UNITED NATIONS ECONOMIC COMMISSION FOR EUROPE **Working paper 3**
English only

CONFERENCE OF EUROPEAN STATISTICIANS

OECD Meeting of National Accounts Experts
UNECE Special session on issues of concern to economies in transition
(OECD, Paris, 10 October 2003)

GDP FLASH ESTIMATES - EXPERIENCE OF STATISTICS LITHUANIA

Gailute Juskiene,
Statistics Lithuania

1. Introduction

Release of the quarterly GDP flash estimates was introduced by Statistics Lithuania in 2000 according to the Government Resolution No 334 adopted on 23 March 2000. GDP at current prices in billion LTL and the growth rate of GDP at constant prices in comparison to the corresponding period of previous year are published. Estimations are made over 25-30 days after the end of the quarter.

The Econometric Study Unit is responsible for this work. GDP estimates by production approach – at current and constant prices – are made at this early stage. NACE section level (17 kind of activities) is considered in the estimation process. In order to evaluate the accuracy of the result, deflators for each kind of activity are forecasted. Different econometric (mostly regression) models are used to estimate Output and Value Added for each activity.

2. Availability of data sources

The reality is that statistical data needed for calculations are rather limited at the time when the calculations have to be done. Preferences are given to GDP production approach because information from this point of view is more reliable. Short-term indices (monthly and quarterly if available) or other regressors that could reflect the paths of changes of the particular kind of activity are used. Nevertheless, for some kind of activity information at this time is available only for 2 months of the quarter in question. For some of them there is no on time information. The flash estimate of GDP by expenditure approach is used for comparison and checking purposes only.

A number of assumptions have to be made before starting the estimations. For each kind of activity Value Added is obtained as the difference between estimated Gross Output and Intermediate Consumption. The regressors chosen mainly are related to the Output, and in particular to the main part of it – sales.

The first assumption is related to Intermediate Consumption. It is assumed that the ratio of Intermediate Consumption to Gross Output is constant or slightly changes over a quarter.

The second one is related to institutional sectors. It is assumed that the structure of production created by different institutional sectors is approximately constant over the short-time period. The part of production made by self-employed and small (or unincorporated) enterprises is the same as that of the corresponding period of previous year.

Besides, it is assumed that changes in inventories are also similar. This is important because the short-term statistics mainly relates to the sales not the output.

Different units of the statistical office provide their produced data for flash GDP estimates. First of all National Accounts quarterly time series produced within 90 days are used. The availability of monthly and quarterly information on labour statistics, price

statistics, construction, industry, agricultural, transports, other kinds of services, foreign and domestic trade statistics as well as data from the Central Bank and Ministry of Finance is very important for the estimates. In 30 days after the end of quarter some data required are only available for 2 months but not for the whole quarter. In this case the third month is extrapolated by ARIMA models.

There is no reliable regressor for modeling GDP expenditure components in time. This is why they are all forecasted.

Annex 1 presents the main data used as the regressors (explanatory variables) for the flash GDP estimates by the production approach at current and constant prices.

3. Selecting the model

The software applied for flash estimate is based on SAS (Statistical Analytical System) and Excel. For the estimate of Value Added in a particular kind of activity one of the 15 econometric models is used (See annex 2). In practice, two options are taken into consideration:

If there is no reliable regressor – ARIMA models are used;

When explanatory variables (regressors) are available – several methods (regression, co-integration and others, selected as appropriate) are applied.

It is possible to determine the best model every time when the estimates are made. It is possible also to fix one model for several quarters and, at the same time, to adjust the parameters of this model every quarter. Practical experience revealed there is no sense in fixing the model for a period longer than one year. The reason is the instability of the time series when new observations became available. The best model is considered as that with the smallest average absolute relative error of model for last 1-4 quarters.

4. Results achieved and problems encountered

As mentioned above, there is a target of producing estimates of overall GDP levels (nominal and real), and growth rates, within 30 days after the end of the quarter. The results of estimates by kind of activity are not published (See Annex 3 – the latest Press Release on GDP flash estimate). In general, the difference between GDP flash estimates and preliminary National Accounts' data is not significant (See Table 1). The average absolute relative error of GDP estimates for the period of more than two years (QI 2000 – QIII 2002) was 0.98%, and 0.57% in real terns if the flash and preliminary (within 90 days) estimates of nominal GDP levels were compared.

Table 1

	GDP flash estimates (over 30 days)		NA data (over 90 days)	
	In mill. LTL	Growth, %	In mill. LTL	Growth, %
2001	**47800**	**5.7**	**47968**	**5.9**
1Q	10600	2.8	10684	3.5
2Q	12098	5.6	12093	6.6
3Q	12726	5.1	12745	5.3
4Q	12313	7.9	12446	7.9
2002	**50704**	**5.9**	**50679**	**6.7**
1Q	11055	4.1	11176	4.5
2Q	12597	6.9	12750	6.9
3Q	13428	6.6	13433	6.8
4Q	13345	5.4	13357	7.0
2003				
1Q	12217	9.2	12213	9.4
2Q	13395	6.1		

When the breakdown by kind of activities is analysed, the problematic areas can be identified. Kind of activities that make up the biggest share in the structure of Total Value Added are estimated often with an average absolute relative error less than 3 per cent. Those are mainly: agriculture, industry, construction and trade activities. The estimates of other market and non-market services are less reliable. It is much more difficult to find plausible regressors for them.

Another problematic area is the result of GDP estimates for the fourth quarter. The estimation errors for that period were almost always higher compared with the other quarters. This is hardly explainable to the users. The reason for this is that the results of GDP flash estimates are compared with the already adjusted fourth quarter results for the previous year in the National Accounts.

Annex 1

	GDP breakdown by activities	Estimation regressors
A-B	Agriculture, hunting, forestry and fishing	Volume index of total agricultural production, compared to corresponding period of previous year at constant prices – for constant price estimates; Purchases of animals (live weight) in tones; milk in tonnes and eggs (market places included) in mill. – for current price estimates;
C-D	Mining, quarrying and manufacturing	Sales of mining, quarrying and manufacturing (VAT and excise excluded), at current prices, mill. LTL – for current price estimates; Volume indices by branches as compared to corresponding period of previous year at constant prices – for constant price estimates;
E	Electricity, gas and water supply	Sales of electricity, gas and water supply (VAT and excise excluded), at current prices in mill. LTL – for current price estimates; Volume indices as compared to corresponding period of previous year at constant prices – for constant price estimates;
F	Construction	Sales of own-account works carried out by construction enterprises, at current prices in mill. LTL (only 2 months data) – for current price estimates;
G	Wholesale and retail trade	Turnover of retail trade (VAT excluded), mill. LTL (only 2 months data) – for current price estimates; Volume index as compared to corresponding period of previous year at constant prices – for constant price estimates;
H	Hotels and restaurants	Turnover of restaurants, bars and other catering enterprises (VAT excluded), mill. LTL (only 2 months data) – for constant price estimates; Volume index as compared to corresponding period of previous year at constant prices – for constant price estimates;
I	Transport and communication	Loading and unloading of goods in railway stations and in the seaport in tones (2 or 3 months) – for current price estimates; Goods traffic by all modes of transport, t/km
J-P	Other services	There are no good regressors for all components. For some data on monthly average earnings, employment and budget revenues and expenditure are used.
	Additional regressors for general comparisons	All kind of price indices; exports and imports of goods (only 2 months) and some others.

Annex 2

Models used

1. ***TrendSeason*:**

$$Y(t) = a_0 + a_{11}\cdot X_1(t) + a_{12}\cdot X_1(t-1) + a_{13}\cdot X_1(t-4) + \\ + a_{21}\cdot X_2(t) + a_{22}\cdot X_2(t-1) + a_{23}\cdot X_2(t-4) + \\ + a_{31}\cdot X_3(t) + a_{32}\cdot X_3(t-1) + a_{33}\cdot X_3(t-4) + \\ + b_1\cdot T_1(t) + b_2\cdot T_2(t) + c_1\cdot S_1(t) + c_2\cdot S_2(t) + c_3\cdot S_3(t) + \\ + d_1\cdot Y(t-1) + d_2\cdot Y(t-4) + e(t)$$

here and later

$T_1(t)$ = 1, 2, 3, ..., *n*,

$T_2(t) = 1, 1/2, 1/3, ..., 1/n$,

$S_1(t), S_2(t), S_3(t)$ – coefficient of identification of the season, that is equel to 1 for the first, second and third quarters, and to 0 for the next.

2. ***LogTrendSeason*:**

$$\ln(Y(t)) = a_0 + a_{11}\cdot \ln(X_1(t)) + a_{12}\cdot \ln(X_1(t-1)) + a_{13}\cdot \ln(X_1(t-4)) + \\ + a_{21}\cdot \ln(X_2(t)) + a_{22}\cdot \ln(X_2(t-1)) + a_{23}\cdot \ln(X_2(t-4)) + \\ + a_{31}\cdot \ln(X_3(t)) + a_{32}\cdot \ln(X_3(t-1)) + a_{33}\cdot \ln(X_3(t-4)) + \\ + b_1\cdot T_1(t) + b_2\cdot T_2(t) + c_1\cdot S_1(t) + c_2\cdot S_2(t) + c_3\cdot S_3(t) + \\ + d_1\cdot Y(t-1) + d_2\cdot Y(t-4) + e(t)$$

3. ***SimpleRegression*:**

$$Y(t) = a_0 + a_{11}\cdot X_1(t) + a_{12}\cdot X_1(t-1) + a_{13}\cdot X_1(t-4) + \\ + a_{21}\cdot X_2(t) + a_{22}\cdot X_2(t-1) + a_{23}\cdot X_2(t-4) + \\ + a_{31}\cdot X_3(t) + a_{32}\cdot X_3(t-1) + a_{33}\cdot X_3(t-4) + e(t)$$

4. ***LogSimpleRegression*:**

$$\ln(Y(t)) = a_0 + a_{11}\cdot \ln(X_1(t)) + a_{12}\cdot \ln(X_1(t-1)) + a_{13}\cdot \ln(X_1(t-4)) + \\ + a_{21}\cdot \ln(X_2(t)) + a_{22}\cdot \ln(X_2(t-1)) + a_{23}\cdot \ln(X_2(t-4)) + \\ + a_{31}\cdot \ln(X_3(t)) + a_{32}\cdot \ln(X_3(t-1)) + a_{33}\cdot \ln(X_3(t-4)) + e(t)$$

5. ***Residuals*:**

$$Y(t) = b_0 + b_1\cdot T_1(t) + b_2\cdot T_2(t) + c_1\cdot S_1(t) + c_2\cdot S_2(t) + c_3\cdot S_3(t) + \\ + d_1\cdot Y(t-1) + d_2\cdot Y(t-4) + e_Y(t)$$

$$X(t) = b_0 + b_1\cdot T_1(t) + b_2\cdot T_2(t) + c_1\cdot S_1(t) + c_2\cdot S_2(t) + c_3\cdot S_3(t) + \\ + d_1\cdot Y(t-1) + d_2\cdot Y(t-4) + e_X(t)$$

$$e_Y(t) = a_1\cdot e_X(t) + a_2\cdot e_X(t-1) + a_3\cdot e_X(t-4) + d(t).$$

6. ***RegressionAuto*:**

$$\begin{aligned} Y(t) = a_0 &+ a_{11}\cdot X_1(t) + a_{12}\cdot X_1(t-1) + a_{13}\cdot X_1(t-4) + \\ &+ a_{21}\cdot X_2(t) + a_{22}\cdot X_2(t-1) + a_{23}\cdot X_2(t-4) + \\ &+ a_{31}\cdot X_3(t) + a_{32}\cdot X_3(t-1) + a_{33}\cdot X_3(t-4) + \ . \\ &+ \sum_{j=1}^{p} b_i \cdot Y(t-j) + e(t) \end{aligned}$$

7. ***LogRegressionAuto*:**

$$\begin{aligned} \ln(Y(t)) = a_0 &+ a_{11}\cdot \ln(X_1(t)) + a_{12}\cdot \ln(X_1(t-1)) + a_{13}\cdot \ln(X_1(t-4)) + \\ &+ a_{21}\cdot \ln(X_2(t)) + a_{22}\cdot \ln(X_2(t-1)) + a_{23}\cdot \ln(X_2(t-4)) + \\ &+ a_{31}\cdot \ln(X_3(t)) + a_{32}\cdot \ln(X_3(t-1)) + a_{33}\cdot \ln(X_3(t-4)) + \ . \\ &+ \sum_{j=1}^{p} b_i \cdot \ln(Y(t-j)) + e(t) \end{aligned}$$

8. ***Cointegration*:**

$$Y(t) = a_0 + a_1\cdot X_1(t) + a_2\cdot X_2(t) + a_3\cdot X_3(t) + e(t),$$

$$\Delta Y(t) = a_1\cdot \Delta X_1(t) + a_2\cdot \Delta X_2(t) + a_3\cdot \Delta X_3(t) + b_0\cdot \Delta Y(t-1) + \sum_{k=1}^{4} b_k\cdot e(t-k) + d(t).$$

9. ***GrowthRates*:**

$$\frac{Y(t)}{Y(t-4)} = a_0 + a_1\cdot\frac{X_1(t)}{X_1(t-4)} + a_2\cdot\frac{X_2(t)}{X_2(t-4)} + a_3\cdot\frac{X_3(t)}{X_3(t-4)} + e(t).$$

10. ***LogGrowthRates*:**

$$\ln\left(\frac{Y(t)}{Y(t-4)}\right) = a_0 + a_1\cdot\ln\left(\frac{X_1(t)}{X_1(t-4)}\right) + a_2\cdot\ln\left(\frac{X_2(t)}{X_2(t-4)}\right) + a_3\cdot\ln\left(\frac{X_3(t)}{X_3(t-4)}\right) + e(t).$$

11. ***InverseX*:**

$$\begin{aligned} Y(t) = a_0 &+ a_{11}\cdot\frac{1}{X_1(t)} + a_{12}\cdot\frac{1}{X_1(t-1)} + a_{13}\cdot\frac{1}{X_1(t-4)} + \\ &+ a_{21}\cdot\frac{1}{X_2(t)} + a_{22}\cdot\frac{1}{X_2(t-1)} + a_{23}\cdot\frac{1}{X_2(t-4)} + \quad . \\ &+ a_3\cdot\frac{1}{X_3(t)} + a_{32}\cdot\frac{1}{X_3(t-1)} + a_{33}\cdot\frac{1}{X_3(t-4)} + e(t) \end{aligned}$$

12. ***InverseXY*:**

$$\frac{1}{Y(t)} = a_0 + a_{11}\cdot\frac{1}{X_1(t)} + a_{12}\cdot\frac{1}{X_1(t-1)} + a_{13}\cdot\frac{1}{X_1(t-4)} + \\ + a_{21}\cdot\frac{1}{X_2(t)} + a_{22}\cdot\frac{1}{X_2(t-1)} + a_{23}\cdot\frac{1}{X_2(t-4)} + \\ + a_3\cdot\frac{1}{X_3(t)} + a_{32}\cdot\frac{1}{X_3(t-1)} + a_{33}\cdot\frac{1}{X_3(t-4)} + e(t)$$

13. ***InverseY*:**

$$\frac{1}{Y(t)} = a_0 + a_{11}\cdot X_1(t) + a_{12}\cdot X_1(t-1) + a_{13}\cdot X_1(t-4) + \\ + a_{21}\cdot X_2(t) + a_{22}\cdot X_2(t-1) + a_{23}\cdot X_2(t-4) + \\ + a_{31}\cdot X_3(t) + a_{32}\cdot X_3(t-1) + a_{33}\cdot X_3(t-4) + e(t)$$

14. ***ChangesQtoQ*:**

$$\Delta Y(t) = a_0 + a_{11}\cdot\Delta X_1(t) + a_{12}\cdot\Delta X_1(t-1) + a_{13}\cdot\Delta X_1(t-4) + \\ + a_{21}\cdot\Delta X_2(t) + a_{22}\cdot\Delta X_2(t-1) + a_{23}\cdot\Delta X_2(t-4) + \\ + a_{31}\cdot\Delta X_3(t) + a_{32}\cdot\Delta X_3(t-1) + a_{33}\cdot\Delta X_3(t-4) + e(t)$$

15. ***LogChangesQtoQ*:**

$$\Delta\ln(Y(t)) = a_0 + a_{11}\cdot\Delta\ln(X_1(t)) + a_{12}\cdot\Delta\ln(X_1(t-1)) + a_{13}\cdot\Delta\ln(X_1(t-4)) + \\ + a_{21}\cdot\Delta\ln(X_2(t)) + a_{22}\cdot\Delta\ln(X_2(t-1)) + a_{23}\cdot\Delta\ln(X_2(t-4)) + \\ + a_{31}\cdot\Delta\ln(X_3(t)) + a_{32}\cdot\Delta\ln(X_3(t-1)) + a_{33}\cdot\Delta\ln(X_3(t-4)) + e(t)$$

Annex 3

GDP PROVISIONAL ESTIMATE

NO 3/123
29 JULY 2003

GROSS DOMESTIC PRODUCT OVER 1ST HALF YEAR GREW BY 7.7 PER CENT

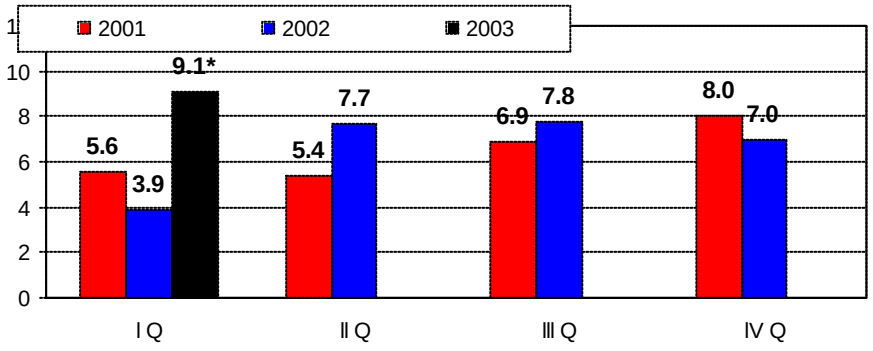
According to the provisional estimation over the 1st half year of 2003 the generated GDP amounted to **LTL 25608 million at current prices** and against the same period in 2002 grew by **7.7 per cent.** A higher gross value added was generated in all economic activities. The biggest upturn against the reference period was observed in economic activities related to manufacture and consumption: manufacturing, electricity, gas and water supply, construction, wholesale and retail trade.

Over the **2nd** of 2003 GDP totalled LTL **13395** million at current prices. It increased by **6.1** per cent against the respective period of 2002 and by **7.2** per cent compared to the previous period. The bulk of gross value added grew in agriculture, mining and quarrying, supply of electricity, gas and water and construction, wholesale and retail trade and hotels and restaurants. The value added change in manufacturing was smaller than in the whole country.

The per capita GDP over the 2nd quarter of the current year is LTL **3873**, i.e. by **6.5** per cent more than during the same period of the previous year.

Diagram 1. **GDP changes at constant prices**

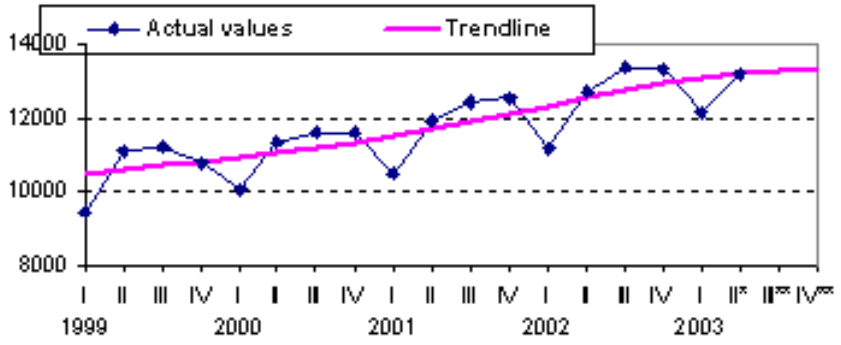
Against the same period of the previous year, per cent



*Provisional estimate of GDP change

Diagram 2. GDP at current prices

Billion LTL



**Provisional GDP estimate