Imposing Frequency-domain Restrictions on Time Domain Forecasts of Economic Variables

with E. Reschenhofer
Typical Spectral Shape of Economic Variable
Granger (1966), Seasonally adjusted real US GDP

Seasonally adjusted real GDP (2000=100)

Original GDP

Diff.log GDP

Log GDP

Periodogram of diff.log GDP
Co-spectrum of two Economic Variables
Reschenhofer and Chudy (2015), GDP growth and lagged spread of interest rates

Motivation
Panel of Macroeconomic Time Series

Variables (quarterly time series from 1960Q2 through 2008Q4):

<table>
<thead>
<tr>
<th>Group</th>
<th>Total</th>
<th>For PCs</th>
<th>Transformations</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDP components</td>
<td>16</td>
<td>9</td>
<td>$\Delta$ (log)</td>
</tr>
<tr>
<td>Industrial production</td>
<td>14</td>
<td>9</td>
<td>$\Delta$ (log)</td>
</tr>
<tr>
<td>Employment</td>
<td>20</td>
<td>16</td>
<td>$\Delta$ or $\Delta$ (log)</td>
</tr>
<tr>
<td>Unemployment</td>
<td>7</td>
<td>7</td>
<td>$\Delta$ or $\Delta$ (log)</td>
</tr>
<tr>
<td>Housing</td>
<td>6</td>
<td>4</td>
<td>log</td>
</tr>
<tr>
<td>Inventories</td>
<td>6</td>
<td>6</td>
<td>-</td>
</tr>
<tr>
<td>Prices</td>
<td>37</td>
<td>24</td>
<td>$\Delta^2$ (log)</td>
</tr>
<tr>
<td>Wages</td>
<td>6</td>
<td>5</td>
<td>$\Delta$ (log)</td>
</tr>
<tr>
<td>Interest rates</td>
<td>13</td>
<td>9</td>
<td>- or $\Delta$</td>
</tr>
<tr>
<td>Money</td>
<td>7</td>
<td>7</td>
<td>$\Delta^2$ (log)</td>
</tr>
<tr>
<td>Exchange rates</td>
<td>5</td>
<td>5</td>
<td>$\Delta$ (log)</td>
</tr>
<tr>
<td>Stock prices</td>
<td>5</td>
<td>5</td>
<td>$\Delta$ or $\Delta$ (log)</td>
</tr>
<tr>
<td>Consumer expectations</td>
<td>1</td>
<td>1</td>
<td>$\Delta$</td>
</tr>
<tr>
<td><strong>∑</strong></td>
<td>143</td>
<td>107</td>
<td></td>
</tr>
</tbody>
</table>
Periodograms of dynamic factors
Chudy and Reschenhofer (2016), obtained from panel of 107 variables by principal components

Motivation
New predictors are derived based on:

- own past with focus on the spectral density,

- single explanatory variable with focus on the cospectrum,

- many explanatory variables with focus on cospectrum with principal components.
Focus on the spectral density

Spectra of quarterly macroeconomic series can be well described by a step function, e.g., spectrum of GDP growth rates,

...we use a subset selection approach to determine the number and location of the steps.
A Subset Selection Problem

Method:
linear regression model with dependent variable
\( I = (I(\omega_1), \ldots, I(\omega_m))^\top \) and design matrix:

\[
X = \begin{pmatrix}
1 & 0 & \cdots & 0 \\
1 & 1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
1 & 1 & \cdots & 1
\end{pmatrix}, \in \mathbb{R}^{m \times m}.
\]

Problem: Select appropriate subset of columns of \( X \).

Notation:
- \( K \), the number of selected columns,
- \( S(K) \subseteq (1, \ldots, m) \), determining the selected columns,
- \( X(S(K)) \), corresponding submatrix of \( X \).
A Modified Subset Selection Criterion

→ $K$ is obtained by minimizing a subset version of FPE criterion,

$$\text{FPE}_{\text{sub}}^{*}(\hat{S}(K)) = \frac{\text{RSS}}{\text{penalty}} = \frac{\left\| \hat{J} - \hat{J}(K) \right\|^2}{E \left\| \hat{J} - \hat{J}(K) \right\|^2},$$

...where $\hat{J}^*$, $\hat{J}$ are independent samples from the same exp. distribution.

→ The penalty term was obtained by simulations for different values of $K$ and $m$ as

$$\sum_{j=1}^{f} \frac{\left\| \hat{J}_j^* - \hat{J}_j(\hat{S}(K)) \right\|^2}{\sum_{j=1}^{f} \left\| \hat{J}_j - \hat{J}_j(\hat{S}(K)) \right\|^2}.$$
A New Time-Domain Predictor

Estimated spectral density has the form:

\[ \hat{f}(\omega) = \sum_{k=1}^{K} \hat{\lambda}_k I(\hat{\omega}_k - \hat{\omega}_{k-1}) \].

...from the spectral representation of the autocovariance function:

\[ \tilde{\gamma}(j) = \begin{cases} 
2/j \sum_{k=1}^{K} \hat{\lambda}_k (\sin(\hat{\omega}_kj) - \sin(\hat{\omega}_{k-1}j)) & \text{for } j = 1, \ldots, p, \\
2 \sum_{k=1}^{K} \hat{\lambda}_k (\hat{\omega}_k - \hat{\omega}_{k-1}) & \text{for } j = 0.
\end{cases} \]

The parameters \( \phi_1, \ldots, \phi_p \) for the predictor

\[ \hat{y}_{t+1} = \phi_1 y_t + \ldots + \phi_p y_{t-p} \]

...can be obtained using Durbin-Levinson algorithm.
Empirical Comparison with ARMA models

Cumulative absolute forecast errors of US GDP growth rate

Predictors Based on Own Past
Starting with two jointly stationary series $x$ and $y$, we are looking for a predictor of the form:

$$\hat{y}_{n+1} = x_n \beta$$

... the OLS estimator for $\beta$ can be rewritten as

$$\hat{\beta} = \frac{\sum_{j=1}^{m} \text{Re} \left( I_{xy}(\omega_j) \right)}{\sum_{j=1}^{m} I_{xx}(\omega_j)}.$$  \hfill (1)

Based on our experience with quarterly time series, we argue, that the co-spectrum $I_{xy} \approx 0$ for high frequencies. Incorporating this into (1), we obtain a new estimator for $\beta$

$$\tilde{\beta}_r = \frac{\sum_{j=1}^{r} \text{Re} \left( I_{xy}(\omega_j) \right)}{\sum_{j=1}^{m} I_{xx}(\omega_j)}.$$
The improved estimator

\[ \tilde{\beta}_r = \frac{\sum_{j=1}^{r} \text{Re} \left( I_{xy}(\omega_j) \right)}{\sum_{j=1}^{m} I_{xx}(\omega_j)} \]

...can be obtained from the band regression estimator

\[ \hat{\beta}_r = \frac{\sum_{j=1}^{r} \text{Re} \left( I_{xy}(\omega_j) \right)}{\sum_{j=1}^{r} I_{xx}(\omega_j)} \]

...by multiplication with a shrinkage factor

\[ 0 \leq \frac{\sum_{j=1}^{f} I_{xx}(\omega_j)}{\sum_{j=1}^{m} I_{xx}(\omega_j)} \leq 1. \]

A selection procedure for number of frequencies \( r \leq m \):

...same idea as, e.g., AIC \( \rightarrow \) by replacing number of regressors \( k \) by fraction \( k \frac{r}{m} \) in the penalty term.
Established method in case of parameter instability. For a predictor of the form

\[ \hat{y}_{n+1} = \hat{\alpha} + \hat{\beta} x_n, \]

where

\[
\begin{pmatrix}
\hat{\alpha} \\
\hat{\beta}
\end{pmatrix} = \left( \sum_{j=0}^{n-1} \lambda_{j,n} \begin{pmatrix} 1 \\ x_{n-j} \end{pmatrix} (1, x_{n-j}) \right)^{-1} \sum_{j=0}^{n-1} \lambda_{j,n} \begin{pmatrix} 1 \\ x_{n-j} \end{pmatrix} y_{n-j+1}.
\]

Get “robust optimal” (RO) weights. \( \lambda_{j,n} \) for one-step-ahead forecasts.

\[
\lambda_{j,n}^{(RO)} \propto \left\{ \begin{array}{ll}
\frac{\log(n)}{n-1}, & \text{if } j = 0,
\frac{\log(n) - \log(j)}{n-1}, & \text{if } j > 0,
\end{array} \right.
\]

... by integrating out the effects of uncertainty about the breaks with respect to some given distribution.
Empirical Comparison with AR(1) and weighted OLS

Cumulative absolute forecast errors with restriction: $r/m = 20\%$

Predictors Based on a Single Explanatory Variable
The framework is the multiple regression model:

\[ y_{t+1} = \beta^T p_t + \epsilon_{t+1}, \quad t \in \{0, \ldots, n-1\}, \quad PP^T/n = I_K, \]

(2)

...with orthonormal predictors \( p_t = (p_{t1}, \ldots, p_{tK})^T \) and parameters \( \beta = (\beta_1, \ldots, \beta_K)^T \).

They compute the one-step-ahead forecast \( \tilde{y}_{n+1|n} \) at time point \( n \) using

\[ \tilde{y}_{n+1|n} = \sum_{k=1}^{K} \psi(\kappa t_k) \hat{\beta}_k p_{nk}, \]

(3)

...where \( \hat{\beta}_k \) is the OLS estimator of \( \beta_k \) from (??) and \( t_k \) is the corresponding t-statistic and \( \psi \) is the shrinkage function specific to the forecasting method. The Formula (??) the generalized shrinkage representation, since \( 0 \leq \psi(x) \leq 1 \).
Predictors Based on Many Regressors
Adjusting Band Regression Estimator

→ Idea from one regressor case, now incorporated in the multivariate case
... projecting both design matrix $X \in \mathbb{R}^{n \times k}$ and $y$ into space spanned by columns of matrix

$$H_r = \sqrt{\frac{2}{n}} \begin{pmatrix}
  \cos(\omega_1 1) & \sin(\omega_1 1) & \cdots & \cos(\omega_r 1) & \sin(\omega_r 1) \\
  \cos(\omega_1 2) & \sin(\omega_1 2) & \cdots & \cos(\omega_r 2) & \sin(\omega_r 2) \\
  \vdots & \vdots & \ddots & \vdots & \vdots \\
  \cos(\omega_1 n) & \sin(\omega_1 n) & \cdots & \cos(\omega_r n) & \sin(\omega_r n)
\end{pmatrix}.$$ 

→ Denote the projection matrix by

$$P_r = H(H^TH)^{-1}H^T = HH^T$$

... and the corresponding projections

$$y_r = P_r y, \quad X_r = P_r X, \quad \bar{y}_r = (I - P_r)y, \quad \bar{X}_r = (I - P_r)X,$$
Adjusting Band Regression Estimator

The band regression estimator is

\[ \hat{\beta}_r = (X_r^T X_r)^{-1} X_r^T y_r. \] (4)

...and the conventional OLS estimator can be rewritten as

\[ \hat{\beta} = \left( X_r^T X_r + \overline{X}_r \overline{X}_r \right)^{-1} \left( X_r^T y_r + \overline{X}_r \overline{y}_r \right). \] (5)

Setting:

\[ \overline{X}_r \overline{y}_r = 0 \]

...in (??) offers a certain balance between (??) and (??),

...the new estimator is given by

\[ \tilde{\beta}_r = \left( X_r^T X_r + \overline{X}_r \overline{X}_r \right)^{-1} \left( X_r^T y_r \right). \]