

Norin, Anna

**Conference Paper**

## Nowcasting of the Gross Regional Product

50th Congress of the European Regional Science Association: "Sustainable Regional Growth and Development in the Creative Knowledge Economy", 19-23 August 2010, Jönköping, Sweden

**Provided in Cooperation with:**

European Regional Science Association (ERSA)

*Suggested Citation:* Norin, Anna (2010) : Nowcasting of the Gross Regional Product, 50th Congress of the European Regional Science Association: "Sustainable Regional Growth and Development in the Creative Knowledge Economy", 19-23 August 2010, Jönköping, Sweden, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/119087>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Nowcasting of the Gross Regional Product

Anna Norin<sup>1</sup>

<sup>1</sup>CERUM, Umeå University, SE-90187 Umeå, Sweden, [anna.norin@cerum.umu.se](mailto:anna.norin@cerum.umu.se)

***Abstract:***

Timely values of the gross regional product are rarely available. Annual GRP are often published with up to two years lag. The present paper evaluates a method of obtaining values of the GRP as soon as monthly and quarterly business cycle indicators become available. Building on earlier work on using bridge equations to obtaining quarterly values of GDP growth, a method is proposed where annual GRP growth is estimated using a large number of business cycle indicators. In a GRP setting, the time series span is far shorter than in a GDP setting. This is addressed by using principal components to summarize the information available in the business cycle indicators.

## Introduction

Information on regional business cycles and regional economic growth is often available only at substantial delay. Data on the gross regional product can be published with as much as a two year lag. For regional decision makers it is vital to have access to timely information on the state of the regional economy. Relying only on national business cycle data may be misleading as regions may differ from the national economy regarding timing, frequency, and amplitude of business cycles.

The present paper examines the possibility of “nowcasting” the gross regional product (GRP) to obtain accurate and up-to-date estimates. Nowcasting has to a large extent focused on using bridge equations where ancillary indicator information is “bridged” to the variable of interest<sup>1</sup>. Indicator information that is published at a higher frequency and/or with a shorter publication lag is then used to get estimates of current GRP.

Previous research on bridge modelling has been used to obtain flash estimates of GDP and GDP growth both at the national and the supranational level. (Rünstler et al., 2009) use data from ten European countries and (Golinelli and Parigi, 2007) include the European union, the euro area and the G7 countries in their analysis. The euro area has also been analyzed by for instance (Giannone et al., 2009), (Frailé et al., 2010), (Castle et al., 2009), and (Baffigi et al., 2004), while others have focused on individual countries (see e.g. (Barhoumi et al., 2010), (Nunes, 2005) and (Schumacher and Breitung, 2008)). Here, the focus is instead regions within a country and the issues that arise in that context. The method is applied to the 21 Swedish regions.

In a regional application, the degrees of freedom problem is even more acute than on national level. GRP data are at best published annually resulting in very short time series. This limits the number of variables that can be allowed in the bridge equation. Different approaches have been suggested and tested empirically. One strand of literature have focused on forecast combination where many bridge equations with one, e.g. (Monaco, 2003), or a small set of indicators, see e.g. (Baffigi et al., 2004), are estimated and then combined or pooled. Another strand of the literature have focused on dynamic factor models where factors are extracted from large numbers of indicators. Some studies using more than 400 indicators (see

---

<sup>1</sup> The approach can also be used for forecasting. One way this can be done is by obtaining future values of the indicator variables through univariate methods. A different way is to formulate the model in terms of past values and make h-step ahead predictions.

(Barhoumi et al., 2010) for a table summary of some of these studies). Recent papers suggest that indicator sets as small as 20 may be just as good or better than larger indicator sets (see (Schumacher and Breitung, 2008) for a contradictory result).

Here, the static factor formulation of (Stock and Watson, 2002) is used. By using a static formulation the factors can be extracted using principal components. This is an advantage as it reduces the complexity of the analysis. Furthermore, the number of indicators must not be large for the approach to be able to provide accurate nowcasts. (Barhoumi et al., 2010) showed that the static formulation of (Stock and Watson, 2002) outperformed different dynamic factor formulations.

One important issue concerns the choice of indicator variables. Which indicator variables are of interest in a regional setting? To large extent the indicator variables are the same as in national nowcasting but the effects are allowed to vary for different regions. On a regional level, there are some indicator data that is published more promptly than GRP. A comparison is made to see whether including the region specific information improves on the nowcasting performance.

The static factor nowcasts are compared to the pure forecast from a benchmark AR(1) model. The models are evaluated by estimating the models using a recursive sample scheme and calculating the root mean squared error.

The following section describes the estimation approach. Subsequently is a section describing the data set used. Results concerning nowcasting accuracy are presented next, followed by some concluding remarks.

## **Nowcasting Estimation**

Consider then the bridge equation where the gross regional product,  $Y_t$ , is linearly related to the factors  $F_t$  according to

$$Y_t = \beta'(L)F_t + \phi(L)Y_{t-1} \quad (1)$$

and where  $\beta'(L)$  and  $\phi(L)$  are lag polynomials in nonnegative powers of  $L$ . The autoregressive formulation for the factors is to introduce some dynamics into the static factors (Stock and Watson, 2002).

The factors are extracted from the indicator variables using principal component analysis. In essence this reduces the dimension of the indicator set, i.e. it reduces the number of variables, by a finding a new smaller set of uncorrelated variables containing most of the variation in the original data (Jolliffe, 2002). The approach was suggested for forecasting by (Stock and Watson, 2002) and has also been applied by (Barhoumi et al., 2010) and (Schumacher and Breitung, 2008) in this context. The factors are related to the indicators according to

$$X_t = \Lambda F_t + e_t \quad (2)$$

where the  $\Lambda$  are the factor loadings. In general, the factors in the above expression will not theoretically be the same as the principal components.

After factors have been extracted, Eq. (1) can be estimated using ordinary least squares regression. Decisions must be made concerning the number of factors and lags of factors and GRP to include. Several earlier studies have relied on different automatic selection procedures such as autometrics (Castle et al., 2009) or some other implementation of the general-to-specific methodology, e.g. (Barhoumi et al., 2010). For an a recent introduction to automatic modelling, see (Doornik, 2008). It is also possible to use some type of information criterion for choosing between alternative model specifications. Here, the latter is used combined with exploratory data analysis due to the small sample that is available. An alternative for future research would be to use the algorithm for general-to-specific modelling that can handle more variables than observations.

Benchmark model is the autoregressive AR(1) model where the GRP is related to its past values according to

$$Y_t = \varphi Y_{t-1} + \epsilon_t \quad (3)$$

For the benchmark model the obtained predictions will be pure forecasts. The forecasts are obtained as  $h$ -step ahead predictions.

The nowcasts are evaluated using a recursive scheme for out-of-sample nowcasts. Initially, the model is estimated for the years 1993-2003. Nowcasts are then calculated up to 2007. The model is then reestimated after adding one observation (year) at a time and new nowcasts are obtained. The nowcasting accuracy is evaluated using the root mean squared error (RMSE) given by

$$RMSE = \sqrt{\frac{1}{n_h} \sum_{t=1}^{n_h} (Y_{t+h} - \hat{Y}_{t+h})^2} \quad (4)$$

where  $h$  is the nowcasting horizon. Given the extraordinariness of the 2009 downturn, the models are also evaluated according to whether they predict negative growth in 2009.

## Data Description

Indicator data sets previously used in forecasting with dynamic factor models are often very large with some studies using more than 400 indicator variables. As has been mentioned this does not necessarily improve on forecasting performance. Very large indicator data sets also entail a data collection cost as data are continuously revised and updated. Earlier applications of bridge models relied on the other hand on a few select indicators. Here, macroeconomic variables are included on the basis of availability. All indicators published no later than three months into the new year are included in the indicator set. In a second bridge model, the indicator set contain the same variables but with region specific data whenever possible. Data is available since 1993, including official GRP data from Statistics Sweden. Data source for all variables is Statistics Sweden.

The indicator sets are available such that a nowcast can be made for year  $t$  at  $t+3$  months. This departs from earlier studies including all the available information at a particular time point. Indicators not yet reported are then either forecasted using univariate methods or a kalman filter approach where the missing observations are updated using the filter. Early estimates based on less information tend to be unreliable. The trade-off is then between timeliness and accuracy. By choosing to base nowcasts on complete indicator sets the nowcasts can be released one quarter into the following year, well ahead of official data which is preliminarily released at  $t+18$  months.

All variables are evaluated with respect to stationarity. Time series plots together with the augmented Durbin-Watson test are used to determine the appropriate transformations to achieve stationarity. Most of the variables are tranformed by taking logarithms and then differencing, i.e. a growth formulation. The extraction of the principal components further require standardizing the indicators to mean zero and variance 1.

**Table 1: Nowcasting accuracy of the bridge and benchmark models - recursive RMSE and 2009 indicator**

Region	AR(1)		BM1		BM2	
	RMSE	09↓	RMSE	09↓	RMSE	09↓
01 Stockholm	0.77		1.00	x	1.02	x
03 Uppsala	0.60		1.00		0.94	
04 Södermanland	1.50		1.00		0.89	
05 Östergötland	0.41		1.00	x	0.75	x
06 Jönköping	2.15		1.00	x	0.89	x
07 Kronoberg	1.28		1.00	x	0.97	x
08 Kalmar	0.42		1.00	x	1.16	x
09 Gotland	0.11		1.00	x	0.50	x
10 Blekinge	0.60		1.00	x	1.04	x
12 Skåne	2.83		1.00	x	0.91	x
13 Halland	0.54		1.00		0.98	
14 Västra Götaland	0.39		1.00		0.95	x
17 Värmland	0.66		1.00	x	0.93	x
18 Örebro	0.79		1.00	x	1.00	x
19 Västmanland	1.11		1.00	x	1.33	x
20 Dalarna	1.71		1.00		1.11	
21 Gävleborg	1.70		1.00	x	1.72	x
22 Västernorrland	0.30		1.00		1.03	
23 Jämtland	0.42		1.00		0.99	
24 Västerbotten	0.85		1.00		1.17	
25 Norrbotten	1.44		1.00		0.65	

Note: The RMSE values are reported relative to the RMSE of the BM1 model. The BM1 model is the benchmark model including all available indicators in the factors. The BM2 model considers the same indicators set as BM1 but replaces with regional specific information for 4 of the indicators. The x in the 09↓ columns indicate that the specification predicts negative growth in 2009.

## Estimation Results

The bridge model is estimated using the different indicator sets discussed above and the nowcasting accuracy is presented in table x. Each specification is estimated separately for all Swedish regions. In all there are 21 regions, 2 model specifications (the bridge model and the AR(1) benchmark), and for the bridge model three different indicator sets are used. The RMSE values are presented relative to the RMSE for the BM1 model. The BM1 model is the bridge model including all available indicators three months into the present year. The BM2 model uses the same indicator set but replaces four of the indicators with regional specific data. The results indicate that the BM1 model performs better than the benchmark for eight of the 21 regions. Adding regional specific data does improve forecasting accuracy according to the

RMSE for six of the eight regions were a bridge model specification has better accuracy than the benchmark model. The evidence for the BM specifications is not compelling although they are suitable for some regions. In 2009 however, Sweden experienced substantial negative growth at a national level. Swedish GDP fell about 5% according to both Statistics Sweden and Konjunkturinstitutet. Likely, this negative growth is experienced by all regions in varying degree. The benchmark AR(1) model fail to detect any negative growth in 2009 while the bridge model specifications predict negative growth for 12 and 13 of the regions, respectively. This indicates an advantage of the bridge models in detecting large unexpected changes.

**Table 2: Informative content of model predictions (RMSE relative to sample standard deviation)**

Region	AR(1)	BM1	BM2
01 Stockholm	0.68	0.89	0.91
03 Uppsala	1.07	1.79	1.68
04 Södermanland	1.64	1.09	0.97
05 Östergötland	0.57	1.42	1.06
06 Jönköping	1.68	0.78	0.70
07 Kronoberg	1.72	1.35	1.30
08 Kalmar	0.77	1.84	2.13
09 Gotland	0.26	2.50	1.25
10 Blekinge	0.52	0.87	0.91
12 Skåne	1.55	0.55	0.50
13 Halland	0.70	1.29	1.27
14 Västra Götaland	0.74	1.88	1.79
17 Värmland	0.69	1.05	0.97
18 Örebro	1.19	1.50	1.50
19 Västmanland	1.47	1.33	1.76
20 Dalarna	1.38	0.81	0.90
21 Gävleborg	0.92	0.54	0.92
22 Västernorrland	0.43	1.42	1.46
23 Jämtland	0.67	1.61	1.59
24 Västerbotten	1.09	1.28	1.51
25 Norrbotten	1.66	1.15	0.75

Note: RMSE estimates reported relative to the standard deviation in the sample.

By reporting RMSE relative to the sample standard deviation an indication of the informative content of the nowcast can be obtained (Schumacher and Breitung, 2008), see table 2. The benchmark model have values below one for 11 of the regions indicating informative forecasts for these 11 regions. The regional specific bridge model turn out to be more informative than the bridge model with only national indicators, with 9 regions having values below one for BM2 compared to six regions for BM1.

## Concluding Remarks

The usefulness of applying bridge models to obtain accurate and up-to-date estimates of the gross regional product is examined for the 21 Swedish regions. The results are mixed. For a number of regions the bridge modelling does provide informative nowcasts which improves on the pure forecasts from an AR(1) model, but far from all regions. However, when it comes to considering the ability of predicting the substantial downturn of 2009, the benchmark

AR(1) specification failed to predict negative growth for any of the regions. On the other hand, both bridge model specifications did so for a majority of the regions indicating an advantage for bridge modelling to detect unexpected and large deviations.

The bridge models examined here are based on indicators available three months into the present year. Two routes to further the analysis is to postpone nowcasting until more data becomes available or to use univariate forecasts of additional indicators. Given the improvement in nowcasting accuracy by including regional specific information the inclusion, if possible, of additional regional specific information could be an interesting extension. However, earlier research does not necessarily favor additional data as a means of improving forecasts. For instance (Barhoumi et al., 2010) find that the smallest indicator set performed best in their analysis. A third approach might then be to be more selective in choosing indicators for the principal component analysis.

An interesting area for further research could be to use even more disaggregated data, but not necessarily on a geographical level. By disaggregating across sectors or industries perhaps more informative nowcasts can be obtained. The public sector has been known to increase output in times of recession for the manufacturing sector.

## References

- BAFFIGI, A., GOLINELLI, R. & PARIGI, G. 2004. Bridge models to forecast the euro area GDP. *International Journal of Forecasting*, 20, 447-460.
- BARHOUMI, K., DARNÉ, O. & FERRARA, L. 2010. Are disaggregate data useful for factor analysis in forecasting French GDP? *Journal of Forecasting*, 29, 132-144.
- CASTLE, J. L., FAWCETT, N. W. P. & HENDRY, D. F. 2009. NOWCASTING IS NOT JUST CONTEMPORANEOUS FORECASTING. *National Institute Economic Review*, 210, 71-89.
- DOORNIK, J. A. 2008. Encompassing and Automatic Model Selection\*. *Oxford Bulletin of Economics and Statistics*, 70, 915-925.
- FRALE, C., MARCELLINO, M., MAZZI, G. L. & PROIETTI, T. 2010. Survey data as coincident or leading indicators. *Journal of Forecasting*, 29, 109-131.
- GIANNONE, D., REICHLIN, L. & SIMONELLI, S. 2009. NOWCASTING EURO AREA ECONOMIC ACTIVITY IN REAL TIME: THE ROLE OF CONFIDENCE INDICATORS. *National Institute Economic Review*, 90-97.
- GOLINELLI, R. & PARIGI, G. 2007. The use of monthly indicators to forecast quarterly GDP in the short run: an application to the G7 countries. *Journal of Forecasting*, 26, 77-94.
- JOLLIFFE, I. T. 2002. *Principal component analysis*, New York, Springer-Verlag.
- MONACO, J. K. A. R. 2003. Real-Time Forecasting in Practice. *Business Economics*, 10-19.
- NUNES, L. C. 2005. Nowcasting quarterly GDP growth in a monthly coincident indicator model. *Journal of Forecasting*, 24, 575-592.
- RÜNSTLER, G., BARHOUMI, K., BENK, S., CRISTADORO, R., REIJER, A. D., JAKAITIENE, A., JELONEK, P., RUA, A., RUTH, K. & NIEUWENHUYZE, C. V. 2009. Short-term forecasting of GDP using large datasets: a pseudo real-time forecast evaluation exercise. *Journal of Forecasting*, 28, 595-611.
- SCHUMACHER, C. & BREITUNG, J. 2008. Real-time forecasting of German GDP based on a large factor model with monthly and quarterly data. *International Journal of Forecasting*, 24, 386-398.
- STOCK, J. H. & WATSON, M. W. 2002. Macroeconomic Forecasting Using Diffusion Indexes. *Journal of Business & Economic Statistics*, 20, 147-162.