



OECD Economics Department Working Papers No. 364

Indicator Models of Real  
GDP Growth in Selected  
OECD Countries

**Franck Sédillot,  
Nigel Pain**

<https://dx.doi.org/10.1787/275257320252>

**Unclassified**

**ECO/WKP(2003)18**



Organisation de Coopération et de Développement Economiques  
Organisation for Economic Co-operation and Development

**22-Jul-2003**

**English text only**

**ECONOMICS DEPARTMENT**

**ECO/WKP(2003)18**  
**Unclassified**

**INDICATOR MODELS OF REAL GDP GROWTH IN SELECTED OECD COUNTRIES**

**ECONOMICS DEPARTMENT WORKING PAPERS NO. 364**

**by Franck Sédillot and Nigel Pain**

All Economics Department Working Papers are now available through OECD's Internet Web at  
<http://www.oecd.org/eco>

**JT00147657**

Document complet disponible sur OLIS dans son format d'origine  
Complete document available on OLIS in its original format

**English text only**

## **Indicator models of real GDP growth in selected OECD countries**

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. The research described in this paper develops a set of econometric models that provide estimates of GDP growth for a number of major OECD countries and zones in the two quarters following the last quarter for which official data have been published. These models exploit the considerable amount of monthly conjunctural information that becomes available before the release of 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. An automated procedure is developed allowing the indicator models to be run at any point in time at which new monthly data are released and the performance of a number of different models is examined in a recursive out-of-sample exercise over the period to 1998-2002. A clear overall conclusion 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.

*JEL classification:* C52, C53, E37

*Keywords:* Indicator models, bridge equations, monthly data, short-term economic forecasts

\*\*\*\*\*

## **Modèles de prévision de la croissance du PIB réel dans certains pays de l'OCDE, à l'aide d'indicateurs conjoncturels**

Disposer d'information précise et à jour sur la situation courante de l'économie est une exigence fondamentale dans le processus de décision économique. Les délais dans la publication des statistiques officielles signifient qu'un tableau complet des évolutions économiques au cours d'une période particulière n'est seulement disponible que quelques temps après la fin de cette période. Les travaux décrits dans cet article présentent un ensemble de modèles économétriques permettant d'estimer la croissance du PIB dans certaines grandes économies ou zones de l'OCDE à un horizon de deux trimestres après le dernier trimestre publié. Ces modèles utilisent une masse importante d'information conjoncturelle mensuelle disponible avant la publication des données officielles de comptes nationaux. Cette information provient à la fois d'indicateurs 'mous' comme les enquêtes d'opinion ou d'indicateur 'durs' comme la production industrielle ou les ventes de détail. Différentes fréquences et méthodes d'estimations sont utilisées. Une procédure automatique est mise en place afin de pouvoir simuler les modèles d'indicateurs conjoncturels à la publication de chaque nouvelle information mensuelle et les performances des différents modèles sont analysées à l'aide d'une procédure récursive utilisant des données hors période d'estimation sur la période 1998-2002. Une conclusion générale des travaux est qu'il est sous-optimal d'utiliser un unique modèle, à coefficients fixes pour chaque pays à chaque point dans le temps. En revanche, il est préférable d'utiliser une succession de modèles d'indicateurs conjoncturels qui peut être mise à jour rapidement à la publication de nouvelles données, le choix du modèle variant dans le temps et dépendant de l'ensemble d'information disponible.

*Mots clefs :* Modèles d'indicateurs conjoncturels, équations de passage, données mensuelles, prévisions économiques à court terme

*Classification JEL :* C52, C53, E37

**Copyright: OECD, 2003**

**Applications for permission to reproduce or translate all, or part of, this material should be made to:  
Head of Publications Service, OECD, 2 rue André-Pascal, 75775 PARIS CEDEX 16, France**

**TABLE OF CONTENTS**

1.	Introduction and summary.....	4
2.	The modelling strategy.....	7
3.	Empirical results for the G7 countries.....	16
4.	Using the indicator models in real-time.....	29
5.	Further issues.....	31
BIBLIOGRAPHY.....		33
Annex 1. Data.....		37
Annex 2. Empirical results.....		38
Annex 3. How important is consumer confidence?.....		45

## Indicator Models of Real GDP Growth in Selected OECD Countries

Franck Sédillot and Nigel Pain<sup>1</sup>

### 1. Introduction and summary

1. 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 of time 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.

2. The needs of policymakers have led recently to the development of a number of new statistical indicators. For instance, for the euro area, the European Commission now releases short-run quarterly GDP estimates each month,<sup>2</sup> the CEPR produces a monthly coincident indicator (EuroCOIN) and the EUROFRAME group has developed a quarterly GDP growth indicator which is published regularly in the Financial Times.<sup>3</sup> A number of national statistical agencies have also reduced considerably the lag between the end of the quarter and the time of the first official publication of national accounts estimates for that quarter. Advance quarterly GDP estimates are now produced in the United States and the United Kingdom within four weeks of the end of the quarter, and monthly GDP estimates are made in Canada and Finland. For the euro area, Eurostat has recently started producing a flash estimate available six weeks after the end of the quarter.<sup>4</sup> Nonetheless there are obvious limits to the extent to which this lag can be shortened, given the need to strike a balance between the timeliness and the accuracy of the preliminary estimates.

3. The research described in this paper develops a set of econometric models that provide estimates of GDP growth for a number of major OECD countries and zones in the two quarters following the last quarter for which official data have been published. These models seek to exploit the considerable amount of 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.<sup>5</sup>

---

1. Macroeconomic Analysis and Systems Management Division, OECD Economics Department. The authors are grateful to Jørgen Elmeskov, Michael Feiner, Michael Kiley and Pete Richardson for helpful comments and suggestions and to Rosemary Chahed and Jan-Cathryn Davies for their assistance in preparing the document. All views expressed are those of the authors and should not be taken as representing those of the OECD.

2. See the European Commission:  
[http://europa.eu.int/comm/economy\\_finance/indicators/euroareagd\\_en.htm](http://europa.eu.int/comm/economy_finance/indicators/euroareagd_en.htm)

3. See <http://www.euro-frame.org>

4. See EUROSTAT: <http://europa.eu.int/comm/eurostat/Public>

5. 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 such as GDP or, more typically, industrial production. Detailed information on the Composite Leading Indicator series is available at <http://www.oecd.org/std/cli>

4. There are a number of ways in which it is possible to use monthly data when making projections of a quarterly aggregate. If a quarterly forecasting model already exists, then one option would be to derive an optimal pooled forecast by combining the predictions from that quarterly model with the predictions from higher frequency models aggregated up to a quarterly basis (Corrado and Greene, 1988).<sup>6</sup> This approach provides a means by which conjunctural monthly information can be taken into account, but 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.

5. A second possible approach that is sometimes used to assess short-run developments is to construct coincident (and leading) indicator series by combining and summarizing the information in a large number of different monthly data series, an idea initiated and popularized by Stock and Watson (1989, 1991 and 2002) and developed further by Forni *et al.* (2000). The National Activity Index for the United States, issued monthly by the Chicago Fed, provides a real-time example of this approach.<sup>7</sup> But such indicators, which have the virtue of seeking to maximize 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.

6. The methodology followed in this paper therefore concentrates on a smaller range of potential indicator variables, although it also exploits 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. Although this approach has already been utilized in other empirical studies for some OECD member economies (see, for example, Bovi *et al.* (2000), Ingenito and Trehan (1996) and Parigi and Schlitzer (1995)), we build on these existing studies by seeking to address four 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.

7. The methodology used 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 the techniques and variables employed are similar across countries, allowing 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. By contrast, most of the existing indicator models in the literature are based on quarterly aggregates of monthly indicators and therefore can be run only once data are available for the entire quarter. Finally, the selected models for each country have been estimated recursively over the period 1998Q1-2002Q4, allowing 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.

8. Seven 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

---

6. See also Miller and Shin (1996), Ruey and Chung (1996) and Stark (2000).

7. <http://www.chicagofed.org/economicresearchanddata/national/index.cfm>

significant in most countries. This suggests there are clear gains from developing empirical indicator models.

- 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.
- 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 outturn estimate for GDP. This finding is in line with other empirical studies that have been conducted on this topic.<sup>8</sup>
- For one-quarter ahead projections, the performance of the estimated indicator models does not appear noticeably better than that of time series models until one to two 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.
- 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 and the euro area, whereas some form of combined model, either through estimation or through a consensus of the different model forecasts, appears more suitable for Germany, France, Italy 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 and Italy.
- 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 per cent confidence band (approximately 1 standard error) around any point estimate for GDP growth in that quarter is found to range from 0.4 to 0.8 percentage points, 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 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.

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.

9. One potential caveat that should be noted is that the forecast evaluations undertaken in this paper have all been carried out with data series as currently estimated by statistical offices. In real time,

---

8. Examples for the United States include Corrado and Greene (1988), Preston and Miller (1989), Rathjens and Robin (1993), Miller and Shin (1996) and Stark (2000). Salazar and Weale (1999) report a similar result for the United Kingdom. For the Euro area, see Rünstler and Sédillot (2003).

practitioners use data which is 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 the United States, Stark and Croushore (2000) conclude that findings on the relative forecasting performance of different models based on final data releases do not necessarily carry over to real time data. Koenig *et al* (2001) argue that forecasting equations should be estimated from real-time data rather than from final data releases.

10. The structure of this paper is as follows. Section 2 describes the modeling strategy employed, discussing the choice of data frequency, the types of models estimated, the procedures used to select the indicator variables used and the tests employed to evaluate out-of-sample forecast performance. The main empirical results are summarized in Section 3, which reports the indicator variables selected for each country and the comparative performance of each of the different types of indicator models estimated. Section 4 provides a discussion of a number of practical problems that can arise when seeking to use the indicator models on a real-time basis. Some further concluding comments and possible extensions to the present work are contained in Section 5.

## 2. The modelling strategy

11. This section describes the overall modeling strategy employed in the study. There are four key features of this strategy. First, the modeling framework developed is comprehensive, allowing forecasts to be generated and compared from a large 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. Second, the variables employed are selected according to a strictly defined process. Third, the models used are estimated with a procedure that determines automatically the optimal combination of current and lagged variables. Finally, various statistical tests are employed to gauge the forecasting capabilities of each estimated model. We discuss each of these features in turn.

### 2.1 The data set

12. 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 Annex 1. All the data used are seasonally adjusted and in many countries both the 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. It is therefore perhaps 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.<sup>9</sup> 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).

---

9. Examples of studies that utilize survey indicators include Brochner-Madsen (1993), Irac and Sédillot (2002), Mourougane and Moreno (2002), Rahiala and Teravirsta (1993) and Reynaud and Scherrer (1996). Authors who have found financial market measures, such as the yield spread, to have predictive power include Davis and Fagan (1997), Estrella and Hardouvelis (1991), Estrella and Mishkin (1998), Harvey (1991) and Plosser and Rouwenhorst (1994). Britton and Pain (1992) use both types of variables.



13. 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 two 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 the publication of industrial production figures for the third month of quarter T. Second, the information they convey is generally relevant only for the quarter to which they relate. Thus hard indicators are more likely to be of use at times when monthly information is not available for the whole quarter. This implies that a statistical approach which combines monthly and quarterly data may potentially be of use. Such an approach is exploited in this paper.

## 2.2 The modeling approach

14. An initial choice which has to be made in any exercise of the kind undertaken here, 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 has been employed in the subsequent empirical exercise.

### *Quarterly Bridge Models*

15. Bridge models are ones in which high frequency indicators are recast into quarterly GDP figures using single equation techniques. Models of this type have been utilized in empirical studies for both the largest OECD countries, as well as for the euro area as a whole. Some recent examples 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), Grasmann and Keereman (2001) and Rünstler and Sédillot (2003). These models are typically derived from an initial unrestricted Autoregressive Distributed Lag (ARDL[p,q]) 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:

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

where  $y_t$  and  $x_{j,t}$  denote the logs of real GDP and the selected indicators respectively,  $\Delta$  denotes the first difference operator and  $\rho(L)$  and  $\delta_j(L)$  denote lag polynomials of order p and  $q_j$ , respectively.<sup>10</sup> 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 additional information into account explicitly.

---

10. This is a restricted version of the standard ARDL model. Normally, variables enter in level terms. The interest of the generalised version is that it can be recast in the usual error correction form.

16. 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:

#### Conditional models

- A pure survey equation to project current quarter real GDP growth. By design, this equation can be used once t survey data are known for the entire quarter whilst GDP growth is not yet available.
- A pure conditional hard indicator equation. This equation produces a current quarter GDP forecast conditional on the release of data on the hard indicators for that quarter.
- A conditional combination of survey data and hard indicator variables.

#### Non conditional models

- A pure non-conditional survey equation projecting current quarter GDP using only lagged values of survey data. This equation allows an estimate of current quarter GDP growth to be made once a complete set of survey data are available for the preceding quarter. In contrast, the conditional model cannot be used until a complete set of survey data are available for the current quarter.
- A pure non-conditional hard indicators equation. In this equation only lagged values of the hard indicators are included.
- A non-conditional combination of survey data and hard indicators.

17. The attraction of this modelling approach lies in its simplicity. Nonetheless, it suffers from a number of drawbacks. First, the models do not exploit the fact that the indicators are released on a monthly basis, which, in principle, should allow predictions to be updated on a rolling basis during the quarter. Second, the predictions from the models are not particularly timely, 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 (mostly hard indicators). Finally, they do not attempt to test whether there are 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.

#### *Vector Autoregressive Models*

18. One straightforward way of attempting to circumvent the delays from waiting for new 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:

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

where  $y_t$  is a vector  $(\Delta \ln(\text{GDP}_t), I_t)'$ , with  $\Delta \ln(\text{GDP})$  denoting the growth rate of real GDP and  $I_t$  a vector containing indicator series.

19. This approach has a number of advantages compared to the single equation bridge model approach. 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 as data for 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) 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  $\mu_t$  in [2] can be contemporaneously correlated (Doan *et al*, 1984). Intuitively, a conditional forecast would be expected to produce a more accurate outcome because it takes additional information into account explicitly. 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.<sup>11</sup>

### *Monthly Auxiliary Models*

20. 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. Most are for the United States, with the pioneering work of Fitzgerald and Miller (1989), followed by Rathjens and Robins (1993), Ingenito and Trehan (1996) and Robertson and Tallman (1999) amongst others. For a recent application to the euro area, see Rünstler and Sédillot (2003). 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.

21. One possible way to exploit the most recent monthly set of conjunctural 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.<sup>12</sup> 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 Reed, 2000 and the initial estimate of euro area GDP now produced by Eurostat). A combined model of this sort has a number of useful properties. First, it is transparent because the information set underlying the projections is defined precisely. Second, the forecasting ability of the equations can be gauged easily when the information set is restricted. Third, it enables the quarterly GDP numbers to be produced using an estimated bridge equation and hard indicator series that are a combination of published monthly data and projected values for the remainder of the quarter. As shown below, the forecasting ability of the estimated hard indicator equation for the current quarter 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. More precisely, Bayesian VAR 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.

12. 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 minimize the out-of-sample forecast errors over the period 1998Q1-2002Q4. As Rünstler and Sédillot (2003) show, the method used to forecast the monthly indicators (naïve, autoregressive, VAR or BVAR models) does not have much influence on the current quarter projection.

22. 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 either the conditional estimated bivariate model or the conditional VAR model.

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. In contrast to the single equation case, 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.

23. To illustrate the practical implementation of the three different modeling approaches and the timing at which they can be employed, consider the following hypothetical example for the United States. Here, the first GDP estimate 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.

24. 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.

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

25. 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 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 progressively available.

### **2.3 The variable selection process**

26. Once the model framework has been designed, the second choice to be made concerns the number of potential variables to consider. Generally, the number of possible monthly indicators is large

compared to the number of quarterly observations on GDP available for estimation.<sup>13</sup> One option is to choose indicators on the basis of *ad hoc* or in-sample 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 [see for instance Bovi *et al.* (2001), Parigi and Schitzler (1995), Baffigi *et al.* (2002), and Rünstler and Sédillot (2003)]. 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.

27. An alternative, and less restrictive approach, is to consider all possible indicators and to summarize the information they contain in a small number of composite variables derived by means of a static or a dynamic principal factor analysis. There are a number of recent studies that have sought to exploit such techniques. Examples include Angelini *et al.* (2001), Doz and Lenglart (1999), Camba-Mendez *et al.* (2001), Kapetanios (2002), Forni *et al.* (2000 and 2001) and Stock and Watson (2002).

28. This approach enables a relatively large set of information to be summarized 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.<sup>14</sup> The computational burden may also be relatively heavy, particularly for dynamic factor analysis, because an iterative algorithm has to be employed.<sup>15</sup> This raises the costs of using the method regularly to obtain up to date estimates. It can also be difficult to identify the reasons for fluctuations in the derived indices because of the large number of variables involved.

29. Neither of the two approaches outlined above is unambiguously better than the other. In this paper we adopt an automated approach to data and model selection which combines elements of both the 'in-sample criteria approach' and the 'factor analysis approach'. We utilise a procedure that permits quarterly and monthly data to be combined in an optimal manner, but do not attempt an exhaustive analysis of every single monthly indicator that exists for each of the separate economies examined. For most countries we typically focus 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.<sup>16</sup>

---

13. For instance, Ingenito and Trehan (1996) consider 34 variables that might be included in their quarterly model to predict real GDP growth.

14. This has been addressed under specific assumptions, see Bay and Ng (2002) and Doz and Lenglart (1999).

15. Kapetanios (2002) proposes an alternative method, which does not require an iterative process (a maximum likelihood estimation). This method is based on a subspace algorithm.

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, this is a fairly recent development, and there is insufficient data in many cases to allow the series to be included in estimation.

30. The set of variables listed in Annex 1 is split into soft /financial and hard indicators. For each of these categories, individual series are ranked according to their explanatory power as measured by the  $\bar{R}^2$  statistic from a bivariate regression between GDP growth and the particular indicator (denoted  $x_{i,t}$ ) in which the retained lag of the indicator is selected automatically according to the Schwarz criterion<sup>17</sup>:

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

This pre-selection ranking is undertaken for each horizon considered in the bivariate models and for each period of the out-of-sample exercise.<sup>18</sup> The current quarter selection process amounts to ranking the indicators according to their contemporaneous causality. The one quarter-ahead selection, 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 real GDP growth.

#### 2.4 Variable combination / model estimation

31. In general, a subset of the four variables with the highest ranking was selected for each horizon.<sup>19</sup> The next step was to identify an optimal combination of these variables. There are a large number of possible specifications. For example, when estimating a single equation bridge model there are 15 possible combinations of these four variables even when no lags are allowed in the regression equation.<sup>20</sup> Once real GDP growth is allowed to enter with up to 4 lags, there are 75 possible models. If lagged values of each of the explanatory variables are also included, subject to the requirement that the same lag length is used for each of them, then the total number of possible specifications rises to 375 [ $15 * (\text{maxlag} + 1) * (\text{maxlag} + 1)$ ] when maxlag is set to 4 quarters. Out of this group of models the 'best' model is selected using the Schwarz criterion.<sup>21</sup> An important point to note is that some of the four pre-selected variables may not be included in the chosen specification. For instance some survey responses may convey similar information, so that a single measure can suffice (see, for instance, Doz and Lengart (1999)).

32. This process was repeated over a number of different sample periods in order to generate out-of-sample predictions. For the period 1998Q1 to 2002Q4 this involves the estimation of a total of 7,500 equations (375 equations per quarter for each of 20 quarters). Such a procedure implies that the specification of the model and the estimated coefficients can change from one quarter to the next. The whole process was repeated for each of the different types of model employed (single equation, VAR and monthly auxiliary). For each conditional and non-conditional quarterly VAR model (pure survey, pure hard

---

17. The Schwarz criterion was used because it tends to result in more parsimonious autoregressive models than the Akaike criterion does.

18. 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.

19. It would be possible to include a larger number of variables, but then the computational burden becomes very heavy, as the example below illustrates.

20. All possible combinations of  $q$  variables lead to  $\sum_{i=1}^q \frac{q!}{(q-i)!}$  models, when a lagged variable is considered as an additional variable. For 8 variables, with no lags, there are 255 possible combinations. When the number of variables increases, the number of possible models becomes rapidly disproportionate. In the case of 16 variables, 65,535 models have to be estimated.

21. Standard diagnostic tests (normality, autocorrelation at order 1 and 4, Arch, RESET and the Chow predictive failure test) were also carried out on this equation to assess its suitability.

indicator, combined survey and hard indicators), the variables present in the selected equivalent single equation bridge model for the same sample period were included. 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.

## 2.5 Comparing different models

33. Once each different type of model has been estimated, the question arises as to how their performance might best be compared. There are a number of different ways in which the forecast accuracy of different models can be evaluated formally. In this paper we use out-of-sample prediction errors and consider two statistical measures – forecast encompassing tests and directional accuracy tests.

34. One widely used measure of forecast accuracy is the Root Mean Square Forecast Error (RMSFE) of a model. This simplifies to the standard deviation of the forecast error when the set of forecasts are unbiased. 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 investigated using forecast encompassing tests. The modified version of the Diebold and Mariano (1995) test proposed by Harvey *et al.* (1997) uses the squared prediction errors to make pair-wise comparisons of different models as follows. Suppose there are vectors of squared prediction errors from two alternative models, denoted  $g(e_{1,t})$  and  $g(e_{2,t})$ . The null hypothesis of equality in the expected forecast performance of two different models, implies that:

$$H_0 : E [g(e_{1,t}) - g(e_{2,t})] = 0 \quad [4]$$

The test compares a statistic  $S$  to a critical value drawn from a student's  $t$ -distribution with  $(n-1)$  degrees of freedom ( $n$  is the number of independent point forecasts in the out-of-sample period) where  $S$  is given by:

$$S = \left[ \frac{n+1+2h+n-1h(h-1)}{n} \right]^{-\frac{1}{2}} \hat{V}(\bar{d})^{-\frac{1}{2}} \bar{d} \quad [5]$$

where  $\bar{d} = \frac{1}{n} \sum_{t=1}^n d_t$ ;  $d_t = g(e_{1,t}) - g(e_{2,t})$ ;  $\hat{V}(\bar{d}) = n^{-1} \left( \hat{\gamma}_0 + 2 \sum_{i=1}^{h-1} \hat{\gamma}_i \right)$ ;  $\hat{\gamma}_i$  are the estimated auto-covariances of the series of squared prediction errors and  $h$  the forecast horizon considered.<sup>22</sup>

35. As there are a large number of different models, it is possible that several of them may perform better than a given benchmark model. Thus it is useful 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 statistic  $S$  now becomes:

---

22. Here a rectangular kernel is used. There is a small possibility that a negative estimate may result. In this case the rectangular window can be replaced by a weighted window as in Newey and West (1987). When  $h=1$  (the current quarter case),  $\hat{V}(\bar{d}) = n^{-1} \hat{\gamma}_0$ .



$$S = (k-1)^{-1}(n-1)^{-1} \bar{d}' \hat{V}^{-1} \bar{d} \quad [6]$$

where  $\bar{d} = [\bar{d}_1 \quad \bar{d}_2 \quad \dots \quad \bar{d}_{k-1}]$  and  $\bar{d}_i = \frac{1}{n} \sum_{t=1}^n d_{i,t}$ , and  $k$  is the number of different models used. This statistic is distributed  $F_{k-1, n-k+1}$ .

36. 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. Pesaran and Timmerman (1992) propose a non-parametric statistic to evaluate forecast directional accuracy over any given sample period. This approach 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 corresponding statistic is:

$$s_n^2 = \frac{(\hat{p} - \hat{p}_*)^2}{\text{var}(\hat{p}) - \text{var}(\hat{p}_*)} \approx \chi_1^2 \quad [7]$$

where  $\hat{p}$  is the sample estimate of the probability of a correctly signed forecast and  $\hat{p}_*$  is the estimate of its expectation obtained under the null that the forecast and the outcome are independent.

37. The detailed tables of empirical results for each country reported in Section 3 and Annex 2, include, in addition to the forecasts from the individual models, two sets of 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. Their inclusion 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.

### 3. Empirical results for the G7 countries

38. 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.

### 3.1 *Indicators selected for each country*

39. 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 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.<sup>23</sup> An example showing the comparative importance of business and consumer survey measures for the euro area is discussed in detail in Annex 3. Financial indicator series were also found to be relatively unimportant.<sup>24</sup> The two composite business climate index measures evaluated – the National Association of Purchasing Managers (NAPM) index for the manufacturing sector in the United States (now produced by the Institute of Supply Management) and the IFO index for Germany – were both found to matter. Indeed the NAPM index dominated all the other business survey variables tested for the United States. The IFO index was found to be important, not only for Germany, but also for Italy, the latter being the only example of the use of a country-specific variable in a model for more than one country.

40. For the euro area as a whole, measures of the level of order books in industry and the level of stocks 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. Finally, 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.<sup>25</sup>

---

23. Consumer confidence may, for example, matter for 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. Desroches and Gosselin (2002) also find that consumer confidence contains useful information in the United States.

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. 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 NAPM was found to matter. This may possibly reflect the nature of the series. The NAPM 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).

Table 2. Selected variables

	Indicator(s) selected	Sources
<b>Surveys</b>		
United States	NAPM	ISM
Euro area	Level of order books and level of stocks	European Commission
- Germany	IFO business climate index, production tendency	IFO
- France	Production tendency and future production tendency	INSEE
- Italy	Production tendency, Production future tendency and IFO business climate	ISAE and IFO
United Kingdom	Production future tendency	CBI
<b>Hard indicators</b>		
United States <sup>26</sup>	Industrial production, consumption in volume, new construction put in place, monthly export volumes, total monthly level of stocks	FED, BEA and Census Bureau
Euro area	Industrial output, construction output, retail sales volumes	Eurostat, authors' calculations
- Germany	Industrial production, construction output, retail sales volumes	Bundesbank, Bundesamt
- France	Industrial production, consumption of manufactured goods	INSEE
- Italy	Industrial production	ISTAT
United Kingdom	Industrial production, retail sales volumes	ONS

41. From the range of hard indicators evaluated, measures of industrial production were found to matter for all countries. Indicators of retail sales were also found to be of use for all countries, with the exception of Italy. For the United States, the coverage of this measure is more comprehensive than in other countries as the Bureau of Economic Analysis produce monthly estimates of total consumers' expenditure. However this is less timely than the retail trade survey measures for other countries as there is a lag of a month before the information appears. Similar considerations apply to the series for total monthly inventory accumulation and, to a lesser extent, the series for monthly merchandise export volumes in the

26. For the United States, because of the large amount of timely monthly indicators available, the variable selection process has been extended to include five indicators. The inclusion of the accumulation of inventories series has only a small impact on the forecasting ability of the hard indicator bridge equation. Nonetheless, it has been decided to retain this variable both because it can be important at turning points and because it allows the US hard indicator bridge equation be to almost fully consistent with a GDP based component computation.

United States. Finally, measures of construction output were also found to contain useful information for the United States, the euro area and Germany.<sup>27</sup>

42. 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 principle 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.<sup>28</sup>

### 3.2 *Individual country results*

43. A complete set of Tables for each country detailing the forecast accuracy results can be found in Annex 2 of the paper. 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.<sup>29</sup> An important point to note is that the timing of the information set is defined according to the timing of data availability for the complete set of included indicators. In some cases, such as the United States, 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 these forecast errors are discussed in the main text. The out-of-sample period used to evaluate the performance of the separate models is relatively short and coincides with a period of comparatively subdued economic growth. Use of a longer evaluation period would however raise the risk of selecting a model with a good average performance but a relatively poor performance over the most recent quarters.

44. 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 growth has been most volatile in the United States and Germany, and the models for these two economies have the highest errors. For most countries the RMFSE is, on average, around 40 per cent 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.

---

27. Preliminary analysis suggested that measures of house prices were important for the United Kingdom with their inclusion reducing the forecast errors. However they are not included in the models discussed below, reflecting the difficulties in obtaining readily available and comparable time series data across countries. The impact of relaxing this requirement might usefully be explored in any future work.

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 quarterly bridge equation and the monthly auxiliary models use quarterly values of the indicator series forecast by the monthly models. The conditional projections made using only the quarterly bridge or VAR models use the actual quarterly value of the indicator series.

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

	United States	Germany	Euro area	France	Italy	United Kingdom
Standard deviation of real GDP growth	0.60	0.58	0.37	0.45	0.40	0.30
Mean of real GDP growth	0.70	0.30	0.50	0.65	0.40	0.60
Consensus forecast RMSFE (current quarter, all indicators available))	0.37	0.36	0.19	0.24	0.24	0.23

### *The United States*

45. The overall results are summarized in Tables 4 and 5 below and Table A1 of Annex 2. 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 labeled monthly auxiliary models. When there is no monthly information available for the current quarter (i.e. the first four to five 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 normality, this implies that there is approximately a 70 per cent chance that the true outturn for GDP growth (at a quarterly rate) will be within  $\pm 0.59$  percentage points 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.<sup>30</sup> Once an additional month of information becomes available, the RMSFE falls again to 0.43 percentage points. 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 points. 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.

46. 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.

47. 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

---

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

current quarter, the gains compared to the autoregressive GDP and the pure survey models, as measured by the ratio of the respective RMSFEs, are about 35 per cent and 26 per cent respectively (see Table 4 below). 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 NAPM 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.<sup>31</sup> 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 to 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 per cent.

48. 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 per cent) and the gains in terms of accuracy compared to 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.

49. A number of the absolute differences apparent from the detailed table of results are statistically significant. Table 4 reports the pairwise Diebold-Mariano test statistics for selected forecast models for GDP growth in the current quarter. For the model using hard indicators, the null of comparable forecast accuracy with the autoregressive models is rejected at the 5 per cent 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 per cent level.

50. However, the tests for multiple forecast encompassing reported in Table 5 provide stronger evidence in favour of the hard indicator model relative to the autoregressive and pure survey models.<sup>32</sup> 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 numeraire.

---

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. Because of the small number of post-sample observations available, only three models are compared for each test. It is not feasible to estimate a regression including all possible models simultaneously.

Table 4. Selected RMSFE ratios for the United States

	Benchmark model					
	AR	Survey 0m	Survey 1m	Survey 2m	Survey 3m	Hard indicators 3m
<b>Alternative model</b>						
Hard indicator 1m	0.77 [0.03]	0.89 [0.51]	0.84 [0.38]	0.88 [0.49]	0.88 [0.50]	1.18 [0.26]
Hard indicator 2m	0.70 [0.02]	0.80 [0.22]	0.81 [0.10]	0.77 [0.17]	0.80 [0.20]	1.08 [0.33]
Hard indicator 3m	0.65 [0.00]	0.75 [0.10]	0.71 [0.08]	0.74 [0.10]	0.74 [0.10]	1.0

Reading: the upper figures represent the RMSFE of the alternative model divided by the RMSFE of the benchmark model. A ratio below 1 means that former is lower than the latter. Figures in brackets are the p-values of the Diebold-Mariano statistics for the null of equal forecast accuracy.

Table 5. Selected forecast encompassing tests for the United States

	Numeraire		
	Autoregressive	Surveys 1m	Hard indicators 1m
GROUP 1.			
P value	0.02	0.06	0.26
GROUP 2.			
P value	0.00	0.01	0.99

Figures are the p-values for the null that a model encompasses the two others.

51. 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 as reported in Table A7 of Annex 2.

### *The euro area*

52. Results for the euro area are summarized in Tables 6 and 7 below and Table A2 of Annex 2. 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).

53. 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 points compared to 0.33 percentage

points from the monthly auxiliary models with one month of information).<sup>33</sup> 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.

54. The pairwise modified Diebold-Mariano tests in Table 6 and the multiple forecast encompassing tests in Table 7 provide 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 per cent level. The encompassing test results reinforce this conclusion. 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. As can be seen from Table A7 of Annex 2, 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.

Table 6. Selected RMSFE ratios for the euro area

	Benchmark model					
	AR	Survey 0m	Survey 1m	Survey 2m	Survey 3m	Hard indicators 3m
Alternative model						
Hard indicator 1m	0.69 [0.11]	0.60 [0.01]	0.68 [0.01]	0.68 [0.05]	0.66 [0.01]	1.24 [0.07]
Hard indicator 2m	0.67 [0.11]	0.65 [0.02]	0.65 [0.02]	0.67 [0.06]	0.60 [0.01]	1.20 [0.08]
Hard indicator 3m	0.55 [0.03]	0.55 [0.10]	0.55 [0.08]	0.55 [0.10]	0.53 [0.01]	1.0

Reading: the upper figures represent the RMSFE of the alternative model divided by the RMSFE of the benchmark model. A ratio below 1 means that former is lower than the latter. Figures in brackets are the p-values of the Diebold-Mariano statistics for the null of equal forecast accuracy.

Table 7. Selected forecast encompassing tests for the Euro area

GROUP 1	Numeraire		
	Autoregressive	Surveys 1m	Hard indicators 1m
P value	0.10	0.01	0.74
GROUP 2	Autoregressive	Surveys 3m	Hard indicators 3m
P value	0.04	0.06	0.66

Figures are the p-values for the null that a model encompasses the two others.

33. 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 points relative to 0.31 percentage points).



## Germany

55. Results for Germany are summarized in Tables 8 and 9 below and Table A3 of Annex 2. 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.<sup>34</sup> 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 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.

56. The forecast encompassing tests reported in Tables 8 and 9 suggest that the predictions from the combined survey/hard indicator model are 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 per cent level. The importance of the information from business surveys is also shown by the results of the Pesaran-Timmerman tests in Table A7. For both the current and one quarter ahead forecasts it is 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.

57. 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.

Table 8. Selected RMSFE ratios for Germany

	Benchmark model					
	AR	Hard indicators 0m	Hard indicators 1m	Hard indicators 2m	Hard indicators 3m	Surveys indicators 3m
<b>Alternative model</b>						
Combination 1m	0.66 [0.00]	0.63 [0.01]	0.85 [0.15]	0.82 [0.11]	0.88 [0.15]	0.96 [0.84]
Combination 2m	0.69 [0.00]	0.67 [0.01]	0.90 [0.36]	0.86 [0.08]	0.93 [0.43]	0.98 [0.63]
Combination 3m	0.68 [0.00]	0.65 [0.02]	0.88 [0.28]	0.85 [0.10]	0.91 [0.34]	0.97 [0.91]

Reading: the upper figures represent the RMSFE of the alternative model divided by the RMSFE of the benchmark model. A ratio below 1 means that former is lower than the latter. Figures in brackets are the p-values of the Diebold-Mariano statistics for the null of equal forecast accuracy.

34. 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).

Table 9. Selected forecast encompassing tests for Germany

	Numeraire		
GROUP 1	Autoregressive	Hard indicators 1m	Combination 1m
P value	0.00	0.09	0.52
GROUP 2	Autoregressive	Hard indicators 3m	Combination 3m
P value	0.00	0.12	0.32

Figures are the p-values for the null that a model encompasses the two others.

### *France*

58. Results for France are summarized in Tables 10 and 11 below and Table A4 of Annex 2. The general pattern of results 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 points). 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 points and directional accuracy of 75 per cent (i.e. there is a three in four chance of correctly predicting the direction of change in the rate of GDP growth).

59. 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.

Table 10. Selected RMSFE ratios for France

	Benchmark model					
	AR	Hard indicators 0m	Hard indicators 1m	Hard indicators 2m	Hard indicators 3m	Surveys indicators 3m
<b>Alternative model</b>						
Combination 1m	0.64 [0.02]	0.62 [0.01]	0.75 [0.07]	0.84 [0.22]	0.90 [0.46]	0.92 [0.64]
Combination 2m	0.61 [0.01]	0.59 [0.01]	0.71 [0.07]	0.80 [0.17]	0.86 [0.34]	0.88 [0.27]
Combination 3m	0.57 [0.01]	0.55 [0.01]	0.67 [0.03]	0.75 [0.08]	0.81 [0.16]	0.82 [0.14]

Reading: the upper figures represent the RMSFE of the alternative model divided by the RMSFE of the benchmark model. A ratio below 1 means that former is lower than the latter. Figures in brackets are the p-values of the Diebold-Mariano statistics for the null of equal forecast accuracy.

Table 11. Selected forecast encompassing tests for France

	Numeraire		
	Autoregressive	Hard indicators 1m	Combination 1m
GROUP 1			
P value	0.01	0.00	0.96
GROUP 2			
P value	0.01	0.03	0.34

Figures are the p-values for the null that a model encompasses the two others.

60. The forecast encompassing test statistics show 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 is especially apparent in the multiple forecast encompassing test using forecasts based on a 1 month within quarter information set.

### *Italy*

61. Results for Italy are summarized in Tables 12 and 13 below and Table A5 of Annex 2. These results are similar in several respects to those for Germany and France. Models based on survey data appear to generate much smaller forecast errors than models based only on hard indicators. For Italy this is true even when making a current quarter projection of GDP growth with the hard indicator model when three months of information are available. There is little change in the size of the errors from the pure hard indicator models as additional monthly information accumulates. In contrast, the expected pattern of a decline over time can be seen in the errors arising from the pure survey model. As Table 1 shows, the only hard indicator found to be statistically important for Italy in the work to date is industrial production. In other countries two or more hard indicators are found to be significant and included in the model.

Table 12. Selected RMSFE ratios for Italy

	Benchmark model					
	AR	Hard indicators 0m	Hard indicators 1m	Hard indicators 2m	Hard indicators 3m	Surveys indicators 3m
<b>Alternative model</b>						
Combination 0m	0.72 [0.08]	0.77 [0.10]	0.80 [0.22]	0.83 [0.21]	0.88 [0.34]	1.01 [0.91]
Combination 1m	0.67 [0.05]	0.71 [0.06]	0.74 [0.13]	0.77 [0.11]	0.81 [0.16]	0.93 [0.56]
Combination 2m	0.67 [0.04]	0.71 [0.04]	0.74 [0.11]	0.77 [0.08]	0.81 [0.11]	0.93 [0.53]
Combination 3m	0.64 [0.03]	0.67 [0.03]	0.71 [0.08]	0.73 [0.06]	0.78 [0.07]	0.89 [0.34]

Reading: the upper figures represent the RMSFE of the alternative model divided by the RMSFE of the benchmark model. A ratio below 1 means that former is lower than the latter. Figures in brackets are the p-values of the Diebold-Mariano statistics for the null of equal forecast accuracy.

Table 13. Selected forecast encompassing tests for Italy

	Numeraire		
	Autoregressive	Hard indicators 1m	Combination 1m
GROUP 1			
P value	0.04	0.10	0.82
GROUP 2			
P value	0.03	0.02	0.79

Figures are the p-values for the null that a model encompasses the two others.

62. The best model, at least in terms of forecast error size, appears to be a model that combines information from surveys and hard indicators. For the current quarter forecasts, the consensus prediction yields the lowest RMFSE. For the one-quarter ahead forecasts, the estimated combined monthly auxiliary model performs best, particularly once two months of information become available for the quarter preceding the one being forecast (four to five months before the first outturn for GDP is published). All these features are confirmed by the results of the Diebold-Mariano tests. The combined model is always better statistically than the autoregressive model and the pure hard indicators model for the same set of within quarter information (at the 10 per cent threshold). The multiple forecast encompassing test indicates 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*

63. Results for the United Kingdom are summarized in Tables 14 and 15 below and Table A6 of Annex 2. The general pattern of the results from the different models 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.<sup>35</sup> 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.<sup>36</sup>

Table 14. Selected RMFSE ratios for the United Kingdom

	Benchmark model					
	AR	Hard indicators 0m	Hard indicators 1m	Hard indicators 2m	Hard indicators 3m	Surveys indicators 3m
<b>Alternative model</b>						
Consensus 0m	0.76 [0.02]	0.88 [0.38]	1.00 [0.92]	1.00 [0.99]	1.03 [0.88]	0.83 [0.11]
Consensus 1m	0.67 [0.01]	0.81 [0.20]	0.93 [0.65]	0.91 [0.71]	0.94 [0.79]	0.77 [0.02]
Consensus 2m	0.67 [0.01]	0.77 [0.14]	0.89 [0.94]	0.87 [0.56]	0.90 [0.82]	0.73 [0.01]
Consensus 3m	0.72 [0.03]	0.83 [0.30]	0.96 [0.99]	0.94 [0.74]	0.97 [0.77]	0.79 [0.09]

Reading: the upper figures represent the RMSFE of the alternative model divided by the RMSFE of the benchmark model. A ratio below 1 means that former is lower than the latter. Figures in brackets are the p-values of the Diebold-Mariano statistics for the null of equal forecast accuracy.

64. 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

35. As the United Kingdom publishes an initial estimate of GDP growth within one month of the end of the quarter, it is possible this will be available before data is published for all three months for each hard indicator. So the relevant comparison is likely to be between a pure survey model with 3 months of information and models with hard indicators with two months of information.

36. 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.

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.

65. The pairwise encompassing tests show that the differences between the errors from the consensus model and the autoregressive model are statistically significant at the 5 per cent 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 to 80 per cent (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 per cent levels (see Table A7 of Annex 2). This is also true of a number of the one-quarter ahead forecast models.

Table 15. Selected forecast encompassing tests for the United Kingdom

	Numeraire		
GROUP1	Autoregressive	Hard indicators 0m	Combination 0m
P value	0.01	0.34	0.35
GROUP 2	Autoregressive	Hard indicators 3m	Combination 3m
P value	0.01	0.10	0.21

Figures are the p-values for the null that a model encompasses the two others.

#### 4. Using the indicator models in real-time

66. There are a number of interesting issues that arise when the indicator models are used 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 utilize a more up-to-date information set. The expected forecast error from the latter may be lower than from the former.

67. For example, consider the results for France reported in Table A4. For any given 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 to 0.30) and directional accuracy (80 per cent compared to 70 per cent). 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.

68. A further issue that can arise when making forecasts for the forthcoming as well as the current quarter is that different models may appear to be preferred for different quarters, especially given the timing of information releases. 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 to 0.45 and a directional accuracy of 65 per cent compared to 45 per cent). 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.

69. 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 outcome 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.

70. 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. One important point to check is therefore that the implicit projection for the residual component of the euro area is plausible. A further issue concerns the most suitable approach in the interval between the publication of estimates for GDP growth in the major economies and that for the euro area as a whole. At this juncture it may be more appropriate to incorporate the country outcome data formally into the euro area projection by weighting together the outcome data with the indicator projection for those countries yet to publish data, rather than simply seek to use the aggregate euro area indicator model.<sup>37</sup>

71. A related consideration is the extent to which judgments may 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.<sup>38</sup> 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.

72. The factors discussed above illustrate some of the practical considerations that need to be taken into account when seeking to use the indicator models in real time. All of the indicator models can be regarded as statistically acceptable models. Variables are included which 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

---

37. This procedure is being used currently to produce projections for the euro area when the national GDP data are available but the zone aggregate is not.

38. In practice, monthly survey information is used in the euro area model to help predict missing monthly hard indicator data.

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.<sup>39</sup>

73. 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.

## 5. Further issues

74. There are a number of additional issues of interest which are not discussed in detail in this paper or, indeed, in much of the extant forecasting literature. One concerns the potential for linkages across countries. In the analysis reported in this paper the sole direct example of this is the inclusion of the German IFO survey in the Italian models. But it is easy to imagine that there might be useful information for a number of countries from any timely national indicator series related to international trade or capital markets. It would also be of use to extend the data set where possible to permit evaluation of additional national variables not considered so far. In particular it would be of benefit to consider a wider range of indicators for the services sectors, although data availability is likely to be a problem here, especially as selection is done empirically rather than through simple inspection of recent cyclical patterns. Judgments may need to be made about the usefulness of seeking to incorporate specific variables that can be obtained readily for one country, but not for others. This could improve the accuracy of the projections in individual countries, but might complicate comparisons of results across countries.

75. There is also a question of whether there are potential benefits to be had from seeking to derive a euro area projection by aggregating the projections for the individual member states rather than making a separate projection for the overall zone. Indicator series for the overall zone are weighted averages of those for the individual economies. The weights appropriate for aggregation purposes may not be the optimal ones for extracting the information of use in the separate national indicator series for projecting short-term economic developments in the zone as a whole. The existing literature provides mixed signals about the potential gains from aggregation of national forecasts. Bodo *et al.* (2000) find that an area-wide model for industrial production outperforms the aggregation of single country models, whereas Marcellino, Stock and Watson (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.

76. Finally, there is the issue of the means by which the uncertainty around any point forecast might usefully be conveyed. Two types of calculations can be made with the range of indicator models developed to date. 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. But there is also information that can be derived from the cross-section of different point forecasts from the different models.

---

39. 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.



77. 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 (in our case, 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. 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.<sup>40</sup>

78. 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.<sup>41</sup> To calculate this, 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 judged to be the most likely outturn, and the difference between the mean and the mode is taken as an estimate of the relative balance of risks around the mode (the skewness of the assumed distribution).<sup>42</sup> If the mean is above the mode than there are likely to be more upside risks than downside risks, and vice versa.<sup>43</sup> 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.

79. 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 only possible for forecasting organizations and official bodies such as the Bank of England to use a single estimate of the forecast error based on the past track record of their forecasts, each of which 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.<sup>44</sup> Equally, it is possible to compute the risks around the modal forecast using information on the past standard errors from modal forecasts.

---

40 . 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).

41 . Two normal distributions with identical means, but different standard deviations are used on either side of the most likely forecast scenario.

42 . See 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.

43 . 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.

44 . 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.

## BIBLIOGRAPHY

- ANGELINI, E., J. HENRY and R. MESTRE (2001), "Diffusion Index Based Forecast for the Euro Area", *ECB Working Paper*, 61.
- BAFFIGI, A., R. GONELLI and G. PARIGI (2002), "Real Time GDP Forecasting in the Euro Area", *mimeo*, Banca d'Italia.
- 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 Factor in Approximate Factors Models", *Econometrica*, 70 (1), pp 191-221.
- BLIX, M. and P. SELLIN (1998), "Uncertainty bands for inflation forecasts", Sveriges Riksbank Working Paper N°65.
- BODO, G.R., R.GONELLI and G. PARIGI (2000) "Forecasting Industrial Production in the Euro Area Economy", *Empirical Economics*, 25, 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.
- BROCHNER MADSEN, J. (1993), "The Predictive Value of Production Expectation in Manufacturing Industry", *Journal of Forecasting*, 12, pp 273-289.
- 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.
- CORRADO, C. and M. GREENE (1988) "Reducing Uncertainty in Short-term Projections: Linkages of Monthly and Quarterly Models", *Journal of Forecasting*, 7, 77-102.
- CROUSHORE, D. and T. STARK (2000), "A Real Time Data Set for Macroeconomists: Does Data Vintage Matter for Forecasting?" *Federal Reserve Bank of Philadelphia Working Paper*, 6 (June).
- 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.
- DESROCHES, B. and M.A. GOSSELIN (2002), "The Usefulness of Consumer Confidence Indexes in the United States", *Bank of Canada Working Paper* No. 2002-22.

- 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. LENGART (1999), "Analyse factorielle dynamique: test du nombre de facteurs, estimation et application a l'enquête de conjoncture dans l'industrie", *Annales d'Économie et de Statistique*, 54, pp 92-127.
- ESTRELLA, A. and G.A. HARDOUVELIS (1991), "The Term Structure as a Predictor of Real Economic Activity", *Journal of Finance*, XLVI (2), pp 555-576.
- 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", *Review 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.
- GRASMANN, P. and F. KEEREMAN (2001), "An indicator-based short-term forecast for quarterly GDP in the euro area", European Commission, *Economic Paper*, 154.
- HARVEY, C.R. (1991), "Interest Rate Based Forecasts of German Economic Growth", *Weltwirtschaftliches Archiv*, 4, pp 701-717.
- 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 Encompassing", *Journal of Applied Econometrics*, 15, 471-482.
- 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.
- KAPETANIOS, R.G. (2002), "Factor Analysis Using Subspace Factor Models: Some Theoretical Results and Application to the UK Inflation", *Banque de France working paper*, 88.
- KOENIG, F.E., S. DOLMAS and J.M. PIGER (2001), "The Use and Abuse of Real Time Data in Economic Forecasting", Federal Reserve Bank of St. Louis, *Working Paper Series*, 15B

- LANGMANTEL, E. (1999), "Das IFO Geschäftsklima 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?", *ECB working paper*, 133.
- NEWBY, W.K. and K.D. WEST (1987), "A Simple Positive Semi Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix", *Econometrica*, 55, pp 703-708.
- 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, 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.
- RATHJENS, P. and R.P. ROBINS (1993), "Forecasting Quarterly Data Using Monthly Information", *Journal of Forecasting*, 12, pp 321-330.
- RAHALA, M. and T. TERASVIRTA (1993), "Business Survey Data in Forecasting the Output of Swedish and Finnish Metal and Engineering Industry: a Kalman Filter Approach", *Journal of Forecasting*, 12, pp 255-271.
- REED, G. (2000), "How the Preliminary Estimate of GDP is Produced", *Economic Trends*, 556, pp 53-61.
- REYNAUD, M. and S. SCHERRER (1996), "Une modélisation VAR de l'enquête mensuelle de conjoncture dans l'industrie", Direction de la Prévision, *Working Paper*, 96-12.
- ROBERSON, 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*, forthcoming.
- 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.

- 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 Macroeconomics 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 (2001), "Forecasting Output and Inflation: the Role of Asset Prices" NBER working Paper, 8180.
- 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.

### Annex 1. Data

80. A complete list of the variables examined during the course of the empirical work is given below. In most countries the business survey information used relates only to the manufacturing sector, reflecting the relatively short time series available in surveys of other parts of the economy. The noticeable exception is Germany for which a broader coverage (retail trade and construction) is available. There are also differences in the coverage of some of the hard indicators. For the United States it is possible to use the monthly information on economy-wide consumers' expenditure published by the Bureau of Economic Analysis. But in other countries, monthly series on consumption, or retail trade and car registrations generally cover only a small part of total private consumption expenditure. Even so, they can still provide a useful indication of recent domestic demand developments. As far as possible the dataset has been constructed so that homogeneity is ensured across countries. This facilitates the practical implementation and makes cross-country comparisons much easier. However, the possibility remains that there are some country specific variables which could improve the accuracy of the estimated models. One example might be indicators of the housing market in the United Kingdom.

#### Variables used

	Country Coverage
<b>Soft indicators</b> <u>Industrial survey</u> <ul style="list-style-type: none"> <li>- Production tendency</li> <li>- Production future tendency</li> <li>- Total order books level</li> <li>- Foreign order books level</li> <li>- Stocks of finished goods level</li> </ul> <u>Composite indicators</u> <ul style="list-style-type: none"> <li>- NAPM</li> <li>- Business climate index</li> </ul> <u>Consumer survey</u> <ul style="list-style-type: none"> <li>Consumer confidence</li> </ul>	US, euro area, Germany, France, UK US, euro area, Germany, France, Italy, UK US, euro area, Germany, France, Italy, UK US, euro area, Germany, France, Italy, UK US, euro area, Germany, France, Italy, UK  US Germany  US, euro area, Germany, France, Italy, UK
<b>Financial variables</b> <ul style="list-style-type: none"> <li>- Long-term interest rate</li> <li>- Short-term interest rate</li> <li>- Real exchange rate</li> </ul>	US, euro area, Germany, France, Italy, UK US, euro area, Germany, France, Italy, UK US, euro area, Germany, France, Italy, UK
<b>Hard Indicators</b> <u>Supply side</u> <ul style="list-style-type: none"> <li>- Industrial output</li> <li>- Construction output</li> <li>- New orders, shipments</li> </ul> <u>Demand side</u> <ul style="list-style-type: none"> <li>- Total consumption in volumes</li> <li>- Retail sales in volumes</li> <li>- Total non farm payroll Employment</li> <li>- New car registrations</li> <li>- Monthly level of stocks</li> <li>- Monthly trade indicators in volume</li> <li>- Unemployment rate</li> </ul>	US, euro area, Germany, France, Italy, UK The US, Euro area, Germany, France US, Germany  US US, euro area, Germany, France, Italy, UK US US, euro area, Germany, France, Italy, UK US US, Germany, France All countries

## Annex 2. Empirical results

Table A1. United States forecast errors, 1998-2002

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

Table A2. Euro area forecast errors, 1998-2002

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



Table A3. Germany forecast errors, 1998-2002

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

Table A4. France forecast errors, 1998-2002

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

Table A5. Italy forecast errors, 1998-2002

	RMFSE		FDA	
	Current	Next	Current	Next
<b>BENCHMARK MODELS</b>				
- Naive model	0.47	0.49	55%	55%
- Autoregressive model	0.41	0.41	55%	55%
<b>SINGLE EQUATION BRIDGE MODELS</b>				
<u>Conditional forecast</u>				
- Pure survey equation	0.29		90%	
- Pure hard indicator equation	0.34		90%	
- Combination hard indicator / surveys	0.26		85%	
<u>Non conditional forecast</u>				
- Pure survey equation	0.39		70%	
- Pure hard indicator equation	0.36		65%	
- Combination hard indicator / surveys	0.40		60%	
<b>MULTIVARIATE MODELS</b>				
<u>Conditional forecast</u>				
- Pure survey equation	0.36	0.41	90%	45%
- Pure hard indicator equation	0.33	0.36	90%	60%
- Combination hard indicator / surveys	0.28	0.39	75%	60%
<u>Non conditional forecast</u>				
- Pure survey equation	0.41	0.42	60%	55%
- Pure hard indicator equation	0.37	0.40	65%	70%
- Combination hard indicator / surveys	0.39	0.44	60%	60%
<b>MONTHLY AUXILIARY MODELS</b>				
- Pure survey equation				
Zero month of within cq information	0.34	0.40	60%	75%
One month of within cq information	0.33	0.38	70%	65%
Two months of within cq information	0.33	0.36	70%	60%
Three months of within cq information	0.28	0.33	85%	65%
- Pure hard indicator equation				
Zero one month of within cq information	0.38	0.41	65%	60%
One month of within cq information	0.37	0.40	70%	55%
Two months of within cq information	0.36	0.40	85%	50%
Three months of within cq information	0.34	0.38	90%	50%
- Combination hard indicator / surveys				
Zero one month of within cq information	0.30	0.42	70%	70%
One month of within cq information	0.28	0.37	85%	65%
Two months of within cq information	0.27	0.31	85%	75%
Three months of within cq information	0.26	0.29	85%	65%
<b>CONSENSUS FORECAST</b>				
Zero one month of within cq information	0.33	0.44	55%	50%
One month of within cq information	0.31	0.43	60%	55%
Two months of within cq information	0.31	0.41	55%	55%
Three months of within cq information	0.24	0.33	75%	65%
<b>MONTHLY AUXILIARY MODELS CONSENSUS</b>				
Zero one month of within cq information	0.32	0.41	70%	65%
One month of within cq information	0.30	0.38	80%	60%
Two months of within cq information	0.29	0.34	85%	65%
Three months of within cq information	0.28	0.32	85%	70%

Table A6. United Kingdom forecast errors, 1998-2002

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

Table A7. Results of the Pesaran and Timmerman Test, 1998-2002

	US		Euro area		Germany		France		Italy		UK	
	CQ	NQ	CQ	NQ	CQ	NQ	CQ	NQ	CQ	NQ	CQ	NQ
<b>BENCHMARK MODELS</b>												
- Naive model												
- Autoregressive model	0.02	0.57	0.01	0.76	0.13	0.08	0.71	0.24	0.64	0.76	0.11	0.59
<b>SINGLE EQUATION BRIDGE MODELS</b>												
<u>Conditional forecast</u>												
- Pure survey equation	0.01		0.03		0.01		0.01		0.00		0.01	
- Pure hard indicator equation	0.00		0.00		0.00		0.01		0.00		0.02	
- Combination hard indicator / surveys	0.00		0.00		0.00		0.04		0.00		0.06	
<u>Non conditional forecast</u>												
- Pure survey equation	0.19		0.15		0.01		0.44		0.07		0.01	
- Pure hard indicator equation	0.02		0.03		0.27		0.84		0.18		0.02	
- Combination hard indicator / surveys	0.02		0.04		0.13		0.24		0.37		0.02	
<b>MULTIVARIATE MODELS</b>												
<u>Conditional forecast</u>												
- Pure survey equation	0.00	0.96	0.01	0.19	0.01	0.08	0.02	0.54	0.00	0.76	0.01	0.02
- Pure hard indicator equation	0.00	0.70	0.00	0.56	0.00	0.64	0.02	0.28	0.00	0.33	0.07	0.52
- Combination hard indicator / surveys	0.00	0.38	0.00	0.38	0.00	0.33	0.06	0.34	0.01	0.42	0.07	0.02
<u>Non conditional forecast</u>												
- Pure survey equation	0.07	0.96	0.25	0.80	0.00	0.42	0.57	0.65	0.18	0.64	0.00	0.36
- Pure hard indicator equation	0.02	0.96	0.05	0.56	0.26	0.64	0.84	0.71	0.19	0.04	0.01	0.06
- Combination hard indicator / surveys	0.07	0.64	0.13	0.34	0.01	0.88	0.44	0.84	0.37	0.37	0.01	0.35
<b>MONTHLY AUXILIARY MODELS</b>												
- Pure survey equation												
Zero month of within cq information	0.01	0.96	0.01	0.54	0.00	0.09	0.24	0.54	0.07	0.13	0.01	0.15
One month of within cq information	0.06	0.57	0.01	0.68	0.01	0.04	0.00	0.43	0.00	0.76	0.01	0.02
Two months of within cq information	0.01	0.57	0.03	0.54	0.01	0.04	0.08	0.19	0.00	0.19	0.01	0.06
Three months of within cq information	0.01	0.37	0.03	0.82	0.01	0.06	0.01	0.06	0.00	0.13	0.01	0.35
- Pure hard indicator equation												
Zero one month of within cq information	0.02	0.64	0.01	0.05	0.13	0.79	0.54	0.43	0.19	0.33	0.02	0.06
One month of within cq information	0.00	0.33	0.01	0.05	0.00	0.17	0.24	0.84	0.07	0.57	0.02	0.06
Two months of within cq information	0.01	0.57	0.01	0.05	0.00	0.19	0.09	0.84	0.00	0.96	0.02	0.06
Three months of within cq information	0.00	0.33	0.00	0.02	0.00	0.19	0.01	0.84	0.00	0.96	0.02	0.16
- Combination hard indicator / surveys												
Zero one month of within cq information	0.01	0.96	0.03	0.56	0.02	0.09	0.08	0.84	0.07	0.05	0.02	0.06
One month of within cq information	0.00	0.33	0.01	0.01	0.01	0.00	0.06	0.57	0.00	0.19	0.02	0.06
Two months of within cq information	0.01	0.64	0.00	0.03	0.00	0.00	0.06	0.75	0.00	0.02	0.03	0.16
Three months of within cq information	0.00	0.33	0.00	0.01	0.00	0.00	0.04	0.03	0.00	0.17	0.06	0.35
<b>CONSENSUS FORECAST</b>												
Zero one month of within cq information	0.02	0.37	0.03	0.91	0.13	0.79	0.84	0.68	0.38	0.15	0.02	0.06
One month of within cq information	0.02	0.70	0.03	0.56	0.04	0.57	0.54	0.85	0.07	0.19	0.02	0.06
Two months of within cq information	0.02	0.70	0.03	0.82	0.04	0.17	0.84	0.84	0.07	0.37	0.02	0.06
Three months of within cq information	0.00	0.33	0.01	0.01	0.00	0.13	0.02	0.24	0.00	0.17	0.02	0.02
<b>MONTHLY AUXILIARY MODELS CONSENSUS</b>												
Zero one month of within cq information	0.02	0.57	0.05	0.31	0.01	0.26	0.28	0.54	0.07	0.17	0.02	0.07
One month of within cq information	0.02	0.33	0.01	0.09	0.00	0.04	0.06	0.43	0.01	0.37	0.00	0.17
Two months of within cq information	0.01	0.64	0.01	0.05	0.00	0.00	0.08	0.54	0.00	0.19	0.00	0.06
Three months of within cq information	0.00	0.17	0.01	0.01	0.00	0.02	0.02	0.09	0.00	0.06	0.06	0.36

Reading: this table reports the P value of the null of independence between forecasts and outcomes.

### Annex 3. How important is consumer confidence?

81. An important feature of the empirical results obtained in this study is that measures of consumer confidence are not amongst the group of variables chosen in the final indicator models. In order to see why this might be, and whether the omission matters, this Annex reports briefly on an experiment in which the out-of-sample forecasting capabilities of an indicator-based equation for real GDP growth in the euro area using consumer confidence are compared to one using business confidence. The conclusion is that, for the euro area at least, the forecasting performance of an equation with business confidence is, in the recent past, statistically superior to one using consumer confidence.

82. Intuitively, this result is not surprising. Between 1985 to 2002, the correlation between real GDP growth (measured by the quarter-on-quarter growth rate) and changes in business confidence is 0.70. The correlation between real GDP growth and changes in consumer confidence is 0.55. The higher correlation with business confidence remains when the correlations are computed either recursively or with a five year rolling window. Recent movements in the respective series are shown in Figure 1. One possible explanation for the weaker link between GDP growth and consumer confidence is that, as can be seen in Figure 2, the level of consumer confidence is closely correlated with changes in the unemployment rate, which tends to lag changes in real GDP growth.

Figure 1

GDP growth, changes in business and in consumer confidence

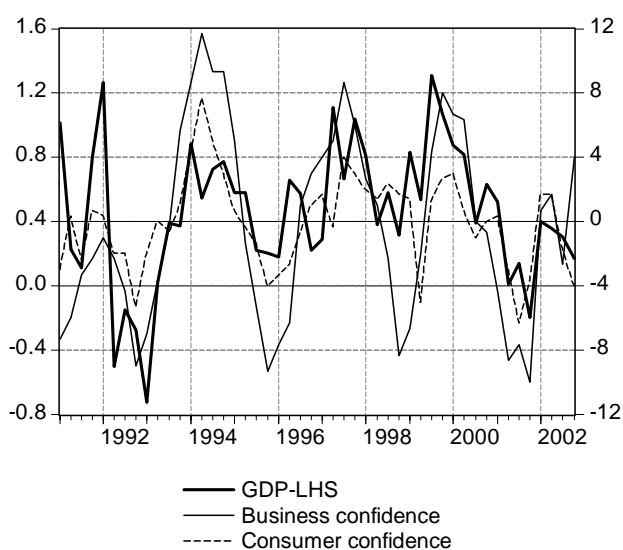
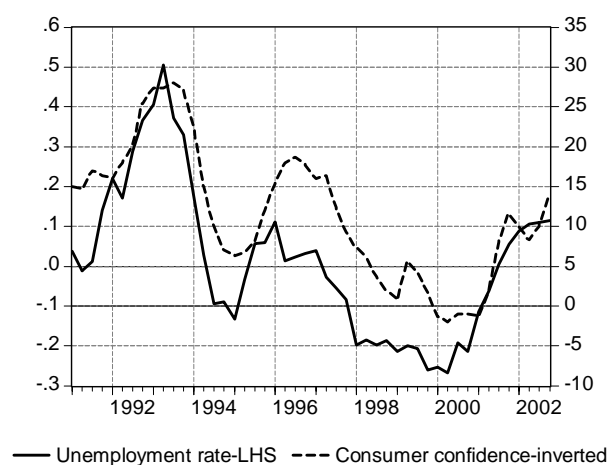


Figure 2

Level of consumer confidence and changes in the unemployment rate



83. The experiment to gauge the usefulness of different survey measures has been carried out through a standard out-of-sample exercise. Three survey indicators are considered – business confidence as measured by the level of domestic order books from the business survey, consumer confidence from the consumer survey and an aggregation of the two derived from a principal component analysis<sup>45</sup>. A separate

45. The aggregated variable is recomputed each quarter (one additional data point is added). The weights of both variables are broadly similar. In order to check the robustness of the conclusions from this experiment, other measures of business intentions were also added. As might be expected, this resulted in less weight being given to consumer confidence, making the principal factor closer to business confidence.

model was estimated for each indicator. In order to replicate a real time forecast exercise as closely as possible, each of the equations was estimated recursively from 1997Q4 to 2002Q1 keeping a fixed start point of 1985Q1. The equation estimated up to 1997Q4 is used to generate a prediction for 1998Q1 and so on. Thus, for each model, 18 out-of-sample forecasts were generated.

84. Within each class of model, two cases are considered. In the first case, the equations have the following format:

$$qoq\_gdpv_t = \alpha + \sum_{i=0}^I survey_{t-i} + \sum_{k=1}^K qoq\_gdpv_{t-k} + \varepsilon_t$$

[A1]

where  $qoq\_gdpv_t$  denotes real gdp growth and  $survey$  represents either business confidence, consumer confidence or an aggregation of both measures.<sup>46</sup> In this equation current quarter values of the surveys are included. This means that an equation estimated until quarter T-1 can be used to produce a conditional forecast for GDP growth in period T using outturn data for the survey value in quarter T.

85. In the second case, the forecasting equation has the format:

$$qoq\_gdpv_t = \alpha + \sum_{i=1}^I survey_{t-i} + \sum_{k=1}^K qoq\_gdpv_{t-k} + \varepsilon_t$$

[A2]

In equation [A2], only lagged values of the surveys are included. This means that an equation estimated up to T-1 generates a non-conditional forecast for quarter T using survey information available only up to quarter T-1.

86. From these out-of-sample forecasts, two summary statistics are examined: the Root Mean Square Forecast Error (RMSFE) and a Forecast Directional Accuracy statistic (FDA), the latter giving the probability of correctly predicting the direction of change in the quarterly rate of real GDP growth. These statistics are reported in Table A8 for both the conditional and non-conditional projections. In terms of the RMSFE (columns 2 and 3), it can be seen that over the period 1998-2002, the errors in the conditional forecasts made using the business confidence model are lower than in the other two models (0.28 percentage points against 0.42 percentage points for consumer confidence or 0.30 percentage points for the aggregation). For the non-conditional forecasts, the forecast errors are similar across models, although the ones from the business confidence model are marginally smaller than those from the other two models. Overall there appear to be no grounds for concluding that consumer confidence indicators regularly contain useful information that would help to improve the predictions of an business confidence based model.

87. This conclusion is reinforced by an examination of the FDA outcomes (columns 4 and 5 of Table A8). These suggest that the business confidence model is much more likely to predict correctly an acceleration (or deceleration) of real GDP growth in both the conditional and non-conditional cases. For the conditional model the rate of success for the business confidence equation is 83 per cent, compared to 55 per cent for the consumer confidence equation. It can also be seen from the principal component model that the inclusion of consumer confidence diminishes the forecast capability of a model with business

---

But, even in this case, the inclusion of consumer confidence resulted in an equation with poorer predictive power than one excluding consumer confidence entirely.

46. The lags of both the endogenous variables and the regressors are selected according to the Schwarz criterion. These lags are automatically computed each time the equation is estimated.

confidence considerably (the rate of success declines from 83 to 55 per cent). For the one-quarter ahead forecasts, there is a clear distinction between the directional accuracy achieved using business confidence alone (67 per cent) and that using consumer confidence, where the directional accuracy is correct only half of the time.

**Table A8. Out-of-sample forecast accuracy (1998-2002)**

	RMSFE (in pp)		FDA (in %)	
	Conditional	Non conditional	Conditional	Non conditional
Business confidence	0.28	0.37	83	67
Consumer confidence	0.42	0.39	55	50
Aggregation	0.33	0.40	50	55

88. Of course, it could be argued that a pure comparison of point estimates does not take into account sampling uncertainty or, in other words, the possibility that the business confidence models provide a lower RMSFE simply due to luck. In order to discriminate among the three competitive models, two formal statistical tests have been carried out using the modified version of the Diebold and Mariano (1995, 1997) for the RMSFE measures (equation [5] in the main text) and the non-parametric test of predictive performance from Pesaran and Timmerman (1992) for the FDA measures (equation [7] in the main text).

89. For the Diebold and Mariano test, the probabilities of rejecting the null hypothesis of equally good forecast performance, expressed in terms of the RMSFE, are shown in Table A9. For the conditional forecasts, the null that the forecast errors generated by the consumer confidence model are not different to those generated by the business confidence model can be rejected at the 6 per cent threshold. In contrast, for the non-conditional one-quarter ahead forecasts it appears impossible to single out one particular model.

**Table A9. P values of the modified Diebold and Mariano test**  
(the reference model is business confidence, 1998-2002)

	RMSFE (in pp)	
	Conditional	Non conditional
Consumer confidence	0.06	0.33
Aggregation	0.61	0.59

90. The test proposed by Pesaran and Timmerman is a distribution free procedure for testing the accuracy of forecasts when the focus of the analysis is on the correct prediction of the direction of change of the variable under consideration. This test is based on the usual  $\chi^2$  test of independence of two variables in a single contingency table. The results are displayed in Table A10. For the conditional forecasts, the null that changes in real GDP and those forecast by the business confidence model are independent is strongly rejected by the data. This is not the case for the forecasts from the other two models. For the non-conditional forecasts, the null of independence is rejected at the 10 per cent level for the business confidence model but not for the other two models. Overall, the Pesaran and Timmerman test results confirm the impression from Table A8 that models using measures derived from industrial surveys are more likely to be able to forecast the direction of change of GDP growth correctly than are models using consumer confidence.

**Table A10. P values of the Pesaran and Timmerman test**

	RMSFE (in pp)	
	Conditional	Non conditional
Business confidence	0.00	0.08
Consumer confidence	0.15	0.20
Aggregation	0.18	0.28



**WORKING PAPERS**

The full series of Economics Department Working Papers can be consulted at [www.oecd.org/eco/Working\\_Papers/](http://www.oecd.org/eco/Working_Papers/)

363. *Pos-Crisis Change in Banking and Corporate Landscapes – the Case of Thailand*  
(July 2003) Margit Molnar
362. *Post-Crisis Changes in Banking and Corporate Landscapes in Dynamic Asia*  
(June 2003) Margit Molnar
361. *After The Telecommunications Bubble*  
(June 2003) by Patrick Lenain and Sam Paltridge
360. *Controlling Public Spending in Iceland*  
(June 2003) Hannes Suppanz
359. *Policies and International Integration: Influences on Trade and Foreign Direct Investment*  
(June 2003) Giuseppe Nicoletti, Steve Golub, Dana Hajkova, Daniel Mirza, Kwang-Yeol Yoo
358. *Enhancing the Effectiveness of Public Spending in Finland*  
(June 2003) Philip Hemmings, David Turner and Seija Parviainen
357. *Measures of Restrictions on Inward Foreign Direct Investment for OECD Countries*  
(May 2003) Stephen S. Golub
356. *Tax Incentives and House Price Volatility in the Euro Area: Theory and Evidence*  
(May 2003) Paul van den Noord
355. *Structural Policies and Growth: A Non-technical Overview*  
(May 2003) Alain de Serres
354. *Tax Reform in Belgium*  
(May 2003) David Carey
353. *Macroeconomic Policy and Economic Performance*  
(April 2003) Pedro de Lima, Alain de Serres and Mike Kennedy
352. *Regulation and Investment*  
(March 2003) Alberto Alesina, Silvia Ardagna, Giuseppe Nicoletti and Fabio Schiantarelli
351. *Discretionary Fiscal Policy and Elections: The Experience of the Early Years of EMU*  
(March 2003) Marco Buti and Paul van den Noord
350. *The US Health System: An Assessment and Prospective Directions for Reform*  
(February 2003) Elizabeth Docteur, Hannes Suppanz and Jaejoon Woo
349. *The Effectiveness of Public Expenditure in Portugal*  
(February 2003) Chiara Bronchi
348. *Comparative Analysis of Firm Demographics and Survival: Micro-Level Evidence for the OECD Countries*  
(February 2003) Eric Bartelsman, Stefano Scarpetta and Fabiano Schivardi
347. *Regulation, Productivity and Growth: OECD Evidence*  
(January 2003) Giuseppe Nicoletti and Stefano Scarpetta
346. *Public Expenditure Management in Poland*  
(December 2002) Andrew Burns and Kwang-Yeol Yoo

345. *Enhancing the Effectiveness of Public Expenditure in Sweden*  
(December 2002) Deborah Roseveare
344. *The Decline in Private Saving Rates in the 1990s in OECD Countries: How Much Can Be Explained by Non-Wealth Determinants*  
(November 2002) Alain de Serres and Florian Pelgrin
343. *Enhancing the Effectiveness of Public Expenditure in Norway*  
(October 2002) Isabelle Joumard and Wim Suyker
342. *Productivity and Convergence in a Panel of OECD Countries: Do Regulations and Institutions Matter?*  
(September 2002) Stefano Scarpetta and Thierry Tresselt
341. *Managing Public Expenditure: The UK Approach*  
(August 2002) Paul van den Noord
340. *The Brazilian Pension System: Recent Reforms and Challenges Ahead*  
(August 2002) Marcos Bonturi
339. *Challenges in the Mexican Financial Sector*  
(August 2002) Marcos Bonturi
338. *Coping with Population Ageing in Hungary*  
(August 2002) Andrew Burns and Jaromir Cekota
337. *Next Steps for Public Spending in New Zealand: The Pursuit of Effectiveness*  
(July 2002) Dave Rae
336. *Strengthening the Management of Public Spending in Hungary*  
(July 2002) Jaromir Cekota, Rauf Gonenc and Kwang-Yeol Yoo
335. *Automatic Stabilisers and Market Flexibility in EMU: Is There a Trade-Off?*  
(July 2002) Marco Buti, Carlos Martinez-Mongay, Khalid Sekkat and Paul van den Noord
334. *The Economic Consequences of Terrorism*  
(July 2002) Patrick Lenain, Marcos Bonturi and Vincent Koen
333. *Investment in human capital through post-compulsory education and training: Selected efficiency and equity aspects*  
(July 2002) Sveinbjörn Blöndal, Simon Field and Nathalie Girouard
332. *Enhancing the Effectiveness of Public Spending in Switzerland*  
(June 2002) Isabelle Joumard and Claude Giorno
331. *Competition and Efficiency in Publicly Funded Services*  
(June 2002) Jens Lundsgaard
330. *Policy Pre-Commitment and Institutional Design: A Synthetic Indicator Applied to Currency Boards*  
(May 2002) Marie-Thérèse Camilleri Gilson
329. *The Role of Policy and Institutions for Productivity and Firm Dynamics: Evidence from Micro and Industry Data*  
(April 2002) Stefano Scarpetta, Philip Hemmings, Thierry Tresselt and Jaejoon Woo
328. *Improving the Efficiency and Sustainability of Public Expenditure in the Czech Republic*  
(April 2002) Andrew Burns and Kwang-Yeol Yoo