Leading indicators in a globalised world

DRAFT VERSION

This version: 26 November 2008

Ferdinand Fichtner[1]

ferdinand.fichtner@ecb.int

Rasmus Rüffer

rasmus.rueffer@ecb.int

Bernd Schnatz

bernd.schnatz@ecb.int

Abstract

Using OECD composite leading indicators (CLI), we assess empirically whether the ability of the country-specific CLIs to predict economic activity has diminished in recent years due to rapid advances in globalisation. Overall, we find strong evidence that the CLI encompasses very useful information for forecasting industrial production, particularly over horizon of four to eight months ahead. The evidence is particularly strong when taking cointegration relationships into account. At the same time, we find indications that the forecast accuracy has indeed declined over time for several countries. Augmenting the country-specific CLI with a leading indicator of the external environment and employing forecast combination techniques further improves the forecast performance for several economies.

JEL codes: C53, E32, E37

[1] All authors: European Central Bank. The views expressed are solely our own and do not necessarily reflect those of the European Central Bank.

1 Introduction

Monetary policy decisions affect the economy only with long and varying lags. It is therefore crucial to have an educated judgement about the economic conditions prevailing at that time. Leading indicators constitute an important tool in applied business cycle analysis to form such a judgement.[2] A main potential shortcoming of country-specific leading indicators is, however, that their components are commonly based mainly on *domestic* variables. The ability of such leading indicators to predict economic activity might have diminished due to the rapid advances in globalisation, as reflected in the deepening of international financial and trade linkages. Notwithstanding the significant structural changes that have taken place at the global level over recent years, to our knowledge, the possible implications for the leading indicator properties have not been addressed in a systematic fashion as yet.

This paper aims at filling this gap. It analyses (1) whether the empirical relationships established in the past are still appropriate for forecasting purposes, and (2) whether augmenting the available leading indicators with information on the outlook for important trading partners can improve the forecasting performance. This paper looks into this issue based on the OECD composite leading indicators (CLI) across 12 countries and uses, in line with common practice, industrial production as the reference series. The OECD CLI has the advantage (1) to be widely monitored by practitioners (2) to be constructed on the basis of criteria which are consistent with those proposed by Marcellino (2006) and (3) to be available for a wide variety of countries, on a monthly basis and over a long time span (see also Camba-Mendez et al., 1999).

The objective of the paper is two-fold. Firstly, we assess the properties of the OECD CLI and check whether its accuracy has declined over time as the process of globalisation has evolved. In line with standard practice, the assessment is made on the basis of mean absolute forecast errors calculated – as suggested in Marcellino, Stock and Watson (2005) – based on iterated multi-step forecasts up to one year ahead. The analysis is based on a data sample between January 1975 and April 2008, using the first fifteen years for estimation purposes and the remainder of the sample for the recursive estimation and forecasting using only information available at the time the forecast is carried out. We find clear evidence that the forecast properties of the commonly used CLI are significantly better than naïve autoregressive forecast benchmarks. We also document strong evidence that taking cointegration relationships between the CLI and industrial production into account is crucial as it further improves the accuracy of the forecasts of industrial production. Regarding the evolution of the forecast accuracy over time, for some countries, we find indications that it indeed declined over time, while for others it is more difficult to discern a clear pattern.

Secondly, we analyse whether adding information on the international business cycle improves the forecast accuracy of the OECD indicator. To this end, we augment the basic forecast equation, which includes the country-specific leading indicators (and lagged values of domestic industrial production) with a leading indicator for the external environment, con-

[2] The academic interest in empirical analysis of leading indicators has been revived almost two decades ago when Stock and Watson (1989) improvements in leading-indicator theory. This stimulated a vivid debate on the different aspects of leading indicators, ranging from the choice of the leading indicators and various prediction frameworks to the presentation of different evaluation benchmarks and strengthened the capacity to assess these indicators (see Marcellino (2005) for a comprehensive review).

structed from the CLI of the other OECD countries. The results appear quite promising. In spite of the differences in the results across models and countries, we find that the inclusion of external leading indicators and using forecast combination techniques improves the forecast performance quite substantially for most economies.

2 Data and Stylised Facts

The OECD CLI is published on a monthly basis with a lag of two months. This means that, for example, in December, the CLI for October is published. Accordingly, the CLI incorporates elements of both a coincident and a leading indicator for economic activity, but even its potential coincident indicator properties are useful because at the time of the publication of the December CLI, actual business cycle data is available only up to September for most OECD countries.

Compared to a single indicator variable, composite indicators eliminate the noise of individual variables and reduce the risk of false signals. More specifically, the OECD CLI is constructed on the basis of criteria which are broadly consistent with those proposed by Marcellino (2006). They are based on (seasonally-adjusted) time series – available for most OECD countries over a long time span – with potential leading indicator properties, which are aggregated into a composite indicator. The component series of CLI cover a wide range of short-term indicators (see annex), which are selected for each country according to a number of criteria:

Firstly, there must be an economic reason for a leading relationship with the reference series. The components may include variables which can cause business cycle fluctuations (e.g. short-term interest rates), express market expectations, measure economic activity at an early stage (e.g. housing starts) or adjust quickly to changes in economic activity (e.g. overtime work). *Secondly*, the cycles of these series should lead those (and only those) of the reference series. *Thirdly*, data quality needs to be ensured in terms of availability and timeliness and revisions should be small.

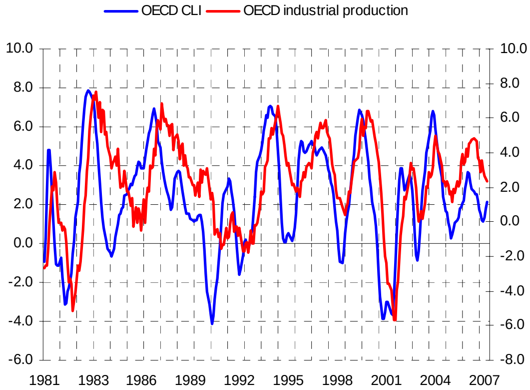
The composition and number of series (five to eleven) included in the country-specific CLI varies across countries. These variables are aggregated in a detrended and scaled form into a composite indicator (following some transformations) using equal weights.[3] While Camba-Mendez et al. (1999) as well as Emerson and Hendry (1996) have criticised such a weighting scheme as suboptimal, it is in line with the literature on forecast pooling which suggests that equal weights work pretty well in practice (Stock and Watson, 2003).

The final set of component series is selected in order to maximise the performance of the CLI in terms of detecting economic cycles. While the CLI was originally designed to detect early signals for turning points (Nilsson and Guidetti, 2008), these indicators are also used to evaluate the cycle as a whole.

[3] For details, see: http://www.oecd.org/dataoecd/4/33/15994428.pdf.

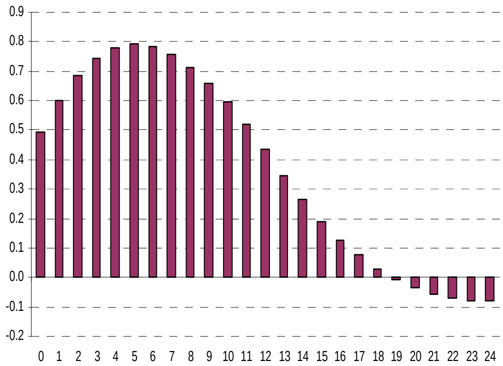
Chart 1: Leading indicator and industrial production in OECD countries

(monthly data, annual rates of growth)



Source: OECD, MEI

Chart 2: Cross-correlation between OECD CLI and industrial production annual growth



We analyse the relationship between the CLI and industrial production for 11 industrialised countries (Canada, Denmark, the United Kingdom, Japan, Sweden, the United States, Germany, Spain, France, Greece and Italy), eight of which are members of the European Union.

In the first step of the analysis, we also include an OECD aggregate indicator (OTO). Industrial production data (covering all sectors excluding construction) is available on a monthly frequency for a broad range of countries. Although these data have the drawback to represent only a small and declining fraction of economic activity, many service activities, such as transport, are likely to be directly linked to industrial activity.

Visual inspection suggests that there is indeed a close correlation between the OECD aggregate CLI and the annual rate of change in industrial production with some time shift. The CLI anticipates movements in industrial production growth (see Chart 1). Cross correlations reveal that the leading indicator properties are most pronounced at a lead time of around half a year at the OECD aggregate and diminishes quickly thereafter, which implies that the leading indicator is likely to become less precise as the lead periods exceed half a year (see Chart 2). The same pattern over time is found for most OECD countries. Looking at these countries in more detail suggest that the correlation between the CLI and industrial production is strongest for the United States, Canada, Japan and Germany, while it appears to be weaker for smaller open economies such Portugal, Greece and Ireland, which might be more affected by the global environment.

3 Econometric analysis

A comprehensive empirical evaluation of the properties of the OECD CLI has many dimensions, which are addressed in turn.

Firstly, the question arises of whether the analysis should be carried out with respect to an in-sample or with respect to an out-of-sample forecast. In this context, Carriero and Marcellino (2007) stress that it is always possible to explain past economic growth reasonably well when a set of parameters is carefully chosen, but that there is no reason to expect that such equations will necessarily be good forecasting tools. Accordingly, the aim must be to describe and evaluate a forecasting strategy rather than simply to find an equation which happens to fit the data.[4] The forecasting performance of the respective models is compared based on the mean absolute forecast error of the forecast h periods ahead (up to twelve months).

Secondly, within the wide range of methodologies that can be chosen to assess the link between the leading indicators and the business cycle, we focus on linear models, the most widely followed avenue in the literature.

Specifically, we use the unrestricted VAR model of the form

$$\Delta z_t = \sum_{j \in J} A_j \Delta z_{t-j} + \varepsilon_t \ , \qquad (1)$$

where Δz_t is a vector of year-on-year differenced endogenous variables to be described in detail below, ε_t is a vector of white-noise innovations, and the A_j are matrices of coefficients (including the intercept) to be estimated. J denotes the set of lags included in the estimation. Lag selection is described in detail below.

Based on this model, the 1-step ahead forecast of Δz_t is the expectation

$$\mathrm{E}_t \left\{ \Delta z_{t+1} \right\} = \sum_{j \in J} A_j \Delta z_{t+1-j} \qquad (2)$$

conditional on past and present information available in period t. Simple forward-iteration of this equation then allows to derive arbitrary h-step ahead forecasts according to

$$\mathrm{E}_t \left\{ \Delta z_{t+h} \right\} = \sum_{j \in J \cap \{1 \ldots h-1\}} A_j \, \mathrm{E}_t \left\{ \Delta z_{t+h-j} \right\} + \sum_{j = J \cap \{h \ldots p\}} A_j \Delta z_{t+h-j} \ . \qquad (3)$$

Forecasting beyond the next period $(h > 1)$ requires a choice between employing iterated multi-step forecasts or direct forecasts. In theory, iterated forecasts as discussed above are more efficient if correctly specified but direct forecasts are more robust to model misspecification (see Bhansali 2002 for an overview). We therefore additionally employ direct forecasts of Δz_t. This obviously requires a specific model to be estimated for each forecast horizon. The h-step ahead forecast of Δz_t based on this model is then

$$\mathrm{E}_t \left\{ \Delta z_{t+h} \right\} = \sum_{j \in J_h} A_{h,j} \Delta z_{t-j} \ , \qquad (4)$$

where, notably, the matrices $A_{h,j}$ and the set of included lags J_h now vary with the forecast horizon h.

Given that the leading indicator variables and industrial production are integrated time series in levels (see the appendix for the relevant ADF tests), we focus on log-differenced representations of the respective series. Since this implies an information loss for the estimation,[5] we also consider cointegration frameworks of the following form for our forecasting models:[6]

---

[4] Note that our results confirm the latter notion. The results with respect to our out-of-sample forecasts presented in this paper are crucially different compared to the results with respect to in-sample forecasts we additionally worked out.

[5] Clements and Hendry (1999) and Emerson and Hendry (1996) provide arguments applicable in the forecasting context.

[6] The appendix presents Johansen cointegration tests for the different combinations of variables. Note that the CLI and industrial production are cointegrated by construction however, since the trend-restored CLI underlying the

$$\Delta^1 z_t = \Pi \cdot z_{t-1} + \sum_{j \in J} A_j \Delta^1 z_{t-j} + \varepsilon_t \ , \tag{5}$$

where Π can be decomposed as $\Pi = \alpha\beta'$. α are the factor loadings and β the long-term cointegration coefficients. Note that $\Delta^1 z_t$ indicates the first difference, z_t the level of the log of endogenous variables. The forecast is carried out again iteratively based on the equation

$$\begin{aligned} \mathrm{E}_t\left\{z_{t+1}\right\} &= z_t + \mathrm{E}_t\left\{\Delta^1 z_{t+1}\right\} \\ &= \left(I + \Pi\right) \cdot z_t + \sum_{j \in J} A_j \Delta^1 z_{t+1-j} . \end{aligned} \tag{6}$$

In order to assess statistically the difference in forecast accuracy between two models (*F*, *B*), we employ the test statistic proposed by Diebold and Mariano (1995) and account for the small sample correction suggested by Harvey et al. (1997). The statistic is based on the mean of the difference of the absolute forecast errors associated with the two forecasts being compared,

$$d_{h,t} = \left|\mathrm{E}_t^F\left\{\Delta z_{t+h}\right\} - \Delta z_t\right| - \left|\mathrm{E}_t^B\left\{\Delta z_{t+h}\right\} - \Delta z_t\right| \tag{7}$$

i. e. the test statistic is calculated for a particular forecast horizon h as

$$\bar{d}_h = \frac{1}{T} \cdot \sum_{t=1}^{T} d_{h,t} \tag{8}$$

where $\mathrm{E}_t^F\left\{z_{t+h}\right\}$ is the h-step ahead forecast of the variable z as derived at time t by means of one of the models outlined above and $\mathrm{E}_t^B\left\{z_{t+h}\right\}$ is the equivalent forecast of an alternative benchmark model. Note that T equals the number of available forecasts, i. e. the length of the full available sample reduced by h and the data used in the first of our recursive estimations. Consequently, $t = 1$ refers to the primary forecasting point.

Dividing the statistic $\bar{d}_h$ by its standard deviation produces a standard normally distributed variable, i. e. the basic Diebold-Mariano statistic of the respective model comparison for a specific forecast horizon h is

$$S_h^{DM} = \frac{\bar{d}_h}{\sqrt{\mathrm{var}\left(\bar{d}_h\right)}} \xrightarrow{d} \mathrm{N}(0,1) . \tag{9}$$

As shown by Diebold and Mariano (1995), possible autocorrelation arising through overlapping forecast windows needs to be taken into account up to an order of $h-1$ when estimating the variance $\mathrm{var}\left(\bar{d}_h\right)$ according to the asymptotic equality

$$\mathrm{var}\left(\bar{d}_h\right) \approx \frac{1}{T} \cdot \left(\hat{\gamma}_0 + 2 \cdot \sum_{i=1}^{h-1} \hat{\gamma}_i\right), \tag{10}$$

where the k^{th} autocovariance of $d_{h,t}$ can be estimated by

$$\hat{\gamma}_k = \frac{1}{T} \cdot \sum_{t=k+1}^{T} \left(d_{h,t} - \bar{d}_h\right) \cdot \left(d_{h,t-k} - \bar{d}_h\right) . \tag{11}$$

---

present analysis is calculated by applying the trend of industrial production index to the amplitude-adjusted CLI. See the OECD's website, http://www.oecd.org/dataoecd/4/33/15994428.pdf for details.

The small sample adjustment suggested by Harvey et al. (1997) multiplies the Diebold-Mariano statistics with a factor which depends on the number of available forecasts T and the forecast horizon h,

$$S_h^{DM,adj} = \left(\frac{T+1-2h+T^{-1}h(h-1)}{T} \right) \cdot S_h^{DM} .$$

This statistic can be tested against the null hypothesis of equality of the forecast performance of the respective models.

4 Empirical strategy and results

4.1 Leading indicator properties of the domestic CLI

The analysis is based on monthly data ranging from January 1975 to April 2008. The first fifteen years of data, from January 1975 to December 1989 are exclusively used for estimation purposes. The second part, from January 1990 to April 2008 is used for conducting recursive rolling-window estimation and forecasting over horizons of up to one year ahead.[7,8]

Lag selection is done for each model and for each forecasting point in time. Since the common information criteria regularly indicate very high lag lengths for the models considered in the present analysis, we choose the lags to be included in the set J based on standard lag exclusion tests. Specifically, we include the 8 lags $j < 24$ in J, which have the highest Wald statistic for testing the joint significance of all j-th lagged endogenous variables in the system.[9]

For the unrestricted VAR models as described by (3) and the direct forecast (4), all series are log-differenced to ensure stationarity prior to estimation.[10] For the error correction approach (6), log-level data is used for estimation and forecasting. All model comparisons are, however, based on log-differences, i. e. forecasts derived from the VEC models are differenced to ensure comparability with the other approaches. In the first step, we follow the traditional approach by assessing the information content of the domestic CLI for domestic industrial production. As such, we compare forecasting models including the national CLI with a simple univariate autoregression (UAR) of industrial production.

---

[7] The complete sample includes 401 observations. The rolling estimation is based on 180 observations. In terms of the above notation, $t=1$ refers to 1990m1. For an $h=6$-step ahead forecast, the total number of available forecasts amounts to $T = 401-180-6 = 215$. We additionally calculated forecasts based on a recursively growing sample. This does not substantially alter our results.

[8] While the employed tests are out-of-sample, the latest data vintage instead of "real time" data was used throughout. As emphasised by Diebold und Rudebusch (1991), such a procedure could yield biased results. However, this issue should be less problematic in the comparative analysis since all methods can be expected to be equally advantaged. This means that we test how well our approach does in a world of random shocks and possible structural changes. We do not examine the separate issue of how well it deals with inaccuracies in earlier vintages of data.

[9] Both selecting precisely 8 lags and setting the highest lag to 24 is to some extent arbitrary. We found, however, that results are robust to changes in these parameters. We also experimented with other selection criteria such as standard information criteria and lag selection based on a certain threshold for the probability indicated by the lag exclusion test and also found our results to be robust to these modifications.

[10] [We employ year-on-year differencing to reduce volatility of the series. Obviously, differencing reduces the length of the available sample by another 12 periods.]

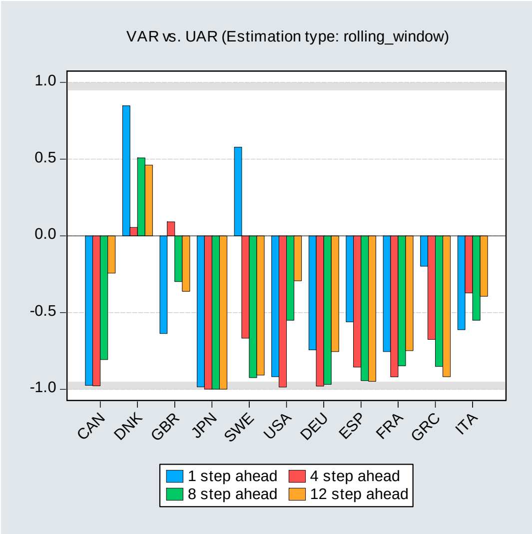


Chart 3: Forecasting performance of the domestic CLI.

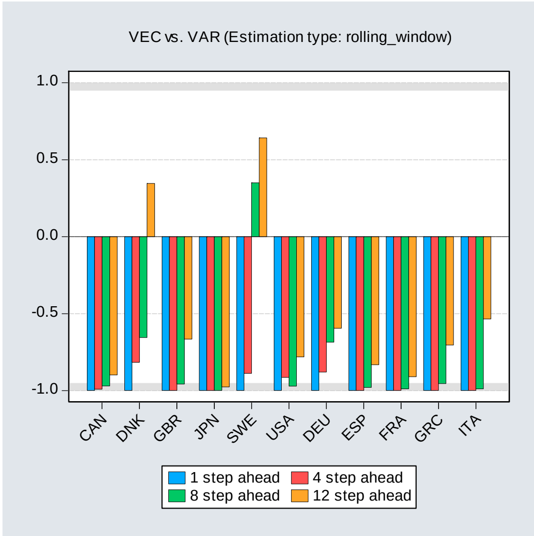


Chart 4: Forecasting performance of an error-correction specification.

In terms of the models previously discussed, this amounts to the following specifications: The simple UAR is clearly a special case of model (3) with the set of endogenous variables being limited to the index of industrial production, $z_t = \left[\mathrm{IP}_t^k\right]'$. We compare this benchmark model with the three specifications outlined above and include the national leading indicator in the analysis, i. e. $z_t = \left[\mathrm{IP}_t^k \ \mathrm{CLI}_t^k\right]'$. Note that $k \in$ {OTO,CAN,DNK,GBR,JPN,SWE,USA,DEU, ESP,FRA,GRC,ITA} refers to the respective country analysed in the specific exercise.

Chart 3, comparing the forecasting performance of the VAR model (3) with the univariate autoregressive model UAR, indeed suggests that the domestic CLI encompasses useful information about the future evolution of industrial production compared with the naïve benchmark model. The graph shows the probability that the VAR model provides us with a better forecast than the benchmark UAR model over selected forecast horizons $h \in \{1,4,8,12\}$. Note that the displayed p-value inherits the sign of the corresponding Diebold-Mariano statistic. As such, negative values signify a better forecast performance of the CLI-based model than the respective benchmark model. The closer the value is to -1 the higher the probability that the CLI-based VAR model outperforms the UAR model.

The inclusion of the national leading indicator allows for a substantial improvement in the performance of our industrial production forecast. For most economies included in our sample and for various forecasting horizons we observe substantial and mostly significant improvements of the VAR model vis-à-vis the UAR model. Some interesting patterns emerge: First, the major gains from the inclusion of the CLI in the forecasting model can be expected in the medium forecast horizon, i.e. for h between 4 and 8 months. This is of course in line with our observation in section 2 that the closest cross-correlation between the CLI and the reference series is found for this lag order. Second, forecast improvements appear to be particularly pronounced for the larger economies in our sample. However, with the exception of Denmark, forecast performance for all countries can be substantially increased through the introduction of leading indicators.

Accounting for cointegration between the CLI and industrial production seems to further improve the overall results. Chart 4 displays the gains in forecasting performance when compar-

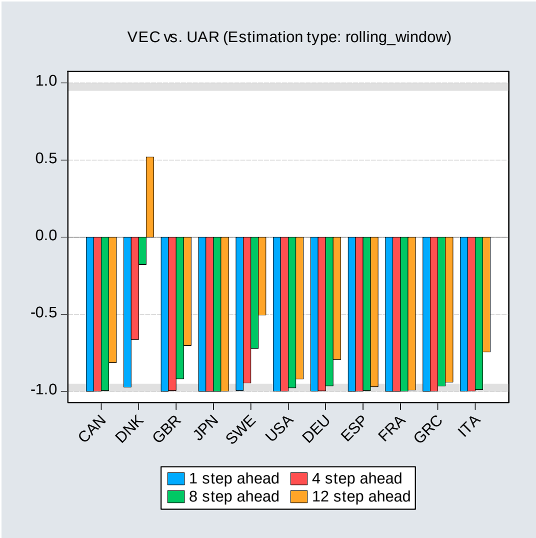


Chart 5: Forecasting performance of the domestic CLI in the VEC model.

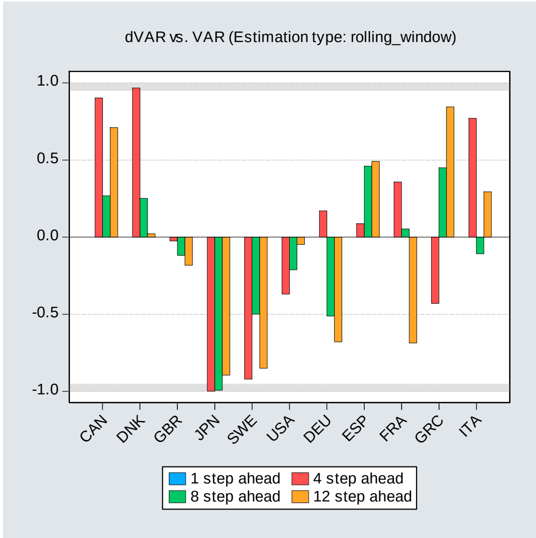


Chart 6: Forecasting performance of the direct forecast.

ing the VEC model as defined in equation (6), $z_t = \left[\mathrm{IP}_t^k \ \mathrm{CLI}_t^k\right]'$, with the VAR model discussed in the previous paragraphs.

The results are unequivocal: Accounting for cointegration between CLI and industrial production improves the results of our forecasting exercise for all economies and effectively over all forecast horizons. Improvements are particularly pronounced for the shorter forecast horizon, where the VAR model was found to be less clearly outperforming the UAR model than in the medium run. Note that the improvements are mostly found to be significant at the 95% or even 99%-level. Given this result, it does not come as a surprise that using the cointegration information contained in the CLI data allows us to beat the univariate autoregressive forecast with a very high degree of confidence (see Chart 5), which indicates an almost perfect dominance of the error correction approach over the naïve univariate autoregression.

To conclude the assessment of the different model specifications, consider Chart 6 for a comparison of the direct forecast model dVAR as defined in equation (4), $z_t = \left[\mathrm{IP}_t^k \ \mathrm{CLI}_t^k\right]'$, with the standard iterative VAR model. Our results indicate no clear dominance of one model specification over the other. Depending on the forecast horizon and the country under observation, the p-value associated with the Diebold-Mariano statistic of the comparison of dVAR with VAR takes values in the full range between -1 and $+1$. [[While forecast performance is, by definition, the same for the 1-step ahead exercises, as a general pattern differences between the dVAR and the VAR model appear more pronounced over the short horizon (4 months), while dying out in the long run.[11]]]

Summing up, the previous analysis suggests that the CLI encompasses useful information about the future evolution of industrial production, particularly over horizons of around half a year. Thereby, it does not systematically matter whether the forecast is calculated iteratively from a VAR representation or whether it is directly derived from a horizon-specific model. A clear improvement can be expected from taking cointegration in the vector autoregressive

[11] The (counter-intuitive) convergence in performance for longer horizons might be related to the fact that long-run forecasts are so poor that the choice of the model does not matter at all.

framework into account. Accordingly, the analysis below relies on cointegration models. Furthermore, given the mixed picture obtained with respect to direct forecasts, we focus on iterative VAR models in view of their wide acceptance in applied forecasting.

4.2 The temporal dimension: Forecasting performance of CLI models over time

If the process of globalisation had an adverse impact on the reliability of the domestic CLI for anticipating developments in economic activity, the forecasting properties should have deteriorated over time. To assess this we compute for each sample period the absolute forecast error averaged across 1 to 12 period forecast horizons. In more detail, forecast errors are averaged by forecasting period, providing us with information about the quality of *all* forecasts calculated in one particular period. Since forecast errors are considerably larger for longer forecast horizons, we normalise the errors with their respective standard deviation in order to avoid that the long-run forecasts dominate the averages (see Chart 7).

The charts do not allow drawing very firm conclusions. Evidently, there is no clear pattern of the development of the forecast errors of the model. While there are some economies, where a substantial and significant improvement of the forecast over time can be observed (e.g. Sweden or Spain), for most economies a clear trend cannot be identified. Overall, it seems that the evolution of the forecast errors seem to be dominated by idiosyncratic shocks and unanticipated economic events. In the United States, for instance, the trend of the forecast error over time seems to be strongly affected by the adverse impact of the September 11 terrorist attacks on confidence indicators and the burst of the new economy stock market bubble at the beginning of the decade. As a result, the forecasting performance of the model appears to worsen over time. On the other hand, the decline of forecast errors over the sample period in some countries might reflect a shift towards more stability-oriented economic policies which contributed to a decline in output volatility across countries and provided a better environment for forecasting future economic activity in general.[12]

To address the impact of a less volatile global environment over time in more detail, we compare the difference between the average absolute forecast error of the VEC model and the average absolute forecast error of the UAR model over time. In Chart 8, negative values indicate a larger forecasting error of the UAR compared to the VEC model for forecasts calculated in the respective period.

[12] See e.g. Blanchard and Simon (2001) or, from a global perspective, Stock and Watson (2005) for this “great moderation” phenomenon.

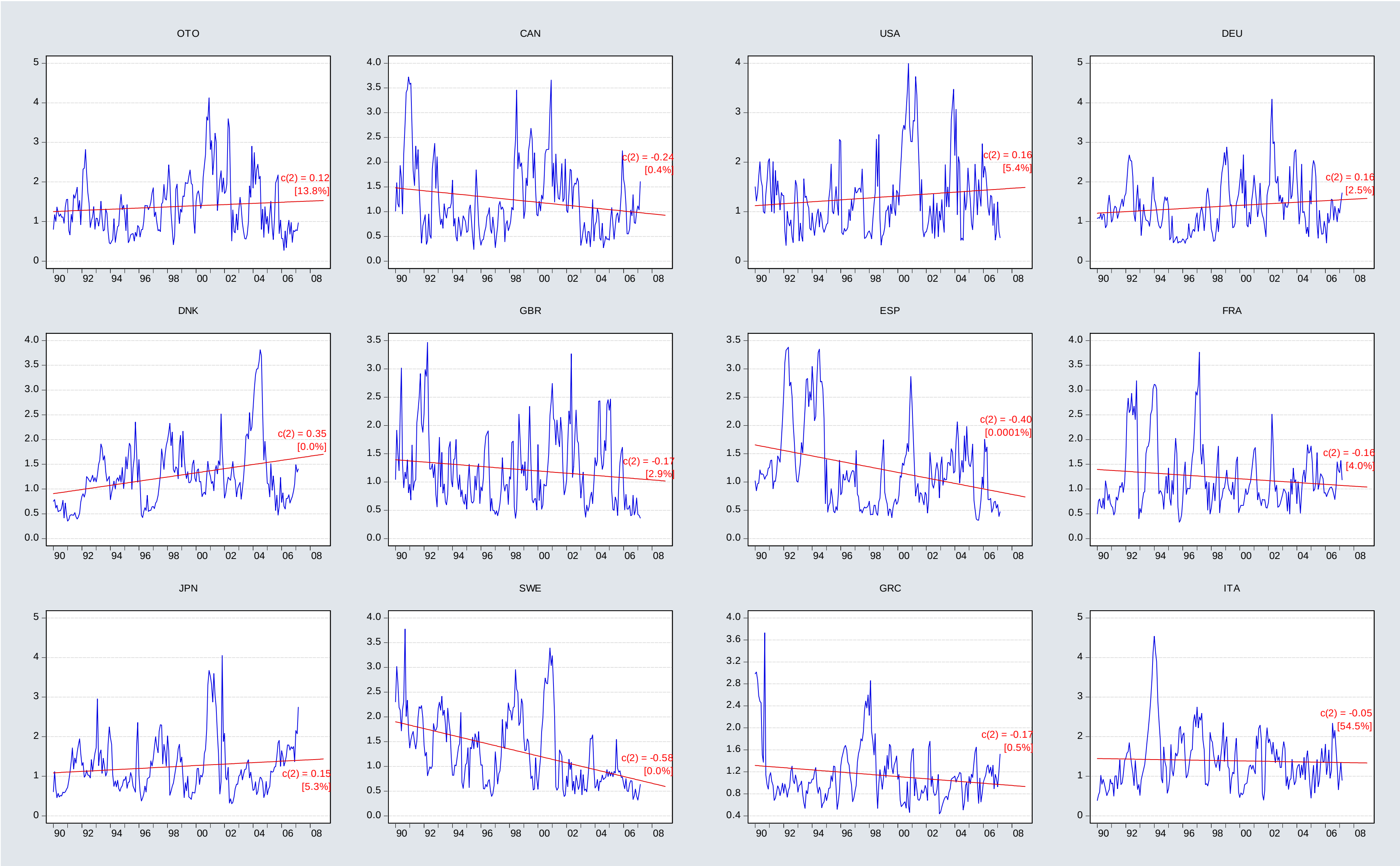


Chart 7: Forecasting performance of the VEC model over time.

Note: Normalized absolute forecast errors averaged by forecasting period. $c(2)$ is the linear trend coefficient, numbers in brackets are p-values of $c(2) = 0$.

The graphs (again) clearly confirm the superiority of the VEC model over the UAR model for most economies and for most of the sample period under observation. This is reflected in an on average negative difference between the forecast error of the VEC model and the forecast error of the UAR. There is, however, a rather substantial change of this pattern over time. For many economies in our sample, the differential has a (significantly) positive trend, i.e. the relative forecasting performance of the VEC model decreases over time. For most of the countries included (the most obvious cases include Denmark and Spain) we find that there is a clear tendency of substantial improvement of the UAR forecast vis-à-vis the CLI-extended VEC model – for Denmark, the US, Germany, Spain and Italy, the naïve univariate autoregression appears to even have surpassed the VEC model for the more recent period.

Overall, therefore, the increase in the forecasting performance of the CLI-augmented model recorded before may be misleading, since it does not document an increasing performance of the respective model but seems simply to be related to a decline in output volatility which contributed to a more stable environment, which is more favourable for forecasting in general. Instead, we find that the *relative* forecasting performance of the CLI-augmented models has actually decreased in previous years for some economies, possibly reflecting a decreasing forecast performance in times of ongoing globalisation. This raises the question whether augmenting the CLI with an international indicator improves the forecast performance over time.

4.3 The external dimension: Can international leading indicators improve domestic forecasts?

We account for country-specific trade patterns and include a (country-specific) external leading indicator ($\mathrm{CLI}^{\mathrm{EXT},k}$), calculated as a bilateral trade-weighted average of the CLIs of the trade partners included in the sample.[13] This external leading indicator is employed in an error correction model similar to the VEC model discussed above, i.e. we include this external leading indicator in the vector $z = \left[\mathrm{IP}^k \; \mathrm{CLI}^k \; \mathrm{CLI}^{\mathrm{EXT},k} \right]$ and estimate the model according to equation (5).[14]

---

[13] Trade weights are based on the import side of the trade matrix for the euro effective exchange rate, which is based on weights for the period 1999-2001.

[14] Alternatively, we use other methods to introduce international leading indicator data in our forecast. Specifically, we employ other aggregation schemes and principal component techniques to derive external leading indicator indices. Additionally, we estimate models including the US and the OECD Total CLI next to the domestic CLI. The results are not systematically different from the results presented here.

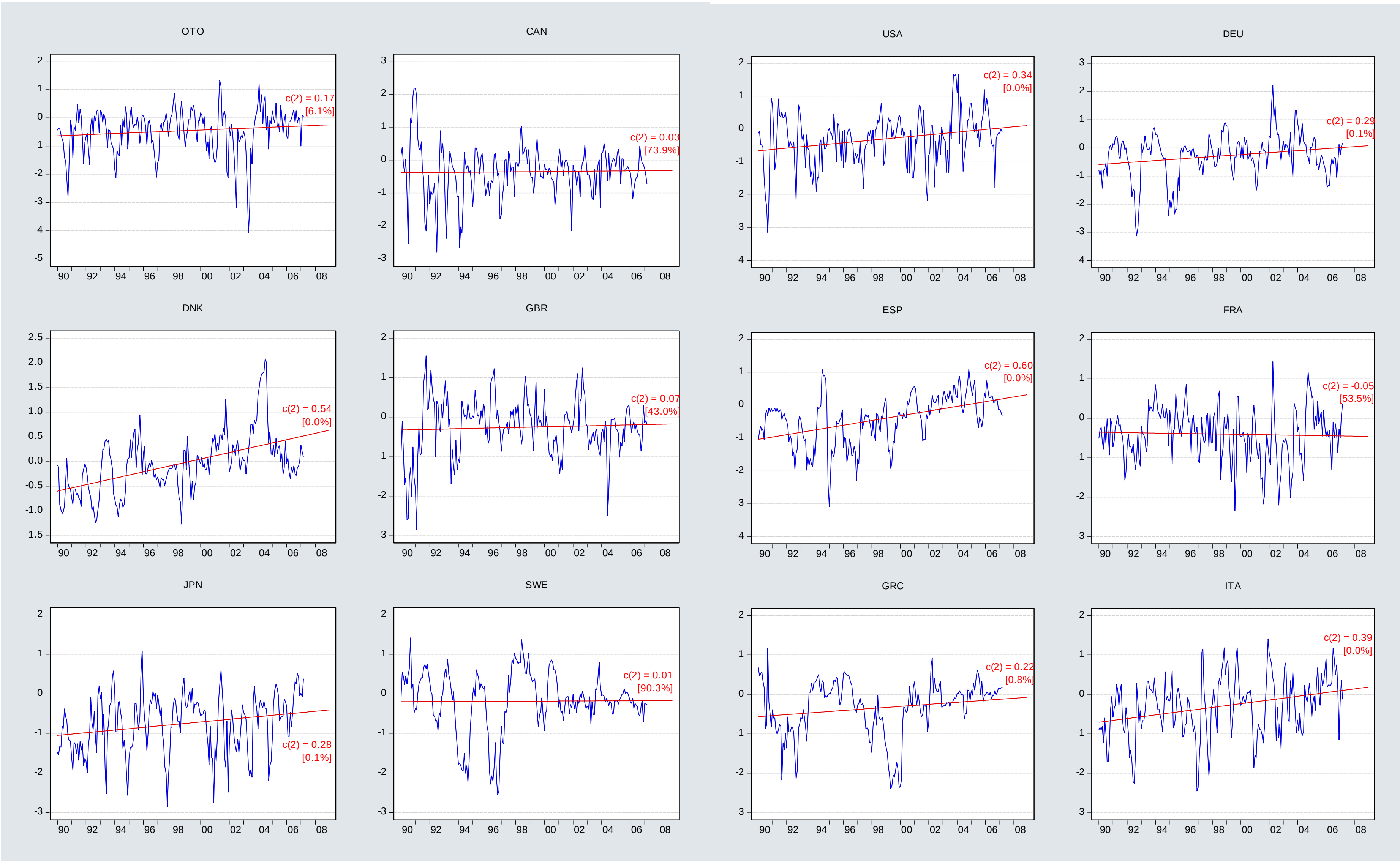


Chart 8: Forecasting performance of the VEC model relative to UAR over time.

Note: Normalized difference of absolute forecast errors averaged by forecasting period. $c(2)$ is the linear trend coefficient, numbers in brackets are p-values of $c(2) = 0$.

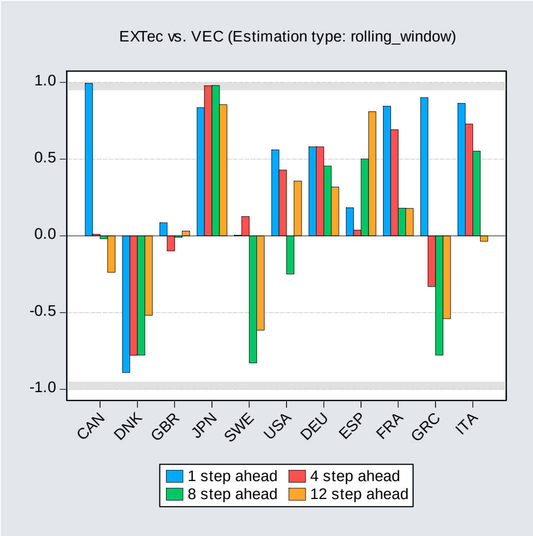


Chart 9: Forecasting performance of the model including external information.

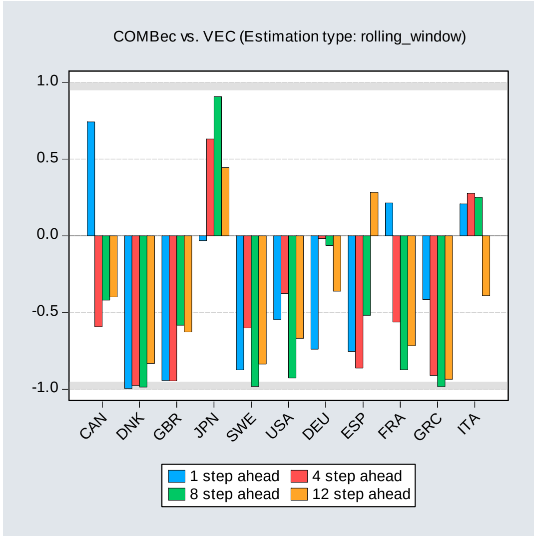


Chart 10: Performance of a combined forecast.

The model is benchmarked against the exclusively domestic CLI-based VEC model underlying the analyses in the previous subsections. The *p*-values of the Diebold/Mariano statistics for the countries in our sample are presented in Chart 9. The results are somewhat ambiguous. While external information appears to improve the forecast for the Danish economy, we find a worsening forecast performance for Japan and Italy, for instance. For the other economies in our sample, forecast performance is only slightly and unsystematically affected by the direct inclusion of external information in the forecasting model.[15]

The forecasting literature regularly points out the benefits of forecast combinations for the projection of macroeconomic developments, as this also insures to some extent against breaks or other non-linearities in the relationships between macroeconomic variables.[16] Indeed, combining the forecast of the closed economy VEC model and the previously discussed model with the external leading indicator EXTec leads to a notable improvement in the projections of industrial production. Chart 10 displays the *p*-value of the Diebold/Mariano statistic of the combined forecast, benchmarked against the purely domestic VEC model.[17] With the exception of Japan, the combined forecast substantially outperforms the forecast based exclusively on domestic information for all economies in our sample. The improvements appear to be particularly pronounced in the medium forecast horizon, while gains over the very short (1-step ahead) horizon are relatively contained due to the weak performance of the EXTec model.

---

[15] Forecast performance over short horizons appears to be negatively affected by the inclusion of external information.

[16] See e.g Elliott and Timmermann (2007) for a recent overview.

[17] For simplicity, we use an unweighted average of the two models' forecasts. Other weighting schemes (e.g. according to historical forecast performance) tend to improve the combined forecast not substantially.

5 Conclusions

This paper assessed empirically whether the ability of the country-specific leading indicators to predict the future economic situation has diminished in recent years due to rapid advances in globalisation. This is done on the basis of the OECD composite leading indicator, which is one of the most well-known composite indicators worldwide.

Overall, we find strong out-of-sample evidence that the CLI encompasses very useful information for forecasting industrial production. For most countries included in the sample and forecasting horizons, we observe substantial and mostly significant improvements of the CLI augmented VAR-based forecasts compared with standard benchmarks. The CLI-based model performs particularly well over horizon of four to eight months ahead and the results are robust to employing iterative forecasts or direct forecasts derived from horizon-specific models. Notably, accounting for cointegration between the CLI and industrial production significantly improves further the results of our forecasting exercise for all economies and horizons.

Turning to the temporal dimension of forecast performance over time, we find indications that the predictive accuracy of the CLI for economic activity has indeed declined over time for several countries. Augmenting the country-specific CLI with a leading indicator of the external environment and employing forecast combination techniques further improves the forecast performance for several economies.

6 Appendix

	IP	CLI
CAN	0.8686	0.8702
DNK	0.9660	0.9660
GBR	0.7438	0.5460
JPN	0.3447	0.2270
SWE	0.9959	0.9698
USA	0.9585	0.9749
DEU	0.9946	0.9924
ESP	0.9679	0.9473
FRA	0.9742	0.8024
GRC	0.6576	0.4913
ITA	0.7231	0.3359

Table 1: ADF test results for industrial production and composite leading indicators.
Note: The table reports the *p*-value of the null that the respective series has at least one unit root.

	Whole Sample	1975m1:1989m12	1990m1:2005m12
CAN	1	2	1
DNK	1	2	1
GBR	2	1	1
JPN	2	2	1
SWE	1	1	1
USA	1	2	1
DEU	1	1	1
ESP	1	1	1
FRA	2	2	1
GRC	2	2	1
ITA	2	2	1

Table 2: Johansen cointegration tests for IP and CLI.
Note: The table reports the number of cointegration relationships indicated by a trace test at the 5% level.

References

Bhansali, R. J. (2002), Multi-Step Forecasting, in: M. P. Clements and D. Hendry (eds.), A *Companion to Economic Forecasting*, New York: Blackwell, 206-221.

Blanchard, O., and J. Simon (2001), The Long and Large Decline in U.S. Output Volatility, *Brookings Papers on Economic Activity*, 2001, 1, 135-164.

Camba-Mendez, G., G. Kapetanios, M. R. Weale (1999), The Forecasting Performance of the OECD Composite Leading Indicators for France, Germany, Italy and the UK, *NIESR Discussion Paper* Nr. 155.

Carriero, A., and M. Marcellino (2007), A comparison of methods for the construction of composite coincident and leading indexes for the UK, *International Journal of Forecasting*, 23, 219-236.

Clements, M. P., and D. F. Hendry (1999), *Forecasting Non-stationary Economic Time Series*, MIT press.

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

Diebold, F. X., and G. D. Rudebusch (1991), Forecasting Output with the Composite Leading Index: A Real-Time Analysis, *Journal of the American Statistical Association*, 86, 603-610.

Elliot, G., and A. Timmermann (2007), Economic Forecasting, *CEPR Discussion Paper* No. 6158.

Emerson, R. A., and D. F. Hendry (1996), An Evaluation of Forecasting Using Leading Indicators, *Journal of Forecasting*, 15, 271-291.

Harvey, D., S. Leybourne and P. Newbold (1997), Testing the Equality of Prediction Mean Squared Errors, *International Journal of Forecasting*, 13, 291-291.

Marcellino, M, J. (2006), Leading Indicators, in: G. Elliott, C. Granger and A. Timmermann (eds.): *Handbook of Economic Forecasting, 1*, Elsevier, 879-960.

Marcellino, M, J. H. Stock and M. W. Watson (2005), A Comparison of Direct and Iterated Multistep AR Methods for Forecasting Macroeconomic Time Series, *Journal of Econometrics*, 135(1-2), 499-526.

Nilsson, R., and E. Guidetti (2008), Predicting the Business Cycle: How good are early estimates of OECD Composite Leading Indicators?, *OECD Statistics Brief* No. 14, February.

Stock, J. H., and M. W. Watson (1989), New Indexes of Coincident and Leading Economic Indicators, *NBER Macroeconomics Annual*, 4, 351-394.

Stock, J. H., and M. W. Watson (2003), Forecasting output and inflation: the role of asset prices”, *Journal of Economic Literature*, 41, 3, 788-829.

Stock, J. H., and M. W. Watson (2005), Understanding Changes in International Business Cycle Dynamics, *Journal of the European Economic Association*, 3, 5, 968-1006.