Nowcasting Manufacturing Value Added for Cross-Country Comparison

Kris Boudt[a] Valentin Todorov[b] Shyam Upadhyaya[b]

[a]*K.U.Leuven and Lessius University College, Naamsestraat 69, 3000 Leuven, Belgium. Email: kris.boudt@econ.kuleuven.be Tel: +32 16 326728 Fax: +32 16 326624*
[b]*Research and Statistics Branch, UNIDO, P.O. Box 300, A-1400 Vienna, Austria.*

Abstract

Manufacturing Value Added (MVA) is the key indicator of a country's industrial production. In order to facilitate international comparisons it is published in UNIDO's International Yearbook of Industrial Statistics for a large set of countries. Because of a time-gap of at least one year between the latest year for which data are available and the year for which MVA data must be reported in the Yearbook, nowcasting methods are used to fill in the missing data up to the current year. We propose a parsimonious methodology that exploits the relationship between MVA and GDP to produce reliable nowcasts of MVA.

Keywords: Manufacturing Value Added, Nowcasting, Robustness, UNIDO.

1 Introduction

The Research and Statistics Branch of the United Nations Industrial Development Organization (UNIDO) is responsible for implementing the international mandate of the Organization in the field of industrial statistics. It maintains a unique industrial statistics database and updates it regularly with data collected from the national statistical offices. A separate database at macro level is also maintained primarily for compilation of statistics related to manufacturing value added (MVA) such as its growth rate and share in gross domestic product (GDP) for various countries and regions. These figures are published in the *International Yearbook of Industrial Statistics* and posted on the statistical pages of the UNIDO web site. For current economic analysis it is crucial that the Yearbook presents MVA data for the most recent years. This paper considers the problem of providing an estimate of the missing values of the current-year MVA. Such an estimate is called a "nowcast", rather than a pure forecast.

Previous research has mainly focused on nowcasting current-quarter GDP growth. Small ([2,9]) or large ([4,6,12]) models are then used to "bridge" the information contained in monthly data with the quarterly growth rate of GDP. One exception is the research of [3,5] who describe how short term interest rates and business survey data can be used to produce nowcasts of the monthly industrial production index for the aggregated euro area. Their analysis (based on different data sets) leads to different conclusions. [3] finds that models with indicators generally do worse than simple autoregressive models. The empirical results of [5] indicate however that the nowcasting method that exploits the information in the survey results outperforms the simple autoregressive models.

Regarding the practical application of nowcasting methods in statistical offices, [4] mentions that the Office for National Statistics (ONS) of the United Kingdom takes a conservative approach to

using external data sources such as business surveys and financial market variables in compiling its statistics. He notes that: "As a National Statistics institution, ONS has an obligation to meet international standards on the formulation of National Accounts, and produce estimates in a transparent way so that users can be confident that quality benchmarks are being maintained. Combining official estimators with indicators would certainly compromise this." ONS therefore uses their nowcast of GDP based on external data only as an informal guide and check for the internal GDP estimate based on timely but incomplete information of the GDP components.

For the research and statistics branch of UNIDO, the only way to obtain the nowcast of MVA is to use external data. To comply with the high quality and transparency requirements for official statistics, it is therefore necessary that the proposed method uses an indicator of high quality and that the method linking the external indicator to the nowcast is simple to understand.

In this paper we propose a nowcasting method for yearly MVA which exploits the economic relationship between MVA and GDP, together with the availability of reliable estimates of GDP growth rates from external sources. The proposed method is based on a parsimonious model of MVA as a function of contemporaneous GDP and past values of MVA. We consider different model specifications. Because the nowcasts of MVA are published by UNIDO in the *International Yearbook of Industrial Statistics*, a good nowcasting method is not only a method with a low mean absolute error, but it also needs to satisfy the following three requirements:

R1. The nowcasts produced by the method are little influenced by revisions of single observations in the data.

R2. The nowcasts should be plausible given the past values of MVA.

R3. The nowcasting method should not only be accurate on average, but also accurate for all countries.

R1 is needed because UNIDO receives new MVA data every year from its different sources. This data contains not only the values of MVA that were missing before, but also revisions of the MVA data delivered the year before. *R2* follows from the fact that UNIDO compares the past values of MVA with the nowcast published in the Yearbook. Finally, *R3* is of major importance since the MVA nowcasts are used for economic policy making and international comparison between countries.

The remainder of the paper is organized as follows. In Section 2 we describe the data and methodology. In Section 3 we compare the nowcasting accuracy for the different models considered. Section 4 presents our conclusions.

2 Data and methodology

The database maintained by UNIDO for nowcasting MVA consists of yearly values of MVA (from 1960 to T-2, where T is the current year) and GDP (1960 to the current year) at constant 2000 prices for around 200 countries. For many countries, observations are missing at the starting points of the series (see [13,14] for more details). The GDP series equals the actual GDP for the earlier years, while for the most recent one or two years they are derived from the nowcasts of GDP growth rates reported in the World Economic Outlook of the IMF (see e.g. [1] for a study on the accuracy of these nowcasts). In contrast with MVA, the GDP data are thus available up to the current year. Note that the MVA is strongly connected to GDP, since on the one hand MVA is a part of the total value added of the country and thus of GDP and on the other hand the production of value added by the manufacturing industry is driven by the demand for its products and thus by GDP. This suggests to nowcast MVA on the basis of the estimated relationship between the contemporaneous values of MVA and GDP**.** Next we consider different modeling approaches and,

because of outliers in the data, we propose to consider not only the OLS estimator, but also a highly robust estimation method called the MM estimator.

Model. Denote by $MVA_{i,t}$ and $GDP_{i,t}$ the MVA and GDP of country i in year t. Up to recently, UNIDO used a nowcasting model in which the log of MVA is modeled as either a linear or a quadratic function of the contemporaneous value of the log of GDP:

$$\log MVA_{i,t} = a_i + b_i \log GDP_{i,t} + e_{i,t} \quad (1)$$

$$\log MVA_{i,t} = a_i + b_i \log GDP_{i,t} + c_i (\log GDP_{i,t})^2 + e_{i,t}, \quad (2)$$

where $e_{i,t}$ is white noise. The nowcast of MVA is then computed as the exponential of the fit of this regression.

The main advantage of this model is the ease of interpretation of its parameters. The parameters give an answer to "what if" questions. Suppose c_i is zero. Then b_i equals the elasticity of MVA in function of GDP. If $c_i > 0$ (< 0), the elasticity of MVA to GDP tends to be larger (smaller) for larger values of GDP than for smaller values of GDP. For $c_i = 0$, the elasticity of MVA to GDP is constant in function of the level of MVA. For our purpose of nowcasting the level of MVA, the interpretability of the parameters is a nice by-product, but it is not a requirement.

The disadvantage of the nowcast method based on this model is that it does not satisfy any of the three requirements enumerated in the Introduction. The first requirement that revisions of single observations have a small influence on the nowcasts, is violated because the nowcast is completely determined by the estimated parameters and the contemporaneous values of GDP. The latter are the GDP nowcasts published by the IMF in its World Economic Outlook. Small changes in their outlook can have a large impact on the MVA nowcasts. Also *R2* is not fulfilled, since under this model the nowcast of MVA does not depend on past values of MVA and therefore the nowcast of

MVA can be very different from the most recently observed value of MVA. The third requirement that the nowcast of MVA should be accurate for all countries, is violated because for many countries (i) the share of MVA in GDP is not stable and hence the parameters in (1)-(2) are likely to be time-varying and/or (ii) the log MVA and log GDP are non-stationary series that are not co-integrated. A final disadvantage of models (1)-(2) is that the forecast based on taking the exponential of the fit for $\log MVA_{i,t}$ is downward biased if the error term is normally distributed [7]. In practice we observed that (without the bias correction) the MVA nowcasts have already an upward bias and hence imposing the bias correction will only amplify this upward bias. For our application, the bias correction is thus not useful.

As an alternative for the econometric models in (1)-(2) we consider models based on the following general representation of MVA:

$$MVA_{i,t} = MVA_{i,t-1}(1 + gMVA_{i,t}), \tag{3}$$

where the MVA growth rate $gMVA_{i,t} = (MVA_{i,t} - MVA_{i,t-1}) / MVA_{i,t-1}$ is a time series for which we consider next four possible specifications. The advantage of modeling the growth rate rather than the log of MVA is that the growth rate is stationary. Moreover, by construction and provided the GDP growth rate is not extreme, the nowcast of the present values of MVA will not deviate much from the most recently observed value of MVA. Requirement *R2* is thus satisfied.

We only consider simple linear time series models for the MVA growth rate. This choice is motivated by the objective of having a parsimonious nowcasting model. Moreover the extensive comparison in [10] of sophisticated nonlinear time series models for GDP growth and inflation with simple linear time series models has shown that in general simple linear time series models can be hardly beaten if they are carefully specified.

We first of all consider the possibility that the growth of MVA is stationary around its mean a_i, which we allow to be different for each country:

$$gMVA_{i,t} = a_i + e_{i,t}\,, \tag{4}$$

where $e_{i,t}$ is white noise. In the second, third and fourth model, we incorporate the relationship between MVA and GDP growth and/or the observed persistence in MVA growth.

$$gMVA_{i,t} = a_i + b_i gGDP_{i,t} + e_{i,t} \tag{5}$$

$$gMVA_{i,t} = a_i + b_i gMVA_{i,t-1} + e_{i,t} \tag{6}$$

$$gMVA_{i,t} = a_i + b_i gGDP_{i,t} + c_i gMVA_{i,t-1} + e_{i,t}\,, \tag{7}$$

where $gGDP_{i,t} = (GDP_{i,t} - GDP_{i,t-1})/GDP_{i,t-1}$ and $e_{i,t}$ is white noise. Note that the models (4)-(6) are all special cases of model (7).

Estimation. The standard OLS estimator may be biased because of a violation of the assumption of exogeneity of the regressors $\log GDP_{i,t}$ and $gGDP_{i,t}$ with respect to the error term and because of the presence of outliers in the data.

Indeed, as pointed out by the referees, an endogeneity problem arises because in the regression models the contemporaneous value of the log (growth rate) of GDP is supposed to affect the log (growth rate) of MVA, but the value of GDP depends itself on the value of MVA. Because of this circularity, the OLS parameter estimates can be biased. If one is interested in the behavioural interpretation of the parameter estimates, then an instrumental variable estimator should be used. However, for our purpose of nowcasting MVA, the OLS estimator is still preferable (even if it is biased) since it minimizes the mean squared error of prediction and is therefore the best linear predictor (see e.g. Proposition 2.8 in [8]).

The scatter plots of $gMVA_{i,t}$ versus $gGDP_{i,t}$ indicate the presence of outliers for some countries. Fig. 1 illustrates this for the 1991–2007 data for Poland. In the transition years 1991 and

1992 the MVA and GDP growth rates are extreme. In 1991 and 1992 the MVA (GDP) growth rates equal -16.3% (-7.0%) and 80.2% (2.6%), respectively. For all other years the MVA (GDP) growth rates are between -0.6% (1.2%) and 13.8% (7.1%). The 1991 and 1992 MVA growth rates are clearly univariate outliers, but from a bivariate perspective only the 1992 observation is an outlier with respect to the general correlation pattern observed in the data. The estimation of the regression models using Ordinary Least Squares (OLS) is known to be problematic in the presence of outliers. As can be seen in Fig. 1, the 1992 observation tilts the OLS slope estimate to its position and yields a distorted estimate of the regression line fitting the bulk of the data. The OLS estimator thus does not satisfy the requirement that the influence of single observations on the nowcast should be small. For this reason, we also consider a robust alternative to the OLS estimator, namely the MM estimator.

The robust MM estimator is a two-step estimator. First, it estimates the parameter vector that minimizes the sum of the 50% smallest squared residuals. This 50% Least Trimmed Squares (LTS) estimate then serves as the starting value for the M-estimation, where a loss function is minimized that downweights outliers. The MM estimator has a high efficiency under the linear regression model with normally distributed errors. Because it is initialized at the LTS estimates, it is also highly robust to outliers (see e.g. Chapter 5 in [11]). In Fig. 1 the OLS and robust MM regression estimates are compared. We see that, in contrast with the OLS estimator, the robust MM estimate is rather insensitive to the outlying observations and produces an accurate fit of the bulk of the data.

3 Results: nowcast accuracy comparison

In Section 2 we have presented 6 econometric models and 2 choices of the estimator. This yields thus a total of 12 possible methods for nowcasting MVA. As an additional benchmark strategy, we also consider the nowcast based on a random walk, i.e. the strategy of setting the nowcast of MVA equal to the observed (or predicted, if not available) value of MVA of the previous year. Table 1 presents the results of a pseudo out-of-sample nowcast accuracy comparison between the methods for the years 2004-2007. For each of these years and for every country, we do as if the MVA data were missing and construct the one step ahead nowcast. For each nowcasting method, we then compute the Mean Absolute Percentage Error (MAPE) and the proportion of observations for which the Absolute Percentage Error (APE) exceeds 10% and 20%. The MAPE measures the average accuracy of the estimator while the proportion of observations for which the APE exceeds 10% and 20% indicate in how many cases we get a very bad estimate (cfr. requirement *R3*).

The analysis is based on the 200 countries in the dataset used to produce the nowcasts published in the 2009 edition of the *International Yearbook of Industrial Statistics*. We present the performance measures not only as an aggregate over all countries, but also after splitting up the sample into the countries for which the share of MVA in GDP in 2003 was below (resp. above) its median value of 12.6%. This allows us to check whether the choice of nowcasting method should depend on the country's share of MVA in GDP.

Let us first focus on the criteria evaluated on all countries. We have the following results.

1. According to all three criteria, the choice of dependent variable in the econometric model is the most important decision when designing the nowcasting method. The MAPE of the nowcast based on the regression of the growth rate of MVA is 3 to 4 times smaller than the nowcast based on the

regression of the log of MVA. The proportion of observations for which the APE exceeds 10% (20%) is approximately 5 (10) times smaller.

2. For both the OLS and MM estimator, the nowcast based on the regression model for the MVA growth rate with intercept and contemporaneous GDP growth has the lowest MAPE.

3. According to all criteria the most accurate MVA nowcasting procedure is the one that uses the model for $gMVA_{i,t}$ with an intercept and contemporaneous GDP growth and with parameters estimated by the robust MM estimator. Its MAPE equals 4.1% and in only 2.7% of all cases, its APE is above 20%. This is our preferred method.

The separate evaluation of the performance criteria for the countries with share of MVA in GDP below and above the median shows that the performance of all nowcasting methods deteriorates (improves) for countries with a low (high) share of MVA in GDP. For both groups of countries, our preferred nowcasting approach based on a robust estimate of the regression model for $gMVA_{i,t}$ with an intercept and contemporaneous GDP growth has the lowest MAPE. For this method, the countries for which the APE exceeds 20% are not key economies. Moreover the extreme prediction errors are almost all isolated in time, indicating that these errors are not due to model misspecification but rather to large shocks in the MVA of that country.

In practice, countries for which the APE has once exceeded 10% are put on an "intensive care list". For the countries on that list, the research and statistics branch of UNIDO evaluates case by case the plausibility of the nowcast produced by the regression method and, if not, expert knowledge is used to adjust the nowcast.

4 Conclusion

To allow cross-country comparison of the current industrial economic situation, nowcasts of manufacturing value added are published in the *International Yearbook of Industrial Statistics*. We consider nowcast methods that exploit the relationship between MVA and GDP and the fact that accurate nowcasts of current GDP are available from external sources. The nowcast accuracy comparison made in this paper shows that the approach using (i) the econometric model that specifies the conditional expectation of the yearly MVA growth rate as a linear function of the contemporaneous GDP growth rate and (ii) a robust estimation method has the best performance of all considered methods.

Acknowledgements

We thank Christophe Croux, Markus Froehlich, two anonymous referees and the editor for helpful comments and suggestions.

References

[1] Artis, M.J. (1996) How accurate are the IMF's short-term forecasts? Another examination of the world economic outlook. IMF working paper 96/89.

[2] Baffigi, A., Golinelli, R. and Parigi, G. (2004) Bridge models to forecast the euro area GDP. *International Journal of Forecasting*, **20**, 447-460.

[3] Blake, A., Kapetanios, G. and Weale, M.R. (2002) Nowcasting EU industrial production and manufacturing output. National Institute of Economic and Social Research.

[4] Chamberlin, G. (2007) Forecasting GDP using external data sources. *Economic & Labour Market Review*, **1**, 18-23.

[5] Ferrara, L. (2007) Point and interval nowcasts of the Euro area IPI. *Applied Economics Letters*, **14**, 115-120.

[6] Giannone, D., Reichlin, L. and Small, D. (2008) Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, **55**, 665-676.

[7] Granger, C.W.J. and Newbold, P. (1976) Forecasting transformed series. *Journal of the Royal Statistical Society B*, **38**, 189-203.

[8] Hayashi, F. (2000) Econometrics. Princeton University Press, Princeton.

[9] Kitchen, J. and Monaco, R.M. (2003) Real-time forecasting in practice: the U.S. treasury staff's real-time GDP forecast system. *Business Economics*, **38**, 10-19.

[10] Marcellino, M. (2009) A comparison of time series models for forecasting GDP growth and inflation. *Journal of Forecasting*, forthcoming.

[11] Maronna, R.A., Martin, R.D. and Yohai, V.J. (2006). Robust Statistics: Theory and Methods. Wiley, New York.

[12] Schumacher, C. and 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.

[13] UNIDO (1996) Industrial Statistics Database: Methodological notes. UNIDO, Vienna.

[14] Upadhyaya, S. and Todorov, V. (2008) UNIDO Data Quality. UNIDO, Vienna.

Table 1. Nowcast accuracy over the years 2004-2007 and 200 countries for six regression specifications and the OLS and MM estimation methods. The results are presented once for all 200 countries together, and once distinguishing between the countries with share of MVA in GDP in 2003 below (above) its median value of 12.6%.

Dependent variable	$\log MVA_{i,t}$	$\log MVA_{i,t}$	$gMVA_{i,t}$	$gMVA_{i,t}$	$gMVA_{i,t}$	$gMVA_{i,t}$
Regressors (besides intercept)	$\log GDP_{i,t}$	$\log GDP_{i,t}$ $(\log GDP_{i,t})^2$		$gGDP_{i,t}$	$gMVA_{i,t-1}$	$gGDP_{i,t}$ $gMVA_{i,t-1}$
Estimator	Mean absolute percentage error					
OLS (all countries)	16.8	17.3	5.1	4.4	4.7	4.4
(MVA/GDP≤12.6%)	21.6	21.8	5.9	5.8	5.9	5.9
(MVA/GDP>12.6%)	12.2	11.7	4.4	3.3	3.6	3.3
MM (all countries)	16.2	18.7	4.7	4.1	4.4	4.7
(MVA/GDP≤12.6%)	21.1	25.0	5.8	5.5	5.9	6.3
(MVA/GDP>12.6%)	11.2	11.0	3.9	2.9	3.5	3.2
Proportion of observations for which the absolute percentage error exceeds 10%						
OLS (all countries)	54.5	40.9	13.9	12.1	10.4	11.8
(MVA/GDP≤12.6%)	65.3	48.6	18.5	19.4	13.9	18.2
(MVA/GDP>12.6%)	43.2	32.9	9.1	4.5	6.6	5.1
MM (all countries)	52.1	38.7	11.7	10.5	10.8	13.3
(MVA/GDP≤12.6%)	64.5	46.2	17.1	16.8	15.0	19.9
(MVA/GDP>12.6%)	39.3	30.8	6.0	3.9	6.0	5.4
Proportion of observations for which the absolute percentage error exceeds 20%						
OLS (all countries)	29.5	18.3	3.2	3.1	2.8	3.5
(MVA/GDP≤12.6%)	40.2	23.4	5.5	5.2	4.6	6.4
(MVA/GDP>12.6%)	18.4	13.0	0.9	0.9	0.9	0.6
MM (all countries)	26.6	19.1	2.7	2.7	2.7	3.2
(MVA/GDP≤12.6%)	39.0	25.7	4.9	5.2	4.3	6.1
(MVA/GDP>12.6%)	13.6	12.1	0.3	0	0.9	0.6

Note: the benchmark strategy of setting the MVA nowcast equal to the value of the MVA of the previous year yields a MAPE equal to 6.0% and 16.5% (2.8%) of observations for which the APE exceeds 10% (20%).

Fig. 1 Scatter plots of the 1991-2007 yearly GDP and MVA growth rates for Poland, together with the OLS and robust MM regression fit