Dynamic Cointegration Modelling with Applications in Commodities

Prof. Gareth W. Peters (YAS-RSE, FIOR)
Chair Professor for Statistics in Risk and Insurance
Director of the Scottish Financial Risk Academy.

HERIOT WATT UNIVERSITY
United Kingdom Dubai Malaysia

This work is based on the recent paper:

``Bayesian Inference for Dynamic Cointegration Models: a case study on the Soybean Crush Spread''

in revision JRSSC

Marowka M. and Peters G.W. and Kantas N. and Bagnarosa G.
Bayesian Inference for Dynamic Cointegration Models with Application to Soybean Crush Spread.
Available at SSRN: **https://ssrn.com/abstract=2960638**

Cointegration Modelling

Talk Structure:

- Context of Cointegration Modelling for Futures on Soy
- Cointegration Dynamic Bayesian Models
- Estimation and Sampling
- Results and Analysis

Context of the Study for Soy

Prof. Gareth W. Peters URL: http://garethpeters78.wixsite.com/garethwpeters
EMAIL: g.peters@hw.ac.uk

❑Cointegration Models

- We investigate how vector autoregressive (VAR) models can be used to study the soybean crush spread.

  Crush Spread a time series marking the difference between a weighted combination of value of **soymeal** and **soyoil** to the value of the original **soybeans**.

- Industry practitioners often use fixed prescribed values for the weights of long and short portfolio positions in these futures contracts:

  → based purely on physical refining and extraction context

- Can we instead take into account any time varying effects or any market dynamical features present in the data to form such a spread dynamic?

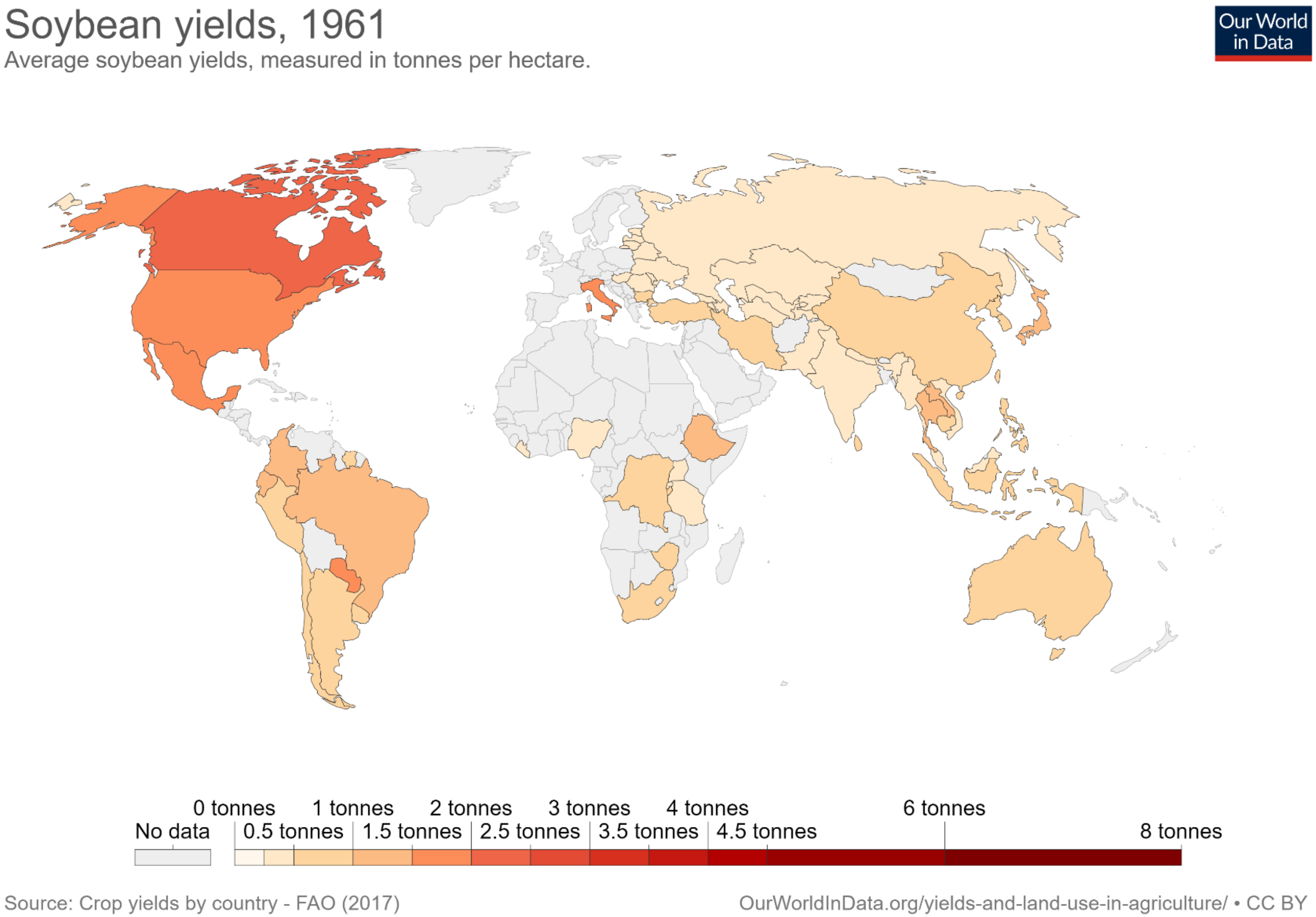
In this work we address this issue by proposing an appropriate time series model with cointegration.

❑Cointegration Models

Why is soybean crush so significant to focus a study around?

- Soybeans : **fourth most produced crop in the world** and if one includes the trading of by-products:
  - *most traded agricultural commodity representing in value 10% of the worldwide agricultural trade.*
- About 85% of the soybeans are either consumed once processed or transformed into **soymeal**, **soyoil** and **soyhull** (waste).
- The soymeal being very rich in protein is mainly used for animal feed (hog, livestock, fish, etc.) and is thus directly related to the worldwide meat and fish consumption.
- Soyoil is used as edible oil for cooking and more recently for biodiesel.

❑Cointegration Models

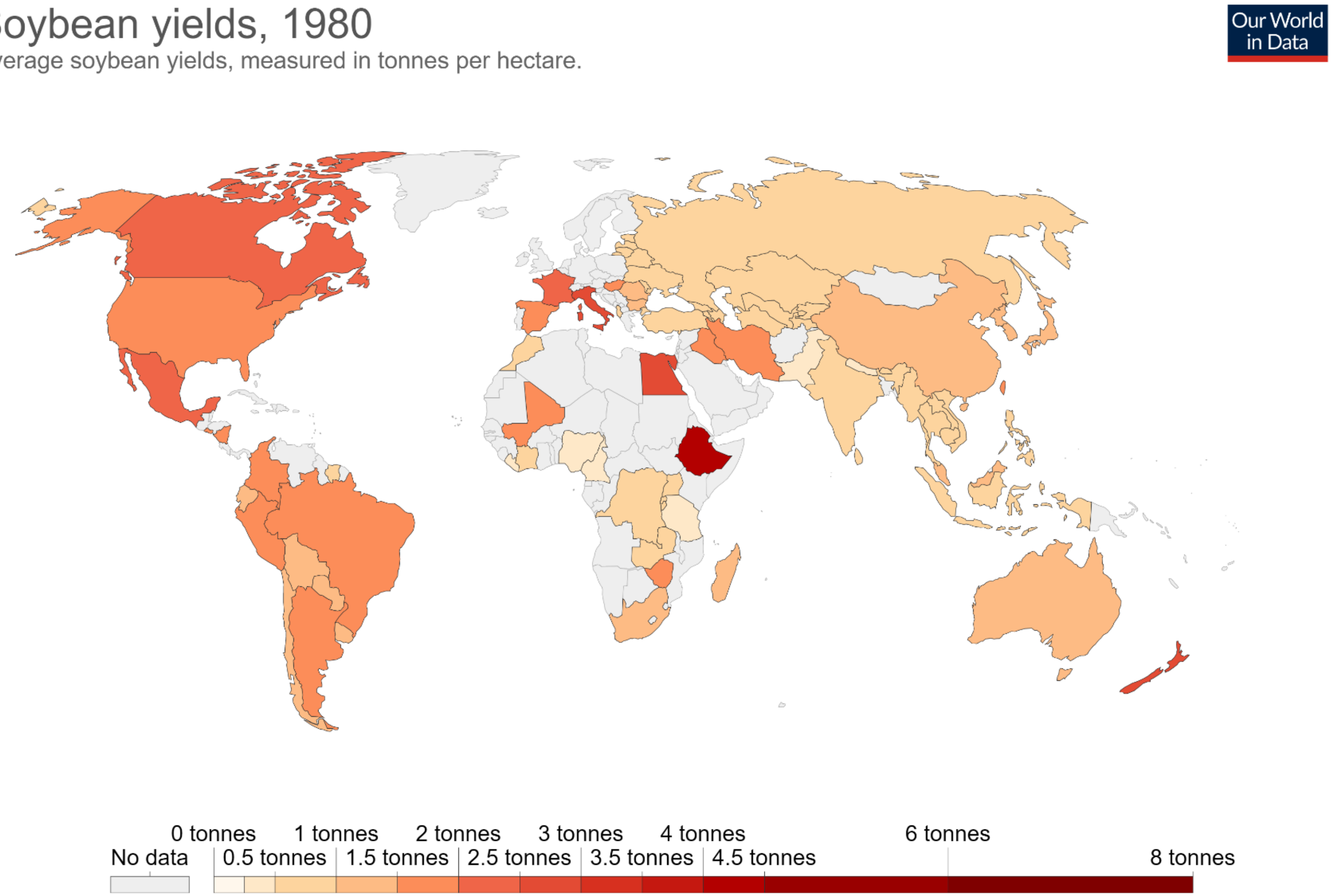


Prof. Gareth W. Peters URL: http://garethpeters78.wixsite.com/garethwpeters
EMAIL: g.peters@hw.ac.uk

❑Cointegration Models

Soybean yields, 1980

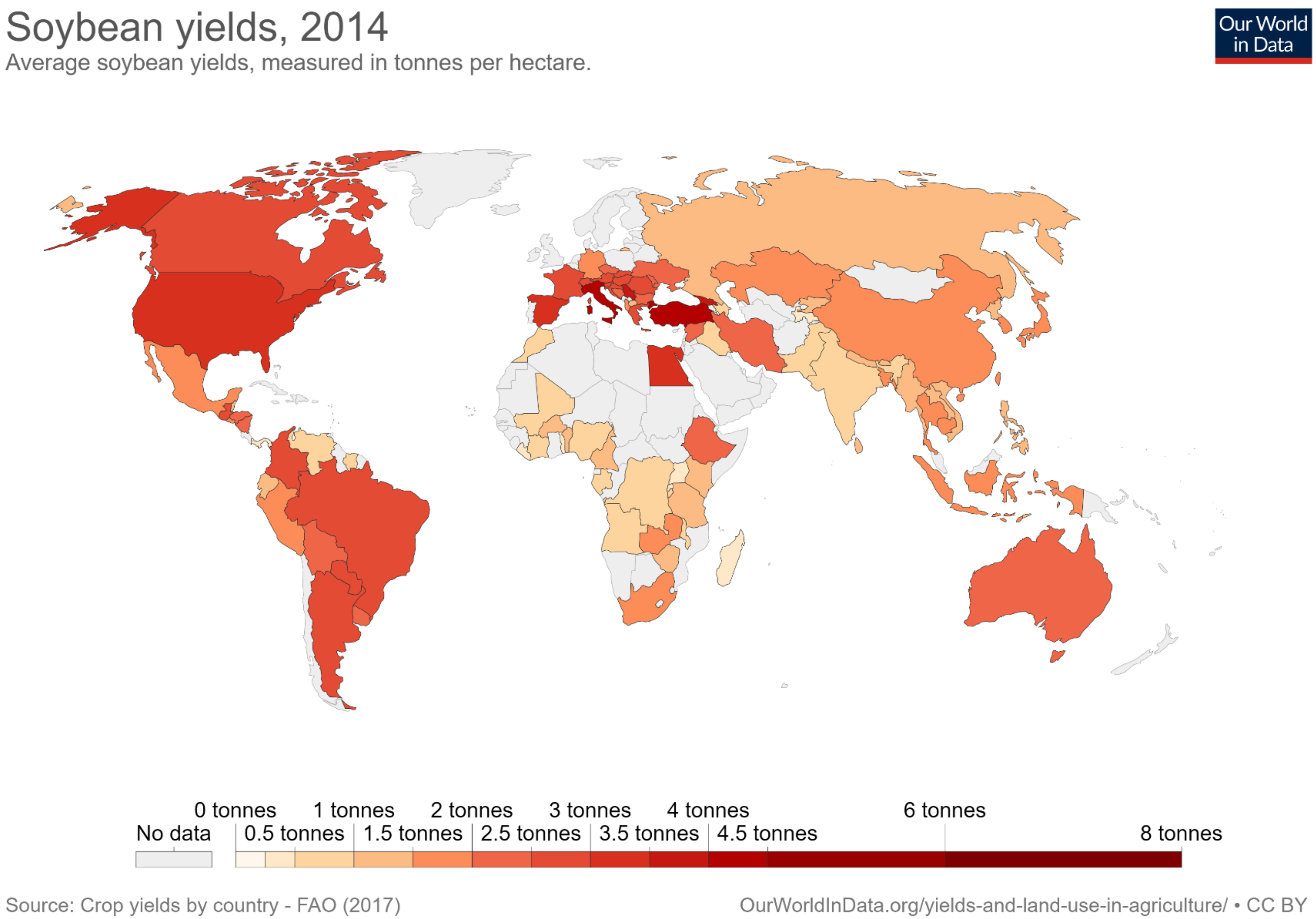
Average soybean yields, measured in tonnes per hectare.



Source: Crop yields by country - FAO (2017)

Prof. Gareth W. Peters URL: http://garethpeters78.wixsite.com/garethwpeters
EMAIL: g.peters@hw.ac.uk

❑Cointegration Models



❑Cointegration Models

Why is soybean crush so significant to focus a study around?

- Both soybean meal and soybean oil are a product of the same processing procedure

  → the total production of each will always remain constant relative to the other

- Therefore, ceteris paribus, a net increase in demand for one of the soy by-products will necessarily correlate with an increase in the demand for the underlying beans

  see Barrett and Kolb [1995], Simon [1999], Mitchell [2010].

❑Cointegration Models

Why is soybean crush so significant to focus a study around?

- **Soybean market:** unlike most other commodities, **derivative components of soybeans are also traded on commodity futures markets** by producers and speculators.

- Prices of soybean, soymeal and soyoil are generally strongly correlated with each other…

  - however, price shifts for meal and oil against the other can be a result of a shock on the demand or supply of one of the aforementioned markets.

Understanding the SOY Board Crush Spread Process

Prof. Gareth W. Peters URL: http://garethpeters78.wixsite.com/garethwpeters
EMAIL: g.peters@hw.ac.uk

❑Cointegration Models

Soybean industry: term **'crush'** refers both to a **physical process** as well as a **value calculation**.

- **Physical crush** is the process of converting soybeans into the by-products of soybean meal and soybean oil.

- **Crush spread** is a dollar value quoted as the difference between the combined sales values of the products and the cost of the raw soybeans.

  → *This value is traded in the:*
  - ***Spot cash (Gross Processing Margin)*** *and*
  - ***Futures markets (Board Crush)*** *based on expectations of future price movement of soybeans versus the components.*

  (We focus on futures case)

❑Cointegration Models

Crush value traded in the futures market is an inter-commodity spread transaction in which Soybean futures are bought (or sold) and Soybean Meal and Soybean Oil futures are sold (or bought).

The Board Crush spread is often used by processors to hedge the margin between the purchase price of soybeans and the combined selling price of the soybean meal and oil.

- The **November-December Board Crush (buying November Soybean futures and selling December Soybean Meal futures and December Soybean Oil futures)** is used to hedge new-crop gross processing margins **(cash spot crush)**.

❑Cointegration Models

November/December futures prices reflect the market's perception of conditions in the new soybean crop year.

Seasonal, cyclical and fundamental factors affect the Soybean Crush spread.

For example: soybean prices are typically lowest at harvest and trend higher during the year as storage, interest and insurance costs accumulate.

Other Driving Factors Include:

- Changes in demand for high protein feed over the course of the year
- Depletion of South American soybean stocks during the late fall and winter
- Crop size and yields,
- World demand,
- Carryover stocks,
- Third World purchases of edible oils, Malaysian palm oil production, European meal demand,
- Government programs
- Weather.

Prof. Gareth W. Peters URL: http://garethpeters78.wixsite.com/garethwpeters
EMAIL: g.peters@hw.ac.uk

❑Cointegration Models

What is a Soy Crush Physical Spread Relationship?

Industry practitioners typically use values based on physical considerations.

Chicago Board of Trade (CBOT) quotes the crush spread as (Nelson Low [2015])

$$y_t(1) - 11y_t(2) – 2.2y_t(3)$$

$y_t(1)$, $y_t(2)$ and $y_t(3)$ correspond to the quoted prices of futures contract associated respectively to soybean, soyoil and soymeal at a given time t.

The intuition here that typically when a bushel of soybeans weighing 60 pounds is crushed, it results in:

- **11 pounds of soybean oil,**
- **44 pounds of 48% protein soybean meal,**
- **4 pounds of hulls, and**
- **1 pound of waste.**

Caution with Contract Units

Board Crush spread includes:

- Soybean futures, which are traded in **cents per bushel**,
- Soybean Meal futures priced in **dollars per short ton**, and
- Soybean Oil futures traded in **cents per pound**.

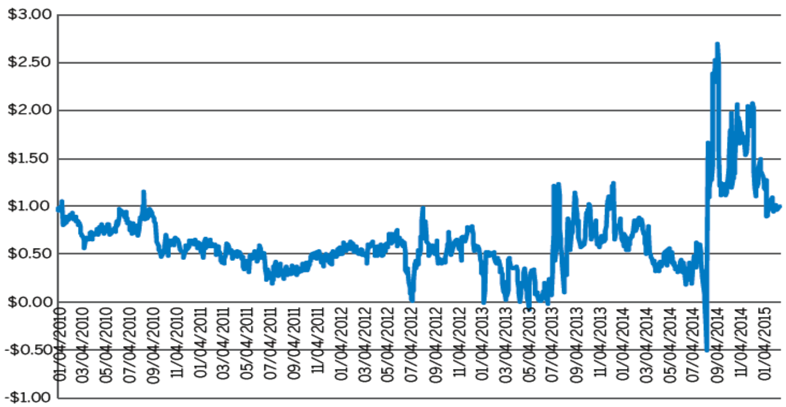
→ **necessary to convert soybean meal and soybean oil prices to cents per bushel** because of their different pricing units.

To Convert Prices into Cents per Bushel

Soybeans:	No conversion required Soybean futures are quoted in cents per bushel
Soybean Meal:	0.022 x price of soybean meal 44 lbs/2000 lbs = 0.022
Soybean Oil:	11 x price of soybean oil 11 pounds of oil per 60 lb. bushel

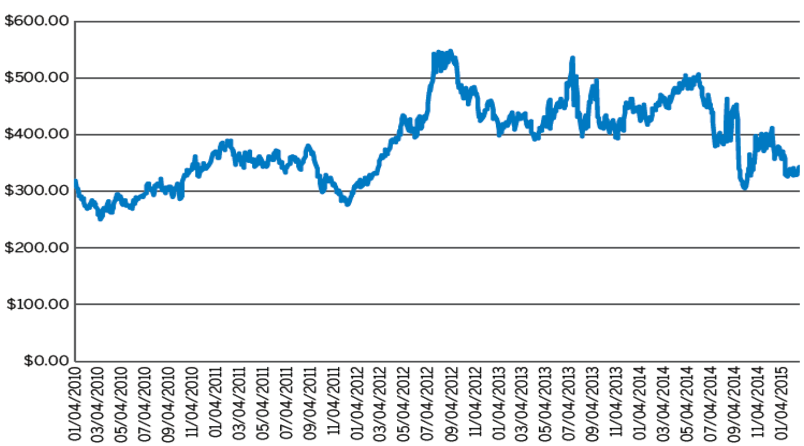
❑Cointegration Models

SOYBEAN CRUSH

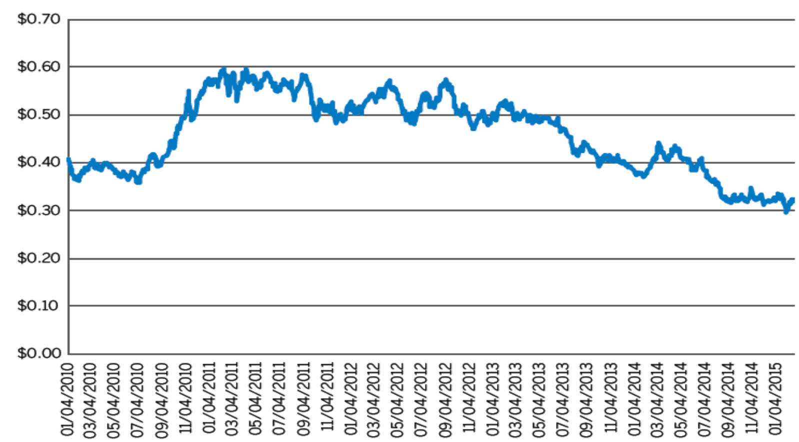


SOYBEAN

SOYBEAN MEAL



SOYBEAN OIL



[1] *Similar to other commodity spreads, the Soybean Board Crush is generally less risky than the outright components but there can be periods of time where the crush spread volatility can be relatively high.*

Preliminaries on Cointegration

Cointegration Preliminaries

A *vector autoregressive* (VAR) process is the multivariate analogue of a univariate autoregressive (AR) process. The VAR model of order p, or VAR(p), can be written:

$$\boldsymbol{x}_t = \boldsymbol{\mu} + A_1 \boldsymbol{x}_{t-1} + \cdots + A_p \boldsymbol{x}_{t-p} + \epsilon_t,$$

with

- $\boldsymbol{x}_t = (x_{1t}, \ldots, x_{nt})'$ is a $(n \times 1)$ random vector
- the A_i are fixed $(n \times n)$ coefficient matrices
- $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)'$ is a fixed $(n \times 1)$ vector of intercept terms and
- $\epsilon_t = (\epsilon_{1t}, \ldots, \epsilon_{nt})'$ is an n-dimensional *white noise* or *innovation process* [1].

---

[1] A white noise process has: $\mathbb{E}(\epsilon_t) = 0$, $\mathbb{E}(\epsilon_t, \epsilon_t') = \Sigma_\epsilon$, and $\mathbb{E}(\epsilon_t, \epsilon_s) = 0$ for $s \neq t$.

Prof. Gareth W. Peters URL: http://garethpeters78.wixsite.com/garethwpeters
EMAIL: g.peters@hw.ac.uk

❑Cointegration Preliminaries

- Ordinary VAR models are stationary and thus do not admit trends or shifts in the mean or the covariances, nor deterministic seasonal patterns.
- We will consider nonstationary processes of a very specific type; these will be allowed to have *stochastic trends* and are then called *integrated* $\Rightarrow$ *however we wont know these trend structures apriori*!
- If we find constraints that show
    1. Some of the variables move together in the long-run via these stochastic trends and
    2. they are in fact driven by a common source of stochasticity

  $\Rightarrow$ then such models are called *cointegrated*.

Prof. Gareth W. Peters URL: http://garethpeters78.wixsite.com/garethwpeters
EMAIL: g.peters@hw.ac.uk

❑Cointegration Preliminaries

- Cointegration represents the presence of equilibrium relationships between sets of time series variables.
- Writing the set of variables as a vector $x_t = (x_{1t}, \dots, x_{nt})'$ the long-run equilibrium can be expressed as

$$\beta' \boldsymbol{x}_t = \beta_1 x_{1t} + \cdots \beta_n x_{nt} = 0,$$

  where $\beta = (\beta_1, \dots, \beta_n)'$.
- **NOTE:** *For a particular time period the relationship may not be satisfied exactly*, but we may have $\beta' \boldsymbol{x}_t = \boldsymbol{e}_t$, where $\boldsymbol{e}_t$ is a stochastic variable representing deviations from the equilibrium.
- It is possible in this arrangement that the variables wander as a group, i.e. they are driven by a common stochastic trend.

Prof. Gareth W. Peters URL: http://garethpeters78.wixsite.com/garethwpeters
EMAIL: g.peters@hw.ac.uk

❑Cointegration Preliminaries

- *Granger representation theorem* states that a multivariate integrated process is cointegrated if and only if it can be represented in the ECM form with certain restrictions.

The general form for the *vector error correction model* (VECM) representation of an *n*-dimensional VAR process of order *p* is

$$\Delta \boldsymbol{x}_t = \boldsymbol{\Pi} \boldsymbol{x}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\Gamma}_i \Delta \boldsymbol{x}_{t-1} + \phi \boldsymbol{d}_t + \epsilon_t,$$

where:

$\Delta \boldsymbol{x}_t \in \mathbb{R}^n \quad = \boldsymbol{x}_t - \boldsymbol{x}_{t-1},$

$\boldsymbol{\Pi}$ is the $(n \times n)$ long-run multiplier matrix,

$\boldsymbol{\Gamma}_i$ is the $\mathrm{i}^{th}(n \times n)$ lag matrix,

$\boldsymbol{d}_t$ is an $(n \times 1)$ vector of deterministic terms (polynomial in t) ,

ϕ is an $(n \times n)$ matrix,

ϵ_t is an $(n \times 1)$ i.i.d. multivariate, correlated vector of errors.

❑Cointegration Preliminaries

The cointegration properties of the vector model depend upon the rank, r, of the *long-run multiplier matrix*, $\mathbf{\Pi}$.

- If the rank, r is equal to zero, then $\boldsymbol{x}_t$ has a stable VAR$(p-1)$ representation and exhibits no cointegrating relationships.
- If $r = n$, that is the matrix $\mathbf{\Pi}$ is full rank, then the VAR operator has no unit roots and $\boldsymbol{x}_t$ is a stable VAR(p) stationary process.
- If, however, the rank, r, is intermediate, $0 < r < n$, then the VECM process is cointegrated.

Prof. Gareth W. Peters URL: http://garethpeters78.wixsite.com/garethwpeters
EMAIL: g.peters@hw.ac.uk

❑Cointegration Preliminaries

- In the case that $0 < r < n$ then $\mathbf{\Pi}$ can be written as a matrix product $\mathbf{\Pi} = \alpha\beta'$ with α and β both of dimension $(n \times r)$ and both of rank r.
- The matrix β is still called the *cointegrating matrix* and its columns are the *cointegrating vectors* of the process.
- The matrix α is called the *loading matrix*, and the $\beta' \boldsymbol{x}_t$ are called the *common trends*.

❑Cointegration Preliminaries

HERIOT WATT UNIVERSITY United Kingdom Dubai Malaysia

We can now re-express the ECM model in a multivariate regression format:

$$Y = X\Gamma + Z\beta\alpha' + E$$

$$\underbrace{\overbrace{\begin{bmatrix} \Delta \boldsymbol{x}_p \\ \Delta \boldsymbol{x}_{p+1} \\ \vdots \\ \Delta \boldsymbol{x}_T \end{bmatrix}}^{Y}}_{t\times n} = \underbrace{\overbrace{\begin{bmatrix} 1 & \Delta \boldsymbol{x}'_{p-1} & \cdots & \Delta \boldsymbol{x}'_1 \\ 1 & \Delta \boldsymbol{x}'_p & \cdots & \Delta \boldsymbol{x}'_2 \\ \vdots & \vdots & \cdots & \vdots \\ 1 & \Delta \boldsymbol{x}'_{T-1} & \cdots & \Delta \boldsymbol{x}'_{T-p+1} \end{bmatrix}}^{X}}_{t\times a} \underbrace{\overbrace{\begin{bmatrix} \boldsymbol{\mu}_0 \\ \boldsymbol{\Psi}_1 \\ \vdots \\ \boldsymbol{\Psi}_{p-1} \end{bmatrix}}^{\Gamma}}_{a\times n} + \underbrace{\overbrace{\begin{bmatrix} \boldsymbol{x}_{p-1} \\ \boldsymbol{x}_p \\ \vdots \\ \boldsymbol{x}_{T-1} \end{bmatrix}}^{Z}}_{t\times n} \underbrace{\overbrace{\begin{bmatrix} \boldsymbol{\beta}_1 \\ \vdots \\ \boldsymbol{\beta}_r \end{bmatrix}'}^{\beta}}_{n\times r} \underbrace{\overbrace{\begin{bmatrix} \boldsymbol{\alpha}_1 \\ \vdots \\ \boldsymbol{\alpha}_r \end{bmatrix}}^{\alpha'}}_{r\times n} + \underbrace{\overbrace{\begin{bmatrix} \boldsymbol{\epsilon}_p \\ \boldsymbol{\epsilon}_{p+1} \\ \vdots \\ \boldsymbol{\epsilon}_T \end{bmatrix}}^{E}}_{t\times n}$$

where t is the number of rows of Y, hence $t = T - p + 1$ giving X dimension $t \times (1 + n(p-1)) = t \times a$, Γ with dimension $a \times n$.

Prof. Gareth W. Peters URL: http://garethpeters78.wixsite.com/garethwpeters
EMAIL: g.peters@hw.ac.uk

❑Cointegration Preliminaries

NOTE: working with ECM form is optimal for forecasting!

- In situations where cointegration exists, the VECM representation will generate superior forecasts than the corresponding first-differenced form representation, especially over medium and long time horizons.

- The reason for this is that under cointegration, $\boldsymbol{z}_t = \beta' \boldsymbol{x}_t$, will have finite forecast error variance, whereas any other linear combination of the forecasts of the series in $\boldsymbol{x}_t$ will have infinite variance [Engle, 1987].

❑Cointegration Preliminaries

Practically

- Any system of financial asset prices with a mean-reverting spread will have some degree of cointegration.
- *Mean reverting* is the term we use in finance to refer to stationary processes.
- If spreads are mean-reverting, assets will be tied together in the long-term by a common stochastic trend.

Cointegration Modelling and Latent Dynamics Structures

❑Cointegration Models

The observable vector $y_t \in \mathbb{R}^{n\times 1}$ is integrated of order 1, $I(1)$, with r linear cointegration relationships.

$$\Delta y_t = \alpha\beta^T y_{t-1} + \mu_t + g(t) + \epsilon_t$$

The parametric seasonal model we consider in this paper is given by:

$$g(t) = \xi_1 + \sum_{i=2}^{12} \xi_i \mathbb{I}_{\{t\in D_i\}}$$

$\mu_t \in \mathbb{R}^{n\times 1}$ is a mean level of the implied spread series and a latent stochastic process; model it according to the two dimensional dynamic factor specification as in the state equations:

$$\mu_t = Hx_t$$
$$x_t = Bx_{t-1} + \delta_t$$

❑Cointegration Models

Identification Challenges

Cointegration Parameter Space

vs

Manifold Characterization.

Prof. Gareth W. Peters URL: http://garethpeters78.wixsite.com/garethwpeters
EMAIL: g.peters@hw.ac.uk

❑Cointegration Models

Identification Challenges in Likelihood:

For any orthogonal matrix $U \in \mathbb{R}^{K \times K}$ consider the transformation

$$T(H, B, Q) = (HU^{-1}, UBU^{-1}, UQU^{T})$$

It is straightforward to show that the marginal likelihood satisfies

$$p(Y|\alpha, \beta, H, B, Q, R, \xi) = p(Y|\alpha, \beta, T(H, B, Q), R, \xi).$$

→ Just integrate out the latent state dynamic!

Prof. Gareth W. Peters URL: http://garethpeters78.wixsite.com/garethwpeters
EMAIL: g.peters@hw.ac.uk

❑Cointegration Models

Restriction Based Identifications:

For any orthogonal matrix A, both pairs of parameters (α, β) and $(\alpha A, \beta A)$ are indistinguishable under the cointegration model.

The following identification restriction can be used

$$\beta^T \beta = I_r$$

where r is the cointegration rank.

This restricts β to belong to the following Stiefel manifold

$$\mathbb{V}_{n,r} := \{V \in \mathbb{R}^{n \times r} : \; V^T V = I_r\}$$

which is compact, so the uniform distribution is a proper prior.

❑Cointegration Models

Restriction Based Identifications:

In the linear normalization, $\beta = [I_r, \beta^*]$, if β^* follows matrix variate t-distribution, then $col(\beta)$ has uniform distribution on the Grassman manifold.

- *a space parametrizing all k-dim linear subspaces of the n-dim vector space*
- *It has a natural manifold structure as an orbit-space of the [Stiefel manifold]()*

A possible drawback of working in this parameter space and specifying such a prior:

→**Class of distributions is largely unexplored for r > 1**;

→ **their second and larger moments do not exist.**

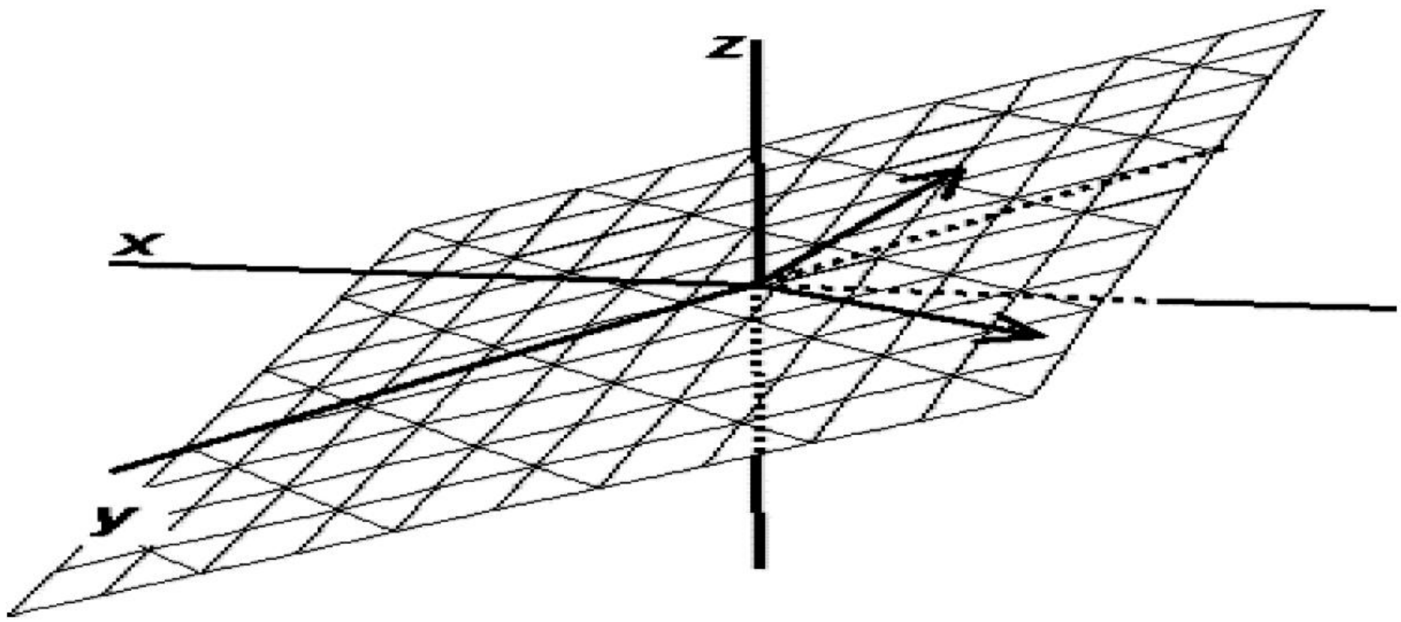
In fact, it is only the space spanned by the columns of cointegration vectors $sp(\beta)$ which is identified.

Prof. Gareth W. Peters URL: http://garethpeters78.wixsite.com/garethwpeters
EMAIL: g.peters@hw.ac.uk

❑Cointegration Models

Identification Challenges:

The cointegration space is an **r -dimensional hyperplane** in a n-dimensional space and **the cointegrating vectors identify this plane.**



❑Cointegration Models

Priors on Manifolds and Haussdorf measure

Defining a prior on the cointegration space requires working with a Haussdorf measure on an appropriate manifold
→ **Grassman manifold.**

In particular, the Haussdorf measure is related to the Lebesque measure according to the expression:

$$\mathcal{H}(dM) = \sqrt{G(M)}\mathcal{L}(M)$$

which is a Lebesque measure scaled by the volume element G(M) which is a Riemanian metric on the manifold.

Prof. Gareth W. Peters URL: http://garethpeters78.wixsite.com/garethwpeters
EMAIL: g.peters@hw.ac.uk

❑Cointegration Models

Bayesian Model Formulation – Priors on Manifolds

Prior distribution for β is a matrix angular central Gaussian distribution, i.e.

$$dp(\beta) \propto |P_\tau|^{-r/2} |\beta^T (P_\tau)^{-1} \beta|^{-n/2} d\mathcal{H}(\beta)$$

where $P_\tau = HH^T + \tau H_\perp H_\perp^T$; $\tau \in [0,1]$, and the distribution is with respect to Haussdorf measure on Stiefel manifold.

P_τ determines the central location of $sp(\beta)$ and τ amount of the dispersion around the central location. If $\tau = 1$, then $P_\tau = I_r$ and defines uniform prior on the Grassman manifold;

Prof. Gareth W. Peters URL: http://garethpeters78.wixsite.com/garethwpeters
EMAIL: g.peters@hw.ac.uk

❑Cointegration Models

For α, we choose a shrinkage prior with zero prior mean:

$$\alpha|\beta \sim N_{n\times r}(0, (v\beta^T P_{1/\tau}\beta)^{-1}, G),$$

where v controls degree of shrinkage; So that $Vec(\alpha)|\beta \sim N(0, \Sigma_\alpha)$ where $\Sigma_\alpha = (v\beta^T P_{1/\tau}\beta)^{-1} \otimes G$.

The semi-orthogonality restriction of β implies that the conditional posterior of β is non-standard.

$$\beta\alpha^T = (\beta\kappa)(\alpha\kappa^{-1})^T = \mathcal{B}\mathcal{A}^T,$$

where

$$\kappa = (\alpha^T\alpha)^{1/2} = (\mathcal{B}^T\mathcal{B})^{1/2},$$
$$\beta = \mathcal{B}(\mathcal{B}^T\mathcal{B})^{-1/2}.$$

and $\mathcal{A}$ is semiorthogonal now, while $\mathcal{B}$ is unrestricted.

Prof. Gareth W. Peters URL: http://garethpeters78.wixsite.com/garethwpeters
EMAIL: g.peters@hw.ac.uk

❑Cointegration Models

Bayesian Model Formulation
→ Hierarchical Priors on Manifolds

Given the hierarchical prior structure on (α, β) the prior for $\mathcal{A}$ and $\mathcal{B}$ is given by:

$$
\begin{aligned}
dp(\mathcal{A}) &\propto |G|^{-r/2}|\mathcal{A}^T G^{-1}\mathcal{A}|^{-n/2} d\mathcal{H}(\mathcal{A}), \\
p(Vec(\mathcal{B})|\mathcal{A}) &\sim N(0, \Sigma_{\mathcal{B}})
\end{aligned}
$$

where $\Sigma_{\mathcal{B}} = \mathcal{A}^T G^{-1}\mathcal{A})^{-1} \otimes \nu P_{\tau}$

Prof. Gareth W. Peters URL: http://garethpeters78.wixsite.com/garethwpeters
EMAIL: g.peters@hw.ac.uk

❑Cointegration Models

Bayesian Model Formulation

Parameters	Prior
B	$N_{K,K}(0, \sigma_B^2 I_K, \sigma_B^2 I_K)$
Q	$\mathcal{W}^{-1}(\nu_Q, I_K)$
H	$N_{n,K}(0, \sigma_H^2 I_n, \sigma_H^2 I_K)$
R	$\mathcal{W}^{-1}(\nu_R, \sigma_R^2 I_n)$
ξ	$N_{n,m}(0, \sigma_\xi^2 I_n, \sigma_\xi^2 I_m)$

Hyper parameters	Prior
σ_H^2	$G(\alpha_H, \beta_H)$
σ_B^2	$IG(\alpha_B, \beta_B)$
σ_R^2	$IG(\alpha_R, \beta_R)$

❑Cointegration Models

Sampling and Bayesian Estimation:

- Two Advanced MCMC Approaches Studied

1. Geodesic Hamiltonian Monte Carlo

2. Rao-Blackwellized Partially Collapsed Gibbs Sampler

→ Kalman Filter Forward Backward Sampler

→ Collapsed Gibbs Sampler

Prof. Gareth W. Peters URL: http://garethpeters78.wixsite.com/garethwpeters
EMAIL: g.peters@hw.ac.uk

Sampling and Bayesian Estimation:

This work is based on the recent paper:

``Estimation of Cointegrated Spaces: Numerical Case Study on Efficiency, Accuracy and Influence of the Model Noise''

Marówka, M., Peters, G. W., Kantas, N., & Bagnarosa, G. Estimation of cointegrated spaces: A numerical case study on efficiency, accuracy and influence of the model noise. 2017.
Available at SSRN: https://ssrn.com/abstract=2918511

NOTE: Geodesic Samplers on Manifolds in Hamiltonian Structure is involved → see paper... next we explain briefly the second class of method developed

❑Cointegration Calibration Results

Bayesian Estimators for Posteriors with Support on Manifolds

Prof. Gareth W. Peters URL: http://garethpeters78.wixsite.com/garethwpeters
EMAIL: g.peters@hw.ac.uk

❑Cointegration Models

Bayesian Point Estimators on a Manifold → requires some attention!

Obtaining point estimates of the cointegration space needs attention when the posterior distribution is defined on a Stiefel manifold.

We use the Frobenius norm as the loss function

$$l(\beta, \beta^*) = ||\beta\beta^T - \beta^*(\beta^*)^T||_F$$

where β and β^* are semi-orthogonal.

Prof. Gareth W. Peters URL: http://garethpeters78.wixsite.com/garethwpeters
EMAIL: g.peters@hw.ac.uk

❑Cointegration Models

HERIOT WATT UNIVERSITY
United Kingdom Dubai Malaysia

Bayesian Point Estimators on a Manifold → requires some attention!

To provide a Bayesian point estimate of the posterior distribution for $col(\beta)$ we use the Posterior Mean Cointegration Space Estimator (PMCS) proposed in Bernardo and Smith [2001] :

$$\hat{\beta} \stackrel{def}{=} \arg\min_{\tilde{\beta} \in \mathbb{V}_{n,r}} \mathbb{E}\left[l(\beta, \tilde{\beta}) \middle| y_{1:T} \right].$$

Which can be computed via $\hat{\beta} = (v_1, \ldots, v_r)$

where v_i is the eigenvector of $\mathbb{E}[\beta\beta^T | y_{1:T}]$ corresponding to i-th largest eigenvalue.

❑Cointegration Models

HERIOT WATT UNIVERSITY United Kingdom Dubai Malaysia

Bayesian Point Estimators on a Manifold → Precision

As a tool to assess the variation of posterior cointegration space distribution from the output of MCMC sampler, we will use the projective Frobenius span variation, Villani [2006].

The projective Frobenius Span Variation (FSV) is defined as:

$$\tau_{sp}^2 = \frac{\mathbb{E}\left[d(\beta, \hat{\beta}) \middle| y_{1:T} \right]}{r(p-r)/p},$$

$$\widehat{\tau_{sp}^2} = \frac{r - \sum_{i=1}^{r} \lambda_i}{r(p-r)/p},$$

where λ_i is the i-th largest eigenvalue of $\frac{1}{N}\sum_{i=1}^{N} \beta_i \beta_i^T$.

❑Cointegration Models

HERIOT WATT UNIVERSITY United Kingdom Dubai Malaysia

Diagnostic Measures for Estimation Performance of Cointegration Space Sampler

Two point estimates of the cointegration matrix β_1 and β_2

We can decompose β_2 as follows:

$$\beta_2 = \beta_1 \gamma_1 + \beta_{1\perp} \gamma_2,$$

where $\gamma_1 = \beta_1^T \beta_2$ and $\gamma_2 = \beta_{1\perp}^T \beta_2$.

A distance measure between β_1 and β_2 is

$$d_s(\beta_1, \beta_2) = tr(\beta_2 \beta_{1\perp} \beta_{1\perp}^T \beta_2)^{1/2}.$$

❑Cointegration Models

Diagnostic Measures for Estimation Performance of Cointegration Space Sampler

Measuring a dissimilarity of $col(\beta_1)$ and $col(\beta_2)$. We have, $d_s(\beta_1, \beta_2) \leq min\left(r, (n-r)\right)$, which is useful for interpretation.

It can be found that the 1% quantile of the simulated distribution of $d(\beta^{fixed}, \beta)$ with $\beta \sim Unif$ on the Stiefel Manifold is 0.14

Prof. Gareth W. Peters URL: http://garethpeters78.wixsite.com/garethwpeters
EMAIL: g.peters@hw.ac.uk

❑Cointegration Calibration Results

Calibration Results for Board Crush Spread Analysis

❑Cointegration Models

Study Set-up

We use daily observations of adjusted prices over a period between 2001-2015 and perform estimation separately for each year.

→*takes into consideration the four seasons and the associated soybean cycle of production and processing*

Allows for a comparison of our model and estimation results under different years each of which could be influenced differently by events such as weather or market dynamics.

❑Cointegration Models

Study Set-up

Raw data set is a panel of futures curves for soybean, soyoil and soymeal reported daily together with the corresponding volumes.

Before doing any statistical analysis we need to determine which commodity contracts to use for $\mathbf{y}_t$.

We use a commonly used strategy in the industry called the **rollover strategy**:
→ assigns $\mathbf{y}_t$ to the quoted contract that day with the highest volume, i.e. the most active contract across the futures curve is being traded. Lucia and Pardo [2010], Carchano and Pardo [2009].

Cointegration Models

Results for Cointegration Vectors PMCS.

We present the PMCS estimates, $\hat{\beta} \stackrel{def}{=} \underset{\tilde{\beta} \in \mathbb{V}_{n,r}}{\arg\min}\, \mathbb{E}\left[l(\beta, \tilde{\beta}) \middle| y_{1:T} \right]$ for each year

→ For convenience we normalised first entry to 1.

We also present results for the posterior mean of the spectral radius of state transition matrix B and A given by:

$$A = I_r + \beta^T \alpha$$

each denoted as $\rho(B)$ and $\rho(A)$ respectively.

Note that both B and A are autoregression matrices for x_t and z_t respectively, so we expect the spectral radii to be less than 1 so that a posteriori suggests μ_t and z_t are stationary ($z_t = \beta^T y_t$)

❑Cointegration Models

Results for Cointegration Vectors PMCS.

An interesting observation is that under our model for 2011 the signs $\hat{\beta}$ differ from the β^C, which from a trading perspective suggests whether a long or short position should be taken.

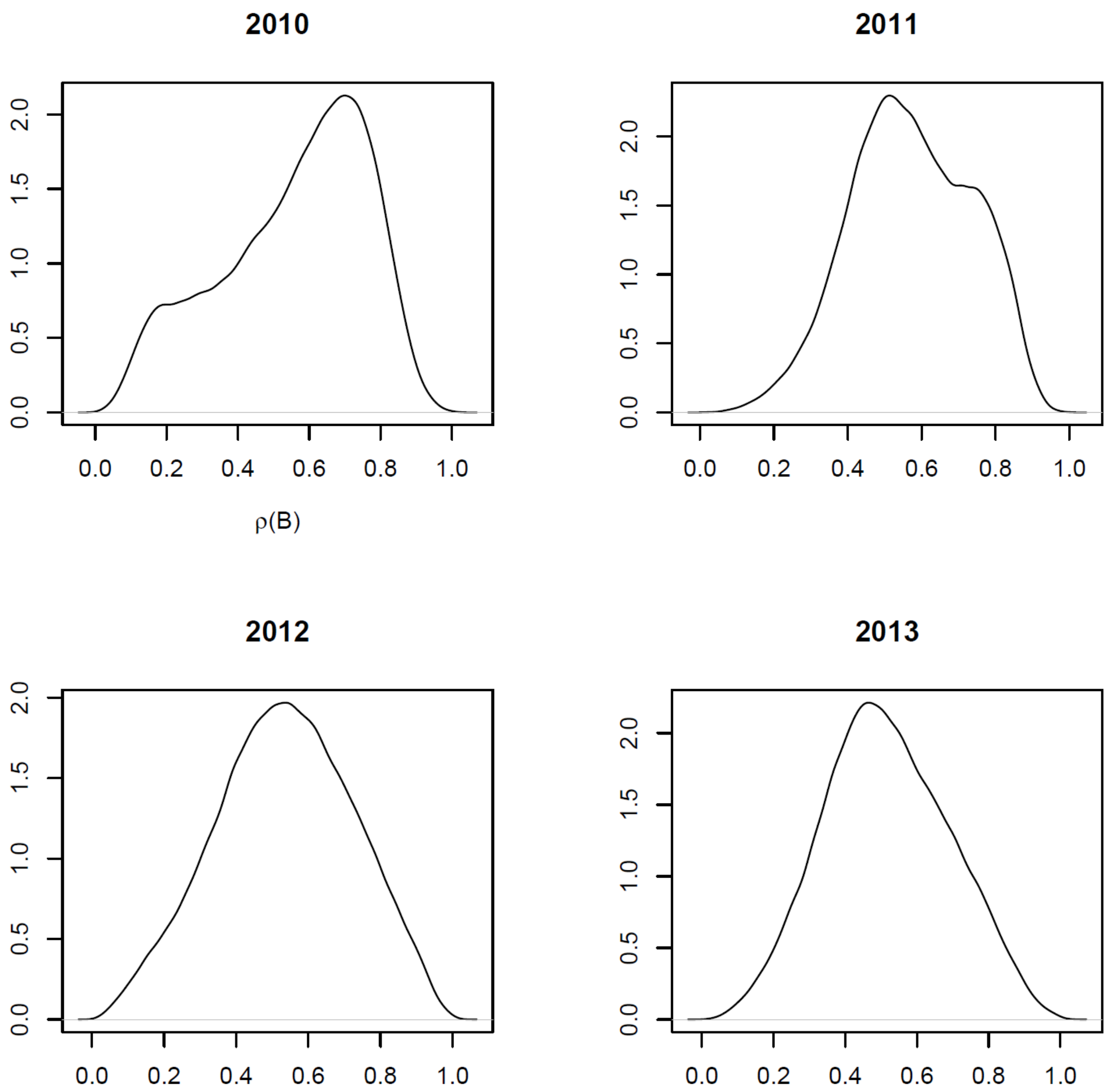
Posterior densities for $\rho(B)$ typically have support in (0, 1), which is evidence for using μ_t in our model.

❑Cointegration Models

MCMC estimates for $\hat{\beta}$, $\rho(B)$ and $\rho(A)$.

$Year$	$(\hat{\beta}(2), \hat{\beta}(3))$	$\hat{\rho}(A)\ (Var(\rho(A)\|y))$	$\hat{\rho}(B)\ (Var(\rho(B)\|y))$
2001	$(-21.20, \quad -1.41)$	$0.91(0.01)$	$0.51(0.19)$
2002	$(-20.58, \quad -3.23)$	$0.92(0.02)$	$0.45(0.16)$
2003	$(-54.53, \quad -0.18)$	$0.86(0.02)$	$0.21(0.11)$
2004	$(-15.91, \quad -2.21)$	$1.02(0.01)$	$0.33(0.11)$
2005	$(-17.52, \quad -2.14)$	$0.94(0.009)$	$0.26(0.10)$
2006	$(-28.58, \quad -1.53)$	$0.94(0.012)$	$0.34(0.09)$
2007	$(-12.15, \quad -2.00)$	$0.87(0.02)$	$0.72(0.12)$
2008	$(-11.10, \quad -2.14)$	$0.74(0.04)$	$0.32(0.18)$
2009	$(-12.92, \quad -2.03)$	$0.98(0.01)$	$0.27(0.11)$
2010	$(-11.92, \quad -2.15)$	$0.92(0.01)$	$0.33(0.11)$
2011	$(0.59, \quad -3.29)$	$0.96(0.01)$	$0.41(0.16)$
2012	$(-10.04, \quad -2.96)$	$0.97(0.01)$	$0.34(0.14)$
2013	$(-4.39, \quad -1.80)$	$0.99(0.01)$	$0.44(0.11)$
2014	$(-15.88, \quad -1.41)$	$0.99(0.02)$	$0.39(0.17)$
2015	$(-21.20, \quad -2.69)$	$0.87(0.02)$	$0.72(0.12)$

❑Cointegration Models

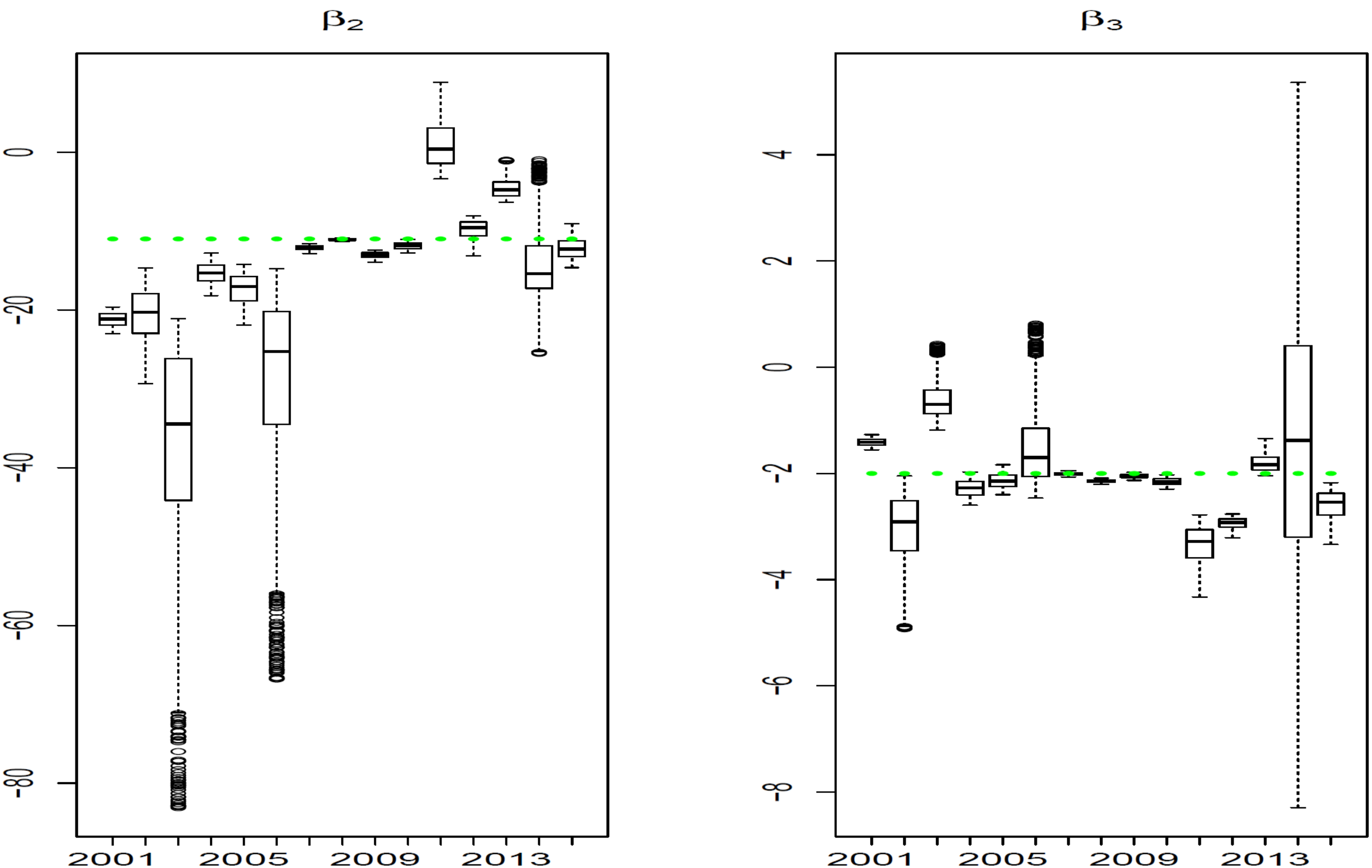


Posterior density of the spectral radius of autoregression matrix B.

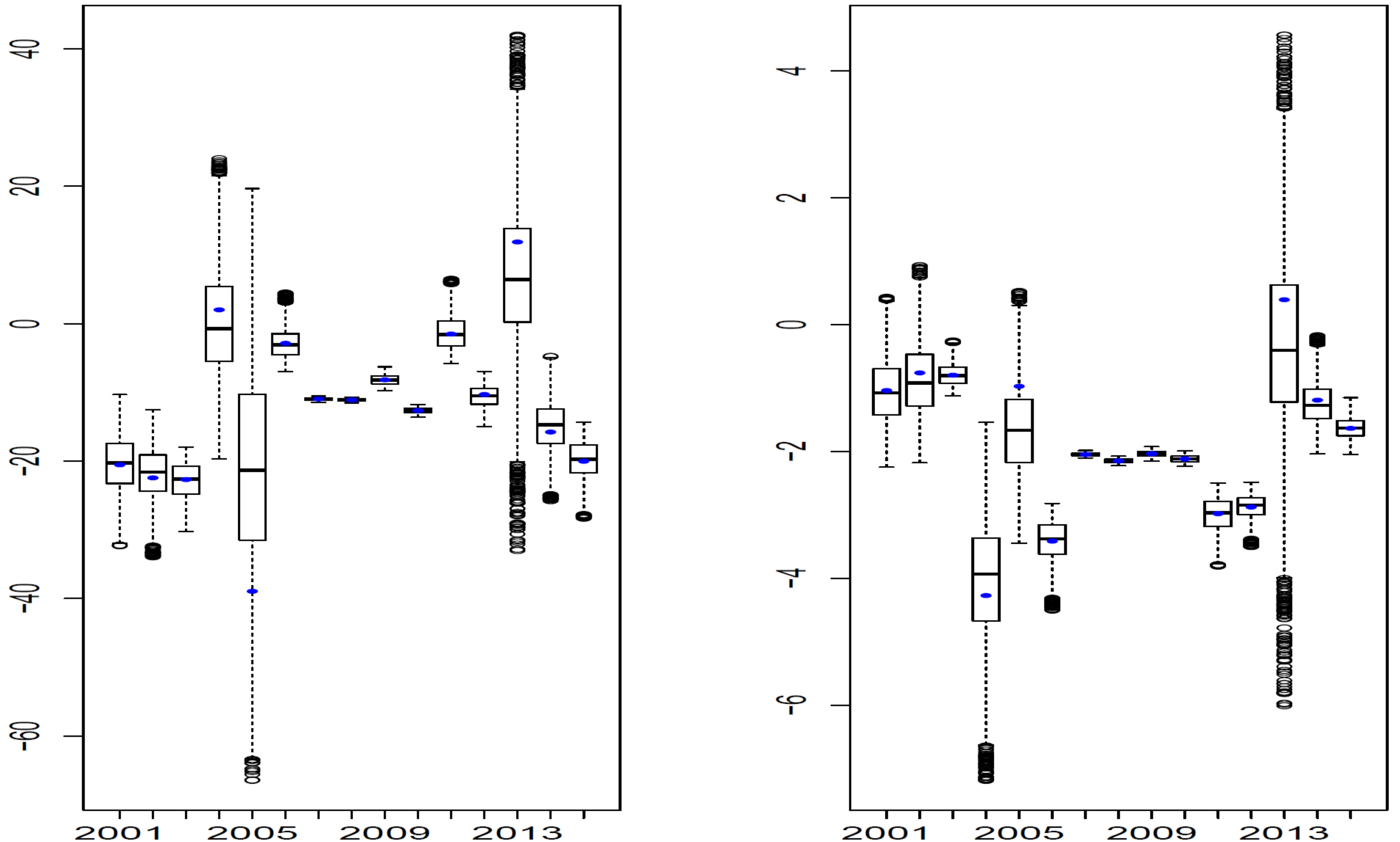
→ Note all have support in (0,1).

box-plots represent posterior credible intervals.

Green dots are $\dot{\beta}^C$



Frequentist Johansen Procedure
Blue dots are Johansen restricted MLE
box-plots represent bootstrap confidence intervals.



❑Cointegration Models

Model Structure Selection

- $\mathcal{M}_1^I$ uses full model for (2) with an informative prior on β using the PMCS estimator of the previous year.
- $\mathcal{M}_1^U$ as above but with a uniform prior on β,
- $\mathcal{M}_2$ as above but with $\mu_t = 0$ and fixed.
- $\mathcal{M}_3$ as above but with both $\mu_t = 0$ and $\alpha, \beta = 0$, i.e. using only a time varying intercept with no cointegration.

Our main motivation here is to seek numerical evidence for including μ_t in our model.

Assuming equal prior probability of each model, we use the Bayes factor to compare the following models:

$$R^{I/U} = \frac{p(y|\mathcal{M}_1^I)}{p(y|\mathcal{M}_1^U)},\ R^{U1/2} = \frac{p(y|\mathcal{M}_1^U)}{p(y|\mathcal{M}_2)},\ R^{2/3} = \frac{p(y|\mathcal{M}_2)}{p(y|\mathcal{M}_3)}.$$

Cointegration Models

Bayes Factors for each year.

$Year$	$R_1^{I/U}$	$R^{U1/2}$	$R^{2/3}$
2001	-	1.70	10.20
2002	1.01	1.64	1.00
2003	1.93	1.45	0.98
2004	1.00	1.29	0.99
2005	0.99	1.60	0.99
2006	1.01	1.56	8.19
2007	0.99	1.88	2.92
2008	1.00	1.28	2.22
2009	1.02	1.38	1.01
2010	1.08	1.37	0.99
2011	0.99	1.45	0.01
2012	0.98	1.19	0.99
2013	0.94	1.30	0.01
2014	0.82	1.25	1.05
2015	1.02	1.26	1.02

$(R^{\cdot})^{1/T}$ Used due to the length of the time Series when presenting estimated Bayes factors

$R^{U1/2}$ provides clear evidence for Including μ_t in every year!

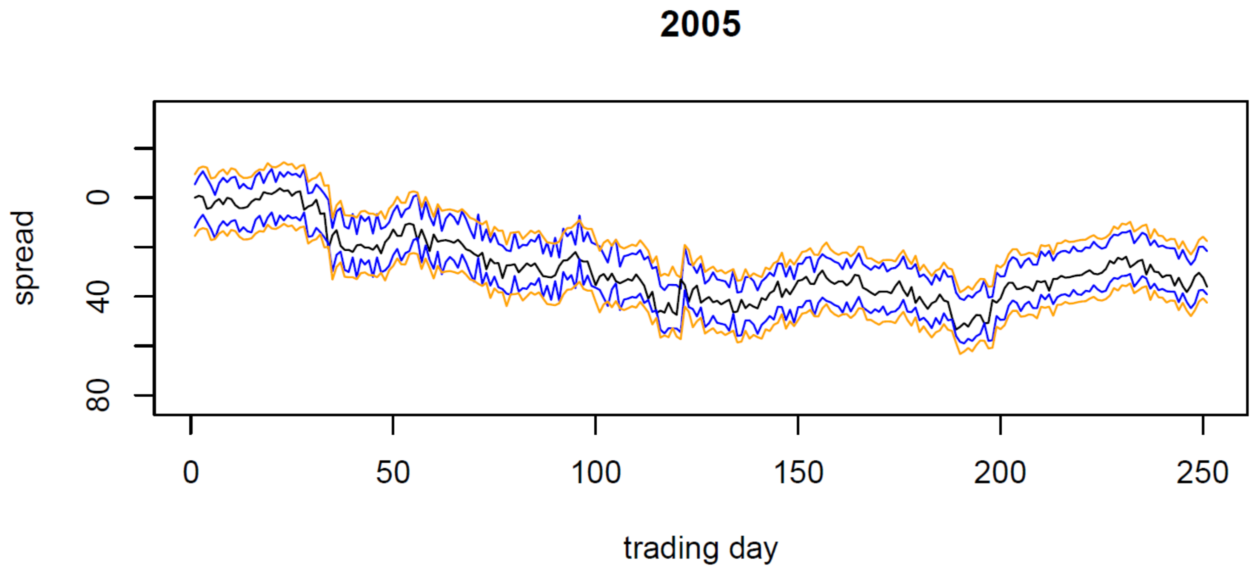
$R^{2/3}$ in some years presence of cointegration relationship is clear, this not always the case when μ_t is not included.

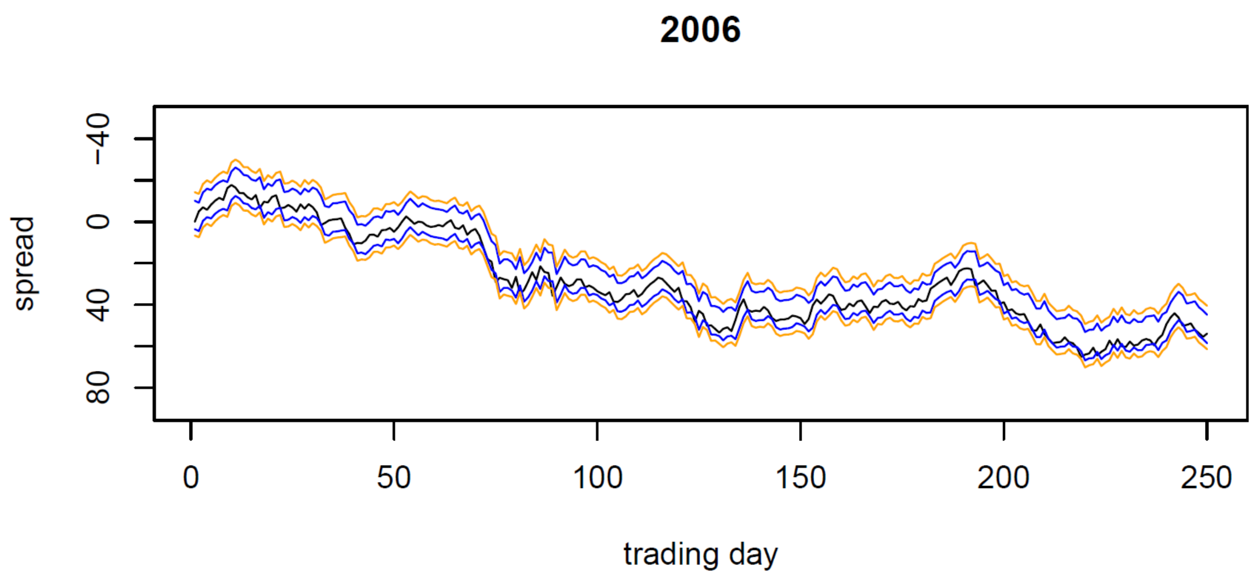
$R^{2/3} \times R^{U1/2}$

always gives values bigger than 1, so not including a term like μ_t in a cointegration model could mask a cointegration relationship here.

Prof. Gareth W. Peters URL: http://garethpeters78.wixsite.com/garethwpeters
EMAIL: g.peters@hw.ac.uk

Predictive Spread Series Forecasts





Out of sample predictive posterior bands for

$$p(z_t | y_{1:t-1}, \tilde{y})$$

under the models

$\mathcal{M}_1^U$ (orange lines)

$\mathcal{M}^2$ (blue lines)

❑Conclusions

There are several advantages in using a CVAR model to estimate the cointegration vectors instead of using physical considerations related to the crush spread.

Our proposed model is more flexible than common CVAR models by including effects related to seasonality and being able to capture time varying behaviour related to aspects that are difficult to model such as market dynamics, climate variation and level of speculation.

Our data analysis for the soy beans crush spread indicates that our proposed model leads to accurate estimation and more confident predictions for the spread series.

❑Conclusions

Bayesian Cointegrated Vector Autoregression models incorporating Alpha-stable noise for inter-day price movements via Approximate Bayesian Computation.
Peters GW, Kannan B, Lasscock B, Mellen C, Godsill S., Bayesian Analysis. 2011;6(4):755-92.

Model selection and adaptive Markov chain Monte Carlo for Bayesian cointegrated VAR models.
Peters GW, Kannan B, Lasscock B, Mellen C., Bayesian Analysis. 2010;5(3):465-91.

Rank estimation in Cointegrated Vector Auto-Regression models via automated Trans-dimensional Markov chain Monte Carlo.
Peters GW, Lasscock B, Balakrishnan K., In Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), 2011 4th IEEE International Workshop on 2011 Dec 13 (pp. 41-44). IEEE.

24

Prof. Gareth W. Peters URL: http://garethpeters78.wixsite.com/garethwpeters
EMAIL: g.peters@hw.ac.uk