A pairwise approach to model and forecast a large set of disaggregates with common trends

Guillermo Carlomagno, Antoni Espasa

Universidad Carlos III de Madrid

London, September 2014

15th OxMetrics User Conference
• **General Objective**: Model and forecast all the components of a macro or business aggregate.

• **Main contribution**: cases with large number of components, when multivariate approaches are not feasible.
Motivation:

1. **Disaggregating is relevant**
   - **Own interest** for economic and business decision making.
   - **Diagnosis and fore. the agg:** increasing agreement that deeper knowledge of the components may lead to better understanding of the agg. (Espasa et.al, 2002; GG, 2004; HH, 2005, 2011; EA, 2007; EMB, 2013)
   - **Comparative analysis** (e.g relative prices)

2. **PW strategy** (EMB, 2013) as a procedure to deal with the **Estimation uncertainty vs informational losses trade-off**, and the problem of non-pervasive common factors in DFM.
Outline

1. Description and asymptotic properties \((T \to \infty)\)
2. (X) Simulation results
3. (X) Outliers and breaks
4. Forecasting strategy and application
   - A note about the forecasting equations
   - Forecasting results for the US CPI
5. Conclusions
How the *pairwise* procedure (restricted to common trends) works...

1. Perform **Johansen cointegration tests** between all the \( N(N - 1)/2 \) pairs (4950 for \( N = 100 \)).
2. Find the largest subset in which every series is cointegrated with all the others (maximum *clique* - **fully cointegrated**).
3. Continue looking for the second largest and so on...
4. For ease of exposition just focus on the largest.
The Pairwise strategy inherits the asy. props. of the Johansen test

However; two specific features deserve special attention:

1. Partial model estimation
2. **Multiple testing;** are we inflating the false rejection probability? **NO**

**Theorem (for pairs inside the FC set):**

Given a set of $n_1$ pairwise cointegrated series, the prob. of finding the same result in all the $n_1(n_1-1)/2$ Johansen’s tests $\rightarrow 1$ as $T \rightarrow \infty$

$\Rightarrow FWER = \alpha$. Even if the hypothesis of interest is the universal one ($FWER$), *p-values should not be corrected.*

*For the other pairs $\Rightarrow$ bounds for error probabilities already ‘tolerable’*
Asymptotic properties...

- The Pairwise strategy inherits the asy. props. of the Johansen test
- However; two specific features deserve special attention:
  1. Partial model estimation
  2. **Multiple testing;** are we inflating the false rejection probability? **NO**

**Theorem (for pairs inside the FC set):**

Given a set of \( n_1 \) pairwise cointegrated series, the prob. of finding the same result in all the \( n_1(n_1 - 1)/2 \) Johansen’s tests \( \rightarrow 1 \) as \( T \rightarrow \infty \)

- \( \Rightarrow FWER = \alpha \). Even if the hypothesis of interest is the **universal** one (FWER), **p-values should not be corrected.**
- For the other pairs \( \Rightarrow \) bounds for error probabilities already ‘tolerable’
Multiple testing: analytical results...

**In summary:**

Size corrections are not needed; multiple testing does not occur (FC series), or bounds provide already tolerable error probabilities”

**Simulation experiments:**

- Confirm the results for the pairs in the FC set
- Bounds are quite loose
- The PW strategy dominates a DFM alternative in terms of *Gauge* and *Potency*
A note about the forecasting equations

1. Description and asymptotic properties \((T \to \infty)\)

2. (X) Simulation results

3. (X) Outliers and breaks

4. Forecasting strategy and application
   - A note about the forecasting equations
   - Forecasting results for the US CPI

5. Conclusions
The effects of alternative normalizations for $\beta$ 

- The CI VAR remains the same if we change $\alpha\beta'$ by $\alpha H^{-1}H\beta'$. Let:

\[
\alpha = \begin{bmatrix}
0 & 0 & 0 \\
-0.2 & 0 & 0 \\
0 & -0.2 & 0 \\
0 & 0 & -0.2
\end{bmatrix}, \quad \beta' = \begin{bmatrix}
-1 & 1 & 0 & 0 \\
-1 & 0 & 1 & 0 \\
-1 & 0 & 0 & 1
\end{bmatrix}, \quad W = \begin{bmatrix}
0.25 \\
0.25 \\
0.25 \\
0.25
\end{bmatrix}
\]

- Define $H$ in order to express CI rels as deviations from the aggregate

\[
\left(\text{equivalence condition; } \alpha^* = \tilde{\alpha}(I_r - \beta_{nf}^* W_{nf})^{-1}\right), \quad \alpha H^{-1} = \begin{bmatrix}
0 & 0 & 0 \\
-0.4 & -0.2 & -0.2 \\
-0.2 & -0.4 & -0.2 \\
-0.2 & -0.2 & -0.4
\end{bmatrix}
\]

- $\text{N}^\circ \text{ of relevant CI rels.}$ in each equation (except for the first one) increased from 1 to 3.
A note about the forecasting equations

The normalization may not be innocuous for the forecasting stage

- $N$ forecast single eq models including CI rels. and other regressors:
  \[
  \Delta x_{i,t} = c + \alpha_{i1} CR_{1i,t-1} + ... + \alpha_{ir} CR_{ri,t-1} + \text{other regressors} + \epsilon_{i,t}
  \]

- No of relevant CI rels. may change depending on $\beta$’s normalization.

  ⇒ Possible ‘free’ reduction in the number of parameters ⇒ ‘free’ estimation uncertainty reduction ⇒ forecasting accuracy improvement.

- **Strategy**: try different normalizations and select over CI rels and the other regressors using Autometrics
Forecasting results for the US CPI

Forecasting: exercise design

Regressors are selected using *Autometrics* in the following initial GUMs:

- **Cointegrated series:**

\[
\Delta x_{i,t} = c + \alpha_{i1} CR_{1i,t-1} + \ldots + \alpha_{ir} CR_{ri,t-1} + \sum_{k_1=1}^{K_1} \phi_{k1} \Delta x_{i,t-k_1} + \ldots \\
\sum_{k_2=1}^{K_2} \theta_{k2} \Delta SubAggi,t-k_2 + \sum_{k_3=1}^{K_3} \delta_{k3} \Delta CPI_{t-k_3} + \sum_{i=1}^{11} \gamma_i S_{i,t} + \epsilon_{i,t}
\]

- **Rest of the series:**

Restricted such that the forecast for the sub-aggregate ‘rest’ is the one coming form a uni-equational model.

- In both cases forecasts are produced with and without IIS.
Forecasting results for the US CPI

Forecasting results: RMSE ($\Delta_{12} \log(CPI)$)

- Recursive forecasting 2011.1 - 2013.12, from $H = 1$ to $H = 12$ steps ahead.
- Baseline: univariate model for the $CPI$.
- Form Diebold-Mariano: PW dominates the baseline in horizons 4 to 8.
- Form Capistrán: considering jointly horizons 1 to 6 PW dominates the baseline (this is not the case when using all horizons, but few observations)
Conclusions

- The pairwise procedure allows to discover blocks of series that share a unique common trend in a large set of series.

- Statistical properties were studied analytically, and confirmed by Monte Carlo.

- A small samples correction procedure was designed and its properties studied by simulations.

- Comparison with DFM: PW clearly dominates when $n_1$ is not large.

- Forecasting accuracy may be improved by changing $\beta$'s normalization.

- PW strategy may improve CI test power

- A strategy for dealing with outliers and breaks was designed.

- Application US CPI: The PW procedure improves forecasting accuracy wrt the baseline at least for some horizons.
Thanks for your attention!