LeadING HUBE: An Effective Leading Indicator for the Performance of the Hungarian Business Sector

András Balatoni
senior macro analyst
ING Bank N.V. Hungary Branch
E-mail: andras.balatoni@ingbank.com

András Chabin
macro analyst trainee
ING Bank N.V. Hungary Branch
E-mail: andras.lukasz.chabin@ingbank.com

In this study a new leading indicator called LeadING HUBE (LeadING Index for the Hungarian Business Economy) is introduced and being calculated on a monthly basis to show the expected trajectory of the output of the private sector (as a proxy of added value) with significant lead and certainty. The authors' aim was to create an indicator which may be used for both policy and business purposes. They present the construction of LeadING HUBE and compare its performance with other, freely accessible leading indicators, demonstrating that this newly developed index outperforms them when it comes to lead time and reliability. Thus, LeadING HUBE does not only add to the extensive literature on leading indicators but also supplements it.

Keywords:
Business sector.
Indicator.
Forecast.

Although it is very important for the economic actors (decision makers, traders, analysts, etc.) to get a clear and prompt picture about the current state and future prospects of the economy, statistical "hard" data, such as GDP, are released with a significant delay. Thus, economists need to monitor other hard and soft indicators too, in order to gain information about the actual business situation and the expected path of the economic activity. A number of variables have some predictive power on the latter, but composite indices often perform quite adversely on a day-by-day forecast basis. In some cases, they contradict each other, making it difficult to derive any clear message from analysis.

In response to this recurrent issue, the main goal of this paper is to develop and present a new composite leading indicator of the Hungarian business economy; one that is constructed to predict business cycles on a monthly basis, in a reference series chosen as a proxy for economic activity. It is called LeadING HUBE (Leading Indicator for the Hungarian Business Economy).

In the construction of such a leading indicator, it is crucial to get early signals of the turning points in economic activity. These signals need to be reliable and minimise the number of false alarms. Besides, the index should be available on a monthly basis in order to offer forecasts regularly for forthcoming periods. It is also assumed that its significant monthly variation and huge ex post revisions are undesirable and thus should be avoided.

After this introduction, the structure of the paper is as follows: in Section 1, the theoretical background of business cycles and the creation of composite leading indicators are presented. Section 2 first summarises the construction of LeadING HUBE along with data issues and transformation of the time series, etc. Then, after giving a detailed demonstration of the index, it compares its performance in recession signalling with that of the OECD[1] leading indicator, SZIGMA[2] and GYIA[3] of ECOSTAT[4].

1. Construction of leading indicators

In the following subsections, we outline the usual way of creating a leading or coincident indicator, relying heavily on *Marcellino*'s [2006] work. First, we investigate

[1] OECD: Organisation for Economic Co-operation and Development.

[2] SZIGMA (Századvég index a gazdasági momentum alakulásáról): index of the "Századvég" Economic Research Institute on the actual state of the economy.

[3] GYIA (gyorsulási irányadó): acceleration index.

[4] ECOSTAT: Institute for Economy and Society Research.

the problems related to choosing a reliable reference series that is calculated on a monthly basis and represents properly the actual phase and dynamics of the business cycle. Then we sketch up two possible groupings of economic variables, which have proved useful in our research. Filtering, data handling and methods of constructing leading indicators are addressed only marginally.

1.1. Choice of reference series

Composite leading indicators aim to signal the performance of an economy and its turning points, hence ultimately trying to predict the future state of the business cycle. The question is how economic performance can be measured. Not surprisingly, the most widely used variable providing a prompt picture of the economic activity (despite its drawbacks) is GDP[5] calculated by national statistical offices. The problem is in that case that GDP is only available on a quarterly basis and with a delay, since its first (flash) estimate for the Hungarian economy is released 45 days after the end of the actual quarter. There are two options to overcome this problem: *1*. transformation of quarterly GDP figures to monthly frequencies (e.g. through interpolation), *2*. choosing or constructing an artificial reference series. In this study, the second approach is followed, therefore, transformation of GDP figures will not be discussed.

It is important that the chosen reference series is produced at a monthly frequency and correlates strongly with GDP. The most popular choice used to be IP[6] in the past few decades, since it met the earlier mentioned requirements. However, due to the changing structure of the economy and the diminishing importance of the sector in developed countries, its correlation with GDP has weakened and the construction of more sophisticated variables has become widespread.

1.2. Groups of variables

In our study, we define three different groups of soft and hard economic variables differentiated on the basis of their relation to business cycles (for more details see *Williamson* [2009]):

– *Acyclical variables'* cyclical movements are independent from the business cycles.

[5] GDP: gross domestic product.
[6] IP: index of industrial production.

– *Procyclical variables* (just like consumption, number of employed persons, level of industrial production, inflation, etc.) have positive correlation with GDP.

– *Anticyclical variables* (e.g. unemployment rate) correlate negatively with GDP.

Acyclical variables are not taken into consideration when a composite leading indicator is created, since they do not contain any information regarding the present or the future value of the reference series. On the contrary, procyclical and anticyclical variables are both useful and possibly worth taken in our leading indicator. (Note that anticyclical variables should be added to the index with a negative sign.)

Another grouping of potential variables showing either strong procyclical/anticyclical correlation with GDP or other reference series takes into account the timing of comovements. According to this categorization, three different groups can be identified: coincident, leading and lagging variables. Leading variables precede the business cycles, while coincident variables either move firmly together with or come shortly before them. Lagging indicators follow the path of the cyclical movements of GDP and as such, contain no relevant information for the construction of a leading index.

1.3. Filtering and data handling

After obtaining a suitable reference series and leading variables, the next step is data transformation. First, given a raw time series including both irrational and seasonal components, the exclusion of high frequency noise and outliers is necessary. After seasonal adjustment of the series, the type of transformation must be chosen. Macroeconomic and financial variables can be described by unit root processes, thus variance is an increasing function of time, while the expected value is non-constant of the time series. These may lead to spurious regression and wrong inference; therefore, the transformation of both leading (explanatory) variables and reference series (dependent variable) is required in order to include them "statistically properly" in the econometric models.[7] The transformation determines the nature of the analysed business cycles. The so-called "classical cycle" refers to fluctuations in the economic activity level (e.g. measured by GDP in volume terms or fluctuations in the output gap), while the "growth cycle" denotes fluctuations in the economic growth around the long-run potential level. Growth cycles may be defined as the difference between the actual growth rate and trend growth (or potential growth). In other words, the

[7] Due to the business cycle focus of the index, we do not deal with the cointegration of variables.

differences of the natural logarithms of the time series are taken to obtain percentage changes of variables, and then the historical average of the series is subtracted. Contrarily, the focus of analysis is the cyclical fluctuations of the level of variables in the case of classical cycles.

1.4. Methods of the construction of composite leading indicators

When creating composite leading indicators, the main aim is to combine and unite information being present in different leading variables in order to get a single index that efficiently predicts the path of GDP. According to *Marcellino* [2006], such a constructed index should have the following features. It

– gives consistent and accurate signals of the turning points of GDP along with steady lead time;
– follows firmly the trajectory of GDP between turning points;
– is based on reliable statistical background;
– is economically interpretable;
– is responding quickly and significantly to both negative and positive impulses;
– can be released regularly and quickly after the actual month/quarter while revisions of previous values are minimal;
– has no large monthly variability, in other words, its "noise" is limited.

In his study, *Marcellino* [2006] reviewed the widely applied methods of construction and differentiated between model-based and non-model-based indices. In the latter group, filtering, transformation and standardisation of time series are followed by a weighting scheme (e.g. coincident indicator of *The Conference Board* [2001]). Model-based indicators may be categorized as either factor models described by *Stock–Watson* [1989] or indicators that are built on Markov models (*Hamilton* [1989]).

2. Creation of LeadING HUBE

In the following subsections, the process of LeadING HUBE development is summarised. First, we introduce our monthly coincident variable and then investigate a large set of potential leading variables.

2.1. Coincident index

As it was mentioned earlier, GDP would be a natural choice as a reference series, but official statistics are released only quarterly, while LeadING HUBE is calculated on a monthly basis. Therefore, it is needed to construct a new reference series or coincident variable that can be derived from official monthly statistics and correlates significantly with GDP.

This coincident variable is determined by the volume of retail sales (denoted by *ret_turn*) (as a proxy for services) and that of production in both the industrial (*ind_prod*) and construction sectors (*con_prod*). Since these variables are available on a monthly basis, it is not necessary to attempt to interpolate GDP, which would raise a number of concerns regarding accuracy and statistical correctness. The first step is to transform the variables to exclude high frequency noise, outliers and seasonal patterns, hence enabling LeadING HUBE to concentrate on the proper periodicities of the time series. In this study the Henderson filter is applied that is derived by minimizing the sum of squares of the third difference of the moving average series (*Henderson* [1916]). The biggest advantages of this filter are the following: it allows the cycles typical of the trend to pass through unchanged and eliminates all the irregular variations that are of very short frequencies. However, just as in the case of the *Hodrick–Prescott* filter [1997], the standard endpoint problem emerges (*Proietti–Luati* [2008]). It means that in the middle of a time series, filter weights are symmetric, while the end filter weights are asymmetric, leading usually to biasedness in the output around the endpoint.

After seasonally adjusting the series using Census (X12) program and deriving their trend cycles by applying the Henderson filter, regression of the volume of retail sales, industrial production and construction output on GDP follows. The estimated coefficients are used as weights in the construction of the coincident variable (the weighted average of the mentioned time series).

Table 1

Regression model for the construction of the coincident index

Dependent variable	dlog(GDP)
Constant	**0.0013**
	1.3501
dlog(*con_prod*)	**0.0435**
	5.2597
dlog(*ind_prod*)	**0.2378**
	7.5053
dlog(*ret_turn*)	**0.1429**
	5.2616

R-squared value	0.793358	Mean of dependent variable	0.004630
Adjusted *R*-squared value	0.781662	Standard deviation of dependent variable	0.009690
Standard error of regression	0.004528	Akaike info criterion	–7.889474
Sum of squared residuals	0.001087	Schwarz criterion	–7.746102
Log likelihood	228.85000	Hannan-Quinn criterion	–7.833754
F-statistic	67.82756	Durbin-Watson statistic	0.970486
Prob(*F*-statistic)	0.000000	Wald *F*-statistic	71.25278
Prob(Wald *F*-statistic)	0.000000		

Note. Sample period: 1st quarter 2000–1st quarter 2014; method: ordinary least squares, Newey–West estimation of the variance-covariance matrix of the coefficients. Estimated parameters are in bold; *t*-statistics are in italics.

Source: Here and in all tables and figures (exluding Figure 2) own calculation.

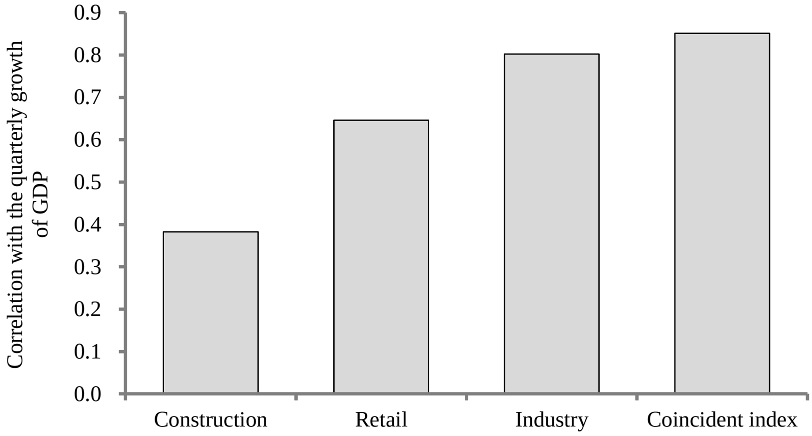
By means of the regression parameters, coincident can be easily calculated. (See equation in /1/.)

$$CI_t = \frac{0.0435 * con\_prod + 0.2378 * ind\_prod + 0.1429 * ret\_turn}{0.0435 + 0.2378 + 0.1429} \qquad /1/$$

where CI stands for the coincident index, while *con_prod, ind_prod* and *ret_turn are* the production of the construction sector, industry and the volume of retail sales, respectively.

Figure 1 demonstrates that the coincident index has the highest correlation with the growth rate of GDP (0.85), outperforming the other variables examined with regard to the strength of comovements.

Figure 1. Correlation between the quarterly growth of GDP and the coincident index along with its components, 2nd quarter 2000–1st quarter 2014



2.2. Variable transformation and selection

In the case of emerging market or transformation economies, it is exceptionally difficult to obtain a reliable estimate of the level of the potential GDP. Therefore, we decided to analyse the growth cycle instead of classical ones. The natural logarithm of the variables and the differences between them were taken in order to get percentage changes. For confidence indicators and consumer and business surveys, only the difference of the variables were calculated without taking the natural logarithms.

In some cases, some additional modifications were implemented. To derive the essence of the expectations of businesses and consumers, the surveys were transformed uniquely. For balance variables, the subtraction of the actual value of a variable at time *t* from its expected value at the same point in time was necessary (i.e. the difference between "Major purchases in the next 12 months" and "Major purchases at present" was calculated). This way, an indicator showing any changes in the consumers' willingness to make major purchases in the coming months could be captured. Another transformation was that the logarithm of the ratio of two variables (the stock of industrial export orders and industrial production) was taken.

In the construction of LeadING HUBE, it was crucial to find numerous relevant time series that lead the business cycles with steady and sufficient lead time. In our view, finding data that are not subject to significant revisions is just as important as being published on a monthly basis, relatively quickly after a given month ends. Early on in the construction of LeadING HUBE, more than eighty time series were collected and analysed. The data sets included soft and hard data on both the German and Hungarian economies. Hard data encompass financial indicators (exchange rates, interest rates, interest rate spreads, monetary aggregates, indices, etc.), industrial figures (production, sale, stock of orders, etc.), construction sector data (number of employed persons, orders, building permits), retail sales data, labour market figures (number of employed persons, number of registered job seekers, number of part-time workers, etc.), and other financial variables (inflation, budget balance, etc.).

We used cross-correlation analysis to separate the leading, coincident and lagging variables (just like *OECD* [2012] for leading indicators). Cross-correlation measures the strength of the comovement between the reference time series (in our case the coincident index) and the potential leading variables at different leads/lags. Formally:

$$r_i = corr\left(x_t ; y_{t-i}\right) \qquad /2/$$

where r is the correlation between variables x and y, when y is delayed by i months. The peak of the cross-correlation defines whether a time series is a leading, lagging or coincident variable. The variables that showed no correlation with the reference series at any lead or lag ("acyclic variables") were dropped from the further examinations. The remainders were divided into three groups according to the location of the peak in the cross-correlation:

– *leading variables* – positive (for anticyclical variables: negative) cross-correlation reaches its maximum (minimum) between –36 and –11 months ($11 < i < 36$);

– *coincident variables* – positive (for anticyclical variables: negative) cross-correlation reaches its maximum (minimum) between –10 and 0 months ($0 < i < 10$);

– *lagging variables* – positive (for anticyclical variables: negative) cross-correlation reaches its maximum (minimum) in +1 month or later ($i < 0$).

Out of the eighty time series that were originally analysed, only a few proved to be sufficiently significant. Where the cyclical profiles of the variable and the coincident index were highly correlated, the indicator was likely to provide a signal, not only of approaching turning points, but also of developments over the whole cycle. In order to investigate the stability of the set of leading variables and lead times, we repeated our model selection method on two sub-samples. The period of both sub-samples started in January 2000; the first ended in December 2008, while the second in December 2011. Table 2 shows the variables with the longest lead time and with at least 0.2 absolute value of correlation.

The volume of industrial export order book level divided by the volume of industrial production (*ind_ord_sh_ex*) has a very long lead time, and the correlation seems to be also stable. However, its lead time dropped from 28 to 14 months when the sample period was reduced. On the contrary, issued non-residential building permits (*con_bpnh_h*) has a stable lead time and correlation as well. The lead time of the volume of industrial order book level (*ind_ord_h*) is highly uncertain, in the two sub-samples the variable acted rather like a lagging variable, thus it was omitted from the model. A few elements of the German and Hungarian households' confidence are stable and good predictors of the business cycle (*cus_pricet_diff_ger*, *cus_majorp_diff_ger*, *cus_majorp_diff_hun*, *cus_save_s_hun_tc*, *cus_mp_s_hun_tc*), while others perform poorly with regard to lead time stability or the strength of comovement. (See Appendix for the explanation of the abbreviations.) The variable "Volume of new orders in construction" (*con_nord_h*) is also a bit unstable, but we assess it as a key variable to the Hungarian business cycle.

Table 2

Cross-correlation between the coincident index and the potential leading variables in different sub-samples

Potential leading variable	Whole sample (ends in June 2014)		Sub-sample whose period ends in December			
			2011		2008	
	Lead time	Correlation coefficient	Lead time	Correlation coefficient	Lead time	Correlation coefficient
ind_ord_sh_ex	–29	0.3119	–28	0.2559	–14	0.3034
con_bpnh_h	–28	0.3675	–28	0.3269	–27	0.3431
ind_ord_h	–28	0.2480	6	0.5739	6	0.4215
cus_pricet_diff_ger	–27	–0.3667	–27	–0.3224	–26	–0.2919
con_nord_h	–27	0.2559	–29	0.2266	0	0.3148
cus_majorp_diff_ger	–26	0.4517	–26	0.5356	–24	0.5482
cus_majorp_diff_hun	–23	0.2876	22	0.3985	–21	0.4223
cus_sav_s_ger	–19	–0.2614	19	–0.2076	–7	–0.2085
cus_save_s_ger	–18	–0.2228		>0.200	–17	–0.2539
cus_pt_s_ger	–13	–0.3141	–14	–0.2412	–6	–0.4005
con_ord_h	–12	0.4193	–12	0.3656	–17	0.3959
cus_pt_s_hun	–12	–0.2621	–24	–0.3683	–7	–0.2632
cus_pte_s_hun	–12	–0.2400		>0.2000	–9	–0.2459
cus_mp_s_hun	–11	0.3863	–29	0.4099	–10	0.3895
cus_save_s_hun	–11	0.2565	–11	0.2034	–11	0.2699
cus_finsite_s_hun	–9	0.2112	–29	0.3242	–24	0.3597

(Continued on the next page.)

(Continuation.)

Potential leading variable	Whole sample (ends in June 2014)		Sub-sample whose period ends in December			
			2011		2008	
	Lead time	Correlation coefficient	Lead time	Correlation coefficient	Lead time	Correlation coefficient
cus_genee_s_hun	–9	0.2734	–29	0.2607	–24	0.2800
cus_mpe_s_hun	–9	0.2222	–28	0.3825	–25	0.4107
cus_gene_s_hun	1	0.4459	–28	0.2848	–24	0.2911
cus_sav_s_hun	4	0.2486	–28	0.2943	–24	0.3048
cus_finsit_s_hun		>0.2000	–23	0.3139	–23	0.3338
con_emp_s	0	0.3015	0	0.3590	–20	0.2594
ret_stock_s		>0.2000	0	–0.1743	–18	–0.3051
con_aob_s	2	0.3674	2	0.4249	–16	0.3727
cus_finsit2_s_hun	–10	0.2983	–5	0.2467	–10	0.2771

Note. See Appendix for the explanation of the abbreviations. In the case of missing values, the cross-correlation does not reach 0.2 at any lead/lag, and it is not possible to define a peak of the lead time.

2.3. Benchmark model

Since this study aims at creating a composite leading indicator, only the variables having leading properties were kept. After determining the lead time of the remaining time series, the next step was to regress them on the reference series. Each variable was set to precede the coincident index (i.e. the dependent variable of the equation) exactly by its peak of cross-correlation. Necessarily, the variables that proved to be insignificant were omitted from the model. The estimator of the variance-covariance matrix of *Newey–West* [1987] was calculated to obtain significance levels that are robust to autocorrelation and heteroskedasticity. The final model is introduced by Table 3.

Although the method proposed by *Stock–Watson* [1989], *Nyman* [2010] and *Rácz* [2012] and applied by *Balatoni* [2014] for the construction of composite leading indicators is a popular "solution", it was found that the principal component analysis and the dynamic factor models do not perform better in signalling the turning points and forecasting the path of the coincident index than an OLS[8] regression in the case of the Hungarian economy. Proponents of the former technique argue that besides losing degrees of freedom, multicollinearity may result in loss of efficiency due to

[8] OLS: ordinary least squares.

the several, possibly weakly correlated regressors. However, after careful examination of VIF[9], it was concluded that multicollinearity of the benchmark model is not a serious problem.

Table 3

Benchmark regression model

Dependent variable	dlog(*CI*)
Constant	**0.0053**
	3.3823
dlog(*con_bpnh_h*(–28))	**0.0371**
	1.8785
dlog(*con_nord_h_tc*(–27))	**0.0362**
	3.1851
dlog(*ind_ord_sh_ex*(–28))	**0.0968**
	6.8738
cus_majorp_diff_ger(–26)	**0.0002**
	2.7517
cus_majorp_diff_hun(–23)	**0.0002**
	4.3173
cus_pricet_diff_ger(–27)	**–0.00004**
	–1.5728

R-squared value	0.764212	Mean of dependent variable	0.001636
Adjusted *R*-squared value	0.752251	Standard deviation of dependent variable	0.005519
Standard error of regression	0.002747	Akaike info criterion	–8.903212
Sum of squared residuals	0.001042	Schwarz criterion	–8.739726
Log likelihood	657.9344	Hannan-Quinn criterion	–8.836784
F-statistic	63.89581	Durbin-Watson statistic	0.247348
Prob(*F*-statistic)	0.000000	Wald *F*-statistic	57.8065
Prob(Wald *F*-statistic)	0.000000		

Note. CI stands for coincident index. See Appendix for the explanation of other abbreviations. Sample period: January 2000–July 2014; method: ordinary least squares, Newey–West estimation of the variance-covariance matrix of the coefficients. Estimated parameters are in bold; *t*-statistics are in italics. Lead times are in parenthesis.

[9] VIF: variance inflation factor.

As Table 3 shows, hard and soft data in the model are balanced in such a way that three of them were used from both types. There are two variables capturing the construction sector, while only one the industrial production. The remaining variables are survey data both from Hungary and from Germany.

To investigate the parameters' robustness, we used recursive estimation of our model. Each time, the sample period started in January 2000, while its end shifted by one month from estimation to estimation. This method shows the evolution of each coefficients' (beta) value and the month when it became significantly different from zero. In our case it was May 2008; no significant change in the model parameters could be detected afterwards. The only exception was the difference between the expected and present price trends (*cus_pricet_diff_ger*(–27)) that became insignificant again for a short period of time (from September 2010 to July 2011).

In sum, the set of leading variables, the lead time and the model parameters are robust enough to use them for further analysis.

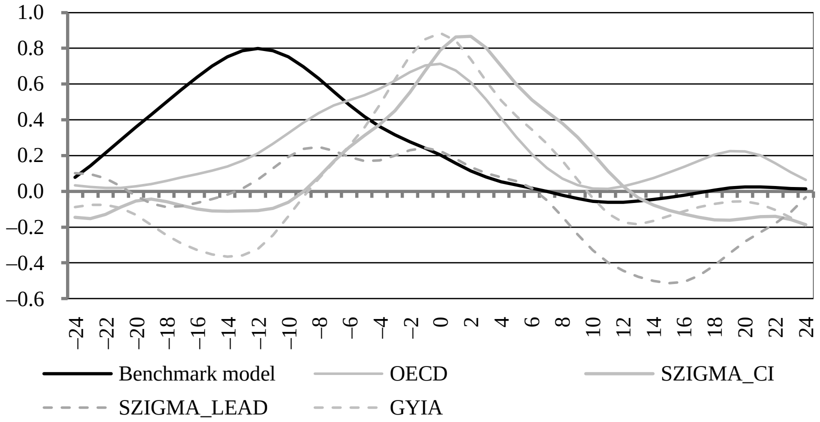
2.4. In-sample forecast performance of the benchmark model and other leading indicators

In this subsection, our benchmark model is compared with several other indices that capture economic momentum. Since it is only an in-sample forecast test, the performance of the benchmark model is compared with only that of indices with changing backcasts (ECOSTAT's GYIA, OECD leading indicator, SZIGMA CI[10] and SZIGMA LEAD[11]). These indices can be also interpreted as an in-sample fit of the respective model to the reference time series.

First, cross-correlation analysis and turning point detection tests are carried out at a monthly frequency. Data for our benchmark model are available from January 2002, thus, this is the maximum time span to be used. The OECD leading indicator is available for the same period as the SZIGMA indicators. GYIA is accessible from January 2006. Figure 2 shows that cross-correlation at zero lead or lag is the highest in the case of GYIA and SZIGMA CI. Therefore, these can be considered as the best coincident indices. However, if the lead time is increased (see left-hand side on the horizontal axis), the cross-correlation coefficient of all leading indices "rapidly fades away". On the contrary, LeadING HUBE's correlation increases significantly and reaches the peak at the 12-month lead time. Hereby, our target to construct an index giving information about the future state of the business economy is achieved.

[10] It summarises the current state of the economy in a single figure.
[11] It provides an overview of the prospective economic growth in nine months (three quarters).

Figure 2. Cross-correlation between various leading indices and our benchmark model
(monthly percentage changes)

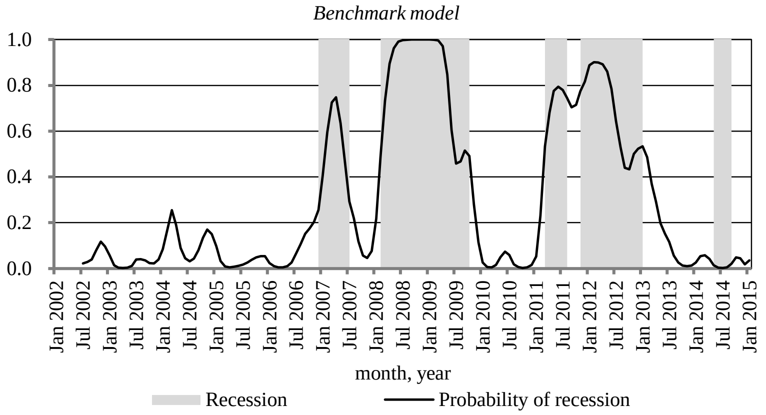


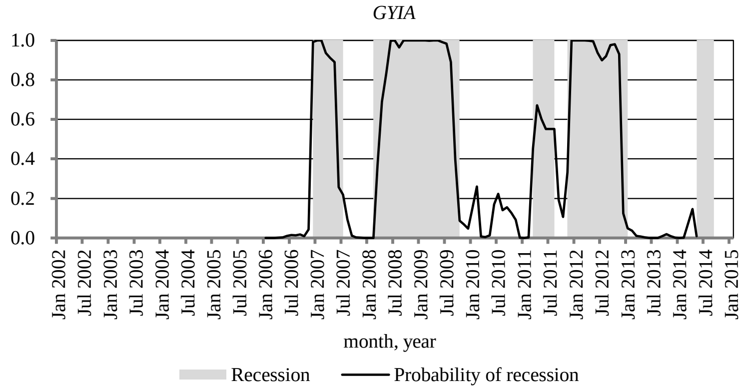
Note. The cross-correlation represents the correlation of the coincident index at time *t* and the other leading variable at time *t* + *n*, where *n* stands for the number of months by which the time series is shifted.

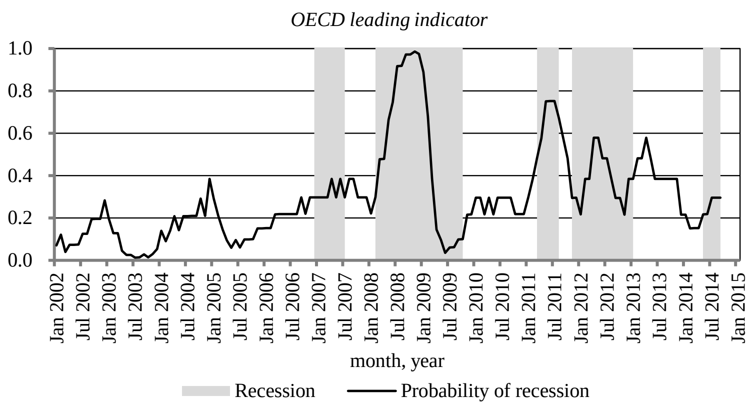
Source: Bloomberg, Századvég and OECD data as well as own calculation.

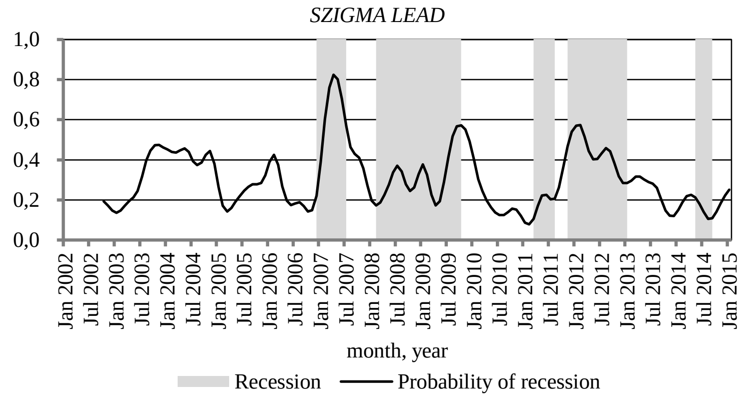
To demonstrate the performance of LeadING HUBE in turning-point detection, an artificial binary (dummy) variable is created that takes the value of 0 if the economy is expanding and 1 if the economy is in recession, according to the coincident index. (Recession is defined here as three consecutive months of decreasing output.) Then a binary outcome model is estimated in which the explanatory variable is our benchmark model with 12 leading months. Since the other indices would correlate poorly with the coincident index with the same lead time, they are included in the regression with no lead time or (in the case of SZIGMA LEAD) with nine-month lead (because the cross-correlation coefficient reaches its maximum nine months earlier than the actual value of the coincident index). The results are shown by Figure 3.

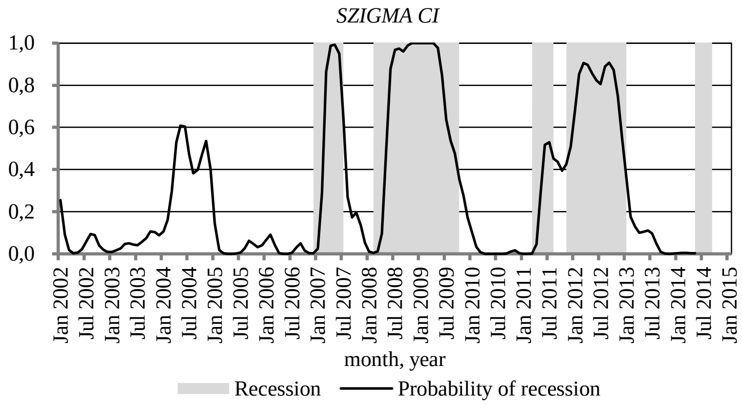
Figure 3. Probability of recession estimated by the benchmark model, GYIA, OECD leading indicator, SZIGMA LEAD and SZIGMA CI











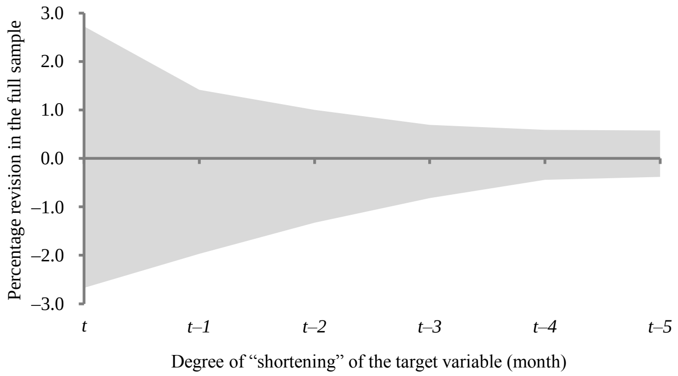
Our benchmark model has outperformed the OECD leading indicator and SZIGMA LEAD, while its performance almost reached the results of the coincident indices (GYIA and SZIGMA CI) that do not have lead time and thus, do not provide additional information about the future state of the business economy.

2.5. Calculation of LeadING HUBE

Our benchmark model was a good in-sample leading variable, but the lack of need for revision is also an important feature when the robustness of an indicator is assessed. As it was revealed in subsection 2.2., our set of leading variables as well as their lead time were “more or less” stable. By means of recursive parameter estimation, it was also presented ,that the regression parameters did not change significantly after 2008, so these elements of our model can be considered as a stable structure. However, we still have to solve the so-called end-point biasedness of the Henderson filter.

To account for the end-point uncertainty of the smoothed time series as a consequence of using this filter (for more details see *Proietti–Luati* [2008]), for each time period, the last four calculated data points of every time series were omitted from the further work (and hence from the calculation of the index). Despite the fact that this "deletion" makes four-month "foresight" or lead time prior to the reference date lost, the stability and reliability of the index improves significantly. Nevertheless, LeadING HUBE still has eight-month lead time, which is a great advantage compared with other leading indices. Figure 4 confirms our decision to omit the last four observations owing to the problem of end-point biasedness. At the end of the sample, the revision can reach even 2.5% (both negative and positive percentage deviation from the underlying trend), but the further (in months) the observations from the endpoint of the sample, the smaller the revision is. Therefore, the revision entailing four-month "deletion" is considered acceptable.

Figure 4. Revision of the coincident index at different distances (in months) from the endpoint



To calculate the final LeadING HUBE, the Henderson filter was used to smooth out both the explanatory variables and the coincident index. Then the last 4 months of the sample were split and a twelve-month forecast for the coincident index was implemented with the benchmark model. (See Table 2.) Next, in each sample period, the forecasted growth figures were linked in a chain fashion (just like in *Kertész–Kucsera–Szentmihályi*'s study [2015]). This chain-linked index is actually the LeadING HUBE, which means, it did not need revision. The parameters of the model became "stable" in 2008, so the index could be calculated from the second half of that year.

Since the final LeadING HUBE is available only for a short period, we cannot test properly its out-of-sample forecast performace with the usual tools (cross-correlation and turning point detection). However, it is still interesting to check at least graphically the comovement of LeadING HUBE and the coincident index in the last few years. Figures 6 and 7 illustrate that their correlation is significant at 4-5 months of lead. LeadING HUBE captures the underlying momentum of the economic growth because unlike the coincindence figures (that had large swings in their monthly changes between 2010 and 2012 and showed a peak in the beginning of 2015), it is not characterized by high frequency volatility.

Figure 5. Percentage changes of LeadING HUBE, August 2008–August 2015

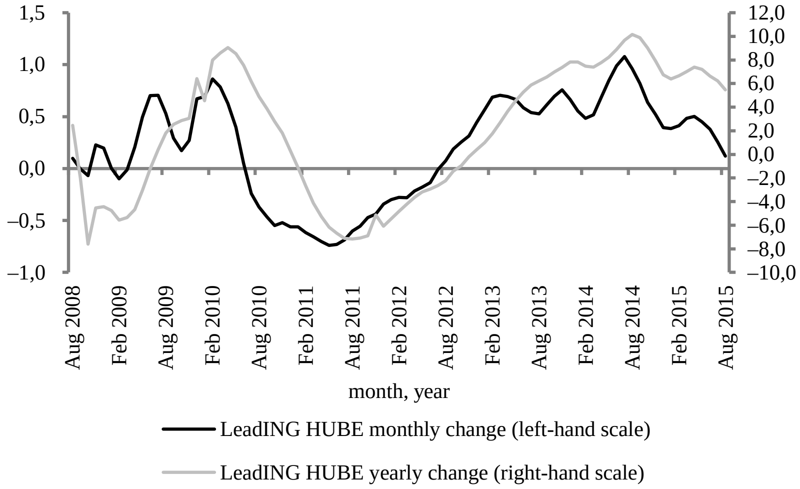
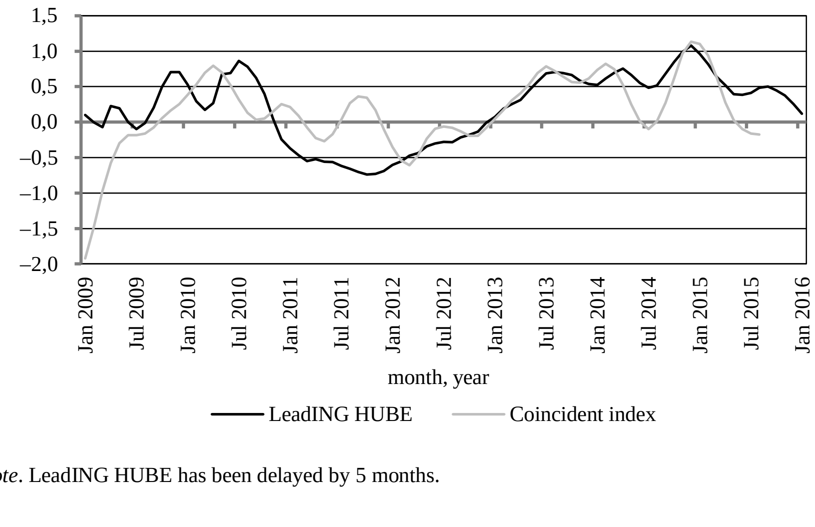
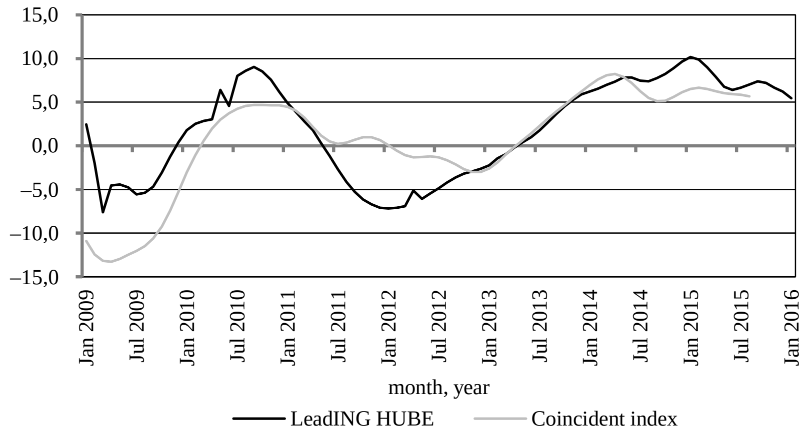


Figure 6. Comovement of LeadING HUBE and the coincident index, January 2009–January 2016
(monthly percentage change)



Note. LeadING HUBE has been delayed by 5 months.

Figure 7. Comovement of LeadING HUBE and the coincident index, January 2009–January 2016
(annual percentage change)



Note. LeadING HUBE has been delayed by 5 months.

In sum, LeadING HUBE performed well in out-of-sample forecast for the last few years, and represented sufficiently (without any revision) the underlying momentum of the Hungarian economic activity. Therefore, by publishing it monthly, we would provide valuable information for decision makers, traders and the public.

3. Summary

In this study, a new leading indicator for the Hungarian business economy was introduced. First, the general methods for the construction of composite leading indicators were described, and then the creation of LeadING HUBE was presented. The main purpose of this new composite leading indicator is to predict the probable path of the private sector performance with significant lead and certainty. In the last part of the study, LeadING HUBE was compared with several other indicators developed for the Hungarian economy, and it was demonstrated that it outperformed them with respect to reliability and lead time. We hope that LeadING HUBE can be a useful tool for both analysts and economic decision makers.

Appendix

Potential leading variables

Variable	Source	Unit	Denoted by
Volume of industrial export order book level divided by the volume of industrial production	HCSO	2010 = 100	*ind_ord_sh_ex*
Issued non-residential building permits	HCSO	m^2	*con_bpnh_h*
Volume of industrial order book level	HCSO	2010 = 100	*ind_ord_h*
Household survey: difference between the expected and present price trends – Germany	Eurostat	balance	*cus_pricet_diff_ger*
Volume of new orders in construction	HCSO	2010 = 100	*con_nord_h*
Household survey: difference between the expected and present major purchases – Germany	Eurostat	balance	*cus_majorp_diff_ger*
Household survey: difference between the expected and present major purchases – Hungary	Eurostat	balance	*cus_majorp_diff_hun*
Household survey: savings – Germany	Eurostat	balance	*cus_sav_s_ger*
Household survey: savings in the next 12 months – Germany	Eurostat	balance	*cus_save_s_ger*
Household survey: price trends – Germany	Eurostat	balance	*cus_pt_s_ger*
Stock of orders in construction	HCSO	2010 = 100	*con_ord_h*
Household survey: price trends – Germany	Eurostat	balance	*cus_pt_s_hun*
Household survey: price trends expectation – Hungary	Eurostat	balance	*cus_pte_s_hun*
Household survey: major purchases at present – Hungary	Eurostat	balance	*cus_mp_s_hun*
Household survey: savings in the next 12 months – Hungary	Eurostat	balance	*cus_save_s_hun*
Household survey: expected financial situation – Hungary	Eurostat	balance	*cus_finsite_s_hun*
Household survey: general economic outlook – Hungary	Eurostat	balance	*cus_genee_s_hun*
Household survey: major purchases in the next 12 months – Hungary	Eurostat	balance	*cus_mpe_s_hun*
Household survey: general economic situation – Hungary	Eurostat	balance	*cus_gene_s_hun*
Household survey: savings – Hungary	Eurostat	balance	*cus_sav_s_hun*
Household survey: financial situation – Hungary	Eurostat	balance	*cus_finsit_s_hun*
Construction survey: employment expectations	Eurostat	balance	*con_emp_s*
Retail survey: stock levels	Eurostat	balance	*ret_stock_s*
Construction survey: order book levels	Eurostat	balance	*con_aob_s*
Household survey: statement on the financial situation of households	Eurostat	balance	*cus_finsit2_s_hun*

References

BALATONI, A. [2014]: SZIGMA: a hazai gazdaságra fejlesztett egyidejű és elõidejű indikátorrendszer. *Statisztikai Szemle*. Vol. 92. No. 2. pp. 109–138.

BAXTER, M. – KING, R. G. [1999]: Measuring the Business Cycle: Approximate Band-Pass Filter for Macroeconomic Time Series. *Review of Economics and Statistics*. Vol. 84. No. 4. pp. 575–593.

HAMILTON, J. D. [1989]: A New Approach of the Economic Analysis of Nonstationary Time Series and Business Cycle. *Econometrica*. Vol. 57. No. 2. pp. 357–384.

HENDERSON, R. [1916]: Note on Graduation by Adjusted Average. *Transactions of the American Society of Actuaries*. Vol. 17. pp. 43–48.

HODRICK, R. J. – PRESCOTT, E. C. [1997]: Postwar US Business Cycles: An Empirical Investigation. *Journal of Money, Credit and Banking*. Vol. 29. No. 1. pp. 1–16.

KERTÉSZ B. – KUCSERA H. – SZENTMIHÁLYI Sz. [2015]: A New Indicator Determing the Medium Term GDP Growth. MNB Working Papers 120. Budapest.

MARCELLINO, M. [2006]: *Handbook of Economic Forecasting*. North Holland. Amsterdam.

NEWEY, W. K. – WEST, K. D. [1987]: A Simple, Positive Semidefinite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*. Vol. 55. No. 3. pp. 703–708.

NYMAN, CH. [2010]: *An Indicator if Resource Utilization*. Economic Commentaries No. 4. Svergies Riskbank. Stockholm.

OECD (ORGANISATION FOR ECONOMIC COOPERATION AND DEVELOPMENT) [2012]: *OECD System of Composite Leading Indicators*. http://www.oecd.org/std/leading-indicators/41629509.pdf

PROIETTI, T. – LUATI, A. [2008]: Real Time Estimation in Local Polynomial Regression, with Application to Trend-Cycle Analysis. *Annals of Applied Statistics*. Vol. 2. No. 4. pp. 1523–1553.

RÁCZ, O. M. [2012]: A gazdaság ciklikus pozíciójának megítélése bizalmi indikátorok segítségével. *MNB Szemle*. Vol. 2. No. 7. pp. 41–46.

STOCK, J. H. – WATSON, M. W. [1989]: New Indexes of Coincident and Leading Indicators. In: *Blanchard, O. – Fisher S.* (eds.): *NBER Macroeconomics Annual*. MIT Press. Cambridge. pp. 351–394.

THE CONFERENCE BOARD [2001]: *Business Cycle Indicators Handbook*. https://www.conference-board.org/pdf_free/economics/bci/BCI-Handbook.pdf

WILLIAMSON, S. D. [2009]: *Makroökonómia*. Osiris. Budapest.