Bayesian Inference in a Cointegrating Panel Data Model

Gary Koop (University of Strathclyde)

Roberto León-González (University of Leicester)

Rodney Strachan (University of Leicester)

This paper (and other related papers) available at:
http://www.le.ac.uk/economics/research/dpseries.html
http://personal.strath.ac.uk/gary.koop/

Dynamics in Panels:

- Growing availability of panel data with large T dimension has stimulated research in panel data models with time series.

- Growing interest in issues of nonstationarity and cointegration.

Baltagi and Kai (2000) identify many applications.

E.g., Jacobson, Lyhagen, Larsson and Nessén (2002) use a multivariate panel cointegration model and demonstrate that, although strong purchasing power parity restrictions are rejected, the location of the cointegrating space is similar for all countries considered. This provides some evidence in support of PPP.

Frequentist methods (a great many):

Residual-based, LM and likelihood tests (Kao (1999), McCoskey and Kao (1998), Pedroni (2004), Larsson, Lyhagen and Löthgren (2001) and *Groen and Kleibergen (2003)).*
Estimation methods used OLS, maximum likelihood and GMM. (Refer to surveys by Phillips and Moon (2000) and Baltagi and Kao (2000)).

Bayesian approaches:

Pesaran, Hsiao and Tahmiscioglu (1999), Li (1999) do not explicitly consider cointegration.

Carmeci (2005) state space model which implies cointegration. This is the only paper explicitly proposing a Bayesian approach to estimation of a cointegrating system in panel data models. Assumes known cointegrating rank is equal in every cross-sectional unit.

No paper presents a fully Bayesian method of inference on cointegration in panels, when the cointegrating rank is unknown and may differ across cross-sectional units.

1 The Models

From the standard VECM for $n \times 1$ vector y_t of $I(1)$ series we generalize to the panel data case. (I.e., we add the subscript i)

i denotes the cross-sectional unit (where $i = 1, .., N$).

The $n \times 1$ vector $y_{i,t}$ of observations for individual i at time t has panel VECM:

$$\Delta y_{i,t} = \Pi_i y_{i,t-1} + \sum_{h=1}^{l_i} \Gamma_{i,h} \Delta y_{i,t-h} + \Phi_i d_{i,t} + \varepsilon_{i,t} \quad (1)$$

where

$$\Pi_i = \alpha_i \beta_i'$$

α_i and β_i (the cointegrating matrix) are $n \times r_i$ full rank matrices and $\mathfrak{p}_i = sp(\beta_i)$ ($\mathfrak{p}_i$ is the object of interest).

Note: The number of cointegrating relationships may vary across individuals.

The covariance matrices for vectors $\varepsilon_{i,t}$ are assumed to be

$$E\left(\varepsilon_{i,t}\varepsilon'_{j,s}\right) = \begin{cases} \Sigma_{ij} & \text{for } t = s \\ 0 & \text{for } t \neq s \end{cases}. \qquad (2)$$

This contrasts with, for instance, Larsson, Lyhagen and Löthgren (2001) who use a more restrictive model assuming $E\left(\varepsilon_{i,t}\varepsilon'_{j,s}\right) = 0$ if $i \neq j$ for all t and s.

We follow the more general model of Groen and Kleibergen (2003).

Note also that our model is more flexible than the one of Groen and Kleibergen (2003) in that we relax the assumption of a common cointegrating rank.

Model features of (possible) interest.

How many stable, equilibrium relationships exist?
Dimension of the cointegrating space:

Does $r_i = r$ for a particular i?

Does $r_i = r$ for all i?

What are the forms of these equilibrium relationships?
Location of the cointegrating space:

Does $\mathfrak{p}_i = \mathfrak{p}$ for a particular i?

Does $\mathfrak{p}_i = \mathfrak{p}$ for all i?

(Or if $\dim(\mathfrak{p}) = r$, but $r_i \leq r$ is $\mathfrak{p}_i \subseteq \mathfrak{p}$?)

A simple illustration of a cointegrating space of interest:

The balanced growth hypothesis in the real business cycle model presented by King, Plosser, Stock and Watson (1991, KPSW).

$$y_{i,t} = \left(c_{i,t}, a_{i,t}, g_{i,t}\right)'$$

$c_{i,t}$ is log consumption for country i, $a_{i,t}$ is log investment for that country and $g_{i,t}$ is log income.

KPSW model suggests $y_{i,t} \sim I\left(1\right)$ and $\beta_i' y_{i,t} \sim I\left(0\right)$

The logs of the 'Great Ratios' $c_{i,t} - g_{i,t}$ and $a_{i,t} - g_{i,t}$ are $I\left(0\right)$ in every country, then

$$\mathfrak{p}_i = \mathfrak{p} = sp \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \end{pmatrix} \text{ for all } i \ .$$

It is possible that $r_i = 2$ for some countries and $r_i = 1$ for others.

In this case $\mathfrak{p}_i = \mathfrak{p}$ for all $i = 1, .., N$ is not reasonable.

So the questions become:

* $\mathfrak{p}_i = \mathfrak{p}$? Are the Great Ratios $c_{i,t} - g_{i,t}$ and $a_{i,t} - g_{i,t}$ stable equilibrium relations?

* $\mathfrak{p}_i \subset \mathfrak{p}$? Or is some combination of them stable?

$$\left(c_{i,t} - g_{i,t}\right) + \beta_1 \left(a_{i,t} - g_{i,t}\right)$$

* $\mathfrak{p}_i \neq \mathfrak{p}$? Or is the equilibrium relationship something else entirely?

In terms of our notation, this involves investigating whether

$$\mathfrak{p}_i \subseteq \mathfrak{p} \text{ for } i = 1, .., N.$$

Some important issues in Bayesian analyses of cointegration.

The VECM suffers from a lack of identification (both locally and globally).

The global non-identification problem: for any nonsingular G

$$\Pi = \alpha\beta' = \alpha G G^{-1}\beta' = \alpha^{*}\beta^{*\prime}$$

It is common to impose the so-called *linear normalization* where

$$\beta = \begin{bmatrix} I_r \\ B \end{bmatrix}.$$

Serious drawbacks to the normalization

$$\beta = \begin{bmatrix} I_r \\ B \end{bmatrix}.$$

Restricts the estimable region of the cointegrating space; Invalid restrictions (JBES: Boswijk 1996, Luukkonen et al. 1999, Strachan 2003).

Possible problems at particular values of α (e.g., local nonidentification, weak exogeneity) (Kleibergen and van Dijk, 1994/1998, Strachan and Villani, 200?).

Such a restriction puts infinite prior weight in the direction of the space orthogonal to the normalisation (Strachan and van Dijk 2004). As a result, a Uniform prior on B in $\beta = [I_r \quad B']'$ implies a *very informative, improper and strange* prior on the cointegrating space $\mathfrak{p}$.

Our approach: A Prior on the cointegrating space

Reconsider the parameter of interest:
– the cointegrating space, $\mathfrak{p} = sp(\beta)$.
– NOT on the cointegrating vectors, β.

Specify the prior on the cointegrating space, $\mathfrak{p}$.

Suggests β being orthogonal, $\beta'\beta = I_r$ (James, 1954).

Prior does not restrict estimable region and has intuitively sensible properties (see related work).

No problems with the posterior at all values α.

As with all other approaches except Kleibergen and van Dijk (1994) and Strachan (2003), is not invariant to rescaling.

Back to the Cointegrating Panel Data Model.

To simplify notation and sampling, we do a bit of rewriting of the model

$$
\begin{aligned}
\Delta y_{i,t} &= \alpha_i \beta_i' y_{i,t-1} + \sum_{h=1}^{l_i} \Gamma_{i,h} \Delta y_{i,t-h} + \Phi_i d_{i,t} + \varepsilon_{i,t} \\
&= \begin{bmatrix} \alpha_i & \Gamma_{i,1} & \dots & \Gamma_{i,l} & \Phi_i \end{bmatrix} \begin{bmatrix} \beta_i' y_{i,t-1} \\ \Delta y_{i,t-h} \\ \Delta y_{i,t-h} \\ d_{i,t} \end{bmatrix} + \varepsilon_{i,t} \\
&= B_i' X_{i,t}' + \varepsilon_{i,t}
\end{aligned}
$$

Collect the $B_i's$ and $\mathfrak{p}_i's$ for all N individuals into

$$
\begin{aligned}
b &= \left[vec\left(B_1\right)', \dots, vec\left(B_N\right)' \right]' \\
\mathfrak{p} &= \left[\mathfrak{p}_1, \dots, \mathfrak{p}_N\right]'
\end{aligned}
$$

The Priors for b, $\mathfrak{p}$, and $\mathbf{\Sigma}$:

Paper has a noninformative prior (see paper for details).

We also use an informative prior which contains what we call "soft homogeneity" restrictions.

$$\begin{aligned} p\left(\mathbf{\Sigma}\right) &\propto \left|\mathbf{\Sigma}\right|^{-(Nn+1)/2} \\ b &\sim N\left(0, \underline{V}\frac{1}{\nu}\right) \end{aligned}$$

The elements of $\underline{V}$ can be used to impose similarity of dynamics $\Gamma_{i,k}$ if desired

Note: this is a hierachical prior where ν is treated as an unknown parameter.

The prior on β_i is implied by the prior on the $\mathfrak{p}_i$.

The prior for the $\mathfrak{p}_i$ has a common location $\mathfrak{p}^H$ across individuals

A dogmatic prior would place all of the prior mass for $\mathfrak{p}_i$ at $\mathfrak{p}^H = sp(H)$.

We want a non-dogmatic prior for the cointegration space that says the cointegration spaces, $\mathfrak{p}_i$, are likely to be close to $\mathfrak{p}^H = sp(H)$ and, thus, farthest from $\mathfrak{p}^{H_\perp} = sp(H_\perp)$ where $H_\perp$ is the orthogonal complement of H.

The $\mathfrak{p}_i$s are weighted averages of $\mathfrak{p}^H$ and $\mathfrak{p}^{H_\perp}$ and we can elicit a hierarchical prior about these weights.

$$\mathfrak{p}_i = \mathfrak{p}^H + \eta_i \mathfrak{p}^{H_\perp}$$

I.e. "weight" η_i is treated as unknown parameter (See paper for details).

Posterior Simulation

See paper for details.

Key idea of posterior simulation: Traditional methods computationally very inefficient due to $\beta'\beta = I_r$.

But we can develop a collapsed Gibbs sampler developed in Liu (1994) and Liu, Wong and Kong (1994)

Simulator 1:

Use the decomposition $\alpha_i = A_i\kappa_i'$ where $A_i'A_i = I_r$ in

$$\alpha_i\beta_i' = A_i\kappa_i'\beta_i' \equiv A_i\beta_i^{*\prime}$$

We will switch between $\beta_i^*|A_i$ and $\alpha_i|\beta_i$ in our MCMC algorithm.

Both of these are (conditionally) *Normal*. Very easy (and very efficient) computation.

Simulator 2:

Alternatively we can augment the posterior with an unidentified parameter D.

$$\alpha\beta = \alpha D D^{-1}\beta = \alpha^*\beta^*$$

The conditional posteriors $\alpha^*|\beta^*$ and $\beta^*|\alpha^*$ are again *Normal*.

We can use an MCMC algorithm which is again very easy (and very efficient) to implement.

Calculating the posterior probabilities of the models.

The vector $r = (r_1, r_2, \ldots, r_N)$ denotes the rank combinations for all individuals. This gives us a model.

Let M_r denote the model with r.
Let M_0 denote the base model $r = (0, \ldots, 0)$ used for comparison.

Bayes factors comparing M_0 to every other model M_r is

$$B_{0,r} = \frac{P\left(M_0|y\right)}{P\left(M_r|y\right)}.$$

and can be calculated using the Savage Dickey density ratio (Sugita, 2004).

Bayes factors can be used to calculate $P\left(M_r|y\right)$ for every model.

Empirical Work

A Monte Carlo exercise is in the paper (approach works well, unless the prior for ν^{-1} is seriously misspecified)

Here, we discuss a simple real-data empirical application using the monetary model of the exchange rate proposed by Groen (2000).

We have: $e_{i,t}$, the log of the exchange rate for country i at time t; $m_{i,t}$, the log of the ratio of the quantity of domestic to foreign money supply; and $x_{i,t}$, the log of the relative real income.

The theory implies the relation

$$e_{i,t} - \beta_1 m_{i,t} - \beta_2 x_{i,t} = \beta_0 + z_{i,t}$$

will be stationary (i.e., $z_{i,t}$ should be an $I(0)$ process) with $\beta_1 = 1$ and $\beta_2 < 0$.

If the variables in the vector $y_{i,t} = \left(e_{i,t}, m_{i,t}, x_{i,t}\right)$ are $I(1)$, this model implies they cointegrate with a particular cointegrating space.

The Data:

Quarterly U.S. dollar exchange rates and the log ratio of money (m) and income (x) for France ($i = 1$), Germany ($i = 2$), and the United Kingdom ($i = 3$) to the U.S. equivalents.

$$1973 : Q1 \text{ to } 1994 : Q4.$$

We expect $r = 1$ and $\beta_1 = 1$ and $\beta_2 < 0$.

The economic model of interest implies such a set of joint restrictions, some of which are linear and some are nonlinear.

Classical inference usually proceeds with a mixture of sequential testing and informal inference to gather evidence for or against the model, with no single statistic with known power to indicate the degree of support in favor of the model.

E.g., Groen (2000) tested sequentially $r = 1$ and the other restrictions, providing only informal evidence about the degree of support for the model.

Bayesian approach is able to provide a formal summary of the evidence for the model via posterior model probabilities.

In the VECM, the monetary exchange rate model implies $r_i = 1$ for all i and $\mathfrak{p}_i = \mathfrak{p} = sp\left(H\varphi_i\right)$ where:

$$H = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & 0 \\ 0 & 1 \end{bmatrix},$$

and $\varphi_i = \begin{bmatrix} \varphi_{1,i} \\ \varphi_{2,i} \end{bmatrix}$, $\varphi_i'\varphi_i = 1$

$$H\varphi_i = \begin{bmatrix} \frac{1}{\sqrt{2}}\varphi_{1,i} \\ -\frac{1}{\sqrt{2}}\varphi_{1,i} \\ \varphi_{2,i} \end{bmatrix}, \text{ with } \frac{\varphi_{2,i}}{\varphi_{1,i}} > 0.$$

The Results: 221 Models estimated.

The posterior model probabilities are spread over the model set: 28 models contain 98% of the probability.

r_1	r_2	r_3	o_1	o_2	o_3	E	Prob
1	1	0	1	1	0	0	0.35
1	1	0	1	0	0	0	0.12
1	1	0	0	1	0	0	0.10
1	1	2	1	1	1	0	0.10
1	1	1	1	1	1	0	0.05

($E = 1$ if $sp(\beta_1) = sp(\beta_2) = sp(\beta_3)$, else $E = 0$)

The 5 most likely models, which get 71.4% of the probability mass.

In these 5 models, $r = 1$ for France and Germany and $sp(\beta_i) \subseteq sp(H)$ in at least one country.

The model with $r_i = 1$ and $sp(\beta_i) \subseteq sp(H)$ $(\beta_1 = 1)$ for every i (the monetary exchange model) has a non negligible probability that is equal to 0.05.

Conditional on this model, $(\phi_{2i}/\phi_{1i} > 0 \equiv \beta_2 < 0)$

$$Pr(\phi_{2i}/\phi_{1i} > 0 \text{ for } i = 1, 2, 3) = 0.12,$$

which means that the probability of all the restrictions implied by the monetary exchange model holding in every country is

$$0.12 * 0.05 = 0.006.$$

Other probabilities of interest:

$$
\begin{aligned}
Pr(r_1 &= r_2 = r_3) = 0.09 \\
Pr(sp(\beta_1) &= sp(\beta_2) = sp(\beta_3)) = 0.004 \\
Pr(r_2 &= 1) = 0.86 \\
Pr(sp(\beta_1) &\subseteq sp(H), r_1 > 0) = 0.79.
\end{aligned}
$$

Finally,

$$
\text{Pr}(sp(\beta_i) \subseteq sp(H) \text{ and } r_i = 1 \\ \text{for at least one country}) = 0.94
$$

which gives further support to the monetary exchange model holding in at least one country.

Conclusion:

- We discuss Bayesian inference in cointegrated panel data models.

- We adopt a very general specification where each individual is characterized by its own vector error correction model.

- We allow for individuals to have common cointegrating rank and/or common cointegrating spaces.

- We develop priors which allows soft homogeneity restrictions for dynamics and equilibrium relations.

- Efficient posterior simulation is carried out using a collapsed Gibbs sampler.