
81:1 90:4	28	32	3.05	1.12 (spread3)	6.14 (ipcd)	
82:1 91:4	6	54	1.92	1.66 (ipcons-g)	5.19 (fexp)	ipi-g, ereff-g, eneff-g, cpifuel, cpifuel-g, wheman-g, wc-g, ccost, ccost-g, fimp-g, gdpgov-g,
83:1 92:4	19	41	2.37	1.87 (gdpc-g)	4.25 (rtvaltot)	ereff-g, cpifuel, cpifuel-g
84:1 93:4	4	56	1.49	1.19 (fs)	3.52 (ccost)	cap, ipi-g, rtvaltot-g, whal-g, spread3, m2g, ereff-g, cpidur-g, cpifuel-g, wheman-g, fimp, fimp-g, fexp-g
85:1 94:4	4	56	1.92	1.80 (ccost)	33.77 (cap)	ipi-g, rtvaltot, whal, rtvaltot-g, rtvoldur-g, whal-g, ff, spread3, m2-g, m3-g, eneff-g, cpifuel-g, whetot-g, wc-g, fbopenet, fexp, fimp-g, fexp-g, gdpc-g
86:1 95:4	11	49	2.08	1.36 (ipcd–g)	4.15 (fexp)	cap, rtvaltot, whal, rtvaltot-g, whal-g, ff, spread3, m2-g, m3-g, ereff-g, eneff-g, cpidur-g, wc-g, fimp, fimp-g,
87:1 96:4	14	46	2.13	1.47 (fy10gov)	4.95 (cpidur-g)	infl, cap, ipi-g, rtvaltot, whal, rtvaltot-g, whal-g, ff, spread3, m2-g, m3-g, lebank-g, ereff-g, eneff-g, wheman-g, wc-g, fbopenet, fimp-g, fexp-g
88:1 97:4	21	39	2.78	1.99 (rtvoldur)	5.57 (cpidur-g)	ipcnd-g, rtvaltot-g, rtvoldur-g, whal-g, cars, ff, lebank-g, ereff-g, eneff-g, cpifuel-g, wheman-g, whetot-g, fbopenet, fimp-g, fexp-g
89:1 98:4	14	46	2.38	2.05 (wc-g)	4.53 (fgoodsnet)	ipc-g, whal-g, lebank-g, eneff-g, cpifuel-g, fexp-g
90:1 99:4*	19	35	1.56	1.89 (ipint-g)	7.09 (fgoodsnet)	ipcnd-g, whal-g, m1-g, cpifuel-g, fexp-g, gdpgov-g

*6 variables less due to data not available up to 2001q4

Table 17: Performance of indicators in forecasting GDP growth up to four quarters ahead

Estimation period	Number of Indicators That Performed		Root-Mean-Squared-Error		
	Better Than AR	Worse Than AR	Autoregression	Best Indicator	Worst Indicator
75:1 84:4	17	43	1.65	0.28 (ipcons-g)	6.22 (ipcd)
75:1 85:4	1	59	0.93	0.68 (spread3)	11.96 (eneff)
75:1 86:4	24	36	2.23	1.13 (m3-g)	4.95 (eneff-g)
75:1 87:4	14	46	1.64	1.17 (ipcnd)	8.93 (fs)
75:1 88:4	30	30	1.56	0.74 (ipcnd-g)	5.07 (m3-g)
75:1 89:4	25	35	4.18	2.02 (m1-g)	7.07 (m3-g)
75:1 90:4	15	45	1.82	1.19 (m1-g)	5.22 (gdpcons-g)
75:1 91:4	9	51	1.36	0.91 (infl)	4.83 (rtvaltot-g)
75:1 92:4	25	35	2.66	1.86 (ipcons-g)	4.49 (whal)
75:1 93:4	10	50	1.60	1.37 (whal-g)	56.01 (cap)
75:1 94:4	15	45	1.71	0.82 (m2-g)	20.50 (cap)
75:1 95:4	6	54	2.06	1.69 (ipi-g)	4.88 (m1-g)
75:1 96:4	18	42	1.47	1.23 (ipint)	2.59 (fs)
75:1 97:4	16	44	2.42	1.62 (ipint)	3.45 (ipi)
75:1 98:4	25	35	2.82	2.07 (ipi)	3.61 (fy10gov)
75:1 99:4*	17	37	2.04	1.24 (ipint)	3.18 (ipman)
75:1 00:4*	25	29	2.91	1.14 (ipi)	3.94 (fcod)

*6 variables less due to data not available up to 2001q4

Note: the list of deleted variables is the same as in Table 12.

Table 18: Ranking the GDP growth Indicators (4- quarter forecasts)

	Number of Times the Indicator			
	Outperforms Autoregression	**Underperforms Autoregression**	**Produces Best Forecast**	**Produces Worst Forecast**
Infl	5	12	1	-
Ip	3	14	-	-
Ipc	6	11	-	-
Ipcd	4	13	-	1
Ipcnd	6	11	1	-
Ipint	10	7	3	-
ipi	2	15	2	1
ipman	4	13	-	1
ipcons	8	9	-	-
cap	2	15	-	2
conf	7	10	-	-
ipg	5	12	-	-
ipcg	6	11	-	-
ipcdg	6	11	-	-
ipcndg	7	10	1	-
ipintg	8	9	-	-
ipig	7	10	-	-
ipmang	6	11	1	-
ipconsg	9	8	2	-
lurat	4	13	-	-
lhman	7	10	-	-
rtvaltot	3	14	-	-
rtvoldur	2	15	-	-
whval	5	12	-	1
rtvaltotg	3	14	-	1
rtvoldurg	6	11	-	-
whvalg	7	10	1	-
cars	4	13	-	-
fs	2	15	-	2
fy10gov	6	11	-	1
fcod	7	10	-	1
ff	9	8	-	-
spread10	7	10	-	-
spread3	8	9	1	-
m1g	3	14	2	1
m2g	5	12	1	-
m3g	2	15	1	2
lebankg	6	11	-	-
ereff	7	10	-	-
ereffg	6	11	-	-
eneff	6	11	-	1
eneffg	5	12	-	1
cpifuel	4	13	-	-
cpidurg	4	13	-	-
cpifuelg	5	12	-	-
whemang	3	14	-	-
whetotg	5	12	-	-
wcg	5	12	-	-

ccost	6	11	-	-
ccostg	4	13	-	-
fbopnet	3	14	-	-
fgoodsnet	2	15	-	-
fimp	6	11	-	-
fexp	5	12	-	-
fimpg	7	10	-	-
fexpg	9	8	-	-
gdpcg	7	10	-	-
gdpgovg	5	12	-	-
gdpconsg	7	10	-	1
gdpig	5	12	-	-

Table 19: Performance of Indicators in Forecasting GDP Growth Four Quarters Ahead, 10-year rolling window

Estimation period	Number of Indicators That Performed		Root-Mean-Squared-Error		
	Better Than AR	Worse Than AR	Autoregression	Best Indicator	Worst Indicator
75:1 84:4	17	43	1.65	0.28 (ipcons-g)	6.22 (ipcd)
76:1 85:4	1	59	0.92	0.62 (gdpcons-g)	13.46 (ereff)
77:1 86:4	25	35	2.37	1.13 (m3-g)	9.55 (ereff-g)
78:1 87:4	12	48	1.70	0.47 (ipcnd-g)	7.34 (fimp)
79:1 88:4	27	33	1.50	0.64 (spread10)	4.84 (m3-g)
80:1 89:4	29	31	4.25	2.68 (ipcons-g)	6.06 (ipcd)
81:1 90:4	30	30	2.75	1.10 (cars)	5.97 (ipcd)
82:1 91:4	10	50	1.32	0.71 (whal-g)	5.83 (fbopnet)
83:1 92:4	35	25	3.05	1.97 (ipcd)	4.31 (whal-g)
84:1 93:4	11	49	1.55	0.70 (gdpi-g)	3.39 (fexp)
85:1 94:4	8	52	1.70	1.52 (eneff)	29.21 (cap)
86:1 95:4	12	48	2.23	1.68 (gdpcons-g)	4.18 (lhman)
87:1 96:4	10	50	1.57	0.84 (ccost-g)	4.27 (fs)
88:1 97:4	25	35	2.63	1.74 (ipcons)	5.43 (cpidur-g)
89:1 98:4	28	32	2.89	2.10 (rtvoldur)	3.77 (eneff)
90:1 99:4*	22	32	2.01	1.41 (ipcons-g)	5.48 (fgoodsnet)
75:1 00:4*	7	47	2.94	1.14 (cpifuel-g)	950.65 (fs)

*6 variables less due to data not available up to 2001q4
Note: the list of deleted variables is the same as in Table 16.

Table 20: Pooled forecasts for GDP growth – RMSE relative to benchmark AR

| | 8q rec | | 4q rec | | 8q roll | | 4q roll | |
Period	*Mean*	*Median*	*Mean*	*Median*	*Mean*	*Median*	*Mean*	*Median*
84:4	2,15	1,79	1,50	1,26	2,15	1,79	1,50	1,26
85:4	1,70	1,31	2,99	2,22	1,84	1,46	3,02	2,24
86:4	1,30	1,15	1,27	1,03	1,43	1,05	1,19	0,97
87:4	1,35	1,12	1,43	1,29	1,49	1,35	1,37	1,24
88:4	1,04	0,97	1,08	0,92	1,03	0,98	1,12	0,96
89:4	1,02	0,99	0,96	0,95	0,96	0,96	0,94	0,94
90:4	1,41	1,29	1,51	1,41	0,97	0,97	1,00	0,93
91:4	1,21	1,11	1,53	1,20	1,30	1,16	1,57	1,23
92:4	1,38	1,02	1,09	1,10	1,10	1,05	0,95	0,96
93:4	1,67	1,19	1,15	0,98	1,24	1,15	1,18	1,01
94:4	1,32	1,09	1,48	1,07	1,41	1,09	1,49	1,07
95:4	1,16	1,08	1,20	1,08	1,13	1,06	1,11	1,00
96:4	1,08	1,04	1,24	1,07	1,08	1,00	1,16	1,00
97:4	1,02	1,00	1,04	1,07	1,00	1,00	0,96	0,98
98:4	0,99	1,00	0,98	1,00	1,06	1,01	0,96	0,97
99:4	0,94	1,00	1,12	1,01	1,92	1,70	1,13	1,02
2000:4			6,78	1,08			6,71	1,07

Note to the table: The table reports the RMSE ratios for 1 to 8 (and 1 to 4) step ahead forecasts of GDP growth using the mean and median of all the single indicator based forecasts, using either recursive (rec) or rolling (10 year window, roll) estimated up to the quarter indicated in the first column.

Table 21: Performance of forecast feasible indicators in forecasting GDP growth

| | RMSE h=8 | | | | RMSE h=4 | | | |
Point in time	AR_{rec}	IND_{rec}	AR_{roll}	IND_{roll}	AR_{rec}	IND_{rec}	AR_{roll}	IND_{roll}
90:4	1.55	2.90	3.05	**2.82**	1.82	3.17	2.75	3.67
91:4	1.91	**0.56**	1.92	**0.66**	1.36	**0.99**	1.32	**1.02**
92:4	2.22	**0.79**	2.37	20.91	2.66	**2.35**	3.05	**0.71**
93:4	1.51	**0.53**	1.49	**1.31**	1.60	**0.71**	1.55	**0.91**
94:4	1.88	**1.61**	1.92	**0.64**	1.71	**0.92**	1.70	**1.13**
95:4	1.87	**1.19**	2.08	**1.42**	2.06	**0.58**	2.23	**0.67**
96:4	2.00	2.85	2.13	**0.67**	1.47	**1.31**	1.57	**1.33**
97:4	2.60	**1.78**	2.78	**1.32**	2.42	**2.00**	2.63	**0.90**
98:4	2.35	**0.91**	2.38	2.57	2.82	**0.52**	2.89	**1.17**
99:4	2.60	**2.13**	1.56	**1.40**	2.04	**0.76**	2.01	**0.66**
00:4					2.91	**2.10**	2.94	**1.81**

Notes: see Table 11.