Seasonal Adjustment and Forecasting of Quarterly Gross Domestic Product: Estonian Experience
Overview

1. Gross domestic product (definitions).
2. Seasonal adjustment of GDP.
3. Forecasting of quarterly GDP: Estonian experience.
4. Results for 2005.
5. Our conclusions
Gross domestic product (1)

The main aggregate indicator of economy in the country is gross domestic product (GDP). GDP at market prices is the final result of the production activity of resident producer (institutional) units.

An institutional unit is resident in a country when it has a centre of economic interests in the economic territory of that country. The institutional units are the following:
— non-financial corporations;
— financial corporations;
— general government;
— non-profit institutions serving households;
— households.
Gross domestic product (2)

As the first estimates of GDP are released after considerable delay, it makes them unsuitable for a person who needs them for operative calculations and analysis. It therefore seems important to provide a reliable and rapid system of preliminary estimates for the GDP. This system has been given the name “flash estimates”.

A flash estimate is defined as the earliest picture of the economy, which is produced and published as soon as possible after the end of the quarter.
Seasonal adjustment of GDP (1)

We started with seasonal adjustment of GDP and its components at the end of 1998. Our time series begin from 1st quarter 1993. The option “four quarters are equal to the year”, means that the sum of seasonally adjusted quarters is equal to the sum of unadjusted four quarters of the same year, was used. Initially we used package X11 (UK version) for seasonal adjustment.
Seasonal adjustment of GDP (2)

Since the end of 2000 we are using interface DEMETRA and applying all its opportunities, where program X12-ARIMA was used for seasonal adjustment of GDP and its components.

Our seasonally adjusted GDP and its components are published in press release, monthly bulletin and are available in the statistical database on our website (http://www.stat.ee).

Since 2005, we started to provide seasonally adjusted quarterly time series on GDP and its components to the Eurostat.
Three types of models for forecasting of trend were used:

a) linear model: \( T_t = a + b \cdot t \);

b) quadratic model: \( T_t = a + b \cdot t + c \cdot t^2 \);

c) logarithmic model: \( T_t = a + b \cdot \log(t) \),

where \( t = 1, 2, 3, \ldots \) denotes time.

For forecasting of seasonal component \( S_t \), \( t = T, T+4, T+8, \ldots, T+4 \cdot \text{YEARS} \), we used the linear model \( S_t = a + b \cdot t \), only.

We used package EViews for regression analysis.
Forecasting of quarterly GDP: ... (2)

To choose the best approximation we used following criteria:
1. $R^2$
2. Akaike Information Criteria (AIC), which is based on the sum of squared residuals. AIC is calculated as:

$$AIC = \log\left(\frac{RSS}{T}\right) + \frac{2k}{T},$$

where $RSS$ is sum of squared residuals and $k$ is number of estimated parameters.

The best specification is chosen with the lowest value of the AIC.
In the end of 2000, when we started to use DEMETRA for seasonal adjustment of GDP, it was possible to do additional forecasting using ARIMA models. Two methods for forecasting were used: the direct ARIMA model and indirect model as sum of ARIMA models for components of GDP by institutional sectors.

We got at least three different forecasts of GDP: at least one forecast is obtained from interpolation of trend and seasonal component, one from direct ARIMA model forecast and one from indirect model.
Forecasting of quarterly GDP: ... (4)

Before 2005 we used only $R^2$, $AIC$ and expert opinion for finding suitable forecast. From 2005 the situation has changed. We decided to formalize the process of choosing the best forecast using concrete econometric criteria for testing.

We divided our time series into two overlapping parts: one is for calculation of the model and second, which is included in or is equal with the first part, is for testing this model. Last nine quarters were used for test models.
We tried four econometric criteria:

Root Mean Squared Error:

\[ RMSE = \sqrt{\frac{1}{T+h} \sum_{t=T+1}^{T+h} (y_t - Y_t)^2} \]

Mean absolute Error:

\[ MAE = \frac{1}{h} \sum_{t=T+1}^{T+h} \left| y_t - Y_t \right| \]

Means Abs. Percent Error:

\[ MAPE = \frac{1}{h} \sum_{t=T+1}^{T+h} \left| \frac{(y_t - Y_t)}{y_t} \right| \]

Theil Inequality Coefficient:

\[ TIC = \frac{\sqrt{\frac{1}{T+h} \sum_{t=T+1}^{T+h} (y_t - Y_t)^2}}{\sqrt{\frac{1}{h} \sum_{t=T+1}^{T+h} y_t^2} + \sqrt{\frac{1}{h} \sum_{t=T+1}^{T+h} Y_t^2}} \]
Results for forecasts of GDP for 1\textsuperscript{st} quarter 2005.

<table>
<thead>
<tr>
<th></th>
<th>RMSE</th>
<th>TIC*</th>
<th>MAE</th>
<th>MAPE*</th>
<th>Expected GDP growth, %</th>
<th>Flash estimate, %</th>
<th>Calculated GDP growth, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trend1</td>
<td>300.9</td>
<td>0.514</td>
<td>254.7</td>
<td>0.855</td>
<td>6.3</td>
<td>7.0</td>
<td>7.2</td>
</tr>
<tr>
<td>Trend2</td>
<td>276.4</td>
<td>0.276</td>
<td>232.6</td>
<td>0.784</td>
<td>6.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARIMA1</td>
<td>442.0</td>
<td>0.754</td>
<td>374.3</td>
<td>1.282</td>
<td>6.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARIMA2</td>
<td>449.9</td>
<td>0.767</td>
<td>380.5</td>
<td>1.309</td>
<td>5.8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* - corresponding data are multiplied by 100
Results for forecasts of GDP for 2\textsuperscript{nd} quarter 2005.

<table>
<thead>
<tr>
<th></th>
<th>RMSE</th>
<th>TIC*</th>
<th>MAE</th>
<th>MAPE*</th>
<th>Expected GDP growth, %</th>
<th>Flash estimate, %</th>
<th>Calculated GDP growth, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trend1</td>
<td>275.2</td>
<td>0.465</td>
<td>181.7</td>
<td>0.654</td>
<td>7.3</td>
<td>10.0</td>
<td>9.9</td>
</tr>
<tr>
<td>Trend2</td>
<td>102.3</td>
<td>0.173</td>
<td>82.1</td>
<td>0.287</td>
<td>8.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARIMA1</td>
<td>303.7</td>
<td>0.514</td>
<td>232.8</td>
<td>0.815</td>
<td>7.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARIMA2</td>
<td>285.7</td>
<td>0.484</td>
<td>237.6</td>
<td>0.817</td>
<td>6.9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* - corresponding data are multiplied by 100
Results for forecasts of GDP for 3rd quarter 2005.

<table>
<thead>
<tr>
<th></th>
<th>RMSE</th>
<th>TIC*</th>
<th>MAE</th>
<th>MAPE*</th>
<th>Expected GDP growth, %</th>
<th>Flash estimate, %</th>
<th>Calculated GDP growth, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trend1</td>
<td>227.4</td>
<td>0.373</td>
<td>150.0</td>
<td>0.479</td>
<td>9.1</td>
<td>10.8</td>
<td>10.6</td>
</tr>
<tr>
<td>Trend2</td>
<td>130.2</td>
<td>0.313</td>
<td>111.9</td>
<td>0.368</td>
<td>9.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARIMA1</td>
<td>355.9</td>
<td>0.585</td>
<td>254.5</td>
<td>0.805</td>
<td>9.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARIMA2</td>
<td>387.4</td>
<td>0.636</td>
<td>282.6</td>
<td>0.892</td>
<td>9.5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* - corresponding data are multiplied by 100
Results for forecasts of GDP for 4\textsuperscript{th} quarter 2005.

<table>
<thead>
<tr>
<th></th>
<th>RMSE</th>
<th>TIC*</th>
<th>MAE</th>
<th>MAPE*</th>
<th>Expected GDP growth, %</th>
<th>Flash estimate, %</th>
<th>Calculated GDP growth, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trend1</td>
<td>252.8</td>
<td>0.407</td>
<td>191.5</td>
<td>0.591</td>
<td>10.0</td>
<td>10.5</td>
<td>11.1</td>
</tr>
<tr>
<td>Trend2</td>
<td>170.1</td>
<td>0.274</td>
<td>150.3</td>
<td>0.484</td>
<td>10.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARIMA1</td>
<td>369.6</td>
<td>0.596</td>
<td>272.1</td>
<td>0.848</td>
<td>10.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARIMA2</td>
<td>399.3</td>
<td>0.644</td>
<td>304.1</td>
<td>0.948</td>
<td>10.8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* - corresponding data are multiplied by 100
Our decision (1)

After testing we finally decided to use two criteria only: \( RMSE \), as most used criteria and \( TIC \), which shows the fit quality.

In case of ideal solution values for both criteria are minimal. Otherwise we could find solution, where \( TIC \) is minimal or close to it, but \( RMSE \) obtains restriction

\[ |RMSE - RMSE_{\text{min}}| < \varepsilon , \]

where \( \varepsilon \) is the fixed value.
In practice this means,

that we choose the optimal solution using no more than 3-5 possible pretenders using criteria $RMSE$ and then choose expert solution using criteria $TIC$ but considering $RMSE$. In that case the values for both criteria may not be minimal. This process applies to all institutional sectors. It improves quality of calculated value added forecasts for all institutional sectors and GDP forecasts altogether. Process of calculation of value of criteria $RMSE$ and $TIC$ can be easily done with EXCEL facilities.
Our decision (2)

We have seen, that the forecasts for the second and the third quarter of 2005 were not so good (the discrepancy from the flash estimated GDP was 1.9 and 1.1 percent points respectively). In this year the best forecast of GDP for the 4th quarter 2005 was close to flash estimate, but the discrepancy from the really calculated GDP was quite big. Therefore in this year we decided to develop a new, simple (one equation) model using econometric methods.
Thank you for attention