MLE of Fractionally Cointegrated Systems

Katarzyna Łasak

CREATES,

Aarhus University, Denmark

Toronto, June 2010

What is cointegration?

- A standard tool in econometrics since Granger (1981)
- Existence of a stationary relation between nonstationary variables
- Robinson and Yajima (2002) revise possible definitions

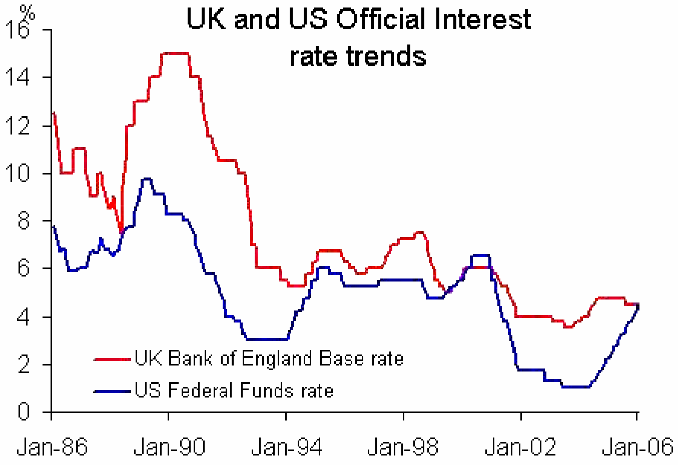
Definition

The p-vector time series X_t is cointegrated if $X_t \in I(\delta)$, but there exists a full rank $p \times r$ matrix β such that $\beta' X_t \in I(\delta - d)$ for $d > 0$. The number r is the cointegration rank and the space spanned by the columns of β is the cointegration space.

- In the standard cointegration setup $\delta = d = 1$

and Johansen's procedure applies, see Johansen (1988, 1991, 1995)

Cointegrated series



source: http://www.hbosplc.com/media/includes/Interest%20rate%20trends.jpg

Why fractional errors?!

Standard analysis: deviations from equilibrium are I(0)

$\Rightarrow$ far too restrictive!

- Fractional cointegration has the same implications,
  exists equilibrium but with slower rate of convergence
- Fractional cointegration parameter is ignored,
  so FI equilibrium error results in a misspecified likelihood
- May imply loss of power for fractional cointegration testing
- May result in inconsistent parameter estimation

Research Projects

Main goal:
to develop methodology to analyze fractionally cointegrated systems

Fractional cointegration - fully parametric approach

- Testing the existence of fractional cointegration, Łasak (JoE, 2010)
- Properties of parameter estimators in fractionally cointegrated system
- Testing the cointegration rank, Łasak and Velasco (2010)

Model and ML estimation

$$\boxed{\Delta X_t \; = \alpha\beta'(\Delta^{1-d} - \Delta)X_t + \varepsilon_t}$$

X_t - vector of $I(1)$, d - fractional cointegration degree,

α - adjustment coefficient matrix, β - cointegration vectors

ε_t - vector of *IID* Gaussian errors with covariance matrix Ω,

(Gaussianity only used to define likelihood, not used for the proofs)

Define $Z_{0t} = \Delta X_t, \; Z_{1t}(d) = (\Delta^{1-d} - \Delta)X_t$.

The log-likelihood is: $L(\alpha, \beta, \Omega, d) =$

$$= -\frac{1}{2} T \log|\Omega| - \frac{1}{2} \sum_{t=1}^{T} [Z_{0t} - \alpha\beta' Z_{1t}(d)]' \Omega^{-1} [Z_{0t} - \alpha\beta' Z_{1t}(d)].$$

Model and ML estimation

Define as well:

$$S_{ij}(d) = T^{-1}\sum_{t=1}^{T} Z_{it}(d)\, Z_{jt}(d)' \quad i,j = 0,1$$

For fixed β and d to estimate α and Ω, regress Z_{0t} on $\beta' Z_{1t}(d)$ and

$$\begin{aligned}\hat{\alpha}(\beta(d)) &= S_{01}(d)\beta(\beta' S_{11}(d)\beta)^{-1},\\ \hat{\Omega}(\beta(d)) &= S_{00} - \hat{\alpha}(\beta)(\beta' S_{11}(d)\beta)\hat{\alpha}(\beta)'.\end{aligned}$$

Plug in estimates into the likelihood and solve the problem:

$$\boxed{\left|\lambda(d)S_{11}(d) - S_{10}(d)S_{00}^{-1}S_{01}(d)\right| = 0.}$$

Cointegrating space is spanned by eigenvectors corresponding to r largest eigenvalues $\lambda(d)$. With this choice of β we can estimate d by

$$\boxed{\tilde{d} = \arg\max_{d\in\mathcal{D}} L_{\max}(d), \qquad L_{\max}^{-2/T}(d) = |S_{00}| \prod_{i=1}^{r}\left(1 - \hat{\lambda}_i(d)\right)}$$

Consistency

We demonstrate for $d_0 \in D \subset (0.5, 1]$, where D is a closed set,

- The estimators $\tilde{d}$, $\tilde{\beta} = \hat{\beta}\left(\bar{\beta}'\hat{\beta}\right)^{-1}$, $\tilde{\alpha} = \hat{\alpha}\hat{\beta}'\bar{\beta}$, $\hat{\Omega}$ are consistent.

- For any fixed d, $d \neq d_0$, $d > 0.5$ the estimator $\tilde{\beta} = \hat{\beta}\left(\bar{\beta}'\hat{\beta}\right)^{-1}$

  remains consistent with a rate $\tilde{\beta} - \beta = o_P(T^{\frac{1}{2}-d})$, but $\tilde{\alpha}$ and $\hat{\Omega}$ not

$=>$ bias and large mean square errors of the estimator of the impact matrix $\Pi = \alpha\beta'$ found by Andersson and Gredenhoff (1998) came from the estimation of α rather than β.

Asymptotic distribution of beta

- The asymptotic distribution of $\tilde{\beta}$ is mixed Gaussian
- Similar to Johansen (1995) for $d_0 = 1$ fixed
- Equal to the distribution of GLS in Robinson and Hualde (2003)
- The convergence rate of $\tilde{\beta}$ is optimal, hence $\tilde{\beta} - \beta \in O_P\left(T^{-d_0}\right)$
- $\tilde{\beta}$ is asymptotically independent of other estimates
- Remains mixed normal, so we can test for the values of cointegration vector using Wald test that will be χ^2 distributed

Asymptotic distribution of alpha and d

- The asymptotic distribution of $\tilde{\alpha}$ is root-T consistent
- It is related with the asymptotic distribution of $\tilde{d}$
- The cointegration degree estimator $\tilde{d}$ is root-T consistent and has asymptotic normal distribution for $d_0 \neq 1$
- The joint asymptotic distribution of $\tilde{\alpha}$ and $\tilde{d}$ is given by

$$\boxed{\begin{bmatrix} T^{\frac{1}{2}}(\tilde{d} - d_0) \\ T^{\frac{1}{2}} vec(\tilde{\alpha} - \alpha) \end{bmatrix} \rightarrow_d N(0, \Psi), \; d_0 \in IntD \subset (0.5, 1]}$$

Monte Carlo (model)

Engle and Granger (1987), Banerjee et al. (1993):

$$\boxed{\begin{array}{c} x_t + 2y_t = u_t \\ x_t + y_t = e_t \end{array}}$$

where $\Delta^{1-d_0} u_t = \varepsilon_{1t}$, $\Delta e_t = \varepsilon_{2t}$, ε_{1t} and ε_{2t} are *IID* Gaussian errors

d_0 - degree of fractional cointegration, $d_0 = 0.55$, 0.65, 0.75, 0.85, 0.95, 1

$T = 50, 100, 200, 500$ observations,

$\alpha = [1, -1]'$ and $\beta = [1, 2]'$, $10,000$ iterations

Ox 6.01 and MaxSQPF procedure

Bias and standard error

- Bias and standard error of $\tilde{d}$, $\tilde{\beta}$ and $\tilde{\alpha}$ are all decreasing with d_0 and with sample size T.
- For β we obtain very good estimates already for moderate values of d_0 in larger samples.
- We can estimate β much better than α even for small values of d_0 where $\tilde{\beta}$ has convergence rate close to $T^{\frac{1}{2}}$.
- Estimates $\tilde{\alpha}$ have in general smaller bias than estimates $\hat{\alpha}^J$ and bigger standard deviation. The significance of the difference in standard deviation seems to be decreasing with the value of d_0.

Wald test for beta

- Wald test has quite distorted size for all values of d_0 in small samples
- However size is getting closer to its nominal value when sample size T increases (the bigger d_0 is, the faster)
- Size distortions of standard Wald test are bigger for small values of d_0 and in fact they seem to diverge, while for values of d_0 relatively close to 1 the standard test has less distorted size.
- Power properties of both versions of the Wald test are comparable and very good.

t-test for d

- t-tests also have quite distorted size in small samples.
- Size distortions seem to be decreasing with the value of d_0 and the sample size T.
- One-sided test (against the alternative $d < d_0$) is performing better than other two.
- It seems that the size distortions of t-tests are caused by bias of $\tilde{d}$.

Empirical Example

- Modelling interest rates with different maturities as Iacone (2009)
- Such a model is necessary both to measure the effects of monetary policy and to price financial assets
- Also an important tool for policy evaluation since the Federal Reserve operates by supplying liquidity on the Federal Funds market by open market operations and discount window loans, being directly present in just one market with contracts with very short maturity. It is therefore necessary to model the conduction of the monetary policy impulses to the rates of contracts with longer maturities.
- Modelling the interactions across rates is also important for the economic agents who would like to forecast the effects of future monetary policy decisions on the price of financial assets.

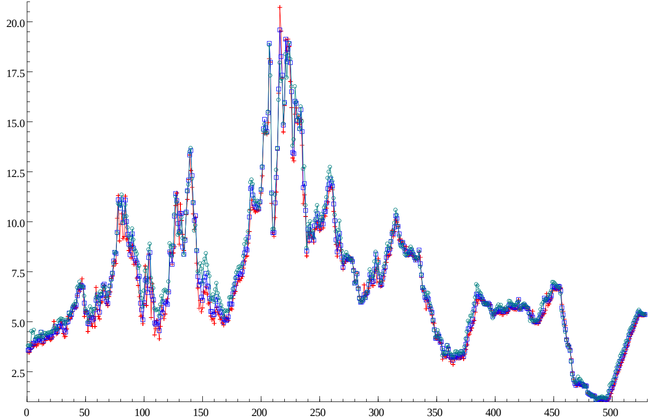
Expectations Hypothesis

- A theoretical model for the term structure of interest rates was discussed by Fisher (1896).
- Given market efficiency and rational expectations, the interest rates of contracts which only differ in maturity, should be linked by a no-arbirtage relation. Therefore, the return from investing in a contract with maturity over multiple periods should be equivalent to the expected return from investing in multiple consecutive contracts, provided that these span jointly the same time.
- If the Fisher equation holds, central banks may also find the information in the term structure of interest rates valuable because long term rates include the market's expectations of future inflation.
- Recent applications of the Expectation Hypothesis approach: Ang and Piazzesi (2003), Favero (2006).

Data description

- Iacone (2009): semiparametric analysis of the US$ interest rates with maturities of 1, 3 and 6 months.
- The offer rate LIBOR has been used, over the period 01/1963-04/2006 of the London interbank deposit.
- LIBOR is not affected by any regulation imposed by the central bank, and thus it is a typical messure of the cost of funds in US$.
- He has found the evidence that 3 considered series share the same order of integration, so it is possible to perform a similar analysis in a VAR framework. The order of the integration has been estimated to be 0.88.
- The 3 series has been found to be cointegrated with cointegration rank $=2$.

Interest rates



My results

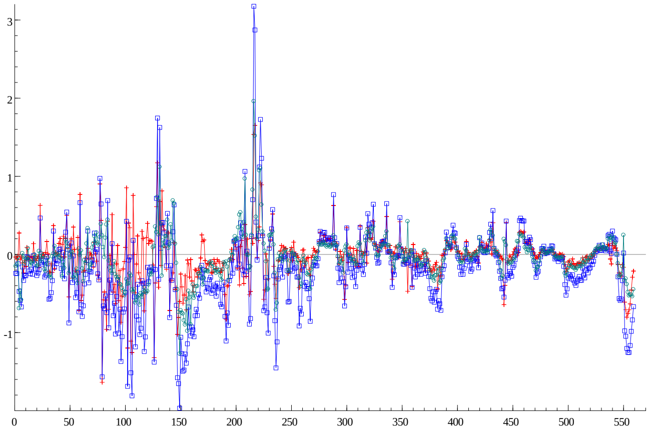
We estimate the model

$$\Delta^{\hat{\delta}} X_t = \alpha\beta'(\Delta^{\hat{\delta}-d} - \Delta^{\hat{\delta}})X_t + \varepsilon_t$$

with $\hat{\delta} = 0.88$ and get the results

model	1,3	1,6	3,6
$\hat{d}$	0.68	0.59	0.88
$\hat{d}_s$	0.32	0.41	-
$\hat{\beta}$	1 -0.98	1 -0.98	1 -0.98
$\hat{\alpha}$	-0.76 0.19	-0.47 0.16	-0.20 -0.04
trace test	194.33	90.42	29.75
λ_{max} test	193.52	86.89	31.76

Cointegration residuals



Comments

- So we find cointegration between each couple of interest rates. The cointegration vector is very close to [1,-1] in each case and the cointegration residuals are always asymptotically stationary.
- Given the high order of integratin in the interest rates, the evidence of cointegration is an important result because it means that transmission of impulses along the term structure is still fast enough to give the central bank to conduct an active monetary policy.
- Transmission is slower the longer the distance (in maturity) from the market where FED is present.
- Given the requirement that the series should be mean reverting, the EH also implies that the spreads are not integrated processes. We find that the hypothesis of cointegration is not rejected by the data, but the spreads are more persistent than they should be.

Main Conclusions

- We generalize the analysis of cointegrated systems to the fractional case
- All estimators of the parameters in FVECM can be estimated consistently
- We show that the asymptotic distribution of $\tilde{\beta}$ remains mixed normal, hence we can test for the values of cointegration vector using Wald test
- We show normal asymptotic distribution of other estimators, so it is possible to adapt standard inference rules

The end