OECD Economic Studies No. 40, 2005/1

INDICATOR MODELS OF REAL GDP GROWTH IN THE MAJOR OECD ECONOMIES

Franck Sédillot and Nigel Pain

TABLE OF CONTENTS

Macroeconomic Analysis and Systems Management Division, OECD Economics Department. The authors are grateful to Jørgen Elmeskov, Michael P. Feiner, Michael Kiley, Giuseppe Parigi, Pete Richardson, Paul Swaim and Martin Weale for helpful comments and suggestions, and to Nicolas Gobalraja and Daria Taglioni for their help in constructing the models for Italy and Japan reported in this paper. All views expressed are those of the authors and should not be taken as representing those of the OECD. The majority of the results in this paper are drawn from Sédillot and Pain (2003), and use data sets available at the end of the first quarter of 2003. The two exceptions are Italy, where the models described in the 2003 paper have been updated, and Japan, for which results are presented for the first time. Since the work on all the indicator models was undertaken there have been a number of important statistical changes in some countries that might alter the reported properties of the models. Examples include the move to chain-linked national accounts in the United Kingdom and the comprehensive revisions of the United States and Japanese national accounts in December 2003 and December 2004, respectively.

INTRODUCTION AND SUMMARY

Accurate and timely information on the current state of economic activity is an important requirement for the policymaking process. Delays in the publication of official statistics mean that a complete picture of economic developments within a particular period emerges only some time after that period has elapsed. Thus considerable resources are, at times, devoted to making an assessment of the immediate past and the current conjuncture as well as projections about future developments. In practice, a regular flow of information is provided by the large number of quantitative and qualitative indicators that appear each month for different sectors of the economy. One challenge for policymakers is to put these together in a consistent manner to obtain a picture of the overall state of the economy.

The needs of policymakers have led to the development of a number of new statistical indicators in the last few years. For instance, for the euro area, the European Commission releases short-run quarterly GDP estimates each month using an indicator model, the CEPR produces a monthly coincident indicator (EuroCOIN) and the EUROFRAME group produces a quarterly GDP growth indicator which is published regularly in the *Financial Times*. A number of countries have also reduced considerably the lag between the end of the quarter and the time of the first official flash estimate of real GDP growth in that quarter. Advance quarterly GDP estimates are now produced in the United States, the United Kingdom and Belgium within four weeks of the end of the quarter, and monthly GDP estimates are made in Canada and Finland. For the euro area, Eurostat now produces a flash estimate six weeks after the end of the quarter. However, there are obvious limits to the extent to which this lag can be shortened further given the need to strike a balance between the timeliness and the accuracy of the preliminary flash estimates.

The purpose of this paper is to illustrate how best to construct indicator models of GDP growth and assess their value for short-term forecasting. A set of econometric models are developed that provide timely estimates of GDP growth for each of the G6 economies, and also the aggregate euro area, in the two quarters following the last quarter for which official data have been published. The models seek to exploit efficiently the considerable amount of monthly conjunctural information that becomes available before the release of the official national accounts data. Information is incorporated from both "soft" indicators, such as business surveys, and "hard" indicators, such as industrial production and retail sales, and use is made of different frequencies of data and a variety of estimation techniques.[1]

There are a number of ways in which monthly data can be used to make projections of a quarterly aggregate. One option is to combine high frequency data into quarterly aggregates that can then be used to forecast GDP before it is published. An extension is to estimate a monthly model for selected indicator series and use this to forecast missing monthly information in the quarter. The resulting quarterly indicator estimates can then be used to forecast the quarterly change in GDP.[2] This approach leaves open the questions as to which indicator series should be used and the means by which quarterly projections can be made from the high frequency models when only partial information is available for the current quarter. It is also often the case that indicator models based on quarterly aggregates of monthly indicators can be run only once indicator data are available for the entire quarter. This limits the lead time between their projections and the first official estimates of GDP for that quarter. In some economies such models could be run only after the first official estimate of quarterly GDP growth has been released.

A second possible approach is to construct coincident (and leading) indicator series by using statistical techniques to combine and summarise the information in a large number of different monthly data series, an idea initiated and popularised by Stock and Watson (1989, 1991 and 2002). The National Activity Index for the United States, issued monthly by the Chicago Fed, provides a real-time example of this approach.[3] But such indicators, which have the virtue of seeking to maximise the largest feasible information set, can be unduly complicated and difficult to understand. This is especially so if there is ever a need to undertake a "post-mortem" exercise to evaluate the factors responsible for past forecast performance.

The methodology followed in this paper therefore concentrates on a smaller range of potential indicator variables, exploiting the information in these series in order to derive a consistent quantitative picture of the state of the economy. High frequency indicators are recast into quarterly GDP figures using univariate or multivariate "bridge" models. This approach has been utilised in a number of empirical studies for OECD member economies. Examples include Parigi and Schlitzer (1995), Ingenito and Trehan (1996) and Bovi *et al.* (2000). We extend such studies by seeking to address four key issues simultaneously: the data selection process, the optimal combination of variables, the treatment of staggered monthly data releases and the optimal means of model selection.

This methodological approach has a number of important features. The procedures used to select the data and the model are automated and identified clearly as, for instance, in Camba-Mendez *et al.* (2001) and Bovi *et al.* (2000), and similar statistical techniques are used across countries. This allows the approach to be replicated readily for other countries. The means by which staggered monthly (or quarterly) data releases can be incorporated is also addressed. This ensures that the indicator models can be run at any point in time at which new monthly data are released, even though the information set available will differ on each occasion. Finally, the selected models for each country have been estimated recursively using a sample ending in each successive quarter over the period 1998Q1-2002Q4. This allows tests of their out-of-sample forecast performance to be undertaken and measures of the uncertainty around their point estimates to be calculated. Tests for multiple forecast encompassing have also been conducted. In addition, forecast directional accuracy is also assessed by investigating the ability of each model to predict whether GDP growth accelerates or decelerates from one quarter to the next.

A potential limitation of many of the model evaluations reported in this paper is that they are conducted using a single vintage of national accounts data. In real time, forecast practitioners use data that are frequently revised or subject to methodological changes in their compilation. It is well known that such changes can have important implications for any assessment of real-time forecasting performance. For example, Croushore and Stark (2003) conclude that findings on the relative forecasting performance of different models based on current vintage data releases do not necessarily carry over to real time data. Koenig *et al.* (2003) also demonstrate the differences that can arise when estimating forecasting equations on different vintages of data. We address this issue in the penultimate section of this paper, which contains a discussion of the actual real time performance of the indicator models.

Eight main findings emerge from the work reported in this paper:

- For current quarter forecasts, that is forecasts made at or after the start of the quarter to which they relate, estimated indicator models appear to outperform autoregressive time series models, both in terms of size of error and directional accuracy. These differences are statistically significant in most countries. This suggests there are clear gains from developing empirical indicator models that use high frequency data.
- Quarterly models do not provide especially timely estimates of current and one-quarter ahead GDP growth, as projections cannot be made until quarterly data on the indicators are available. An approach which combines different frequencies of data and different types of models is shown to be able to provide more accurate near-term projections at any point of time, based on the most recently published monthly conjunctural information, than a forecasting model that uses only data covering periods prior to the quarter being forecast.
- The main gains from the monthly approach start to appear once one month of data is available for the quarter being forecast. This is typically two to three months before the publication of the first official estimate for GDP. This finding is in line with other empirical studies that have been conducted on this topic.[4]

- For one quarter ahead projections in most countries, the performance of the estimated indicator models does not appear noticeably better than that of time series models until 1-2 months of information become available in the quarter preceding the one for which the forecast is being made. However there are some modest gains in terms of directional accuracy from using the indicator models. Japan is the country where it appears most useful to look at the indicator model even when no monthly information is yet available in the quarter prior to the one being forecast.
- The most suitable model for any given information set and any fixed forecast horizon varies both across countries and across time. For the current quarter, models either with hard indicators alone, or with hard indicators combined with survey information, outperform models that use only survey data. The pure hard indicator model appears the most suitable for the United States, the euro area and Italy, whereas some form of combined model, either through estimation or through a consensus of the different model forecasts, appears more suitable for Japan, Germany, France, and the United Kingdom. For the one-quarter ahead forecasts, the inclusion of hard indicator data for the quarter in which the forecast is being made appears to add little to the information provided by surveys. Survey data appear to contain especially useful information in France.
- There are clear limits to the ability of any estimated model to forecast the quarterly rate of GDP growth precisely. Even when a complete set of monthly indicators are available for a quarter, the 70% confidence band (approximately one standard error) around any point estimate for the quarterly rate of GDP growth in that quarter is found to range from 0.4 to 1.0 percentage point, depending on the country or region being forecast. The degree of uncertainty around a point estimate is also found to widen as the forecast horizon lengthens.
- The forecast errors for Japan, Germany and the United States are noticeably larger than for the other countries considered. This appears to stem from the larger standard deviation of quarterly GDP growth in these economies over the post-sample evaluation period.
- The cross-economy pattern of the real-time forecast errors from the indicator models (over 2003Q1-2005Q1) is similar to the cross-economy pattern of the simulated out-of-sample errors (up to 2002Q4) on the single vintage of data used to estimate the indicator models.

A clear overall conclusion that emerges from the work is that it is not optimal to employ a single, fixed coefficient, indicator model for each country at all points in time. Instead, it is preferable to have a suite of indicator models that can be updated automatically as new data appears, with the appropriate choice of model varying over time according to the information set available.

The structure of this paper is as follows. The next section describes the modelling strategy employed, focusing on the choice of data frequency, the types of models estimated, the procedures used to select the indicator variables and the tests employed to evaluate out-of-sample forecast performance. The main empirical results are summarised in the following section, which reports the indicator variables selected for each country and the comparative performance of each of the different types of indicator models estimated. Subsequent sections provide a discussion of a number of practical problems that can arise when seeking to use the indicator models on a real-time basis, the experience to date in using the indicator models, and an illustration of different ways in which estimates of forecast uncertainty can be produced.

THE MODELLING STRATEGY

There are four key features of the overall modelling strategy employed in the study:

- The modelling framework is flexible, allowing forecasts to be generated and compared using a number of different techniques. These range from quarterly predictions using published quarterly data through to quarterly predictions generated from a limited subset of monthly information.
- The variables employed are selected using a strictly defined process.
- The models are estimated with a procedure that determines automatically the optimal combination of current and lagged variables.
- Statistical tests are employed to gauge the forecasting capabilities of each estimated model.

The data set

The set of possible indicators investigated includes "soft" indicators, such as business surveys, "hard" indicators, such as industrial production and retail sales, and financial variables. A comprehensive list of the variables considered is given in Sédillot and Pain (2003). All the data used are seasonally adjusted and in many countries both GDP and the hard indicator series also include working day adjustments. Although many hard indicators are direct components of GDP, this need not mean that other types of indicators do not contain useful information as well. Both soft and financial indicators are released on a timely basis, often two weeks or more before the monthly hard indicators. Moreover, their initial outturn is subject to little, if any, revision. Thus it is not surprising that many empirical studies have found that survey and financial indicators contain useful information that can help to predict real GDP growth in many OECD countries.[5] In some countries, such as France and Germany, survey data are used extensively to provide a "benchmark" for current and one quarter ahead real GDP growth (or industrial production), see, for example, Herkel-Rousse and Prioux (2002) and Langmantel (1999).

Explicit measures of the actual level of economic activity should contain more information than surveys of intended activity. However, compared with survey data (or financial data), "hard" indicators suffer from several major drawbacks. First, they are less timely. For instance, in many countries the first release of GDP figures for quarter T is published soon after, and in some cases even before, the publication of hard indicator data for the third month of quarter T. Second, the information they convey is generally relevant only for the quarter to which they relate. It is also the case that they are more likely to be revised than other types of indicators. This implies that a statistical approach which combines soft and hard indicators may be sensible. Such an approach is followed in this paper.

The modelling approach

An initial choice concerns the type of models that are to be estimated. In this paper three different approaches are outlined – single equation quarterly "bridge" models, Vector Autoregressive (VAR) models and auxiliary models that allow quarterly and monthly data to be mixed. Each of these is then employed in the subsequent empirical exercise.

Quarterly Bridge models

Bridge models are ones in which high frequency indicators are used to produce quarterly GDP estimates using single equation techniques. Models of this type have been utilised in empirical studies for some of the largest OECD countries, as well as for the euro area as a whole, although there appear to be few published examples for Japan.[6] These models are typically derived from an initial unrestricted Autoregressive Distributed Lag (ARDL) model, estimated using quarterly data. Real GDP growth on a quarterly basis is regressed on survey data or other monthly indicators aggregated to a quarterly frequency. The unrestricted model has the form:

$$\rho(L)\Delta y_t = \sum_{j=1}^{k} \delta_j(L)\Delta x_{j,t} + \varepsilon_t \qquad [1]$$

where y_t and $x_{j,t}$ denote the logs of real GDP and the selected indicator(s) respectively, Δ denotes the first difference operator and $\rho(L)$ and $\delta_j(L)$ denote lag polynomials of order ρ and δ_j, respectively.[7] This approach can be used to provide a forecast for growth in the current quarter when data for both GDP and the indicator variables are not available (a non-conditional forecast) or when indicators are available but GDP is not (a conditional forecast). Intuitively, a conditional forecast would be expected to produce a more accurate outcome because it takes explicit account of additional information.

A number of equations of this type have been estimated. They are termed "single equation bridge models" in the tables of empirical results. Three conditional models and three non-conditional models are reported, one using only survey data, one using only hard indicators and one using both soft and hard indicators. The non-conditional models are ones that can be run for the current quarter once a complete set of indicator data are available for the previous quarter. The conditional models can be run only when a complete set of data is available for the current quarter.

The attraction of this modelling approach lies in its simplicity. But it has a number of drawbacks. In particular, the models do not exploit the release of indicators on a monthly basis, which, in principle, should allow predictions to be updated on a rolling basis during the quarter. The predictions from the models are not particularly timely either, especially for countries in which there is a considerable delay after the end of the quarter before a full set of data for that quarter becomes available. A further limitation is that the models do not allow for possible linkages between different types of indicators, notably the potential link between surveys (or financial variables) and hard indicators. If such links exist, then it may be possible to exploit them in order to produce quarterly estimates for the hard indicators even when some monthly information is missing. For these reasons we estimate the conditional and non-conditional models using a number of different information sets.

Vector Autoregressive models

One way of attempting to circumvent the delays from waiting for a full set of quarterly data on the indicator variables is to utilise a quarterly Vector Autoregressive (VAR) model that combines indicator variables and GDP growth. The VAR model can be written as follows:

$$y_t = v + A_1 y_{t-1} + \cdots + A_p y_{t-p} + \mu_t \qquad [2]$$

where y_t is a vector($\Delta\ln(GDP_t)$, x_t)', with $\Delta\ln$(GDP) denoting the growth rate of real GDP and x_t a vector containing indicator series.

This approach has a number of advantages over the single equation bridge model. For example, a single VAR model can be used to produce multi-period forecasts at any point in time whilst the bridge model can be run only for one quarter at a time once data for all the included indicator series become available. It is also easy to condition the VAR model to take into account the release of data at different points in time. Thus it can be used to provide a forecast when (some) quarterly indicators are available but GDP is not (a conditional forecast) or when data for both GDP and the indicator variables are not available (an unconditional forecast), exploiting the fact that the elements of the innovation process μ_t in [2] can be contemporaneously correlated (Doan *et al.*, 1984). For the sake of consistency in the presentation of results we again report three sets of models – a pure survey data VAR, a pure hard indicators VAR and a VAR combining surveys and hard indicators. They are termed "multivariate models" in the tables of empirical results.[8]

Monthly auxiliary models

There are only a few studies that address explicitly the means of incorporating new monthly information that becomes available within the quarter for which GDP growth projections are being made.[9] If monthly information is to be used in a model that produces continuously updated projections of quarterly GDP growth, the missing high frequency information for that quarter (and, potentially, for the following quarter) also has to be projected using monthly auxiliary models.

One possible way to exploit monthly information is to combine the single equation quarterly bridge model described above with separate monthly equations that allow missing high frequency information on the conjunctural indicators to be projected over the appropriate forecast horizon.[10] GDP projections can then be made for both the current and the next quarter. For the current quarter, such an approach based on a partial set of within quarter information is similar to that employed by some statistical offices to produce "flash GDP" estimates (see, for instance, Skipper, 2005). A combined model of this sort has a number of useful properties. It is transparent, because the information set underlying the projections is defined precisely, and the forecasting ability of the equations can be gauged easily when the information set is restricted. It also enables quarterly GDP numbers to be produced using an estimated bridge equation that includes estimates of hard indicators for the current quarter. As shown below, the forecasting ability of such an equation appears better than that of bridge models that incorporate only survey data. It is therefore advantageous to be able to make use of this equation as soon as possible, rather than waiting for the official release of hard indicator data for the entire quarter, even if the resulting projections may be a little less accurate than they would be if official quarterly data for the indicator series were available.[11]

The performance of the auxiliary models is considered using four different information sets:

i) Zero months of within quarter information. Monthly variables are projected for a period of six months (three for the current quarter, three for the next quarter). This is, of course, the most difficult projection to make because of the absence of any conjunctural information, even for the current quarter in which the forecast is being made.

ii) One month of within quarter information. Here monthly variables are projected over five months (two in the current quarter and three in the following quarter).

iii) Two months of within quarter information. Here monthly variables are projected over a four month period.

iv) Three months of within quarter information. Here monthly variables are projected over the three remaining months (*i.e.* the three months in the forthcoming quarter). For the current quarter, this model is equivalent to the conditional quarterly bridge and VAR models.

As before, three types of monthly auxiliary models are examined: pure survey monthly models, pure hard indicator monthly models and models that combine optimally survey data and hard indicators. They are termed "monthly auxiliary models" in the tables of empirical results. As with the quarterly VAR model [2], the combination of survey and hard indicator data in a multivariate framework allows survey data to be included in the monthly auxiliary equations for the hard indicators as well as in the single equation used to derive the quarterly GDP projections.

To illustrate the practical implementation of the three different modelling approaches and the timing at which they can be employed, consider the following hypothetical example for the United States. The first estimate of GDP for any quarter becomes available within one month of the end of that quarter. Monthly information on hard indicators such as industrial production is generally available within two to three weeks after the end of the month. The flow of information and the type of models that can be used are illustrated in Table 1.[12]

Consider the situation in the third month of the quarter labelled Q-1. By the latter half of this month, information for the first two months in the quarter is available. A prediction for GDP growth in quarter Q-1 cannot yet be made using any conditional quarterly models (VARs or single bridge models), although it can be made using the non-conditional quarterly models, as a complete information set exists for quarter Q-2. The monthly auxiliary model can however be used to produce a one step ahead monthly projection (to fill in the missing information in quarter Q-1) and a four step ahead monthly projection (to produce estimates for month 3 in quarter Q-1, and months 1-3 in quarter Q). These monthly estimates can then be aggregated to a quarterly basis, with the resulting series being used to produce estimates of GDP growth in quarters Q-1 and Q.

By the middle of the first month in quarter Q, indicator data for month 3 in quarter Q-1 will be published, giving a complete quarterly observation for Q-1. This allows the conditional models for GDP growth in Q-1 to be run, producing an estimate around two weeks before the first official estimate of GDP growth is published. The non-conditional quarterly models can now be used to provide a forecast of GDP growth in quarter Q. There is no within-quarter information yet

Table 1. **Rolling quarterly GDP estimates for the United States**

Quarter	Month	Data availability		Model availability	
		GDP	Indicator series	Quarterly models	Monthly models
Q-1	3	Q-2 (third estimate)	2 months of information for Q-1	All conditional models for Q-2 All non-conditional models for Q-1	Monthly VAR produces 1 step ahead projection for Q-1 and 4 step ahead projection for Q
Q	1	Q-2 (third estimate) Q-1 (first estimate by end of month)	3 months for Q-1 (after mid-month) 0 months of information for Q	All conditional models for Q-1 (after mid-month) All non-conditional models for Q	Monthly VAR produces 3 step ahead projection for Q and 6 step ahead projection for Q+1
	2	Q-1 (second estimate by end of month)	1 month of information for Q	Same as month 1	Monthly VAR produces 2 step ahead projection for Q and 5 step ahead projection for Q+1
	3	Q-1 (third estimate by end of month)	2 months of information for Q	Same as month 1	Monthly VAR produces 1 step ahead projection for Q and 4 step ahead projection for Q+1
Q+1	1	Q-1 (third estimate) Q (first estimate by end of month)	3 months of information for Q (after mid-month) 0 month of information for Q+1	All conditional models for Q (after mid-month) All non-conditional models for Q+1	Monthly VAR produces 3 step ahead projection for Q+1 and 6 step ahead projection for Q+2

available, but the monthly auxiliary model can be used to produce a three step ahead monthly projection (to fill in the missing information in quarter Q) and a six step ahead monthly projection (to produce estimates for each month in quarters Q and Q+1). These monthly estimates can then be aggregated to a quarterly basis and used to produce estimates of GDP growth in quarters Q and Q+1. Thus, the full set of models provides a comprehensive means of projecting current and next quarter growth as new information becomes available.

The variable selection process

The number of possible monthly indicators is large compared with the number of quarterly observations on GDP available for estimation, so choices have to be made. One option is to select indicators on the basis of particular statistical criteria. For example, a linear combination of variables can be selected which generates a quarterly equation (such as a bridge equation) that satisfies conventional diagnostic tests.[13] The advantages of this approach are simplicity and robustness. The results are easy to interpret because the variables being used to project real

GDP growth are clearly identifiable, as are their coefficients. But such a process is unlikely to make optimal use of all available monthly information. The initial choice of indicators is arbitrary and the models can be used only at particular points in time. Projections can be made only when data are available for a whole quarter, as illustrated in the United States example above, unless a vector autoregressive model is used. In the latter case, projections for the current quarter will not necessarily make use of any information from indicator variables in that quarter.

An alternative, and less restrictive approach, is to consider all possible indicators and to summarise the information they contain in a small number of composite variables derived by means of a static or a dynamic principal factor analysis.[14] This approach enables a relatively large set of information to be summarised into a few components through linear transformations of the series contained in the dataset. Quarterly and monthly variables can also be mixed. However, a decision still has to be made about the number of factors to use.[15] The computational burden may also be comparatively heavy. It can also be difficult to identify the reasons for fluctuations in the derived indices because of the large number of variables involved.

Neither of the two approaches outlined above is unambiguously better than the other. We adopt an automated approach to data and model selection that combines elements of both the "in-sample criteria approach" and the "factor analysis approach", and utilise a procedure which permits quarterly and monthly data to be combined in an optimal manner. However, an exhaustive analysis of every monthly indicator for each economy is not attempted. For most countries the focus is typically on indicators derived from business and consumer surveys, financial market variables, such as interest and exchange rates, and "hard" indicators such as industrial and construction output, retail sales and employment. These variables are the ones that appear most frequently in related studies that have sought to develop conjunctural or leading indicator models. They are also ones which have been published on a regular basis over a long enough time period to make it feasible to include them in an empirical exercise.[16]

The set of variables listed in Annex 1 is split into soft/financial and hard indicators. For each of these categories, individual series were ranked according to their explanatory power as measured by the $\overline{R}^2$ statistic from a bivariate regression between GDP growth (at constant prices) and the particular indicator (denoted $x_{i,t}$) in which the retained lag of the indicator was selected automatically according to the Schwarz criterion:

$$\Delta \ln(GDP)_t = \beta(L)_i x_{i,t} + \varepsilon_{i,t} \qquad [3]$$

The initial ranking was for the period 1980Q1-1997Q4 for all the country models, and 1985Q1-1997Q4 for the euro area model. New rankings were then obtained by extending the sample period successively one quarter at a time to 2002Q4. For the current quarter models this amounts to ranking the indicators based on their contemporaneous causality with GDP growth. The selection process for the one quarter-ahead models, in which indicators enter [3] with a minimum lag of one quarter, is broadly equivalent to ranking the variables according to the extent to which they Granger-cause GDP growth.

Variable combination/model estimation

A sub-set of the variables with the highest ranking was selected for each horizon.[17] The optimal combination of these variables was then identified by searching over all possible combinations of variables from an initial ARDL model with up to four lags on both the dependent and explanatory variables. A related procedure is described in Hendry and Krolzig (2001).[18] The best model was selected on the basis of the Schwarz criterion.[19] An important point to note is that there is no necessary requirement for all of the pre-selected variables to be included in the chosen specification. For instance, some survey responses may convey similar information, so that a single measure can suffice (Doz and Lenglart, 1999).

This whole process was repeated over a number of different sample periods, with the end point of the sample period shifted one quarter at a time from 1998Q1 to 2002Q4, in order to generate out-of-sample predictions. This procedure means that the specification of the model and the estimated coefficients can change from one quarter to the next. These steps were repeated for each of the different types of model employed (single equation, VAR and monthly auxiliary). Each of the quarterly VAR models used only those variables present in the selected single equation bridge model for the same sample period. The lag length of the VAR was selected automatically using the Schwarz criterion. The composition of the monthly VAR models was chosen in a similar way. In a small number of cases, discussed in Sédillot and Pain (2003), some additional series were included in the VAR equations for the individual indicator series used in the bridge equations for GDP.

Comparing different models

There are a number of different ways in which the forecast accuracy of different models can be evaluated formally. One widely used measure is the Root Mean Square Forecast Error (RMSFE) of a model.[20] The RMSFE provides a quantitative estimate of the forecasting ability of a specific model, allowing different models to be ranked, but does not provide a formal statistical indication of whether one model is significantly better than another. This can be done using forecast encompassing tests and directional accuracy tests on the out-of-sample prediction errors of the different models.[21]

Formal forecast encompassing tests between each pair of models can be undertaken using the modified version of the Diebold and Marino (1995) test proposed by Harvey *et al.* (1997). This uses the squared prediction errors to make pair-wise comparisons of different models. When there are a number of different models, many of them may outperform a particular benchmark model. It is also possible to test whether the forecasts from any particular model simultaneously outperform the (joint) forecasts of several rival models. If they do not, and the reverse is true (the individual model forecast is significantly poorer than the combination of competing forecasts) then it is the case that a linear combination of the individual model forecasts is to be preferred to any one particular model. This can be investigated using the multivariate extension of the modified Diebold and Mariano test proposed by Harvey and Newbold (2000).

The RMSFE is not a useful indicator of whether a model performs well at turning points. Sometimes it may happen that one model has a lower RMSFE than another but does not do as well in detecting whether GDP growth accelerates or decelerates from one quarter to the next. Information on the expected direction of movement of GDP growth is likely to be of at least as much interest to policymakers as the point estimate itself. Forecast directional accuracy over any given sample period can be evaluated using the non-parametric statistic proposed by Pesaran and Timmerman (1992). This tests whether there is a significant difference between the observed probability of a correctly signed forecast and the estimate of what the probability would be under the null of independence between forecasts and outcomes.

The detailed tables of empirical results for each country reported in Appendix 1, include, in addition to the forecasts from the individual models, two sets of benchmark forecasts from time series models and two sets of combined forecasts. The first time series forecast is from a naïve model in which GDP growth is assumed to be unchanged from that observed in the most recent quarter for which data are published. The second time series forecast is taken from an estimated autoregressive model of GDP growth. The first combined forecast (termed the "consensus forecast" in the tables) is the simple mean of all the different projections that can be produced from the single and multiple equation bridge models, plus the most up-to-date predictions possible from the monthly auxiliary models. The second combined forecast (termed the "monthly auxiliary models consensus" in the tables) is the mean of those quarterly predictions obtained with the monthly auxiliary models that can be run given the information set available at a particular point in time. This second combined forecast differs from the combined monthly model that includes both hard indicators and survey variables because equal weights are placed on each different type of model forecast rather than varying coefficients derived using estimation techniques. The inclusion of the combined forecasts permits an assessment of whether it is preferable to use the consensus rather than any one single model. The average of the separate model projections might easily generate a lower RMSFE if the error from one model is offset by the error from another model.[22]

EMPIRICAL RESULTS FOR THE G6 COUNTRIES AND THE EURO AREA

This section describes the results of the variable selection process for each country and provides a detailed summary of the main empirical results for each type of model in each country.

Indicators selected for each country

The results from the indicator selection process are given in Table 2 below. For the soft indicators, only business surveys were found to matter consistently

Table 2. **Selected variables**

Surveys	Indicator(s) selected	Sources
United States	ISM manufacturing survey Purchasing Managers Index	ISM
Euro area	Level of order books and level of stocks	European Commission
Germany	IFO current and expected business situation indices	IFO
France	Production tendency and future production tendency	INSEE
Italy	Order book position, order book assessment	ISAE and European Commission
United Kingdom	Production future tendency	CBI
Japan	Current and expected sales, inventory, and cash flow diffusion indices	Japan Finance Corp. for Small and Medium Enterprises
Hard indicators		
United States[1]	Industrial production, consumption in volume, new construction put in place, monthly export volumes, total monthly level of stocks	Federal Reserve, BEA, Census Bureau
Euro area	Industrial output, construction output, retail sales volumes	Eurostat, authors' calculations
Germany	Industrial production, construction output, retail sales volumes, new orders	Bundesbank, Bundesamt
France	Industrial production, consumption of manufactured goods	INSEE
Italy	Industrial production, German industrial production, new car registrations, real bank lending rate	ISTAT, Deutsche Bundesbank and Eurostat
United Kingdom	Industrial production, retail sales volumes	ONS
Japan	Tertiary sector activity, industrial inventory to shipment ratio, consumers' expenditure survey (all households), ratio of job vacancies to applicants	METI, Statistics Bureau

1. The larger number of indicators included for the US reflects the greater amount of timely monthly indicators directly available for the expenditure components of GDP.

across the whole of the sample period. Measures of consumer confidence are not included in the set of selected variables for any country or zone. This does not mean that such indicators do not convey any useful information. It simply indicates that they contain only limited additional information of use for predicting GDP growth once allowance is made for the information contained in other indicator series. An example showing the comparative importance of business and consumer survey measures for the euro area is discussed in detail in Sédillot and Pain (2003).[23] Financial indicator series were also found to be relatively unimportant. The only series included is the real bank lending rate in the Italian model, as measured by the average interest rate charged by 34 banks to their most creditworthy customers (the prime rate) deflated by the current annual rate of consumer price inflation.[24] One business survey variable, the German IFO index, was found to be important not only for Germany but also for Italy. This was the only example of the use of a country-specific variable in a model for more than one country.[25]

For the euro area as a whole, measures of the level of order books in industry and the level of stocks of finished goods were found to be the two most useful survey indicator series. For Germany, the main variable of use in explaining the current situation was the IFO business climate index, with the business situation expected during the next six months being a useful indicator of GDP movements a quarter ahead. For France, the current situation is explained by the current production tendency series, with expectations of future production providing useful information for developments in the following quarter. For Italy, the national current production series and the IFO business climate in Germany were found to be helpful for tracking current quarter developments, whilst the national intended production index and the business climate in Germany are of use for one quarter ahead changes. For the United Kingdom, the future production tendency series was found to be important not only for the one quarter ahead specifications but also for the current quarter.[26]

Japan raises some unusual issues, since the most widely quoted business survey, the Tankan survey issued by the Bank of Japan, is undertaken only once a quarter. Use of this measure would mean that a business survey indicator could not be included in the monthly models. However some monthly surveys do exist, such as the *Monthly Survey of Small Business in Japan*, which has been undertaken by the Japan Finance Corporation for Small and Medium Enterprises since 1963. We use seasonally adjusted data for the sales diffusion index and the cash-flow diffusion index in the current quarter models and the expected sales diffusion index and the inventory index in the one quarter ahead models.

From the range of hard indicators evaluated, measures of industrial production were found to matter for all countries with the exception of Japan, where the monthly tertiary sector index was found to be more important. Indicators of consumer activity, were also found to be of use for all countries. For the United States and Japan, the coverage of the consumption series is more comprehensive than in other countries. However the series are less timely than other survey information on retail trade, as there is a lag of a month or more before the information appears. Similar considerations apply to the monthly data for total inventory accumulation and merchandise export volumes in the United States, and the tertiary activity index in Japan. Finally, measures of construction output were also found to contain useful information for the United States, the euro area and Germany.[27]

The hard indicator series for construction output and retail sales in the euro area are (GDP) weighted aggregates of the corresponding data in France and Germany. They therefore differ from the euro zone data published by Eurostat. There are two main benefits from using the weighted country data series. First, they are available on a more timely basis than the euro area aggregates. Second, the euro area data are available only from the first half of the 1990s, limiting the period available for estimation.[28]

Individual country results

Detailed tables with the findings for each country are reported in Appendix 1. These results report the out-of-sample RMSFE and directional accuracy statistics for each type of model for each country over the period 1998Q1-2002Q4, conditioned on different monthly information sets.[29] Although this out-of-sample period is relatively short and coincides with a period of comparatively subdued economic growth, use of a longer evaluation period would raise the risk of selecting a model with a good average performance but a relatively poor performance over the most recent quarters. An important point to note is that the timing of the information set is defined according to the time at which data become available for the complete set of included indicators. In some countries, a complete quarter of data (termed "three-months of within current quarter (cq) information" in the tables of results) for some hard indicators becomes available only after the publication of the initial GDP estimate for that quarter. The results from the comparative statistical tests carried out on the forecast errors from the separate models are discussed in the main text. Tables with the test statistics can be found in Sédillot and Pain (2003).

To provide a comparative benchmark for the detailed empirical results, Table 3 provides some summary statistics on the quarterly rate of GDP growth over the period in which the performance of the respective indicator models is examined. The final row of the table reports the RMFSE obtained from taking a simple, unweighted consensus forecast from the forecasts made by all the different models once a full set of information for the indicators is available for the quarter being forecast. There is a clear positive correlation between the out-of-sample error and the standard deviation of the quarterly rate of GDP growth. GDP

Table 3. **Real GDP growth: descriptive statistics and RMSFE (1998:1 2002:4)**

	United States	Japan	Euro area	Germany	France	Italy	United Kingdom
Standard deviation of real GDP growth	0.60	0.80	0.37	0.58	0.45	0.41	0.30
Mean of real GDP growth	0.70	0.10	0.50	0.30	0.65	0.38	0.60
Consensus forecast RMSFE (current quarter, all indicators available)	0.37	0.54	0.19	0.36	0.24	0.26	0.23

growth has been most volatile in the United States, Japan and Germany, and the models for these three economies have the highest errors, with Japan being especially difficult to forecast accurately. For most countries the RMFSE is, on average, around 40% lower than the standard deviation of the quarterly rate of GDP growth. The largest relative gains appear to be in the aggregate euro area model, with a RMFSE half the size of the standard deviation of GDP growth.

The United States

The overall results are summarised in Table A1 of Appendix 1. To illustrate the format of the table (and those for other countries), and the relationships between the current quarter and one quarter ahead results as the information set changes, consider the combined hard indicator/survey results in the panel of Table A1 labelled monthly auxiliary models. When there is no monthly information available for the current quarter (*i.e.* the first 4-5 weeks of the quarter), the RMSFE of forecasts made for GDP growth in the following quarter is 0.59 percentage points, based on the average of the out-of-sample forecast errors for 1998-2002. Assuming that the forecast errors have a normal distribution, this implies that there is a 68% chance that the true outturn for GDP growth (at a quarterly rate) will be within ± 0.59 percentage point of whatever the point estimate of the indicator model is. By the time three months of information have become available in the current quarter, the RMSFE for the forecast of GDP growth in the following quarter has declined to 0.50 percentage points. Keeping the same forecast horizon, the model can, at this point, be expressed equivalently as a current quarter forecast model with zero months of within quarter information.[30] Once an additional month of information becomes available, the RMSFE falls again to 0.43 percentage point. By the end of the quarter, with an extra six months of information since the initial estimate was made, the RMSFE has declined to 0.37 percentage point. This illustrates how the range of uncertainty around the point forecast for any given forecast horizon diminishes as the forecast horizon is approached. Optimal use of information implies that forecasts for more distant horizons should have a wider band of uncertainty than forecasts for closer horizons.

Turning to the results themselves, the first feature apparent from the tables is that there is some benefit from seeking to estimate an indicator model. All the estimated models outperform a naïve model based on the assumption that GDP growth in the quarter to be forecast is the same as that in the last quarter for which information is available at the time of the forecast. An autoregressive time series model for GDP performs a little better than the naïve model, but again accuracy, as measured by the RMSFE, is poorer than for any of the estimated models.

Amongst the estimated models themselves, specifications using hard indicators are found to perform much better for the current quarter forecasts than those that use just survey information. For the single equation bridge model with a complete set of monthly information for the hard indicators in the current quarter, the gains compared with the autoregressive GDP and the pure survey models, as measured by the ratio of the respective RMSFEs, are about one-third and one-quarter respectively (see Sédillot and Pain, 2003, Table 4). These gains start to appear once one month of information on hard indicators is available for the quarter, which is approximately 2-2½ months before the first official estimate of GDP growth in the quarter is released. From this point there appear to be few benefits from augmenting the monthly auxiliary hard indicator model with an additional monthly survey observation, and the RMSFE of the model with hard indicators is always consistently lower than that of the model with just survey indicators. This does not necessarily mean that the survey information from the ISM is without useful content, it is just that it is more helpful for predicting the hard indicators themselves (via the linking equations in the monthly auxiliary model) than it is for predicting GDP directly.[31] For the current quarter, the models appear to track the acceleration (or deceleration) of real growth closely, with a three-in-four, or better, chance of predicting correctly the direction of change in GDP growth compared with the previous quarter. Once all monthly data are released for the quarter, the directional accuracy from the combined model with both hard indicators and surveys is 95%.

The results for the one-quarter ahead forecasts are less clear-cut. The directional accuracy of the estimated indicator models is little different from a random outturn (50%) and the gains in terms of accuracy compared with time series models are far smaller. Pure survey models appear to perform at least as well as pure hard indicator models. There appear to be few gains from combining the two into a mixed model, or taking a consensus view, suggesting that both are tracking similar factors.

A number of the absolute differences apparent from the detailed table of results are statistically significant. Pairwise Diebold-Mariano test statistics showed that, for the model using hard indicators, the null of comparable forecast accuracy with the autoregressive models is rejected at the 5% threshold when at least one month of within quarter information is available. Given that the RMFSE of the former model is below that of the latter, the encompassing test statistic implies that latter model is rejected against the former. The hard indicator model is also shown to be better than the pure survey model, but the differences are significant only at the 10% level.

However, the tests for multiple forecast encompassing reported in Table 5 of Sédillot and Pain (2003) provide stronger evidence in favour of the hard indicator model relative to the autoregressive and pure survey models. In particular, they show that it is not possible to reject the hypothesis that the predictions from the hard indicator equation with one or three months of within quarter information cannot be improved by the predictions from the other two models. In contrast, it is possible to reject this hypothesis when either of the other two models is taken as the numéraire.

For the one-quarter ahead forecasts, there were no statistically significant results, even between the indicator models and the pure autoregressive model, from either the encompassing tests or the forecast directional accuracy tests.

The euro area

Results for the euro area are summarised in Table A2 of Appendix 1. One noticeable difference with the results for the United States is that the size of the RMFSE from many of the models for the euro area is a lot lower. This reflects the lower volatility of quarterly GDP growth in the euro area rather than any inherent differences in the quality of the equations for the two regions. But aside from this, the general features of the results are similar to those for the United States. The performance of the estimated indicator models for the current quarter is again noticeably better than that of the time series models of GDP, especially once some monthly information becomes available. For the one-quarter ahead forecasts, the differences are smaller in terms of size of errors, although the estimated indicator models tend to have better directional accuracy (much more so than found for the United States).

Amongst the indicator models themselves, hard indicator models for the current quarter appear to perform best from the point at which one month of within quarter data becomes available, with a noticeably lower RMFSE than pure survey models (0.22 percentage point compared with 0.33 percentage point from the monthly auxiliary models with one month of information).[32] Prior to that point, there is little to choose between the different types of models. The reduction in uncertainty as the information set expands can again be seen quite clearly, both from the monthly auxiliary model results and from the difference between the conditional and unconditional forecast results from the single equation bridge model and the VAR model. Mixing hard indicator and survey data, either directly through estimation or indirectly by taking a consensus projection, does not appear to yield any noticeable benefits.

The pair-wise modified Diebold-Mariano tests and the multiple forecast encompassing tests provided further evidence in favour of the hard indicator models for the current quarter once some monthly data become available. The bi-model differences between the hard indicator model and the other models are generally significant at, or under, the 10% level. This finding is reinforced by the encompassing test results. The hard indicator predictions cannot be improved by either the autoregressive forecast or the survey-based projections, whereas the projections for the other models can be improved by including the hard indicator projections. In terms of the forecast directional accuracy test, only the models including hard indicators seem to convey useful information on a consistent basis. These are the only models for which the null of independence between forecast and outcome can be rejected for both the current and one quarter ahead projections.

Germany

Results for Germany are summarised in Table A3 of Appendix 1. The size of the forecast errors is noticeably larger for Germany than for the aggregate euro area, reflecting the greater quarterly variation in GDP shown in Table 3.[33] Another difference is that survey measures appear to contain far more useful information for Germany than they do in either the United States or the aggregate euro area. The errors from the pure survey models of current quarter growth are lower than those from the hard indicator model, although the directional accuracy of the latter is better. There also appear to be some modest gains from combining information once a complete quarterly set of monthly data is published, either through using an estimated mixed model, or, more notably, through taking a consensus forecast. The directional accuracy of the combined forecasts for the one-quarter ahead GDP forecasts also appears somewhat better than that of any of the individual models, although there is relatively little difference in terms of the size of the errors.

The forecast encompassing tests suggested that the predictions from the combined survey/hard indicator model were significantly better than those from a time series model, and at times significantly better than those from an equivalent pure hard indicator model, at the 10% level. The importance of the information from business surveys was also shown by the results of the Pesaran-Timmerman tests. For both the current and one quarter ahead forecasts it was possible to reject the null of independence between forecasts and outcomes when considering the pure survey monthly auxiliary model. The test statistics are even more significant for the model that combines hard indicators and survey data, possibly reflecting the usefulness of survey data for predicting monthly movements in the hard indicators.

There are some features in common with the results for the United States and the euro area. In particular it can again be seen that the main gains from the monthly auxiliary models start to appear for current quarter GDP forecasts once one month of information is available for that quarter. This is true for both the pure survey and the pure hard indicator monthly models. Prior to that point a pure survey model outperforms a time series model for GDP, but the pure hard indicator model does not.

France

The general pattern of the results for France, summarised in Table A4 of Appendix 1, is similar to that found for Germany. The performance of models using only survey data is at least as good, and at times better, than that of models using only hard indicators. This is true for both current and one-quarter ahead forecasts, the latter suggesting that the future production tendency variable is particularly useful. In both instances the performance of a combined forecast, whether estimated or obtained as a consensus of the individual model forecasts, appears better still. For the current quarter, the conditional forecast from the single equation bridge model combining hard indicators and survey data has the lowest RMSFE (0.24 percentage point). From the forecasts made of quarterly growth prior to the start of the quarter, the best performing model is the combined survey/hard indicator monthly auxiliary model with three months of survey information (for the quarter prior to the one being forecast). This has a RMFSE of 0.31 percentage point and directional accuracy of 75% (*i.e.* there is a three in four chance of correctly predicting the direction of change in the rate of GDP growth).

Again it is clear that the predictions of the indicator models improve as more information becomes available. But there are some notable differences with the pattern of results found in other countries. For the current quarter forecasts, the performance of a pure hard indicator model becomes comparable with that of a pure survey model only once two months of information are available. (The lag is shorter elsewhere.) For the one-quarter ahead forecasts, there is marked improvement in the performance of the pure survey models relative to that of a time series model after two months of information become available in the quarter prior to the one being forecast. In contrast there is little to choose between the one-quarter ahead forecasts from a hard indicator model and a pure time series model.

Forecast encompassing test statistics showed that, for the current quarter forecasts, the predictions from the combined survey and hard indicator model appear significantly better than those from other models. This was especially apparent in the multiple forecast encompassing test using forecasts based on a 1 month within quarter information set.

Italy

The results for Italy, summarised in Table A5 of Appendix 1, are similar in several respects to those found for the aggregate euro area, and hence differ from those found for Germany and France. For current quarter forecasts, the pure hard indicator model generates statistically smaller forecast errors than models based only on survey data once one month of within quarter information is available. (This is not the case for one-quarter ahead forecasts.) Thereafter, there is little change in the size of the errors from the models as additional monthly information accumulates. There appear to be few additional benefits from using a model including both survey and hard indicator variables. It is also noticeable that the performance of the consensus of the individual forecasts is little different from that of the hard indicator model. In contrast, for the one quarter ahead forecasts, the best performing model appears to be the consensus of the monthly auxiliary models, although the difference in the performance of that and the individual models is not large.

All these features were confirmed by the results of the Diebold-Mariano tests. It is not possible to single out a statistically superior model, at least in terms of forecast error size. Models combining information from surveys and hard indicators, or models including only hard indicators, have the same forecast performance for both the current and the next quarter. But these models are always better statistically than the autoregressive model and (in the current quarter) the pure survey model for the same set of within quarter information (at the 5% threshold). The multiple forecast encompassing tests indicated that the forecasts from the combined monthly auxiliary model cannot be improved by adding the information contained in the forecasts made by other models.

The United Kingdom

The general pattern of the results for the United Kingdom, summarised in Table A6 of Appendix 1, appears closer to those found for the United States and the aggregate euro area than to those found for the other individual European countries. Models containing only hard indicators appear to outperform ones containing only survey information, especially for current quarter GDP forecasts.[34] There appear to be few benefits from estimating models that include both surveys and hard indicators. However the combination of variables through a consensus of the individual model estimates does appear to be of some use, generating the lowest RMFSE of all the forecasts. These results suggest that the relationship between measured GDP and the indicator series may possibly have changed over time.[35]

Some features of the UK results are familiar from all the preceding country results, in particular the decrease in the size of the forecast error as the forecast horizon diminishes and the relative difficulties in outperforming a time series model for a one-quarter ahead forecast made at a five or six month horizon. Gains start to become apparent only once the horizon shortens to three or four months. The non-conditional one-quarter ahead forecasts from the VAR model appear to be of limited use, as their errors are higher than those from the autoregressive time series model.

Pair-wise encompassing tests showed that the differences between the errors from the consensus model and the autoregressive model are statistically significant at the 5% level, implying that the latter is rejected against the former. The consensus forecast also outperforms other models, but the differences are not statistically significant. For all models, with the exception of the two benchmark time series models, the current quarter forecast directional accuracy measures are around 75-80% (the latter implying that there is a four in five chance of correctly predicting the direction of change in the rate of GDP growth) and the null of independence between forecasts and outcomes is rejected at the 5% or 10% levels. This is also true of a number of the one-quarter ahead forecast models.

Japan

The results for Japan, summarised in Table A7 of Appendix 1, provide a clear illustration of the difficulties in forecasting the quarterly growth rate of the Japanese economy. The average root mean squared error from all types of model is at least as large, or greater, than that found in all the other economies. Despite this, the directional accuracy of the forecasts appears better than for most other economies, especially for current quarter forecasts. This suggests that the models provide a useful means of predicting the direction of quarterly changes in the rate of growth of GDP, but have difficulty in predicting the full magnitude of the swings. In part this reflects the high volatility of the quarterly growth rate, as shown in Table 3.[36]

For the current quarter projections the results are similar to those for the United States and the United Kingdom, with the models containing hard indicators outperforming the ones containing only survey information. All the models outperform the simple time series models significantly. As in other countries the size of the forecast errors declines as more monthly data become available, but the size of the decline is smaller than that elsewhere. For example, in the euro area the RMSE from the monthly hard indicator model falls by almost one-half as more monthly information emerges through the quarter (from 0.33 to 0.18 percentage points) and that in the United States declines by one-third. But in Japan the drop is only one-tenth (from 0.61 to 0.55 percentage points). This pattern can also be seen in the errors from the conditional and non-conditional bridge models that use hard indicators. There is little difference between them in

Japan, in contrast to the other countries where the errors from the conditional models are considerably smaller than those from the non-conditional models.

However the incorporation of additional monthly information within the current quarter does appear to improve the one-quarter ahead projections markedly, much more so than in other countries. Again taking the monthly hard indicator model as an example, the one-quarter ahead RMSE falls by 25% (from 0.74 to 0.56 percentage points) as information appears over the course of the current quarter. The declines in the United States and the euro area are far smaller (0.58 to 0.55 percentage points and 0.37 to 0.34 percentage points respectively). One possible explanation for these cross-country differences might be that different methods are being used to construct GDP data. For almost all monthly models, pair-wise modified Diebold-Mariano tests find that the errors made when using three months of current quarter data are statistically lower than those made when no information is available. All the indicator models also outperform the pure time series models in terms of forecast accuracy. But the gains in terms of forecast directional accuracy are much more limited and it is not possible to single out one particular model.

USING THE INDICATOR MODELS IN REAL TIME

There are a number of practical considerations which arise when using the indicator models in a real-time analysis. Many of these are not considered in the theoretical literature. Perhaps the most important concerns the choice of which model to use at any particular point in time. For most countries the preferred model when a complete information set is available for an entire quarter is one including hard indicators. Yet because the publication of these indicators typically lags that of survey data, it is quite possible that the practical choice of models during a quarter may lie between models including hard indicators and pure survey models that can utilise a more up-to-date information set. The expected forecast error from the latter may be lower than from the former.

For example, consider the results for France reported in Table A4. For any given complete monthly data set, current quarter forecasts from a monthly auxiliary model that includes both hard indicators and survey information have a lower RMSFE than those from an auxiliary model that includes only survey information. Yet when one month of within-quarter survey information is available, but no information is yet published for the hard indicators, the pure survey model appears preferable to the mixed model, both in terms of RMSFE (0.26 compared with 0.30) and directional accuracy (80% compared with 70%). Thus, if the indicator projections are updated continuously, rather than at fixed-length discrete intervals, it may be optimal to use different models at different points in time.

It may also be the case that different models will appear to be preferred for different quarters, especially given the timing of information releases.[37] Again, taking France as an example, consider the choices between the forecasts made by a pure survey monthly auxiliary model with two months of information for the current quarter and the forecasts made by the mixed auxiliary model with only one complete month of information. For the current quarter, the mixed model appears preferable to the pure survey model, at least in terms of their RMSFEs. But for the following quarter, the pure survey model appears preferable (a RMSFE of 0.39 compared with 0.45 and a directional accuracy of 65% compared with 45%). Yet given that the one-step ahead forecast from this model will be conditioned in part on a GDP forecast for the current quarter that is expected to be poorer than that from other models (reflecting the VAR framework), a decision has to be made as to how much weight should be placed on it. This point can also be seen clearly for the forecasts made at the start of the current quarter by a purely autoregressive model for GDP. The forecasts for current quarter growth are poor, but for the following quarter, their expected accuracy is little different to those of bridge or monthly auxiliary model forecasts.

Even if the point forecasts for growth do not change as the data set expands during the course of a particular quarter, the information conveyed by the indicator model may still evolve since the probability of any particular point outturn can change. This is because the uncertainty around any particular point estimate diminishes as the forecast horizon shortens. The error band may also change if the forecast is produced using different models during the quarter, all of which have different degrees of uncertainty associated with them. Other things being equal, it is likely that if the point estimate from the indicator-based models for a particular event does not change over time, the confidence that can be placed in that forecast will rise.

Another issue which may arise is the consistency of the aggregate euro area projection and those for the three largest individual economies. No attempt has been made to impose any formal restrictions between the aggregate and the individual country models in estimation. In real-time it is important to check that the implicit projection for the residual component of the euro area is plausible. A related question that has not been explored in the work reported here is whether an aggregation of the projections for Germany, France and Italy would provide a better projection for the euro area than the aggregate euro zone model.[38]

Judgments may also need to be made about the weight to place on a piece of country-specific information for one of the European economies which is not reflected, or not used explicitly, in the aggregate euro area model. For instance, the evaluation of the forecast errors from the different models for the euro area, showed that for current quarter GDP forecasts a pure hard indicator model was preferable. In contrast, in Germany, France and Italy a model combining survey and hard indicator data appeared to perform better. So it would be quite possible for an unusual survey observation to affect the indicator forecasts in one or more of these economies without changing anything directly in the euro area aggregate projection.[39] To guard against this possibility it is sensible to obtain projections from all the models at each point in time, even though, on the basis of past performance, there are some which can be expected to be less accurate on average than others.

The factors discussed above all need to be taken into account when seeking to use the indicator models in real-time. Each of the indicator models can be regarded as a statistically acceptable model. The variables included appear to have the strongest consistent relationship with GDP growth over time and the estimated models all pass standard diagnostic tests. Some perform better than others at times, but the observed differences are often not statistically significant. There are clearly gains to be had from using the indicator models, but a role is likely to remain for informed judgment when evaluating their real-time projections. Delays in the release of information for particular indicators and differences in the set of conditioning information mean that conflicting signals can, at times, be given by different models. Much of the information required to make decisions about the relative weight to place on different models can be accumulated only over time as knowledge of the merits of the competing models starts to accrue. Whilst it is convenient to try and find a simple model that outperforms all others at all points in time, it is doubtful that such a model would always make optimal use of the information available.[40]

Until such a model appears, the most appropriate procedure at any point in time is to select the projections made by the particular model (or collection of models if the consensus is chosen) which can be expected to deliver the lowest (average) forecast error based on the performance of forecasts made at a similar juncture in the past. In doing this, it is also sensible to monitor the overall distribution of point forecasts from the full range of indicator models to assess the balance of risks around the point estimate from the selected model. The final section of this paper illustrates how this can be done.

THE REAL-TIME PERFORMANCE OF THE INDICATOR MODELS

The summary forecast errors in Tables A1-A7 are generated from an out-of-sample forecasting exercise over 1998-2002. A limitation of this exercise is that it uses a single vintage of GDP and indicator data, rather than "real-time" data. As mentioned above, this could provide a misleading indication of the likely performance of such models in real-time. However, as many of the indicator models set out in this paper have been in use since 2003, it is possible to begin to assess the importance of this distinction by comparing the actual out-of-sample performance of the indicator models over 2003Q1-2005Q1, with the simulated out-of-sample performance over 1998Q1-2002Q4.

The set of projections considered in this exercise are ones for GDP growth in the current and forthcoming quarters made at the end of the third month of the current quarter. There are slight differences in the timing and information sets for each forecast, but in almost all cases one to two months of hard indicator data and two to three months of survey data are available for each economy. In each case the projections are assessed against the first published outturn estimate for GDP growth.[41] The resulting root mean squared errors (RMSE) are shown in Table 4, with the number of forecasts available reported in parentheses. The table also shows the corresponding errors from Tables A1-A7, taking the combined survey and hard indicator model with two months of information as the reference model. Three main points emerge:

- The cross-country pattern of the actual real-time errors is consistent with that generated in the initial out-of-sample exercise using a single data vintage. The models for United Kingdom, the aggregate euro area and France have comparatively small forecast errors, while Japan and the United States have comparatively larger ones.
- In all economies the real-time errors for the one-quarter ahead forecast are greater than those for the current quarter forecast. The difference between the two is smallest for the euro area.
- In almost all economies the real-time errors are smaller than the simulated out-of-sample errors. The exceptions are in Italy, especially in the current quarter, and the one-quarter ahead forecast errors for Japan. However, in the latter case especially, only a limited sample is available.

There are a number of potential explanations why the real-time performance of the indicator models might be expected to be a little better than that of the simulated models. One factor may be the differences in the information sets, with three months of survey information being used for some of the real-time estimates, whereas only two months are used for the simulation exercise. However, the effect of this is likely to be small, at least over a large enough sample. As can be seen from Tables A1-A7, there is little difference in the performance of most combined survey and hard indicator models whether two or three months of information are available.

A more likely explanation of the differences is that the indicator models developed in this paper are an example of the type of models used by many national statistical offices when estimating missing information in order to produce early estimates of quarterly GDP growth. For example, the Office for National Statistics produces the first estimate of GDP growth in the United Kingdom around 25 days after the end of the quarter, at a time when under half of the actual data

Table 4. **A Comparison of actual and simulated out-of-sample forecast errors**

	Current quarter		One quarter ahead	
	Actual RMFSE	Simulated RMFSE	Actual RMFSE	Simulated RMFSE
United States	0.36 (9)	0.40	0.43 (8)	0.53
Euro area	0.20 (9)	0.21	0.22 (8)	0.35
Germany	0.32 (9)	0.41	0.37 (8)	0.54
France	0.24 (9)	0.26	0.29 (8)	0.40
Italy	0.45 (9)	0.26	0.53 (8)	0.40
United Kingdom	0.12 (9)	0.25	0.22 (8)	0.26
Japan	0.57 (6)	0.57	0.70 (5)	0.61

Note: The number of real-time forecasts is shown in parentheses. These range over the period 2003Q1-2005Q1. The simulated errors are calculated over the period 1998Q1-2002Q4, using the data vintage available at the time the particular country model was constructed.

for the quarter is available (Skipper, 2005). The remainder of the data is forecast using a variety of different approaches, drawing on related indicators for individual components of aggregate (output) GDP and predictions from econometric models. In contrast, subsequent official estimates of GDP growth are able to utilise far more direct information on the behaviour of various GDP components, information that is not available at the time the first estimates are produced. So it is not surprising to find that real-time errors from the initial estimates of GDP are smaller than the simulated errors from predictions of GDP growth measured using all necessary information.

This leaves open the question of why the real-time performance of the indicator model for Italy has to date been markedly inferior to what might have been expected. Care is needed in making comparisons given the relatively small sample of outturn data available, but a possible factor behind this is that the variance of the initial official real-time quarterly outturn estimate over 2003Q1-2005Q1 has been high in relation to the mean quarterly growth rate. The coefficient of variation during this period was 5.2, almost five times the size of that over the period used to calculate the simulated out-of-sample errors (see Table 3),[42] and nearly three times the size of the coefficient of variation for the next highest economy, Germany. The United Kingdom was the economy with the lowest coefficient of variation for the first official real-time outturns during the period considered.

ASSESSING FORECAST UNCERTAINTY

All forecasts are uncertain, and it is clearly of use to policymakers to be able to quantify how great the uncertainty might be (Britton *et al*, 1998; CBO, 2003). Two types of calculations can be made with the range of indicator models developed in this paper. The recursive out-of-sample forecast errors for each model enable the probability of any particular event to be calculated given the point estimate from that model by making an assumption about the underlying probability distribution of the forecast errors. Additional information can also be derived from the cross-section of different point forecasts from the different models.

One simple way of representing the range of uncertainty around the point forecast from a particular model is to derive the probability of other possible forecasts using an estimate of the standard error of past forecasts (for instance, the recursive out-of-sample errors for 1998-2002) from that model. The resulting probability distribution would be symmetric around the point forecast if the forecast errors were assumed (or found in the past) to follow a normal distribution. In this case, for any point forecast γ, the probability of the two outturns $\gamma + \delta$ and $\gamma - \delta$ would be identical. This approach assumes that the model used to produce the forecast is the "true" model. However, just because one model is found to perform better on average over time than others, it may not be the most suitable to use at all points in time. With a wide range of possible models, there is uncertainty about the choice of forecasting model and hence about the possible value of the point estimate itself. The distribution of possible point forecasts can be used to quantify possible asymmetries in the risks around the point estimate thought to be most likely.[43] In doing this, it is useful to include as many different forecasts as possible, as this will help to raise the chances that a particular "shock" will be reflected somewhere in the distribution.

A notable example of this approach to assessing forecast uncertainty is the fan chart produced in the *Inflation Report* of the Bank of England (Britton *et al.*, 1998). This is produced on the assumption that the possible forecast errors follow a two-piece normal distribution, as set out in Appendix 2 below.[44] One feature of this distribution is that it allows the risks around the central forecast to be skewed (*i.e.* it may be thought that the forecast error is more likely to be in one direction than the other). To derive the distribution, use is made of the mean forecast and the forecast judged to be the most likely, along with an estimate of forecast uncertainty derived from past forecasting errors. In practice, the modal forecast is the one often judged to be the most likely outturn. If so, the difference between the mean and the mode provides an estimate of the relative balance of risks around the mode (the skewness of the assumed distribution).[45] If the mean is above the mode than there are likely to be more upside risks than downside risks, and *vice versa*.[46]

The Bank of England calculations of the distributional parameters necessary to quantify the probability distribution use a mix of expert judgment and information from different economic models. It is possible to compute estimates of this kind for the indicator models. For example, use can be made of the distribution of the different point forecasts from each possible model for each country or zone to obtain an estimate of potential skewness, given by the difference between the mode (or median) of the distribution and the mean.

In fact, the information that can be derived from the indicator models developed in this paper gives considerable flexibility over how forecast uncertainty and risks can be quantified. This is because it is possible to quantify the standard deviation of the forecast error from each individual model as well as those based on the mean, median and mode of those forecasts at each point in time. In contrast it is often possible for forecasting organisations and official bodies such as the Bank of England to use only a single estimate of the forecast error based on the past track record of their forecasts, even if each of them may have been produced using a different methodology.

The suite of indicator models also allows the estimate of risk (skewness) to be automated rather than dependent on informed judgment. So, for instance, it is possible to provide an illustration of the risks around any individual forecast using the error range attached to that forecast together with an estimate of the relative balance of risks obtained using the difference between that point forecast and the mean of all the point forecasts from the separate indicator models.[47] Equally, it is possible to compute the risks around the modal forecast using information on the past standard errors from modal forecasts. Another possible way to assess the risks surrounding the short term projection would be to compare the selected mechanical indicator based projection with the judgmental projections made by the country experts within the OECD Economics Department.

To illustrate the ways in which forecast uncertainty can be calculated, consider the projections made for the quarterly rate of GDP growth in France in 2003Q4 and 2004Q1, as published in the November 2003 *Economic Outlook* (Table 1.2). Given the monthly information set available at the time, and the historical out-of-sample performance of the models that could be estimated with that information set, the most suitable indicator model estimate was judged to be the consensus of the projections from the separate monthly models. The point estimates of growth were 0.52% in 2003Q4 and 0.51% in 2004Q1, with respective standard errors of 0.35 and 0.47 percentage points.[48]

These standard errors could be used to calculate a symmetric confidence interval around the point forecast. However, the full range of point estimates from all the different models suggested that the risks around the indicator model forecast could be skewed downwards, especially in 2003Q4, as the modal forecasts were for GDP growth of 0.67% in 2003Q4 and 0.55% in 2004Q1. This difference arose as the hard indicator equations were pointing to weak growth, whereas the equations using survey data were pointing to much stronger growth.

The results from using this information to obtain confidence intervals around the central forecast are shown in Figure 1. The solid line in the middle of the distribution (the darkest shaded interval) covers the central 10% of the distribution, with the distribution being centred on the mode. Each successive interval above

Figure 1. **Forecast uncertainty for France using the two-piece normal distribution**

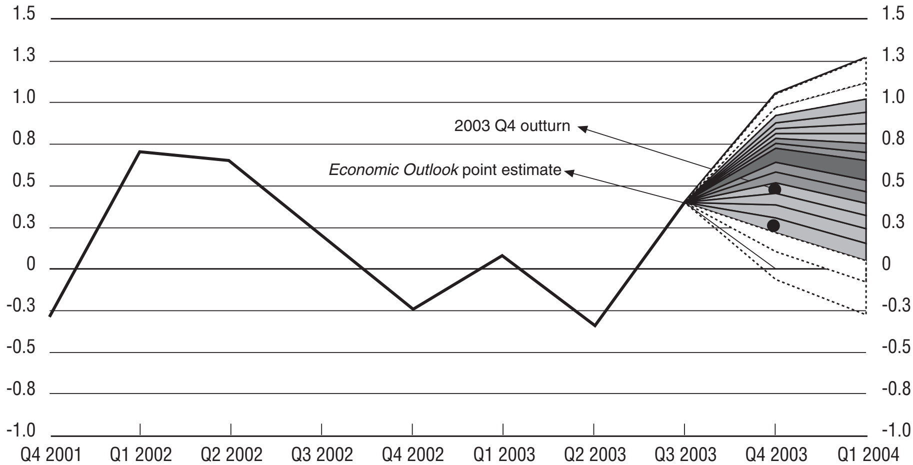
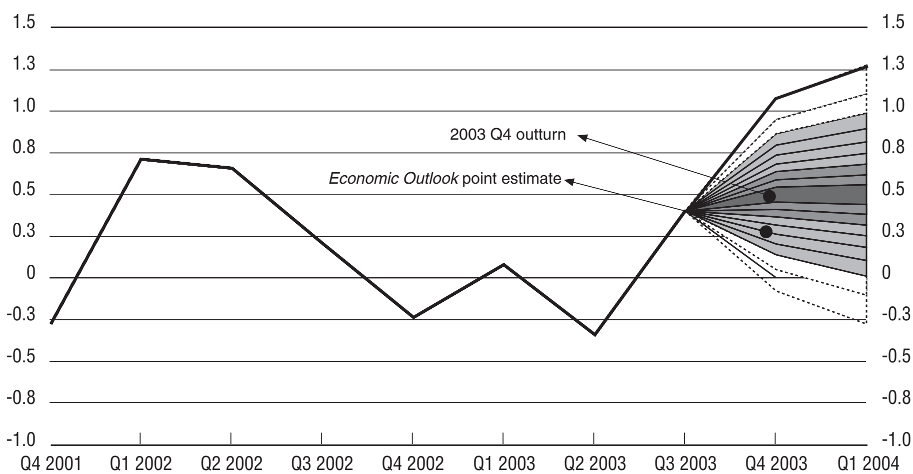


Figure 2. **Forecast uncertainty for France using a single indicator model**



and below this then adds a further 5% of the distribution, with the outer lines defining an interval that covers 90% of the distribution. It is clear from that chart that the risks associated with the forecast were skewed downwards, especially in 2003Q4. The forecast for GDP growth in 2003Q4 published in OECD *Economic*

Outlook, No. 74 was 0.35%, and the first official flash outturn estimate (published in February 2004) was 0.45%.

Using the evidence from the particular indicator model chosen at the time, and assuming a normal distribution around the mean, the probability of the outcome being lower than the OECD forecast was 32%. The chance of the outcome being below the first official estimate was 42%. Using the evidence from the alternative two-piece normal distribution, the chances of these two outcomes being lower than the OECD central projection (the modal projection) were 18% and 26%, respectively. Figure 2 shows a fan chart based solely on the information from the selected indicator model. The differences between the alternative charts are clear and illustrate that it can be possible to take very different views about the probability of particular (point) estimates, according to the way in which uncertainty is assessed.[49]

Notes

1. The use of regression techniques to identify indicator series that are closely related to GDP growth over the economic cycle as a whole differs from the longstanding approach used to produce the OECD Composite Leading Indicator series. The latter are constructed using a set of 5-10 variables for each country that have been observed to be closely related to past turning points in a proxy reference series, typically, industrial production. Detailed information on the Composite Leading Indicator series is available at *www.oecd.org/std/cli*.
2. An early example of forecasting with monthly data is provided by work at the US Federal Reserve (Corrado and Greene, 1988). Subsequent extensions and further discussions can be found in Miller and Shin (1996), Ruey and Chung (1996) and Stark (2000).
3. *www.chicagofed.org/economic_research_and_data/cfnai.cfm*.
4. Examples include Corrado and Greene (1988), Miller and Shin (1996) and Stark (2000) for the United States, Salazar and Weale (1999) for the United Kingdom and Rünstler and Sédillot (2003) for the euro area.
5. Examples of studies that utilise survey indicators include Golinelli and Parigi (2004), Irac and Sédillot (2002), Mourougane and Roma (2002) and Parigi and Schlitzer (1995). Financial market measures, such as the yield spread, have been found to have predictive power in several studies; examples include Davis and Fagan (1997), Estrella and Miskin (1998) and Plosser and Rouwenhorst (1994). Britton and Pain (1992) use both types of variables.
6. Studies of this type include Ingenito and Trehan (1996) for the United States, Bovi *et al.* (2000), Parigi and Schlitzer (1995), Van Rooij and Stokman (2000), and for the euro area, Baffigi *et al.* (2002 and 2004), Grasmann and Keereman (2001) and Rünstler and Sédillot (2003).
7. A possible extension, as in Baffigi *et al.* (2004), would be to estimate a Vector Equilibrium Correction Model. In principle such a model might be more powerful as it also includes information on the long-run co-variation of the individual series. But careful testing would be required, as many of the series contain unit roots and there is no necessary requirement for bridge models to contain structural behavioural relationships. We do not pursue this approach in this paper.
8. Bayesian VAR (BVAR) models are used, estimated with traditional Minnesota priors. The three priors (the rate of decay, overall tightness, the tightness on lag n) are those minimising the out-of-sample forecast errors over the period 1998Q1-2002Q4.
9. Most are for the United States, with the pioneering work of Fitzgerald and Miller (1989) followed by Ingenito and Trehan (1996) and Robertson and Tallman (1999) amongst others. For a recent application to the euro area, see Rünstler and Sédillot (2003).

10. As with the pure quarterly models, all the monthly variables are determined jointly by means of a BVAR approach, with priors selected so as to minimise the out-of-sample forecast errors over the period 1998Q1-2002Q4.
11. This is because the forecasting errors from the auxiliary models of the hard indicators will be included in the quarterly values of the hard indicator series used in the quarterly bridge equation.
12. The publication lags for industrial production are shorter in the United States than in all the other countries for which we estimate models.
13. See, for instance, Parigi and Schitzler (1995), Bovi *et al.* (2001), Baffigi *et al.* (2002 and 2004) and Rünstler and Sédillot (2003).
14. Examples of recent studies that have sought to exploit such techniques include Angelini *et al.* (2001), Doz and Lenglart (1999), Camba-Mendez *et al.* (2001), Forni *et al.* (2000 and 2001), Stock and Watson (2002) and Grenouilleau (2004).
15. This has been addressed under specific assumptions, see Bay and Ng (2002) and Doz and Lenglart (1999).
16. One obvious gap in the variables considered is the lack of monthly information on activity in many service industries. Whilst survey information is available for the non-manufacturing sector in many countries, there is often insufficient data to allow such series to be included in estimation. The IFO survey variables for Germany include the wholesale and retail sectors as well as the manufacturing and construction sectors, but this is not the case for the other country surveys.
17. In most cases the subset consists of four variables. It would be possible to include a much larger number of variables, but then the computational burden becomes very heavy, see Sédillot and Pain (2003, footnote 20).
18. This procedure can be used to solve simultaneously the variable selection process and the optimal combination of variables. Based on a reduction path algorithm, the procedure is able to determine the "best" specification in a limited amount of time, even if the initial set of information is large. Dubois and Michaud (2005) produce a short-term GDP forecast using equations selected by this method.
19. Standard diagnostic tests (for normality, autocorrelation of order 1 and 4, Arch, RESET and the Chow predictive failure test) were carried out on this equation to check its statistical adequacy.
20. This simplifies to the standard deviation of the forecast error when the set of forecasts are unbiased.
21. Whilst it is standard practice to place greater weight on out-of-sample information than in-sample information, this need not be true in all situations (Inoue and Kilian, 2002).
22. Hendry and Clements (2004) show that consensus projections derived by averaging individual projections may out-perform consensus projections formed by using estimated weights to combine the individual projections when forecasting time series that are subject to location shifts.
23. Golinelli and Parigi (2004) report that consumer sentiment indicators can help to forecast the evolution of GDP, but do not consider explicitly whether this is the case when business confidence is also taken into account. It may also be the case that consumer confidence is a useful indicator of individual components of GDP, such as consumers' expenditure. Using a structural model in which private consumption expenditure is conditioned initially on disposable incomes and net wealth, Pain and Weale (2001) find

that consumer confidence contains significant additional information in the United States, but not in the United Kingdom.

24. There was some weak evidence that the real exchange rate might contain useful additional information in some, but not all, of the different euro area models. It is not included in the models reported in this paper.

25. The possibility of cross-country linkages has not been tested systematically in the work to date. Other studies suggest that this issue might be worth exploring further. For example, the US ISM survey is an input into the euro area indicator model of Grasmann and Keereman (2001).

26. In most country models current and lagged values of the survey variables enter the estimated equations. The main exception is in the United States, where only the current quarter value of the ISM manufacturing survey balance was found to matter. This may possibly reflect the nature of the series. The ISM survey makes explicit reference to quarter-on-quarter growth rates whilst in European countries the means of comparison in the survey questionnaire is less clear (year-on-year or quarter-on-quarter).

27. The possibility of using the OECD Composite Leading Indicator was also investigated in each economy. As with consumer confidence, these series were found to contain useful information for predicting GDP growth by themselves, but the additional information content they provided was limited once other indicators were included. They are not therefore incorporated in the models presented here.

28. The correlation between the quarterly rate of growth of the weighted country series and the official euro area data is high. The correlation coefficients are 0.82 for the construction output series over 1990-2002 and 0.59 for retail sales over 1996-2002.

29. The projections made using the combination of the quarterly bridge equation with the monthly auxiliary models use quarterly values of the indicator series forecast by the monthly models. The reported forecast errors for these models therefore include errors made in forecasting the monthly indicator series as well as the errors in the bridge equation. The conditional projections made using only the quarterly bridge or VAR models use the actual quarterly value of the indicator series.

30. The RMSFE for the current quarter forecast with zero information is 0.51 percentage point, rather than 0.50, because one additional quarter is included in the out-of-sample evaluation exercise.

31. In addition to the variables shown in Table 2, monthly series for retail trade and non-farm payroll employment are used in the monthly auxiliary equations for the United States in order to help predict the key indicator series that influence GDP growth directly.

32. Given publication lags, two months of information may become available for the pure survey model by the time at which a complete set of one month of data becomes available for the hard indicator model. But the latter still has a considerably lower RMFSE than the former (0.22 percentage point relative to 0.31 percentage point).

33. Covariances need to be taken into account when combining the variances of components (GDP in the member states of the euro area) to get the variance of their sum (euro area GDP).

34. The United Kingdom publishes an initial estimate of GDP growth within one month of the end of the quarter. This is available before data is published for all three months for each hard indicator. So the relevant comparison is likely to be between a pure sur-

vey model with three months of information and models with hard indicators with two months of information.

35. The consensus implicitly imposes equal weights on the different models based on surveys and hard indicators. The results from doing this appear to differ from those in the empirically estimated model that combines both types of indicators. The coefficients in this latter model will reflect the average effects of the individual series over the entire sample period. Differences between the consensus and estimated model forecast are therefore suggestive of differences between these two sets of weights. Hendry and Clements (2004) suggest that it may be optimal to average the estimates from different models if structural breaks have occurred in the series being forecast.
36. Another illustration of this is provided by the comparatively large gap between the errors from the one-quarter ahead forecasts when three months of current quarter information are available and those from the current quarter forecasts when zero months of current quarter information are available. The only differences between these results are that the sample period employed for estimation and prediction differs by one quarter.
37. Kang (2003) discusses the conditions under which this result might arise.
38. The existing literature provides mixed signals about the potential gains from aggregation of national forecasts. Bodo *et al.* (2000) find that a single euro area model for industrial production outperforms the aggregation of single country models, whereas Marcellino *et al.* (2003) find that the aggregation of country specific forecasts for a number of different macroeconomic variables improves upon the forecasts from single area-wide models.
39. In practice, monthly survey information is used in the euro area model to help predict missing monthly hard indicator data.
40. A similar conclusion is reached by Banerjee *et al.* (2003) in their detailed analysis of different types of models for predicting GDP growth in the euro area.
41. An additional real-time problem is that the first flash estimate of GDP growth occasionally coincides with the introduction of methodological changes in the national accounts that were not available at the time the prediction was made. One example of this concerns changes in the base year of the national accounts. No corrections have been incorporated for this in the results reported here, although in principle a correction factor could be estimated using the difference between the in-sample rates of GDP growth in the old and the new accounts.
42. The coefficient of variation of a sample is given by the ratio of the sample standard deviation to the sample mean. In Table 3, the coefficient of variation for Italy is 1.08.
43. A related strategy to assessing model uncertainty would be to adopt a Bayesian decision-theoretic approach and assign prior probabilities to each model based on their relative goodness-of-fit (Brock *et al.* 2003).
44. Two normal distributions with identical means, but different standard deviations are used on either side of the most likely forecast scenario.
45. See also Blix and Sellin (1998) for a derivation of the relationship between the balance of risks and the different standard deviations of the two piece normal distribution.
46. The Bank of England approach has been criticised by Wallis (1999) who recommends the use of the median rather than the mode as the most likely forecast.

47. In this case, the point forecast from the individual model rather than the mode of the range of different forecasts is taken as the most likely forecast. It is also possible to incorporate information from the point estimates of forecasts made using other methodologies.
48. These errors are marginally larger than the errors of 0.31 and 0.44 percentage points shown in Table A4. The differences reflect changes in the data set used to calculate the recursive out-of-sample errors over 1998-2002.
49. It would also be possible to centre Figure 1 on the median rather than the mode of the distribution (Wallis, 1999). This would result in smaller differences between Figures 1 and 2 in this case, since the median forecast was closer to the selected individual forecast than the modal forecast was.

Appendix 1

Detailed Empirical Results

Table A1. **United States forecast errors, 1998-2002**

	RMFSE		FDA	
	Current	Next	Current	Next
BENCHMARK MODELS				
Naïve model	0.68	0.71	55%	55%
Autoregressive model	0.57	0.59	75%	45%
SINGLE EQUATION BRIDGE MODELS				
Conditional forecast				
Pure survey equation	0.50		80%	
Pure hard indicator equation	0.37		90%	
Combination hard indicator/surveys	0.37		90%	
Non-conditional forecast				
Pure survey equation	0.55		65%	
Pure hard indicator equation	0.53		75%	
Combination hard indicator/surveys	0.55		75%	
MULTIVARIATE MODELS				
Conditional forecast				
Pure survey equation	0.49	0.50	85%	50%
Pure hard indicator equation	0.34	0.52	90%	55%
Combination hard indicator/surveys	0.35	0.54	95%	60%
Non-conditional forecast				
Pure survey equation	0.51	0.52	70%	50%
Pure hard indicator equation	0.50	0.60	75%	50%
Combination hard indicator/surveys	0.53	0.62	70%	45%
MONTHLY AUXILIARY MODELS				
Pure survey equation				
Zero month of current quarter information	0.49	0.55	80%	50%
One month of current quarter information	0.52	0.57	70%	45%
Two months of current quarter information	0.50	0.51	80%	45%
Three months of current quarter information	0.50	0.49	80%	60%
Pure hard indicator equation				
Zero month of current quarter information	0.55	0.58	70%	55%
One month of current quarter information	0.44	0.55	75%	60%
Two months of current quarter information	0.40	0.52	85%	55%
Three months of current quarter information	0.37	0.55	90%	60%
Combination hard indicator/surveys				
Zero month of current quarter information	0.51	0.59	80%	50%
One month of current quarter information	0.43	0.55	85%	60%
Two months of current quarter information	0.40	0.53	80%	45%
Three months of current quarter information	0.37	0.50	95%	60%
CONSENSUS FORECAST				
Zero month of current quarter information	0.49	0.55	75%	40%
One month of current quarter information	0.46	0.54	75%	45%
Two months of current quarter information	0.45	0.52	75%	45%
Three months of current quarter information	0.37	0.49	85%	60%
MONTHLY AUXILIARY MODELS CONSENSUS				
Zero month of current quarter information	0.49	0.56	75%	45%
One month of current quarter information	0.43	0.54	75%	60%
Two months of current quarter information	0.41	0.50	80%	55%
Three months of current quarter information	0.39	0.49	90%	65%

Table A2. **Euro area forecast errors, 1998-2002**

	RMFSE		FDA	
	Current	Next	Current	Next
BENCHMARK MODELS				
Naïve model	0.36	0.38	30%	30%
Autoregressive model	0.32	0.36	70%	45%
SINGLE EQUATION BRIDGE MODELS				
Conditional forecast				
Pure survey equation	0.31		65%	
Pure hard indicator equation	0.18		85%	
Combination hard indicator/surveys	0.18		85%	
Non-conditional forecast				
Pure survey equation	0.36		60%	
Pure hard indicator equation	0.36		65%	
Combination hard indicator/surveys	0.38		60%	
MULTIVARIATE MODELS				
Conditional forecast				
Pure survey equation	0.31	0.37	70%	70%
Pure hard indicator equation	0.18	0.37	85%	45%
Combination hard indicator/surveys	0.19	0.38	85%	50%
Non-conditional forecast				
Pure survey equation	0.36	0.40	55%	60%
Pure hard indicator equation	0.36	0.37	60%	65%
Combination hard indicator/surveys	0.37	0.39	50%	65%
MONTHLY AUXILIARY MODELS				
Pure survey equation				
Zero month of current quarter information	0.31	0.39	70%	60%
One month of current quarter information	0.33	0.41	70%	50%
Two months of current quarter information	0.31	0.36	65%	60%
Three months of current quarter information	0.31	0.31	65%	60%
Pure hard indicator equation				
Zero month of current quarter information	0.33	0.37	70%	45%
One month of current quarter information	0.22	0.36	75%	70%
Two months of current quarter information	0.22	0.36	75%	70%
Three months of current quarter information	0.18	0.34	85%	75%
Combination hard indicator/surveys				
Zero month of current quarter information	0.30	0.39	65%	45%
One month of current quarter information	0.22	0.40	75%	75%
Two months of current quarter information	0.21	0.35	80%	65%
Three months of current quarter information	0.18	0.31	85%	75%
CONSENSUS FORECAST				
Zero month of current quarter information	0.33	0.38	65%	55%
One month of current quarter information	0.31	0.38	65%	65%
Two months of current quarter information	0.30	0.36	65%	60%
Three months of current quarter information	0.19	0.33	80%	75%
MONTHLY AUXILIARY MODELS CONSENSUS				
Zero month of current quarter information	0.30	0.38	60%	60%
One month of current quarter information	0.24	0.38	75%	65%
Two months of current quarter information	0.23	0.35	75%	70%
Three months of current quarter information	0.21	0.31	80%	80%

Table A3. **Germany forecast errors, 1998-2002**

	RMFSE		FDA	
	Current	Next	Current	Next
BENCHMARK MODELS				
Naïve model	0.81	0.73	45%	45%
Autoregressive model	0.59	0.59	65%	45%
BIVARIATE BRIDGE MODELS				
Conditional forecast				
Pure survey equation	0.41		80%	
Pure hard indicator equation	0.44		95%	
Combination hard indicator/surveys	0.40		85%	
Non-conditional forecast				
Pure survey equation	0.61		80%	
Pure hard indicator equation	0.63		60%	
Combination hard indicator/surveys	0.63		65%	
MULTIVARIATE MODELS				
Conditional forecast				
Pure survey equation	0.41	0.64	80%	60%
Pure hard indicator equation	0.40	0.65	95%	55%
Combination hard indicator/surveys	0.37	0.67	80%	60%
Non-conditional forecast				
Pure survey equation	0.62	0.64	80%	60%
Pure hard indicator equation	0.65	0.61	60%	45%
Combination hard indicator/surveys	0.65	0.64	75%	50%
MONTHLY AUXILIARY MODELS				
Pure survey equation				
Zero month of current quarter information	0.51	0.66	85%	65%
One month of current quarter information	0.44	0.61	80%	70%
Two months of current quarter information	0.41	0.57	80%	70%
Three months of current quarter information	0.41	0.53	80%	70%
Pure hard indicator equation				
Zero month of current quarter information	0.61	0.61	65%	50%
One month of current quarter information	0.45	0.62	90%	65%
Two months of current quarter information	0.47	0.61	90%	65%
Three months of current quarter information	0.44	0.59	95%	65%
Combination hard indicator/surveys				
Zero month of current quarter information	0.51	0.65	75%	65%
One month of current quarter information	0.39	0.61	80%	80%
Two months of current quarter information	0.41	0.54	90%	80%
Three months of current quarter information	0.40	0.51	85%	80%
CONSENSUS FORECAST				
Zero month of current quarter information	0.55	0.62	65%	50%
One month of current quarter information	0.51	0.60	70%	55%
Two months of current quarter information	0.50	0.57	70%	65%
Three months of current quarter information	0.36	0.54	95%	65%
MONTHLY AUXILIARY MODELS CONSENSUS				
Zero month of current quarter information	0.50	0.63	75%	60%
One month of current quarter information	0.38	0.59	80%	70%
Two months of current quarter information	0.39	0.54	100%	80%
Three months of current quarter information	0.37	0.51	95%	75%

Table A4. **France forecast errors, 1998-2002**

	RMFSE		FDA	
	Current	Next	Current	Next
BENCHMARK MODELS				
Naïve model	0.45	0.51	40%	40%
Autoregressive model	0.42	0.45	55%	65%
SINGLE EQUATION BRIDGE MODELS				
Conditional forecast				
Pure survey equation	0.29		80%	
Pure hard indicator equation	0.30		80%	
Combination hard indicator/surveys	0.24		70%	
Non-conditional forecast				
Pure survey equation	0.33		60%	
Pure hard indicator equation	0.46		50%	
Combination hard indicator/surveys	0.33		65%	
MULTIVARIATE MODELS				
Conditional forecast				
Pure survey equation	0.29	0.36	75%	60%
Pure hard indicator equation	0.28	0.44	75%	45%
Combination hard indicator/surveys	0.24	0.35	70%	60%
Non-conditional forecast				
Pure survey equation	0.35	0.45	55%	55%
Pure hard indicator equation	0.44	0.45	50%	55%
Combination hard indicator/surveys	0.33	0.47	60%	55%
MONTHLY AUXILIARY MODELS				
Pure survey equation				
Zero month of current quarter information	0.31	0.44	65%	60%
One month of current quarter information	0.26	0.44	80%	45%
Two months of current quarter information	0.30	0.39	70%	65%
Three months of current quarter information	0.29	0.31	80%	70%
Pure hard indicator equation				
Zero month of current quarter information	0.43	0.48	60%	45%
One month of current quarter information	0.36	0.46	65%	55%
Two months of current quarter information	0.32	0.46	70%	55%
Three months of current quarter information	0.30	0.44	80%	55%
Combination hard indicator/surveys				
Zero month of current quarter information	0.30	0.46	70%	55%
One month of current quarter information	0.27	0.45	70%	45%
Two months of current quarter information	0.26	0.40	70%	50%
Three months of current quarter information	0.24	0.31	70%	75%
CONSENSUS FORECAST				
Zero month of current quarter information	0.33	0.44	55%	50%
One month of current quarter information	0.31	0.43	60%	55%
Two months of current quarter information	0.31	0.41	55%	55%
Three months of current quarter information	0.24	0.33	75%	65%
MONTHLY AUXILIARY MODELS CONSENSUS				
Zero month of current quarter information	0.31	0.44	65%	60%
One month of current quarter information	0.27	0.43	70%	45%
Two months of current quarter information	0.26	0.39	70%	60%
Three months of current quarter information	0.25	0.32	75%	70%

Table A5. **Italy forecast errors, 1998-2002**

	RMFSE		FDA	
	Current	Next	Current	Next
BENCHMARK MODELS				
Naïve model	0.48	0.49	45%	45%
Autoregressive model	0.42	0.42	65%	45%
SINGLE EQUATION BRIDGE MODELS				
Conditional forecast				
Pure survey equation	0.38		80%	
Pure hard indicator equation	0.26		90%	
Combination hard indicator/surveys	0.26		85%	
Non-conditional forecast				
Pure survey equation	0.34		65%	
Pure hard indicator equation	0.38		70%	
Combination hard indicator/surveys	0.41		70%	
MULTIVARIATE MODELS				
Conditional forecast				
Pure survey equation	0.42	0.42	75%	60%
Pure hard indicator equation	0.31	0.41	850%	65%
Combination hard indicator/surveys	0.27	0.39	90%	65%
Non-conditional forecast				
Pure survey equation	0.43	0.47	70%	55%
Pure hard indicator equation	0.44	0.43	65%	70%
Combination hard indicator/surveys	0.40	0.46	75%	60%
MONTHLY AUXILIARY MODELS				
Pure survey equation				
Zero month of current quarter information	0.42	0.42	70%	55%
One month of current quarter information	0.38	0.38	80%	80%
Two months of current quarter information	0.38	0.38	80%	60%
Three months of current quarter information	0.38	0.42	80%	55%
Pure hard indicator equation				
Zero month of current quarter information	0.36	0.45	75%	45%
One month of current quarter information	0.26	0.42	80%	65%
Two months of current quarter information	0.27	0.39	95%	65%
Three months of current quarter information	0.26	0.37	90%	70%
Combination hard indicator/surveys				
Zero month of current quarter information	0.42	0.45	70%	50%
One month of current quarter information	0.26	0.40	85%	70%
Two months of current quarter information	0.26	0.40	85%	85%
Three months of current quarter information	0.26	0.42	85%	70%
CONSENSUS FORECAST				
Zero month of current quarter information	0.36	0.41	80%	60%
One month of current quarter information	0.33	0.38	80%	65%
Two months of current quarter information	0.33	0.38	80%	65%
Three months of current quarter information	0.26	0.35	85%	70%
MONTHLY AUXILIARY MODELS CONSENSUS				
Zero month of current quarter information	0.37	0.41	70%	55%
One month of current quarter information	0. 26	0.36	85%	60%
Two months of current quarter information	0.26	0.36	85%	65%
Three months of current quarter information	0.26	0.36	85%	70%

Table A6. **United Kingdom forecast errors, 1998-2002**

	RMFSE		FDA	
	Current	Next	Current	Next
BENCHMARK MODELS				
Naïve model	0.38	0.44	50%	50%
Autoregressive model	0.32	0.34	65%	55%
SINGLE EQUATION BRIDGE MODELS				
Conditional forecast				
Pure survey equation	0.29		80%	
Pure hard indicator equation	0.24		75%	
Combination hard indicator/surveys	0.24		70%	
Non-conditional forecast				
Pure survey equation	0.27		85%	
Pure hard indicator equation	0.25		75%	
Combination hard indicator/surveys	0.25		75%	
MULTIVARIATE MODELS				
Conditional forecast				
Pure survey equation	0.32	0.29	80%	75%
Pure hard indicator equation	0.26	0.32	70%	55%
Combination hard indicator/surveys	0.27	0.30	70%	70%
Non-conditional forecast				
Pure survey equation	0.28	0.35	85%	60%
Pure hard indicator equation	0.32	0.36	70%	70%
Combination hard indicator/surveys	0.29	0.39	80%	60%
MONTHLY AUXILIARY MODELS				
Pure survey equation				
Zero month of current quarter information	0.30	0.34	80%	65%
One month of current quarter information	0.30	0.31	80%	75%
Two months of current quarter information	0.29	0.29	80%	70%
Three months of current quarter information	0.29	0.30	80%	60%
Pure hard indicator equation				
Zero month of current quarter information	0.28	0.31	75%	70%
One month of current quarter information	0.24	0.28	75%	70%
Two months of current quarter information	0.24	0.26	75%	70%
Three months of current quarter information	0.24	0.28	75%	65%
Combination hard indicator/surveys				
Zero month of current quarter information	0.28	0.31	75%	70%
One month of current quarter information	0.25	0.27	75%	70%
Two months of current quarter information	0.25	0.26	75%	65%
Three months of current quarter information	0.24	0.28	70%	60%
CONSENSUS FORECAST				
Zero month of current quarter information	0.24	0.32	75%	70%
One month of current quarter information	0.22	0.30	80%	70%
Two months of current quarter information	0.21	0.29	80%	70%
Three months of current quarter information	0.23	0.25	70%	75%
MONTHLY AUXILIARY MODELS CONSENSUS				
Zero month of current quarter information	0.27	0.31	75%	70%
One month of current quarter information	0.22	0.27	75%	65%
Two months of current quarter information	0.21	0.25	75%	70%
Three months of current quarter information	0.21	0.27	75%	60%

Table A7. **Japan forecast errors, 1998-2002**

	RMFSE		FDA	
	Current	Next	Current	Next
BENCHMARK MODELS				
Naïve model	0.97	1.17	55%	58%
Autoregressive model	0.92	0.93	80%	63%
SINGLE EQUATION BRIDGE MODELS				
Conditional forecast				
Pure survey equation	0.57		90%	
Pure hard indicator equation	0.55		85%	
Combination hard indicator/surveys	0.51		90%	
Non-conditional forecast				
Pure survey equation	0.80		100%	
Pure hard indicator equation	0.52		95%	
Combination hard indicator/surveys	0.54		95%	
MULTIVARIATE MODELS				
Conditional forecast				
Pure survey equation	0.67	0.80	95%	68%
Pure hard indicator equation	0.62	0.83	90%	63%
Combination hard indicator/surveys	0.59	0.72	90%	68%
Non-conditional forecast				
Pure survey equation	0.78	0.91	85%	63%
Pure hard indicator equation	0.64	0.87	95%	63%
Combination hard indicator/surveys	0.67	0.88	95%	79%
MONTHLY AUXILIARY MODELS				
Pure survey equation				
Zero month of current quarter information	0.64	0.79	85%	74%
One month of current quarter information	0.63	0.74	90%	68%
Two months of current quarter information	0.61	0.64	90%	68%
Three months of current quarter information	0.57	0.59	90%	68%
Pure hard indicator equation				
Zero month of current quarter information	0.61	0.74	90%	53%
One month of current quarter information	0.59	0.69	95%	47%
Two months of current quarter information	0.59	0.61	85%	47%
Three months of current quarter information	0.55	0.56	85%	58%
Combination hard indicator/surveys				
Zero month of current quarter information	0.57	0.75	90%	58%
One month of current quarter information	0.59	0.68	90%	53%
Two months of current quarter information	0.57	0.61	95%	63%
Three months of current quarter information	0.51	0.55	90%	68%
CONSENSUS FORECAST				
Zero month of current quarter information	0.60	0.78	100%	63%
One month of current quarter information	0.59	0.74	95%	74%
Two months of current quarter information	0.59	0.70	100%	68%
Three months of current quarter information	0.54	0.59	95%	68%
MONTHLY AUXILIARY MODELS CONSENSUS				
Zero month of current quarter information	0.58	0.74	100%	53%
One month of current quarter information	0.57	0.68	90%	58%
Two months of current quarter information	0.57	0.60	95%	74%
Three months of current quarter information	0.51	0.55	90%	68%

Appendix 2

The Two-piece Normal Distribution

This appendix contains a short summary of the properties of the two-piece normal distribution, and illustrates how this can be used to derive measures of forecast uncertainty around the point estimates from the indicator models. The two piece distribution is one in which two separate normal distributions with identical means, but different standard deviations, are used to describe the range of possible outcomes.

The density of a two-piece normal distribution (John, 1982; Johnson *et al.*, 1994) is written as follows:

$$f(x,\mu,\sigma_1,\sigma_2) = \begin{cases} C\exp\left[-\dfrac{(x-\mu)^2}{2\sigma_1^2}\right] & x \leq \mu \\ C\exp\left[-\dfrac{(x-\mu)^2}{2\sigma_2^2}\right] & x \succ \mu \end{cases} \qquad \text{[A1]}$$

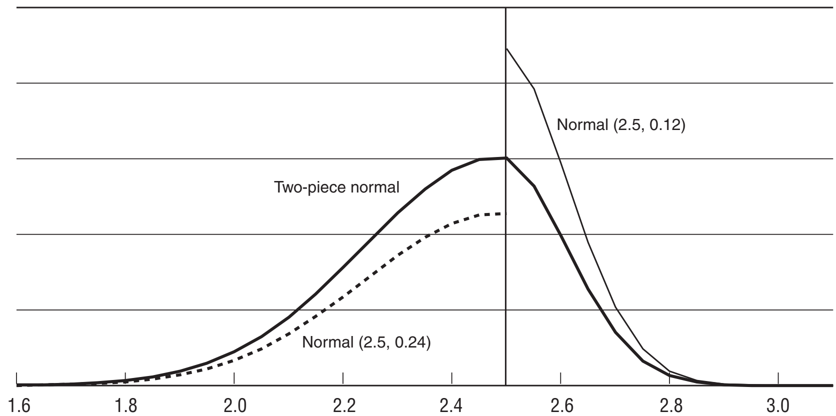
where μ is the distribution mode [or median, see Wallis (1999)], $C = k(\sigma_1 + \sigma_2)^{-1}$ and $k = \sqrt{\dfrac{2}{\pi}}$. Thus the distribution is defined by three parameters, the mode (in this case equal to the mean of the two original distributions) and the two standard deviations (σ_1 and σ_2). The first two moments of the density are:

$$\begin{aligned} \sigma^2(x) &= (1-k^2)(\sigma_2-\sigma_1)^2 + \sigma_2\sigma_1 \\ \bar{\mu} &= \mu + k(\sigma_2 - \sigma_1) \end{aligned} \qquad \text{[A2]}$$

If $\sigma_1 = \sigma_2$, the distribution collapses to the standard normal distribution. The difference between the mean (denoted by $\bar{\mu}$) and the mode (denoted by $\gamma = \bar{\mu} - \mu$) gives a measure of the asymmetry (skewness) of the distribution. When γ is negative, the distribution is skewed to the left, with the probability of downside risks being greater than that of upside risks. Once σ_1 and σ_2 have been estimated the density is fully identified and the confidence intervals associated with a given projection can be computed.

Figure A1 illustrates an example in which the mode equals 2.5 and $\sigma_1 = 0.24$ and $\sigma_2 = 0.12$. The solid line is the left hand side of a normal distribution with mean 2.5 and standard deviation 0.24, and the dotted line is the right hand side of a normal distribution with mean 2.5 and standard deviation 0.14. The bold line is the two-piece normal distribution computed from [A2]. The distribution is skewed to the left: the mode is 2.5% whereas the mean is 2.4%. In this particular case the risks are on the downside.

Figure A1. **Density of the two-piece normal distribution**



BIBLIOGRAPHY

Angelini, E., J. Henry and R. Mestre (2001), "Diffusion Index Based Forecast for the Euro Area", *European Central Bank Working Paper*, No. 61.

Baffigi, A., R. Gonelli and G. Parigi (2002), "Real Time GDP Forecasting in the Euro Area", *Banca d'Italia Discussion Paper* No. 456.

Baffigi, A., R. Gonelli and G. Parigi (2004), "Bridge Models to Forecast the Euro Area GDP", *International Journal of Forecasting*, forthcoming.

Banerjee, A., M. Marcellino and I. Masten (2003), "Leading Indicators for Euro Area Inflation and GDP Growth", CEPR *Discussion Paper* No. 3893.

Bay, J. and S. Ng (2002), "Determining the Number of Factors in Approximate Factor Models", *Econometrica*, 70 (1), pp. 191-221.

Blix, M. and P. Sellin (1998), "Uncertainty Bands for Inflation Forecasts", *Sveriges Riksbank Working Paper* No. 65.

Bodo, G.R, R. Gonelli and G. Parigi (2000) "Forecasting Industrial Production in the Euro Area Economy", *Empirical Economics*, 25, pp. 541-561.

Bovi, B., C. Lupi and C. Pappalardo (2000), "Predicting GDP Components Using ISAE Bridge Equations Econometrics Forecasting Model (BEEF)", ISAE *Working Paper* No. 13.

Britton, A. and N. Pain (1992), *Economic Forecasting in Britain*, National Institute of Economic and Social Research Report Series No. 5.

Britton, E., P. Fisher and J. Whitley (1998), "The Inflation Report Projections: Understanding the Fan Chart", *Bank of England Quarterly Bulletin*, February, pp. 30-37.

Brock, W.A., S.N. Durlauf and K.D. West (2003), "Policy Evaluation in Uncertain Economic Environments", *Brookings Papers on Economic Activity*, forthcoming.

Camba-Mendez, G., R.G. Kapetianos, R. Smith. and M.R. Weale (2001), "An Automatic Leading Indicator of Economic Activity: Forecasting GDP Growth for European Countries", *Econometric Journal*, 4, pp. 856-890.

CBO (2003), *The Uncertainty of Budget Projections: A Discussion of Data and Methods*, Congressional Budget Office Report, The Congress of the United States.

Corrado, C. and M. Greene (1988) "Reducing Uncertainty in Short-term Projections: Linkages of Monthly and Quarterly Models", *Journal of Forecasting*, 7, pp. 77-102.

Croushore, D. and T. Stark (2003), "A Real Time Data Set for Macroeconomists: Does Data Vintage Matter for Forecasting?", *The Review of Economics and Statistics*, 85, pp. 605-617.

Davis, E.P. and G. Fagan (1997), "Are Financial Spreads Useful Indicators of Future Inflation and Output Growth in EU Countries?", *Journal of Applied Econometrics*, 12, pp. 701-714.

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

Doan, T.A., R.B. Litterman and C.A. Sims (1984), "Forecasting and Conditional Projection Using Realistic Prior Distributions", *Econometric Reviews*, 3, pp. 1-100.

Doz, C. and F. Lenglart (1999), "Analyse factorielle dynamique: test du nombre de facteurs, estimation et application à l'enquête de conjoncture dans l'industrie", *Annales d'Économie et de Statistique*, 54, pp. 92-127.

Dubois, E. and E. Michaux (2005), "Étalonnages à l'aide d'enquêtes de conjoncture: de nouveaux résultats", *Économie et Prévision*, forthcoming.

Estrella, A. and F.S. Mishkin (1998), "Predicting US Recessions: Financial Variables as Leading Indicators", *Journal of Business Economics and Statistics*, 16 (2), pp. 198-205.

Fitzgerald, T.J. and P.J. Miller (1989), "A Simple Way to Estimate Current Quarter GNP", Federal Reserve Bank of Minneapolis, *Quarterly Review*, 4.

Forni, M, M. Allin, M. Lippi and L. Reichlin (2000), "The Generalized Factor Model: Identification and estimation", *TheReview of Economics and Statistics*, 82 (4), pp. 540-554.

Forni, M, M. Allin, M. Lippi and L. Reichlin (2001), "Coincident and Leading Indicators for the Euro Zone", *The Economic Journal*, 111, pp. 62-85.

Golinelli, R. and G. Parigi (2004), "Consumer Sentiment and Economic Activity", *Journal of Business Cycle Measurement and Analysis*, 1, pp. 147-170.

Grasmann, P. and F. Keereman (2001), "An Indicator-Based Short-Term Forecast for Quarterly GDP in the Euro Area", *European CommissionEconomic Paper* No. 154.

Grenouilleau, D. (2004), "A Sorted Leading Indicators Dynamic Factor Model for Short-Run Euro Area GDP Forecasting", *European Commission Economic Paper* No. 219.

Harvey, D.I., S.J. Leybourne and P. Newbold (1997), "Testing the Equality of Prediction Mean Square Errors", *International Journal of Forecasting*, 13, pp. 273-281.

Harvey, D.I. and P. Newbold (2000), "Test for Multiple Forecast Enconpassing", *Journal of Applied Econometrics*, 15, pp. 471-482.

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

Herkel-Rousse, H. and G. Prioux (2002), "L'apport des enquêtes de conjoncture dans les différents secteurs d'activité à l'analyse conjoncturelle", INSEE, *Note de conjoncture*, juin.

Ingenito, R. and B. Trehan (1996), "Using Monthly Data to Predict Quarterly Output", Federal Reserve Bank of San Francisco, *Economic Review*, 3.

Irac, D. and F. Sédillot (2002), "Short-Run Assessment of French Economic Activity Using OPTIM", *Banque de France Working Paper*, 88.

Kang, I-B. (2003), "Multi-Period Forecasting Using Different Models for Different Horizons: An Application to US Economic Time Series Data", *International Journal of Forecasting*, 19, pp. 387-400.

Koenig, F.E., S. Dolmas and J.M. Piger (2003), "The Use and Abuse of Real Time Data in Economic Forecasting", *The Review of Economics and Statistics*, 85, pp. 618-628.

Krolzig, H-M. and D.F. Hendry (2001), "Computer Automation of General-to-Specific Model Selection Procedures", *Journal of Economic Dynamics and Control*, 25, pp. 831-866.

Langmantel, E. (1999), "Das IFO Gesxhäfttklima als Indikator für die Prognose des Bruttoinlandsprodukt", IFO *Schnelldienst*, 16/17, pp. 16-21.

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

Miller, P.J. and D.M. Chin (1996), "Using Monthly Data to Improve Quarterly Forecasts", Federal Reserve Bank of Minneapolis, *Quarterly Review*, 2.

Mourougane, A. and M. Roma (2002), "Can Confidence Indicators be Useful to Predict Short-Term Real GDP Growth?", *European Central Bank Working Paper*, No. 133.

Pain, N. and M.R. Weale (2001), "The Information Content of Consumer Surveys", *National Institute Economic Review*, 178, pp. 44-47.

Parigi, G. and G. Schitzler (1995), "Quarterly Forecasts of the Italian Business Cycle by Means of Monthly Economic Indicators", *Journal of Forecasting*, 14, pp. 117-141.

Pesaran, M.H and A.G. Timmerman (1992), "A Simple Non Parametric Test of Predictive Performance", *Journal of Business and Economic Statistics*, 10, pp. 461-465.

Plosser, C.I and K.G. Rouwehhorst (1994), "International Term Structure and Real Economic Growth", *Journal of Monetary Economics*, 33 (1), pp. 133-155.

Robertson, J.C. and E.W. Tallman (1999), "Vector Autoregression and Reality", Federal Reserve Bank of Atlanta, *Economic Review*, 1.

Ruey, W.L. and H.S. Chung (1996), "The Use of High Frequency Data to Improve Macroeconometric Forecast", *International Economic Journal*, 10(2), pp. 65-83.

Rünstler, G. and F. Sédillot (2003), "Short-Term Estimate of Euro Area Real GDP Growth by Means of Monthly Indicators", *European Central Bank Working Paper*, No. 276.

Salazar, E. and M.R. Weale (1999), "Monthly Data and Short-Term Forecasting: An Assessment of Monthly Data in a VAR Model", *Journal of Forecasting*, 18, pp. 447-462.

Sédillot, F. and N. Pain (2003), "Indicator Models of Real GDP Growth in Selected OECD Countries", OECD *Economics Department Working Papers*, No. 364.

Skipper, H. (2005), "Early Estimates of GDP: Information Content and Forecasting Methods", *Economic Trends*, No. 617, pp. 26-35.

Stark, T. (2000), "Does Current Quarter Information Improve Quarterly Forecasts for the US Economy?", *Federal Reserve Bank of Philadelphia Working Paper*, 2.

Stock, J.H. and M.W. Watson (1989), "New Indexes of Coincident and Leading Economic Indicators", NBER *Macroeconomic Annual*, pp. 351-394.

Stock, J.H. and M.W. Watson (1991), "A Probability Model of the Coincident Economic Indicators", *in Leading Economic Indicators: New Approaches and Forecasting Records*, edited by K. Lahiri and G. Moore, Cambridge University Press.

Stock, J.H. and M.W. Watson (2002), "Macroeconomic Forecasting Using Diffusion Indexes", *Journal of Business and Economic Statistics*, 20 (2), pp. 143-162.

Van Rooij, M. and A.C.J. Stokman (2000), "Forecasting GDP Growth in the United States, Japan and EU on the Basis of Leading Indicators", *De Nederlandsche Bank Research Department*, *Research Memorandum* No. 636.

Wallis, K. (1999), "Asymmetric Density Forecasts of Inflation and the Bank of England's Fan Chart", *National Institute Economic Review*, 167, pp. 106-112.

TABLE OF CONTENTS

This paper assesses the effect of strengthened intellectual property rights in developing countries on international licensing activity. The analysis draws on indicators for four dimensions of intellectual property right stringency (covering patent rights, copyrights and trademark rights, as well as enforcement effectiveness) and on firm-level data related to licensing. Overall, the analysis points to a net positive effect of IPR strength on licensing activity, an effect that is strongest with respect to the indicators for patent rights and effective enforcement. Where developing countries have moved to address weaknesses in these areas in recent years, they have tended to experience increased inward licensing of intellectual assets. The overall implication is that intellectual property rights can play an important role in enabling firms in developing nations to access and exploit technologies and know-how through licensing agreements with parties in developed nations.

Traditionally, immigrant stocks have been estimated by the foreign-born population in some countries and the foreign population in others. With the 2000 round of censuses, almost all OECD countries have identified the country of birth of enumerated persons. This allows for a more comprehensive and comparable portrayal of migration movements both within and to the OECD zone over recent decades, with a number of European countries showing immigrant numbers that are as large in relative terms as those observed for the United States. In addition, data on the educational attainment of the population permit, for the first time, direct estimation of the extent of expatriation of highly educated persons to OECD countries for over a hundred countries of origin across the globe. For a number of countries, more than half of all highly educated persons born there are living (and working) in OECD countries. Expatriation of the highly educated on this scale constitutes a significant drain on the human capital capabilities of these countries.

Using micro data for individual firms, this paper finds that non-financial corporations in Japan and the major European countries in 2003 were more vulnerable to a rise in short-term interest rates than they were in 1993 when the previous interest rate tightening cycle began (with a vulnerable firm being defined as one which has a high debt-to-equity ratio and a low ability to service the debt). In contrast firms in the United States and Canada appear more prepared for rising interest rates. Furthermore, looking only at data from 2003 the paper finds that firms in Japan and the major euro area countries are more vulnerable than firms in the United States, Canada and United Kingdom. The micro data are also used to create for each country an economy-wide measure of vulnerability, which turns out to be significantly related to future movements in GDP and investment growth.