Short-run GDP forecasting in G7 countries: temporal disaggregation techniques and bridge models

by

Giuseppe Bruno
Tommaso Di Fonzo
Roberto Golinelli
Giuseppe Parigi

Luxembourg April 7-8, 2005
Needs & Tools

- Understanding of the present economic situation

**Bridge Models (BM)**

Timely monthly indicators

Quarterly GDP data

- Knowledge of the short term evolution of GDP

**Temporal Disaggregation Procedures (TDP)**

Quarterly GDP data + monthly indicators

Monthly GDP data
Bridge Models

- Dynamic *quarterly* models where GDP is *related* to (aggregated) *monthly* indicators.

- Careful *search* of a parsimonious specification both for dynamics and for the most suitable indicators (generally coincident-leading).

- *Validation* through several in-sample and out-of-sample testing procedures.

How to compute a monthly version of GDP?
Temporal Disaggregation Procedures

- It is assumed there exists a relationship between GDP and a set of related variables at monthly level.
- *Ex ante* assumption (*no search*) of the dynamic form of the relationship.
- The choice of the relates series is assumed *a priori* or checked through simple statistics (valid only if the specific dynamic form is valid).

**How to validate the choice of a specific TDP (dynamic form and the related series)?**
Motivation

Is it possible to merge BM and TDP to fulfil both needs?

If yes, what is the best way to combine the relevant information?
Dynamic specification and coincident indicators selected by preliminary analyses and the BM

Monthly data

3 monthly out-of-sample values are computed with known indicators and then aggregated to obtain quarterly data

Out-of-sample forecasting performance

Validation (RMSE or other criteria) one quarter-ahead
- BM estimation (see Golinelli and Parigi, 2004)

- Preliminary analysis of available information at quarterly frequency (levels and log-levels)

- Use of alternative related series and different TDP:
  - Chow and Lin (1971): \( u_{t,u} \mid X_h \sim AR(1) \)
  - Fernández (1981): \( u_{t,u} \mid X_h \sim \text{random walk} \)
  - Litterman (1983): \( u_{t,u} \mid X_h \sim ARIMA(1,1,0) \)
  - ADL model
  - Guerrero (1990): \( ARIMA \) based

- Comparison of the out-of-sample forecasting performance on the basis of RMSE
Preliminary analysis
(quarterly frequency)

- Selection of indicators: BM coincident indicators could be good candidates for TDP
- Integration analysis using the Elliott, Rothemberg and Stock (1996) testing procedure
- Cointegration analysis between GDP and indicators with alternative methods (Engle and Granger (1987), static regression; Ericsson and MacKinnon (2002), ECM; Johansen (1995), VAR)
Results

Model Identification: period 1985.1 - 2000.4

- Almost all the variables are I(1)
  → levels or differences?

- Cointegration outcomes are specific to the TDP (lack of robustness: small sample, lack of a theory...)
  → how to identify the appropriate TDP?

Comparing forecasting performance
Forecast comparisons: period 2001.1 - 2003.4

- Twelve quarters, up to 2003.4, are left for out-of-sample forecasting comparisons

- Three-months ahead GDP forecasts over the period 2001.1-2003.12 are obtained from monthly disaggregated models with a rolling procedure and are converted into the corresponding quarterly figures

- Computation of forecast statistics (e.g. RMSE)
## Results I: RMSE of Alternative TDP

**One indicator: Industrial Production Index**

<table>
<thead>
<tr>
<th>TDP</th>
<th>France</th>
<th>Germany</th>
<th>Italy</th>
<th>United Kingdom</th>
<th>United States</th>
<th>Japan</th>
<th>Canada</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>levels</td>
<td>logs</td>
<td>levels</td>
<td>logs</td>
<td>levels</td>
<td>logs</td>
<td>levels</td>
</tr>
<tr>
<td>Chow-Lin</td>
<td>0.54</td>
<td>0.55</td>
<td>0.87</td>
<td>0.93</td>
<td>0.54</td>
<td>0.56</td>
<td>0.74</td>
</tr>
<tr>
<td>Fernandez</td>
<td>0.53</td>
<td>0.53</td>
<td>0.77</td>
<td>0.83</td>
<td>0.48</td>
<td>0.50</td>
<td>0.59</td>
</tr>
<tr>
<td>Litterman</td>
<td>0.58</td>
<td>0.59</td>
<td>0.74</td>
<td>0.79</td>
<td>0.41</td>
<td>0.43</td>
<td>0.55</td>
</tr>
<tr>
<td>ADL</td>
<td>0.66</td>
<td>0.65</td>
<td>0.67</td>
<td>0.70</td>
<td>0.37</td>
<td>0.38</td>
<td>0.43</td>
</tr>
<tr>
<td>Guerrero</td>
<td>0.37</td>
<td>0.56</td>
<td>0.42</td>
<td>0.70</td>
<td>0.62</td>
<td>0.59</td>
<td>0.93</td>
</tr>
</tbody>
</table>

- The level and the log-level results are basically the same.
- The best results are obtained by the ADL and Guerrero’s specifications.
- In Italy and the UK the ADL seems to be the preferred choice.
- In France and Germany the ADL procedure seems to be affected by an unsatisfactory dynamic representation.
- In Japan and maybe Canada both the ADL and the Guerrero procedures seem to be affected by a bad indicator choice, as suggested by BM results.
## Results II: RMSE of Alternative TDP Several indicators

<table>
<thead>
<tr>
<th>TDP</th>
<th>France</th>
<th>Germany</th>
<th>Italy</th>
<th>United Kingdom</th>
<th>United States</th>
<th>Japan</th>
<th>Canada</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IP, CONS, IPIT</td>
<td>IP, RET, JOB, ICT</td>
<td>IP, R, CLI</td>
<td>IP, R, CPI</td>
<td>IP, R</td>
<td>IP, IPSE, JOB</td>
<td>IP, IPSE</td>
</tr>
<tr>
<td></td>
<td>levels logs</td>
<td>levels logs</td>
<td>levels logs</td>
<td>levels logs</td>
<td>levels logs</td>
<td>levels logs</td>
<td>levels logs</td>
</tr>
<tr>
<td>Chow-Lin</td>
<td>0.58 0.59</td>
<td>0.69 0.72</td>
<td>0.53 0.56</td>
<td>0.78 0.75</td>
<td>0.72 0.81</td>
<td>0.93 1.05</td>
<td>0.38 0.36</td>
</tr>
<tr>
<td>Fernandez</td>
<td>0.53 0.54</td>
<td>0.60 0.63</td>
<td>0.48 0.50</td>
<td>0.70 0.69</td>
<td>0.67 0.71</td>
<td>0.92 0.94</td>
<td>0.39 0.37</td>
</tr>
<tr>
<td>Litterman</td>
<td>0.52 0.53</td>
<td>0.60 0.63</td>
<td>0.42 0.45</td>
<td>0.65 0.63</td>
<td>0.68 0.72</td>
<td>0.78 0.76</td>
<td>0.37 0.36</td>
</tr>
<tr>
<td>ADL</td>
<td>0.57 0.58</td>
<td>0.52 0.56</td>
<td>0.41 0.42</td>
<td>0.56 0.60</td>
<td>0.58 0.55</td>
<td>0.78 0.80</td>
<td>0.36 0.36</td>
</tr>
<tr>
<td>BM-search</td>
<td>0.55 0.55</td>
<td>0.60 0.63</td>
<td>0.31 0.33</td>
<td>0.70 0.69</td>
<td>0.63 0.56</td>
<td>0.72 0.76</td>
<td>0.39 0.37</td>
</tr>
</tbody>
</table>

- In general, there is an improvement over the one indicator case, especially for Japan and Canada, the two countries where the IP index is not a good indicator.

- The ADL and BM-search results suggest that TDPs with richer dynamics and several indictors may perform better than restricted procedures, both on the dynamics (Chow-Lin, Fernandez, Litterman) and/or on the number of indicators (Guerrero).
How to disaggregate euro area GDP

Area wide indicators

- Industrial production
- Qualitative indicators
- Coincident indicators (i.e. Eurocoin)

Single country or sub-aggregate indicators

- "Real" indicators
  - Choice of the countries (weight) and/or the indicators (timeliness, reliability)

- "Generated" indicators
  - Two-steps procedures
    1) Monthly interpolation of Countries GDP
    2) Use these monthly GDPs as indicators of Euro area GDP

Sub-aggregate country or single country indicators?
## Results III: RMSE of Alternative TDP

### EURO AREA

<table>
<thead>
<tr>
<th>TDP</th>
<th>IPFR, IPDE, IPIT</th>
<th>Weighted average of IPFR, IPDE, IPIT</th>
<th>GDPFR, GDPDE, GDPIT</th>
<th>GDPFR, GDPDE, GDPIT</th>
<th>Simple average of GDPFR, GDPDE, GDPIT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chow-Lin</td>
<td>0.47</td>
<td>0.38</td>
<td>0.56</td>
<td>0.38</td>
<td>0.30</td>
</tr>
<tr>
<td>Fernandez</td>
<td>0.38</td>
<td>0.31</td>
<td>0.45</td>
<td>0.30</td>
<td>0.24</td>
</tr>
<tr>
<td>Litterman</td>
<td>0.35</td>
<td>0.29</td>
<td>0.42</td>
<td>0.30</td>
<td>0.24</td>
</tr>
<tr>
<td>ADL</td>
<td>0.38</td>
<td>0.24</td>
<td>0.32</td>
<td>0.25</td>
<td>0.20</td>
</tr>
<tr>
<td>Guerrero</td>
<td></td>
<td>0.31</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- The use of the 3 countries monthly GDP as indicator for the Euro area provides better results than the country index of industrial production.
- Aggregate indicators tend to perform better than disaggregate one.
Main findings - 1

- The forecasting ability of TDP is, as expected, worse than that of the corresponding BM for all the G7 countries.

- Short samples prevent the integration-cointegration tests to act, *per se*, as clear guidelines about the choice of the more appropriate TDP.

- According to the out-of-sample forecasting performance the only robust outcome is the choice of the TDP characterised by a richer dynamics, such as the ADL or Guerrero’s specifications.

- This result may be interpreted as a statistical test for the restricted dynamic specifications implied by the other TDP’s.
Main findings - 2

- The preliminary analysis and the BM may help select the most appropriate monthly variable(s) to be used in the TDP (see for instance the Japan and Canada results).

- Deeper fine tuning of the TDP specification does not seem to improve further the results.

- Results can be extended by using other TDP, such as those based on the state-space model approach, see Harvey and Chung (2000), Casals et al. (2004), and Moauro and Savio (2005).