Marcus Scheiblecker, Sandra Steindl

Austrian GDP Flash Estimates: A Description of Methods

The publication of short-term economic indicators has risen in importance in recent years. Quite in line with this, more and more countries are now producing a reduced data set of national accounts variables right after the end of a quarter. Since summer 2005, the Austrian Institute of Economic Research has been producing such GDP flash estimates for Austria on behalf of the Austrian Federal Ministry of Finance. A variety of methods are being used, fully embedded in the regular national accounts framework.

Marcus Scheiblecker and Sandra Steindl are economists at WIFO. **The authors are grateful to Serguei Kaniovski for his helpful and constructive comments. The data were processed and analysed with the assistance of Christine Kaufmann.** **• E-Mail addresses: Marcus.Scheiblecker@wifo.ac.at, Sandra.Steindl@wifo.ac.at**

The rapid publication of aggregate economic data has gained in importance recently. In order to meet this demand, a number of countries are producing quarterly GDP estimates with a very short delay after the end of the reference period. For the same reason, the European Commission requested its Member States to compile quarterly national accounts flash estimates. These are not intended to replace the regular calculations (which are based on an EC regulation and are to be transmitted no later than 70 days after the end of the reference quarter) but should be supplementary. The delay of those flash estimates is recommended to range between 30 and 45 days after the end of the reference quarter.

Despite the fact that the compilation and publication processes for the basic data required for calculating national accounts have been speeded up considerably in recent years, they are completed far too late to allow a calculation of quarterly GDP in the usual manner at this early point in time. For that reason, different methods have to be employed for flash estimates in order to supply sufficiently reliable economic information sooner.

In an international perspective, the methodological toolbox for calculating flash estimates is large, ranging from methods that involve a large amount of subjectivity to strict formal approaches[1]. There are also notable differences concerning the point of departure and the degree of detail. Some countries embody their calculations in the framework of regular QNA compilations, while others set up external models for generating flash estimates or at least some GDP components[2].

As nowadays many countries produce quarterly national accounts flash estimates, and several already have many years of experience in this field, it was possible to do some pilot studies on these methods. A promising, albeit very labour-intensive approach appears to be to base the estimation on the existing framework of quarterly national accounts. The great advantage of this approach, employed for instance by Italy[3], lies in the possibility of tracing the typical data generation process of the regular estimates. Furthermore, it is also possible to revise the results of the previous quarter during this process, which normally has to wait until the production of the

[1] See *Statistisches Bundesamt Wiesbaden* (2003), *Hartmann – Schmidt – Otmanns* (2005).

[2] Examples for this group of external approaches are the SAFE Model of the *Centraal Planbureau (2003)* and in *Cors – Kouzine (2003)*.

[3] Such a method was presented by *Savio* (2002) at the *Workshop on Flash Estimates of QNA* held in Luxembourg in December 2002.

next regular QNA[4]. This can decisively improve the quality of flash estimates, as they often rely on time series analysis, where great importance is attached to recent observations. Probably the most important reason justifying this approach is that flash estimates do not have to be calculated by regressing on several indexes for which some relation to national accounts figures has to be assumed beforehand but needs yet to be tested. Instead, this concept concentrates on indicators used in the regular QNA calculations. After a forecast of these indicator series to extend their time horizon they can be used within the regular framework in the usual manner. This reduces the danger of data mining and spurious regression. A further advantage of this direct approach is that it improves the efficiency of the estimates, as already published monthly figures covering some part of the respective quarter are included. *Rünstler – Sédillot* (2003) criticised bridge models frequently used in the estimation of QNA aggregates for having to wait for data for all months of a quarter to be available in order to be considered in the estimation process. They presented an alternative method – quite in line with the approach used by WIFO – which completes the missing months of the respective quarter.

Table 1: Variables to be estimated by WIFO flash estimates

	Nominal		Real[1]	
	Unadjusted	Adjusted[2]	Unadjusted	Adjusted[2]
GDP, output approach				
Gross value added NACE A + B	✓	✓	✓	✓
Gross value added NACE C + D + E	✓	✓	✓	✓
Gross value added NACE F	✓	✓	✓	✓
Gross value added NACE G + H + I	✓	✓	✓	✓
Gross value added NACE J + K	✓	✓	✓	✓
Gross value added NACE L + M + N + O + P	✓	✓	✓	✓
Taxes on products	✓	✓	✓	✓
Subsidies on products	✓	✓	✓	✓
GDP, expenditure approach				
Private consumption expenditure	✓	✓	✓	✓
Government consumption expenditure	✓	✓	✓	✓
Gross capital formation	✓	✓	✓	✓
Exports	✓	✓	✓	✓
Imports	✓	✓	✓	✓
GDP, income approach	✓	✓		
Compensation of employees	✓	✓		
Gross operating surplus and mixed income	✓	✓		
Taxes on production and imports less subsidies	✓	✓		
Labour market				
Number of employees in NACE A + B			✓	✓
Number of employees in NACE C + D + E			✓	✓
Number of employees in NACE F			✓	✓
Number of employees in NACE G + H + I			✓	✓
Number of employees in NACE J + K			✓	✓
Number of employees in NACE L + M + N + O + P			✓	✓

[1] Monetary aggregates: real growth rates, chained, reference year 2000. – [2] Adjusted for seasonal and working day effects.

Internationally, significant differences exist also with respect to the details to be calculated and published. Growth rates of real GDP adjusted for seasonal and working day effects are part of all of these estimations, which are further disaggregated either by production or expenditure side components. This disaggregation is the main characteristic that distinguishes flash estimates from coincident business cycle indicators. As the computation of Austrian GDP flash estimates is embedded in the regular QNA framework, it allows calculating a wide variety of aggregates and breakdowns. In view of the data demand of Austrian economic policy and the availability of the data, it was decided to compile flash estimates from all three sides: output, expenditure and income side; estimates of some employment figures are also included. This labour-intensive process can be justified, as there exist many synergies between regular QNA and flash estimates.

[4] Surely, this advantage can only be exploited if the regular calculation and flash estimates are compiled within one institution, as is the case for Austria.

Basically, the approach presented here uses multivariate time series models for extending only those series that are used in the regular calculation of QNA. Unlike univariate time series models, they allow new external information to be considered in estimates for the missing months. This improves the possibility of detecting cyclical turning points at an early stage, which univariate time series models do not offer. Using external leading or coincident data is probably the most important feature, as the detection of turning points can be considered the main reason behind the call for a more timely publication of QNA results[5].

For Austria, the most important external data source that can be resorted to when extending the indicators used in regular QNA estimation, is the monthly business survey conducted by WIFO on behalf of the European Commission. These survey data can be assumed to contain not only information about changing production activities in several sectors but also about turning points of the business cycle.

This applies especially to the survey in the manufacturing sector (NACE D) and in the construction industry (NACE F). As the results in the appendix indicate, calculated test statistics confirm the significantly better performance of forecasts including these data compared to univariate ARIMA forecasts. This makes overall flash estimates quite promising, as these two sectors not only account for around 40 percent of Austrian GDP, but are held responsible for a large part of business cycle volatility. Due to their importance, their estimation procedure is explained separately in the appendix.

For some of the components to be compiled, no special forecast of indicators is necessary. Either they are already available in the required length at the time of compilation of the flash estimates or they are not even available at the time of regular QNA calculation. The first goes for example for employment data released by the Federation of Austrian Social Security Institutions, which are available already a few days after the end of a quarter, and therefore no forecast is necessary. This data constitutes the backbone of the calculation of QNA employment figures, so this component is of the same quality in flash estimates as in the regular QNA. A similar example are wage deflation indexes by sectors underlying QNA gross salaries and other price statistics necessary for deflating certain components.

In the regular QNA, the indicator variable for "quarterisation" of taxes on products comprises the tax revenues of the federal government. This data covers the last two months of a quarter and the first of the following, which is sufficient for most types of taxes on products. The principle of accrual concerning VAT requires instead assigning only the last month of a specific quarter to the reference quarter plus the first two months of the following. For estimating the two missing months, a simple univariate seasonal ARIMA model is used. Furthermore, this type of estimation is used for all aggregates at the production side of national accounts for which neither employment data nor a regression on business survey data turned out to be appropriate.

Like in the regular QNA, the indicator variables for government consumption expenditures are the employment figures from the public administration, education, healthcare and social work sectors. The methodology used for household consumption expenditure is explained in the appendix.

For the export of goods, the trade statistics published by Statistics Austria are extended with univariate time series models. In addition, qualitative information coming from the manufacturing sector and the wholesale and retail trade sector is also considered in the series. In general, these results are not perfectly reliable because the trade statistics from Statistics Austria and the Oesterreichische Nationalbank are usually subject to substantial revisions, which show up in QNA as well. The indicator for estimating tourism exports is the number of overnight stays of foreign guests. The remaining exports of services are estimated with simple time series models. The estimates of imports of goods and services are done analogically.

[5] This advantage must not be overrated in practice, as cyclical turning points will be recognised through monthly observation only after some time. So there will be a time lag between their registration and their recognition.

Investment in construction is estimated using information gathered from the output of the construction industry. The remaining components of fixed capital formation go along with developments in the output of the industries producing them plus net imports.

Generally, results obtained from the production approach of national accounts are statistically more reliable compared to the expenditure approach. This holds true for the annual and quarterly NA as well as for flash estimates. One shortcoming in our approach of flash estimates is that the data used in the regressions for this forecast were already subject to repeated revisions (real time data are missing). This may possibly result in a small systematic bias.

Flash estimation results for the output approach of the SNA are calculated at current and previous year prices. The latter provide a basis for computing real growth rates. In order to generate absolute values, growth rates are chained with the nominal values of the reference year. All values are seasonally and working-day adjusted to allow comparison with the preceding period in order to draw a meaningful conclusion about the state of the business cycle.

The method presented here can be regarded as a first attempt to set up a reliable framework for GDP flash estimates. With growing experience reasonable adjustments concerning the models as well as the indicators used can be expected.

Appendix: estimation methods

Gross value added in manufacturing (NACE D)

The indicator underlying the "quarterisation" of gross value added in manufacturing (NACE D) in the regular QNA framework is the business survey on nominal production of the respective sector published by Statistics Austria on a monthly basis. At the time when flash estimates are compiled, only the first month of the respective quarter is available. For this reason this time series has to be extended by two more months in order to obtain an indicator that can be used in a similar manner as in the regular QNA calculation. It had to be decided whether to set up simple univariate seasonal ARIMA models (SARIMA) or methods that consider external coincident or leading variables. As the latter is more appropriate to potentially capture turning points on the recent margin, this type of approach was favoured. An Autoregressive Distributed Lag Model (ADL), including lagged endogenous as well as lagged and coincident exogenous variables, turned out to be most favourable[6].

Table 2: Statistics for manufacturing (NACE D)

Dependent Variable: $KEDSUMLOG12$
Method: Least Squares
Sample (adjusted): April 1997 to April 2005
Included observations: 97 after adjustments

	Coefficient	Standard error	t statistic	Probability
$(PR3MO\_AB250(_{-1}))$	0.092637	0.017525	5.286049	0.0000
$(KEDSUMLOG12_{-1})$	0.680800	0.058211	11.69536	0.0000
$DUM200401$	0.011234	0.004772	2.354307	0.0206
R^2	0.891856	Mean dependent variation		0.047283
Adjusted R^2	0.889555	Standard deviation dependent variation		0.046090
Standard error of regression	0.015317	Akaike info criterion		– 5.489217
Sum squared residual	0.022055	Schwarz criterion		– 5.409587
Log likelihood	269.2270	Durbin-Watson statistics		1.980995

After comparing several specifications, a discrimination based on several test statistics gave favour to a model using 12th differences in logs of cumulated past three months' WIFO business survey data as external variable $(PR3MO\_AB250(_{-1}))$[7] and a lagged endogenous term $(KEDSUMLOG12_{-1})$. Furthermore, a structural break that happened in first quarter 2004 was detected and included by means of a dummy

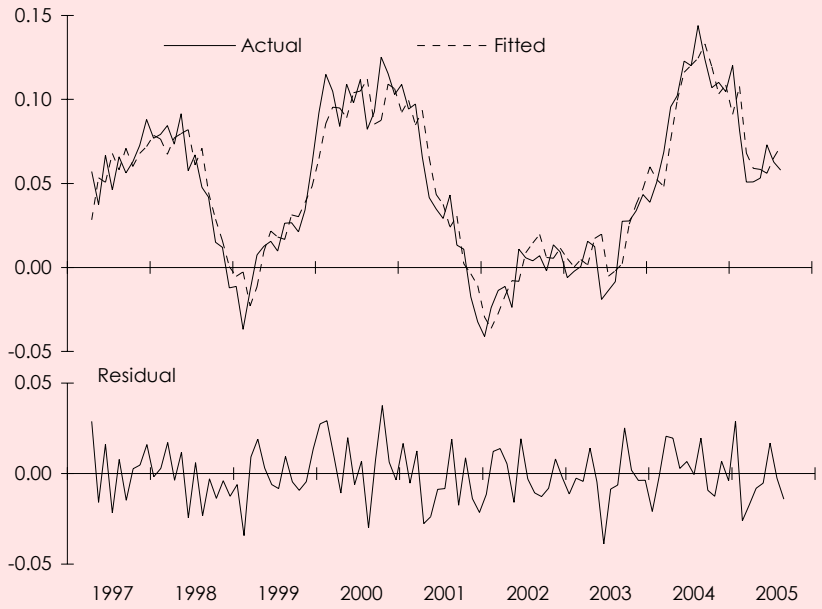
[6] This type of model was also employed by *Rünstler – Sédillot* (2003) in producing flash estimates.

[7] It turned out that considering only enterprises with more than 250 employees improved the significance of parameters. The relevant question from this survey concerned the past three months' production qualitative indicator.

variable (*DUM200401*). Employment figures did not improve the estimation, which seems to make sense as employment should theoretically lag production.

Figure 1: Indicator for the nominal output in manufacturing (NACE D)

Differences in logs from preceding year



Source: WIFO.

Despite the necessarily parsimonious character of the model (data starting at the beginning of 1996), it proved to have a rather good forecast ability. Besides, the model output fitted the original time series quite well, as can be seen in Figure 1.

Gross value added in construction (NACE F)

In the regular QNA, census surveys provided by Statistics Austria are used as an indicator to calculate the nominal gross value added in construction, as is the case in manufacturing. At the time of the compilation of flash estimates, two months of this indicator have to be estimated. This is done by regressing the year-on-year growth rates of the cumulated three months' results of the census surveys on the results of the WIFO business cycle survey (regarding the question about construction output volume in the last three months). Besides, seasonal ARIMA terms of the order (1,0,1) (0,0,1) are included[8]. Again, employment did not show to be of any significant value in improving the model.

Private household consumption

In the regular QNA, household consumption expenditure is classified into 12 groups according to the COICOP system. Sources for the estimation of the value added are usually sales and revenue statistics of the respective retail trade. As at the time when the flash estimates are compiled the quarterly retail sales results published by Statistics Austria are not yet available for the whole reference quarter, the remaining months have to be forecasted. This is done by means of univariate time series models, econometric techniques as well as combinations thereof (Reg-ARIMA models).

When employing time series models, a multivariate approach is used. In the deterministic part calendar effects (like holidays and the Easter effect) and special effects (like certain developments in gross salaries) are considered. Special events are identified as outliers. This approach offers an advantage compared to interventions analysis; the time of the special effects does not have to be known[9]. Outlier detection techniques in an iterative approach simultaneously estimate the parameters of the model as well as the outlier effects. In every step every observation of the time

[8] Unlike in the estimation for the goods producing sector, where an ADL model was used, here the error term is modelled with a seasonal ARIMA process.

[9] See, for example, *Chen – Liu – Hudak* (1990), *Thury – Wüger* (1992) and *Brandner – Schuberth* (2000).

series is tested for an additive outlier (an incident that happens just at a certain time in the series), a level shift (a persistent change in the data generating process), an outlier in the error term (an innovation in the process) or a transitory effect (which dies away as time proceeds). After identifying the form of the outlier, it will be corrected appropriately. These three steps, i.e., identifying the outliers, adjusting for them, and estimating the parameters, are repeated until no more outlier is found.

Table 3: Test statistics for construction (NACE F)

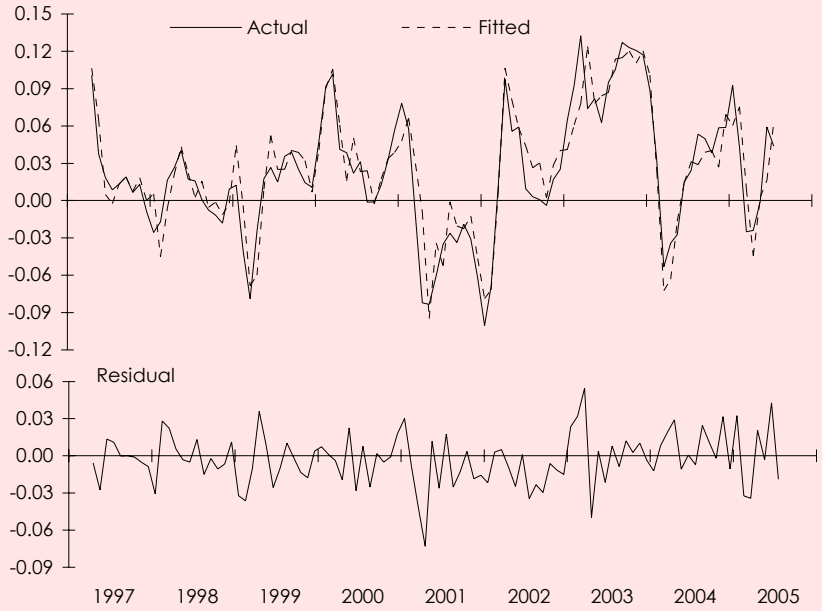
Dependent Variable: *KEDSUMLOG12*
Method: Least Squares
Date: 08/10/05 Time: 15:13
Sample (adjusted): 1997M04 2005M04
Included observations: 97 after adjustments
Convergence achieved after 16 iterations
Back cast: 1996M03 1997M03

Variable	Coefficient	Standard error	*t*-Statistic	Probability
BAUVMO_3MO12	0.059824	0.012257	4.880585	0.0000
C	0.022690	0.004075	5.568213	0.0000
AR(1)	0.712801	0.085291	8.357268	0.0000
MA(1)	0.305608	0.118812	2.572194	0.0117
SMA(12)	– 0.881378	0.032485	– 27.13223	0.0000

R-squared	0.839418	Mean dependent variation	0.021674
Adjusted *R*-squared	0.832436	S.D. dependent variation	0.050656
S.E. of regression	0.020736	Akaike info criterion	– 4.863753
Sum squared residual	0.039557	Schwarz criterion	– 4.731036
Log likelihood	240.8920	*F*-statistic	120.2289
Durbin-Watson statistics	1.955641	Probability (*F*-statistic)	0.000000

Figure 2: Indicator for the nominal output in construction (NACE F)

Differences in logs from preceding year



Source: WIFO.

After removing the deterministic part, the rest of the series is specified as an ARIMA model including a regular and a seasonal part. The regular part includes the trend and the cycle, whereas the seasonal part accounts for the regular effects occurring during the year. Hence, a multivariate time series model uses calendar effects and cyclical, trend, and seasonal factors for the explanation of sales developments.

When econometric causality models are used to forecast sales figures for certain branches of the retail sector, the figures are explained with the total sales statistics for the retail sector as a whole (which is easier to forecast). The differences in the seasonal pattern of a given branch and of the retail sector as a whole are taken into account. Apart from that, appropriate business cycle indicators (retail sales) and already available sales figures from KMU-Forschung[10] are included. If diagnostic checks of the residuals are not satisfactory, they are modelled with an appropriate ARIMA Model (thus a Reg-ARIMA model is constructed).

At the end all the respective results (from time series models, econometric or Reg-ARIMA approaches) are pooled, resulting in a general outcome which now includes information obtained by the different procedures.

References

Brandner, P., Schuberth, H., Analyse der Geldmengenentwicklung im Euro-Gebiet, study of WIFO commissioned by the Austrian National Bank, Vienna, 2000.

Centraal Planbureau, "SAFE – A Quarterly Model of the Dutch Economy for Short-term Analyses", CPB Document, 2003, (42).

Chen, Ch., Liu, L.-M., Hudak, G.B., Outlier Detection and Adjustment in Time Series Modelling and Forecasting, Scientific Computing Associates, Lisle, 1990.

Cors, A., Kouzine, V., "An Approach for Timely Estimations of the German GDP", Allgemeines Statistisches Archiv, 2003, 87.

Hartmann, N., Schmidt, J., Otmanns, E., Schnellschätzung für das Bruttoinlandsprodukt: Ergebnisse einer Machbarkeitsstudie", Wirtschaft und Statistik, 2005, (7).

Rünstler, G., Sédillot, F., "Short-Term Estimates of Euro Area Real GDP by Means of Monthly Data", European Central Bank Working Paper Series, 2003, (276).

Savio, G., "Flash Estimates of QNA in Italy", paper presented at the "Workshop on Flash Estimates of QNA in Luxemburg", Italian National Institute of Statistics, National Accounts Directorate, Rome, 2002.

Statistisches Bundesamt Wiesbaden, "Feasibility Study on Rapid Quarterly GDP Estimates", Study commissioned by the Statistical Office of the European Communities, Wiesbaden, 2003 (mimeo).

Thury, G., Wüger, M., "Outlier Detection and Adjustment – An Empirical Analysis for Austrian Data", Empirica, 1992, 19(1).

Austrian GDP Flash Estimates: A Description of Methods – Summary

In order to meet the rising demand for short-term economic data, the Austrian Federal Ministry of Finance requested the Austrian Institute of Economic Research to calculate in-advance estimates of national quarterly GDP. The first quarterly national accounts data were published in summer 2005 with a delay of no more than 45 days after the end of the reference period. Publication dates are coordinated with Eurostat so as to guarantee the simultaneous publication of the results with the other member states.

As WIFO also produces official Austrian quarterly national accounts data, it was decided to use a method embedded in this framework. Therefore, no separate external model was set up, but indicator series used in the regular production of quarterly national accounts were extended instead. For the manufacturing and construction industries, which account not only for more than 40 percent of the Austrian GDP but also for the largest part of the variance, Autoregressive Distributed Lag models and Reg-ARIMA models were developed for the forecast of the indicator series. By using Austrian business cycle survey data in these models, external information potentially containing turning point information on the recent time margin is considered.

The quarterly GDP flash estimate covers the output approach of GDP as well as some large aggregates of the expenditure and income side, and also includes an estimate on some employment figures.

[10] This is private institute conducts business surveys especially in small and medium sized enterprises.