

MACRO AND FINANCIAL MARKETS:

The memory of an elephant?

Karim Abadir and Gabriel Talmain

Background, REStud 2002:

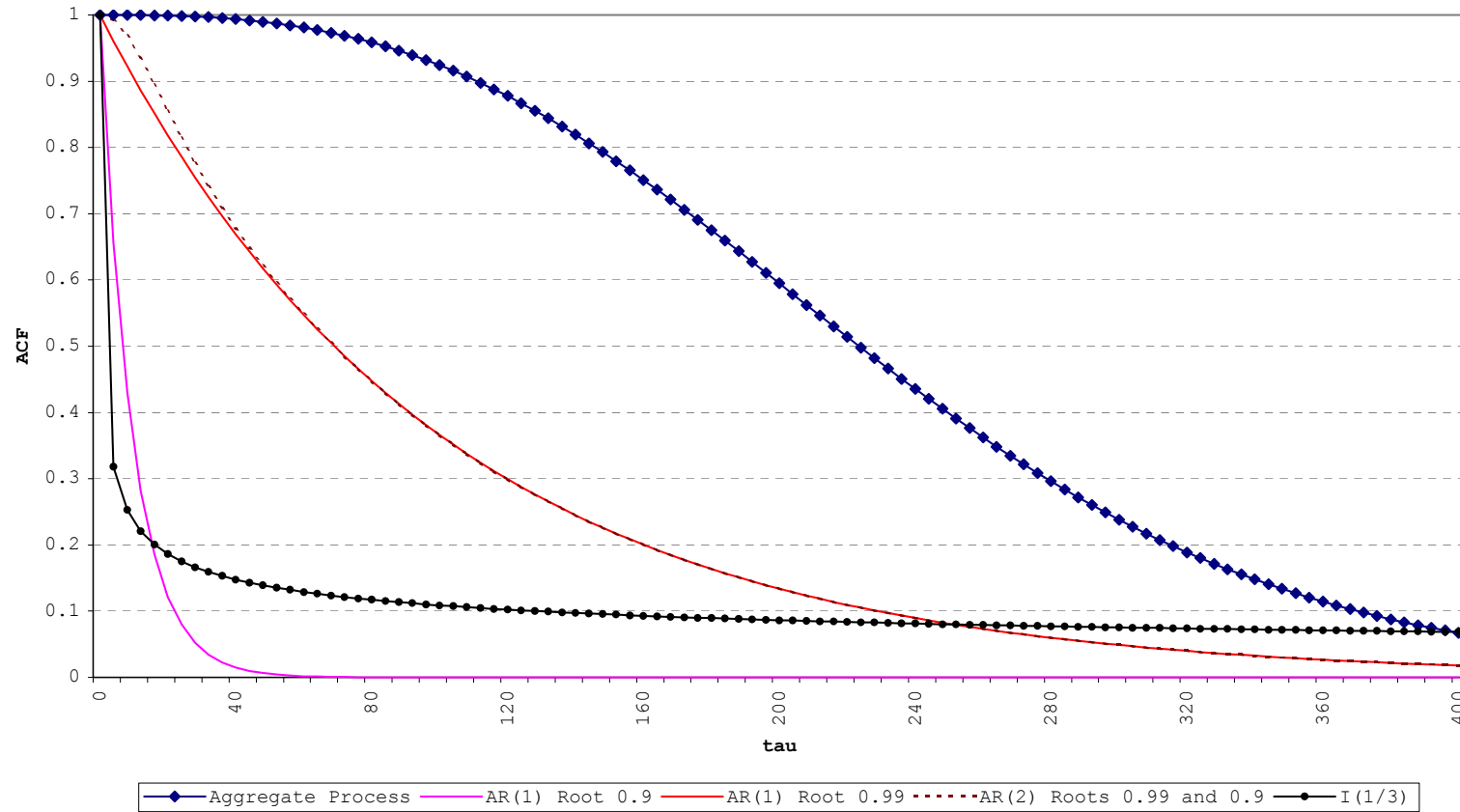
- Model and solution:
 - Micro-founded macro model.
 - Standard RBC + heterogeneity: drop “representative firm” assumption.
 - General Equilibrium yields explicit dynamic equation for GDP etc.

Background, REStud 2002:

- Model and solution:
 - Micro-founded macro model.
 - Standard RBC + heterogeneity: drop “representative firm” assumption.
 - General Equilibrium yields explicit dynamic equation for GDP etc.
- When \exists heterogeneity, aggregation \implies long-memory; e.g. Robinson (1978), Granger (1980), and TS lit.
- But in economics, \exists an inherent nonlinearity. Decompose GDP as

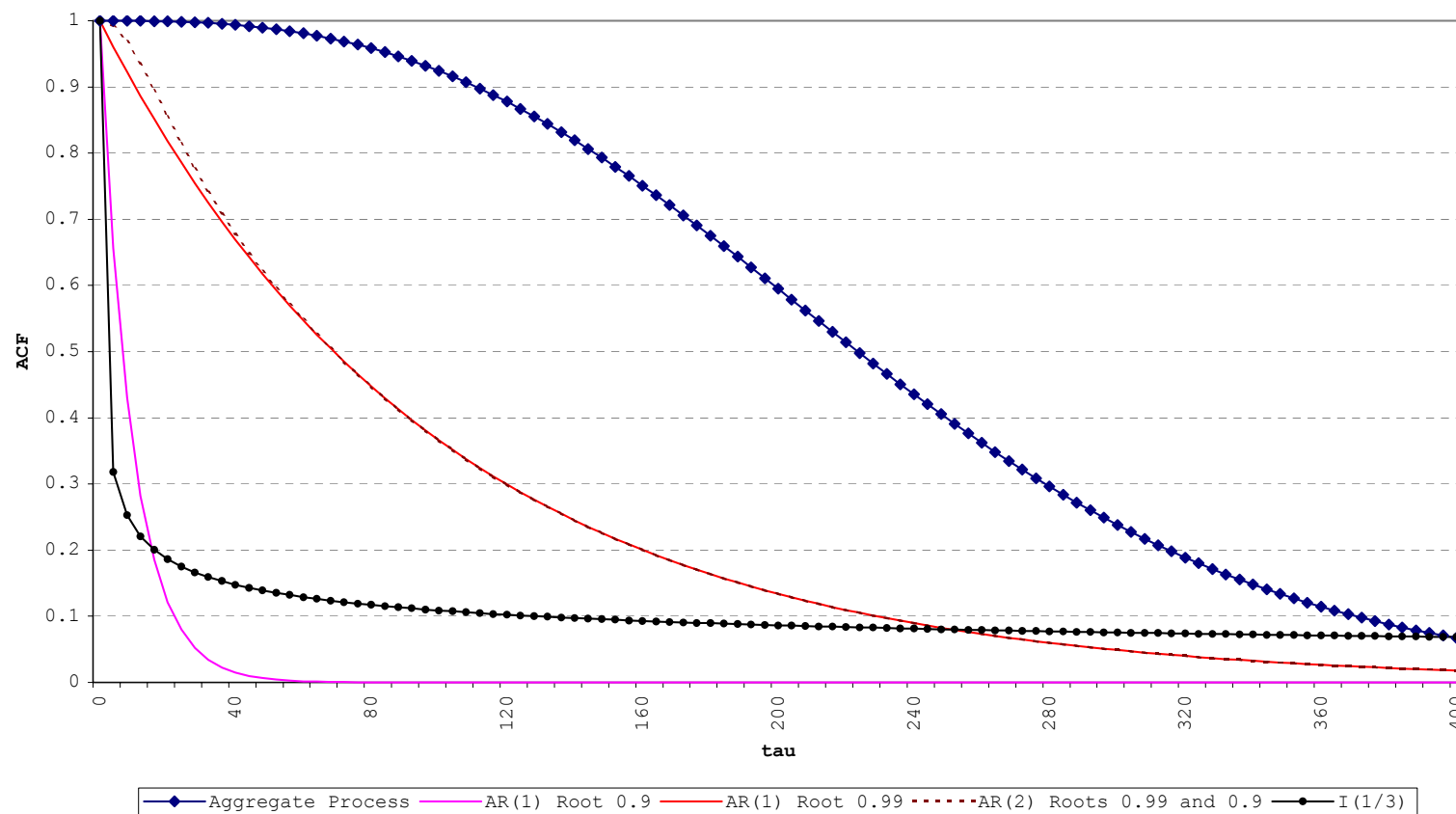
$$Y \equiv Y_1 + Y_2 + \dots = e^{\log Y_1} + e^{\log Y_2} + \dots \neq e^{\log Y_1 + \log Y_2 + \dots}$$

- The result is a new auto-correlation function (ACF) ρ_τ :



with different macropolicy implications.

- The result is a new auto-correlation function (ACF) ρ_τ :



with different macropolicy implications.

- But if \nexists integration, what about the modification of co-integration?

1 UIP and forward premium puzzle

- Fisher (1930): “speculation” equates expected returns after conversion to the same currency (UIP); e.g.
 - £1 invested in domestic (UK) bond yields $\pounds(1 + I_t)$ at maturity;
 - vs.
 - £1 invested in foreign (US) bond converted into $\$1/S_t$, which yield $\$(1 + I_t^*) / S_t = \pounds(1 + I_t^*) S_{t+1}/S_t$ at maturity.
- Define $i := \log(1 + I)$, $s := \log S$, and

$$r_{t+1} := \Delta s_{t+1} + i_t^* - i_t,$$

as the excess return from investing in the foreign asset. Then, $\mathbf{E}_t[r_{t+1}] = 0$ and r_{t+1} should not be predictable.

- Typical empirical implementation: regress r_{t+1} on the forward premium $(f_t - s_t)$, where f and s are the forward and spot rates in logs:

$$r_{t+1} = \alpha + \beta (f_t - s_t) + u_{t+1},$$

where α is the average risk premium and β is the informational content of the forward premium.

- UIP hypothesis implies $H_0 : \alpha = 0$ and $\beta = 0$.
- If one believes $f_t = E_t [s_{t+1}]$, then $(f_t - s_t) < 0$ indicates that the US\$ should depreciate.
- Routine finding: $\beta < 0$. As more US\$ depreciation is *expected*, higher returns are *actually* made on the US\$!

Investors are ready to pay more for an asset which, according to their expectations, should have become less attractive!

- See Cumby and Obstfeld (1981, J Fin) or Engel (1996, J Emp Fin).

- Running this regression with our data (3-month rates on US\$ and UK£ deposits in London and 3-month forward rate)

$$\begin{array}{rcl} \hat{r}_{t+1} & = & -0.0157 \quad -3.26 \quad (f_t - s_t) \\ \text{t-ratio} & & (-3.67) \quad (-6.14) \\ \text{HAC t} & & [-2.85] \quad [-2.90] \end{array}$$

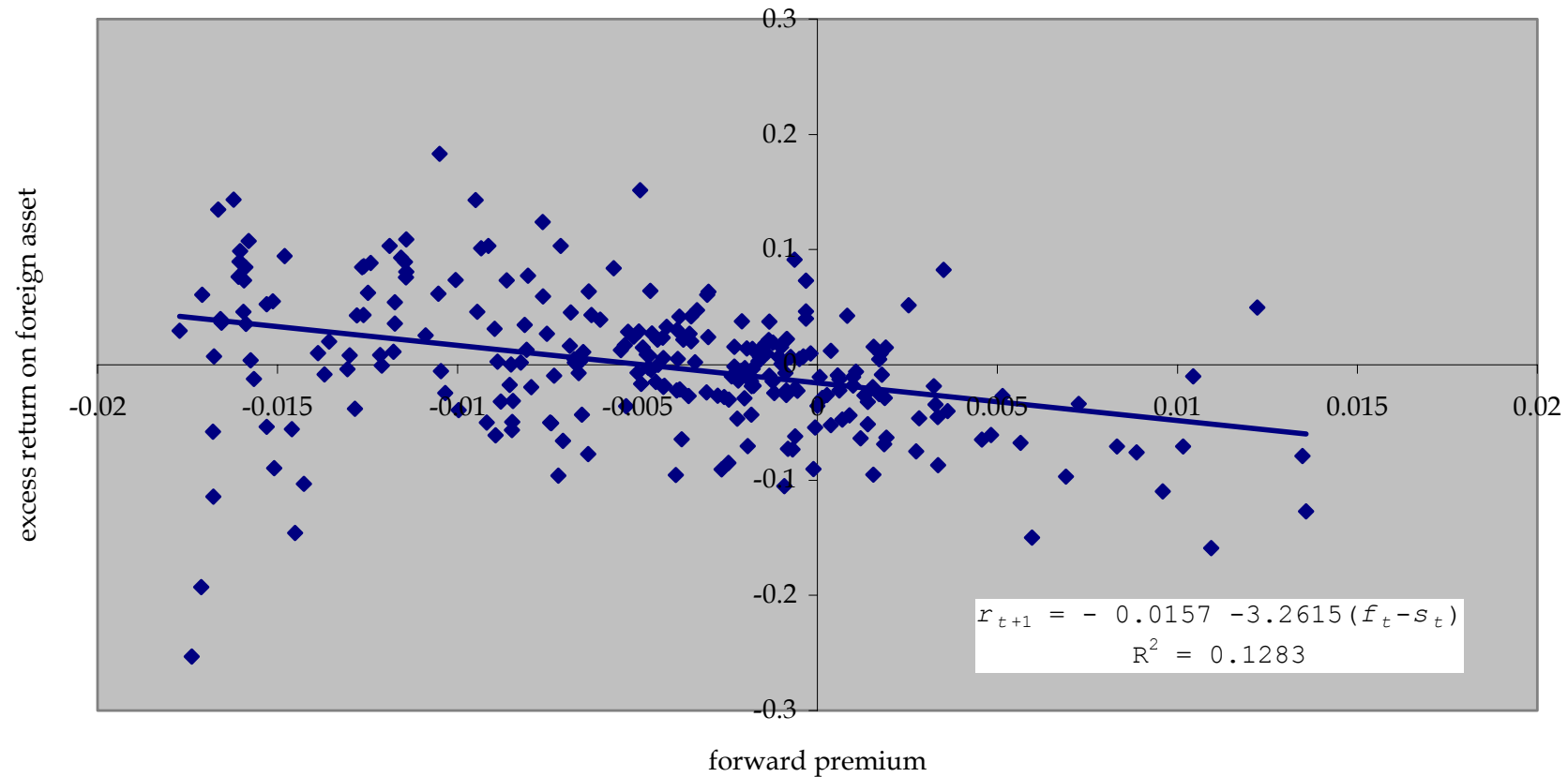
where HAC = heteroskedastic and autocorrelation-consistent,

Durbin-Watson statistic : 0.65

ARCH(7) test, $F(7, 242)$: 24.10 {0.0%}

RESET (omitted nonlinearities), $F(1, 255)$: 9.66 {0.2%}

- Scatter plot of data



- For years, it has defied economic logic to find such a result: how could investors' expectations be so systematically wrong? Or are they?!

2 Solution to the puzzle: coping with non-linear long-memory

- Co-movements vs. own dynamics.
- Incomplete modelling of dynamics can make the estimation of co-movements biased and inconsistent.
- Example (just an illustration): consider autoregressive process

$$y_t = \alpha y_{t-1} + \beta x_t + u_t,$$

with $u_t = \rho u_{t-1} + \varepsilon_t$ and $\varepsilon_t \sim \text{IID} (0, \sigma^2)$.

Running OLS on the first equation only \implies biased and inconsistent estimators; e.g. Maddala and Rao (1973, Ecta).

- Usual approaches:
 - GLS on the augmented first equation; or

– error correction mechanism (ECM), autoregressive distributed lag (ADL)

$$y_t = (\alpha + \rho) y_{t-1} - \alpha \rho y_{t-2} + \beta x_t - \beta \rho x_{t-1} + \varepsilon_t,$$

estimating

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + a_3 x_t + a_4 x_{t-1} + e_t$$

then testing for the restriction

$$a_1 = \frac{a_2 a_3}{a_4} - \frac{a_4}{a_3}.$$

– error correction mechanism (ECM), autoregressive distributed lag (ADL)

$$y_t = (\alpha + \rho) y_{t-1} - \alpha \rho y_{t-2} + \beta x_t - \beta \rho x_{t-1} + \varepsilon_t,$$

estimating

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + a_3 x_t + a_4 x_{t-1} + e_t$$

then testing for the restriction

$$a_1 = \frac{a_2 a_3}{a_4} - \frac{a_4}{a_3}.$$

- Long-memory times series models: very persistent time series (“Joseph effect”).
- Problem of ECM/ADL for data with long memory: too many lags.
- How about the GLS route?

- We would like to estimate the relation

$$z_t = \tilde{z}_t + u_t, \quad t = 1, 2, \dots, T,$$

where the $T \times 1$ vector $\tilde{\mathbf{z}} = \mathbf{X}\boldsymbol{\beta}$ is the “fundamental” value of \mathbf{z} , but \mathbf{z} and $\tilde{\mathbf{z}}$ (hence possibly \mathbf{u}) have long memory which needs to be accounted for. A possible 2-step procedure (à la GLS):

- decompose the autocorrelation matrix of \mathbf{z} as $\mathbf{R} = \mathbf{L}\mathbf{L}'$, where \mathbf{L} is lower-triangular and invertible;
- the transformed data $\mathbf{L}^{-1}\mathbf{z}$ and $\mathbf{L}^{-1}\tilde{\mathbf{z}}$ do not contain long memory and can be regressed by traditional methods.

- Unfortunately, estimating \mathbf{R} requires estimating $T - 1$ parameters: same as infeasible GLS!

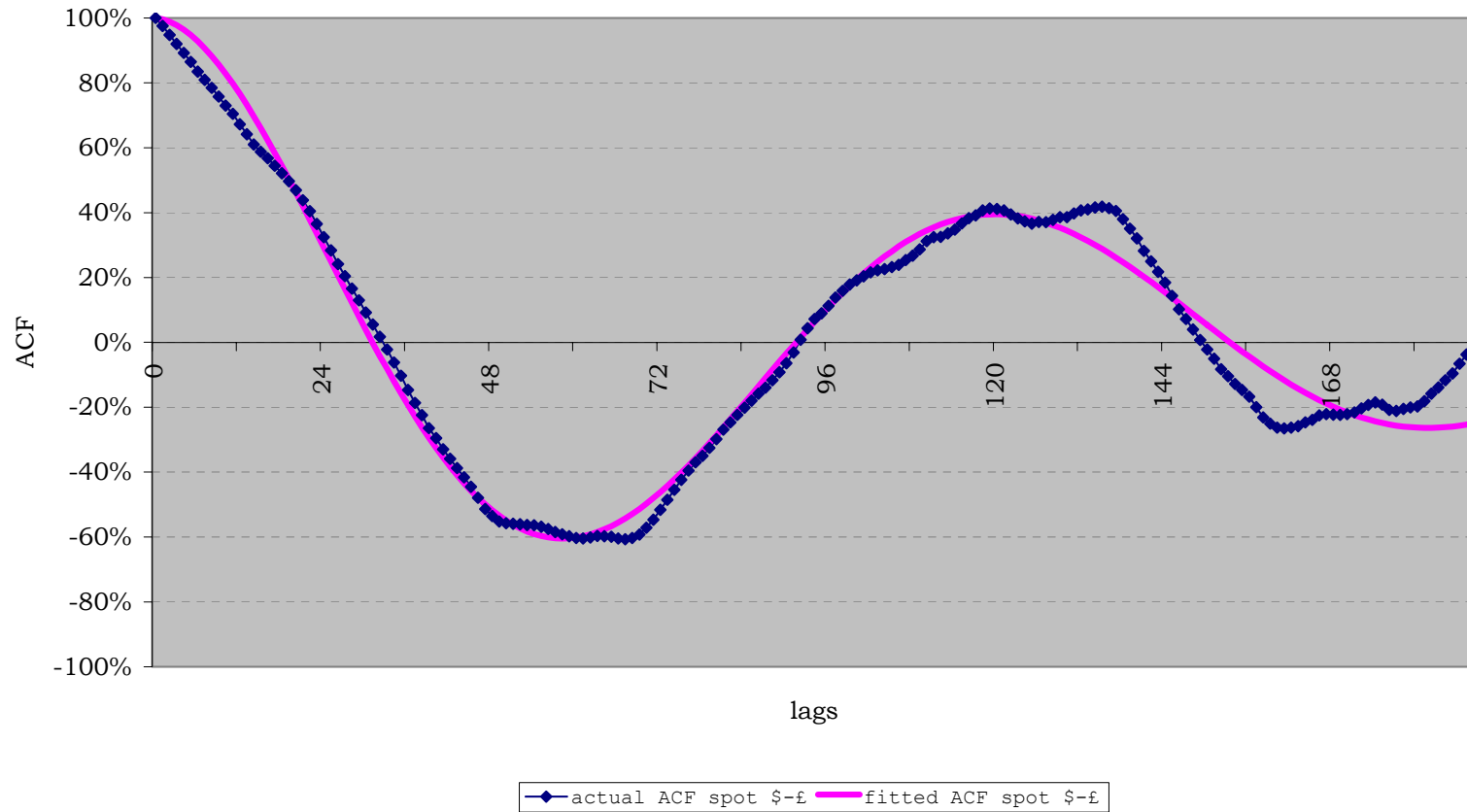
Solution: estimate the ACF of \mathbf{z} , using a variant of the functional form in Abadir and Talmain (2002, R E Stud)

$$\rho_{\tau} \approx \frac{1 - a [1 - \cos(\omega\tau)]}{1 + b\tau^c},$$

with only 4 parameters to fit. (Note: denominator controls decay of memory.)

- UIP example: fit is excellent for s

Actual vs fitted ACF of the spot rate \$-£



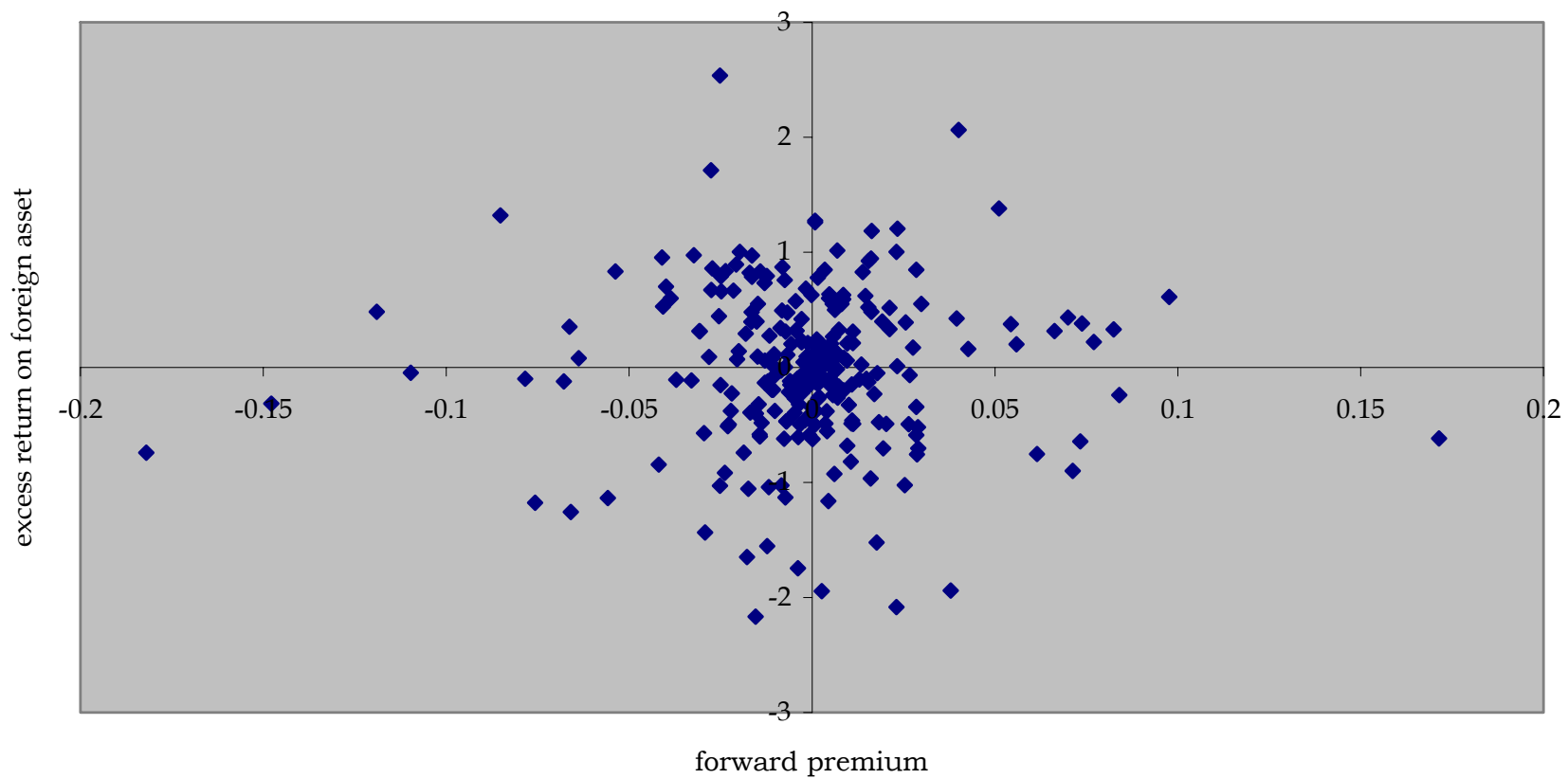
– The estimated $T \times T$ correlation matrix is then

$$\hat{\mathbf{R}} \equiv \begin{pmatrix} 1 & \hat{\rho}_1 & \hat{\rho}_2 & \cdots & \hat{\rho}_{T-2} & \hat{\rho}_{T-1} \\ \hat{\rho}_1 & 1 & \hat{\rho}_1 & \cdots & \cdots & \hat{\rho}_{T-2} \\ \hat{\rho}_2 & \hat{\rho}_1 & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \hat{\rho}_1 & \hat{\rho}_2 \\ \hat{\rho}_{T-2} & \cdots & \cdots & \hat{\rho}_1 & 1 & \hat{\rho}_1 \\ \hat{\rho}_{T-1} & \hat{\rho}_{T-2} & \cdots & \hat{\rho}_2 & \hat{\rho}_1 & 1 \end{pmatrix}.$$

– Find the Cholesky decomposition (matlab) $\hat{\mathbf{R}} = \hat{\mathbf{L}}\hat{\mathbf{L}}'$; and

– calculate $\mathbf{s}^{\text{acf}} = \hat{\mathbf{L}}^{-1}\mathbf{s}$ and $\mathbf{f}^{\text{acf}} = \hat{\mathbf{L}}^{-1}\mathbf{f}$.

– The scatter plot of the transformed data is a nice spherical cloud



– and the regression with transformed data becomes

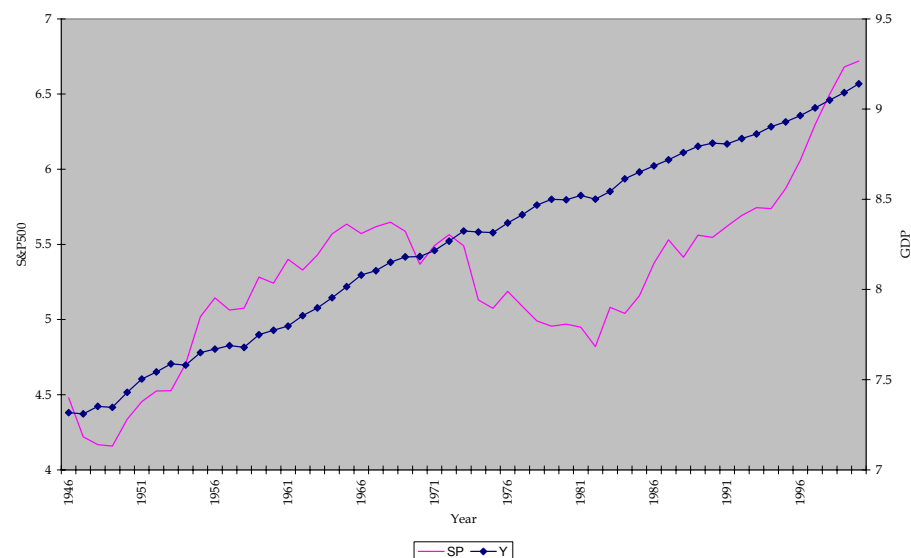
$$\begin{array}{rcl} \hat{r}_{t+1}^{\text{acf}} & = & 0.00582 \quad +0.0604 \left(f_t^{\text{acf}} - s_t^{\text{acf}} \right) \\ \text{t-ratio} & & (0.14) \quad (0.05) \\ \text{HAC t} & & [0.19] \quad [0.04] \end{array}$$

Durbin-Watson statistic : 2.20

ARCH(7) test, $F(7, 242)$: 6.54 {0.0%}

RESET (omitted nonlinearities), $F(1, 255)$: 0.82 {36.6%} .

3 Stock market application



- Unit roots?! (NB: unit root = permanent memory and no mean-reversion to any regular pattern or trend!)
- Lack of unit roots shown for US and UK GDPs in Abadir and Talmain (2002, R E Stud).
- How about Stock indices?

- Jegadeesh and Titman (1993 and 2001, J Fin) find momentum;
- De Bondt and Thaler (1985 and 1987, J Fin) find long cycles;
- Cavaliere's (2001, Ect J) non-parametric test detects long-memory.

- Jegadeesh and Titman (1993 and 2001, J Fin) find momentum;
- De Bondt and Thaler (1985 and 1987, J Fin) find long cycles;
- Cavaliere's (2001, Ect J) non-parametric test detects long-memory.
- GE theory: long-run proportionality relationship between the aggregate real value of firms and GDP. On a balanced growth path:
 - real interest rate is a function of capital/output, which is constant
⇒ rate at which future aggregate profits are discounted is fixed;
 - but share of aggregate profits in GDP is constant;
 - hence the discounted stream of future profits (i.e. capitalized value of the stock market) is proportional to GDP.

- Jegadeesh and Titman (1993 and 2001, J Fin) find momentum;
- De Bondt and Thaler (1985 and 1987, J Fin) find long cycles;
- Cavaliere's (2001, Ect J) non-parametric test detects long-memory.
- GE theory: long-run proportionality relationship between the aggregate real value of firms and GDP. On a balanced growth path:
 - real interest rate is a function of capital/output, which is constant
⇒ rate at which future aggregate profits are discounted is fixed;
 - but share of aggregate profits in GDP is constant;
 - hence the discounted stream of future profits (i.e. capitalized value of the stock market) is proportional to GDP.
- Empirically-testable version: log of stock market index (S&P500) corrected for inflation, s_t , should be in a long-term proportionality with the log of real GDP, y_t .

- Error-correction model of s_t on y_t ,

$$\Delta s_t = \alpha + (\beta_1 \Delta s_{t-1} + \dots + \beta_m \Delta s_{t-m}) + (\gamma_0 \Delta y_t + \dots + \gamma_n \Delta y_{t-n}) - \delta (s_{t-1} - y_{t-1}) + \delta_1 y_{t-1} + \varepsilon_t.$$

- The term $-\delta (s_{t-1} - y_{t-1}) + \delta_1 y_{t-1}$ is the ECM.
- It represents the long-run ‘equilibrium’ relationship between s and y :

$$s_e = \left(1 + \frac{\delta_1}{\delta} \right) y_e, \quad \delta \neq 0.$$

- H_0 : $\delta_1 = 0$ for long-run proportionality between S_e and Y_e .
- Define $d_{t-1} := s_{t-1} - s_e$ as the deviation of s_{t-1} from its long-term value s_e .
- ECM: this deviation will pull s_t back towards its long-term equilibrium value by δd_{t-1} , where $\delta > 0$.

- For S&P500 over 1958-2000, we obtained the regression

$$\begin{aligned}
 \widehat{\Delta s_t} = & -0.728 + 0.539 \Delta s_{t-4} + 0.380 \Delta s_{t-6} \\
 & (-2.33) \quad (4.35) \qquad \qquad (2.83) \\
 & + 3.11 \Delta y_t - 1.72 \Delta y_{t-1} + 1.84 \Delta y_{t-2} - 1.15 \Delta y_{t-6} \\
 & (4.30) \qquad (-2.41) \qquad (2.58) \qquad (-1.77) \\
 & - 0.112 (s_{t-1} - y_{t-1}) + 0.0396 y_{t-1} \\
 & (-2.11) \qquad \qquad \qquad (0.92)
 \end{aligned}$$

where the t-ratios are in parentheses, and we have $R^2 = 57.5\%$,

$$\text{AR}(2) \text{ test, } F(2, 32) : 0.43 \{65.4\%\}$$

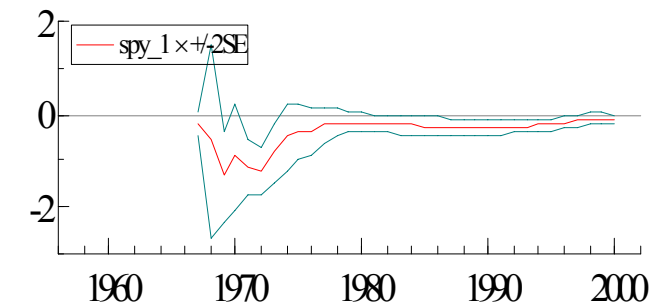
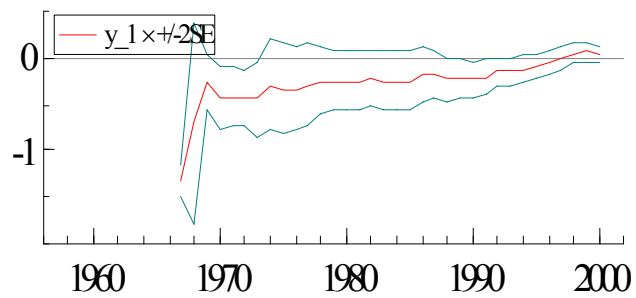
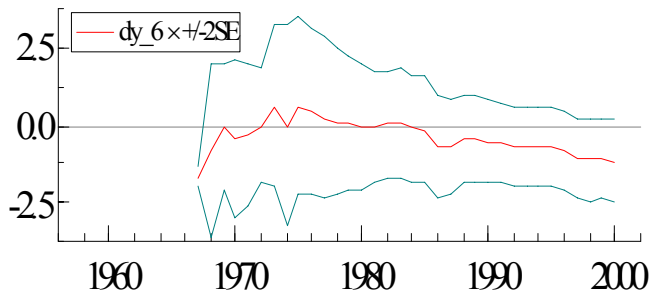
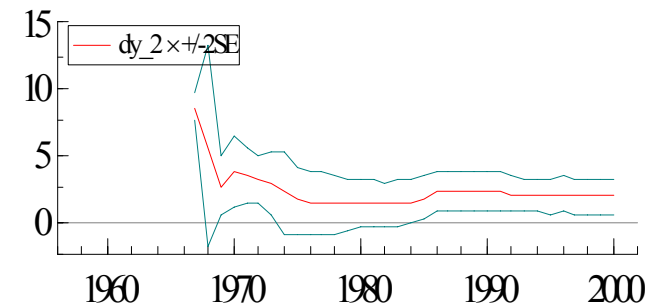
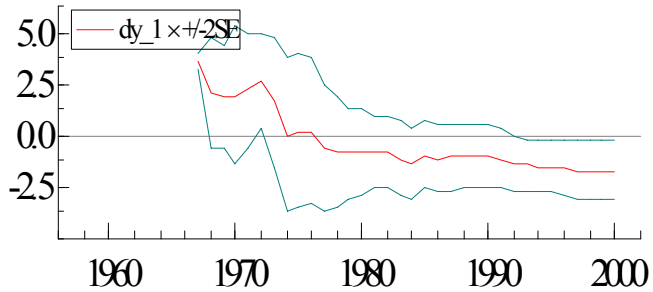
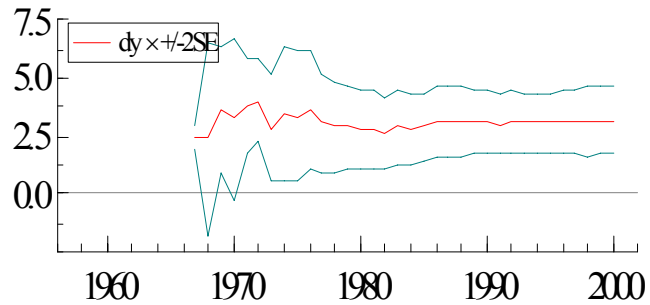
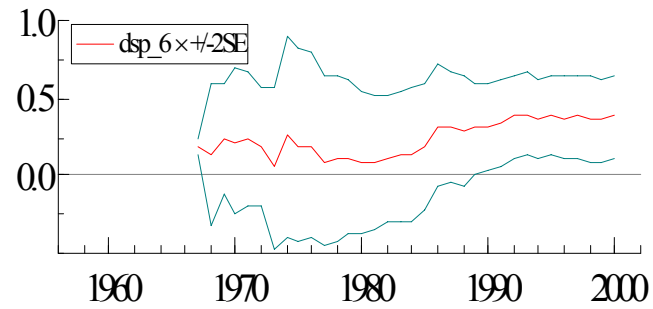
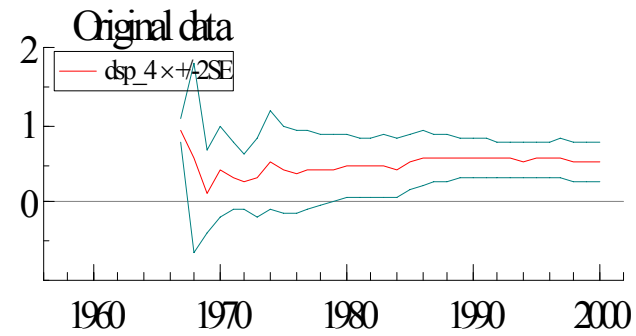
$$\text{ARCH}(1) \text{ test, } F(1, 32) : 0.17 \{68.1\%\}$$

$$\text{RESET, } F(1, 33) : 1.36 \{25.3\%\} .$$

- $H_0: \delta_1 = 0$ is supported, but:

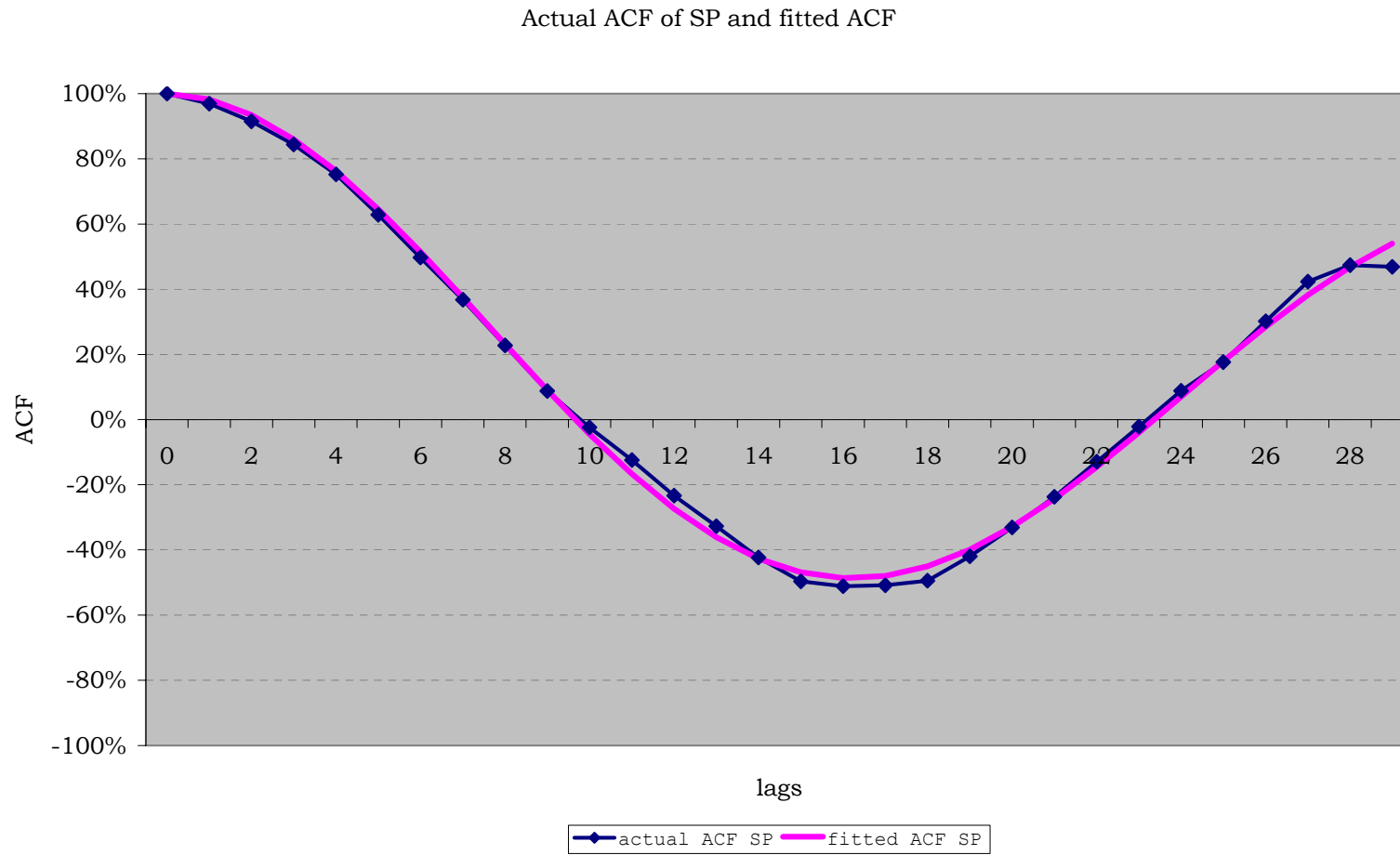
- by the end of the period, 1995-2000, the fit is poor;
- the coefficient of y_{t-1} is unstable and H_0 would be rejected on a sample ending in 1994;
- including more lags of Δs in the regression worsens rather than improves stability, while not improving the fit.

The recursive parameter estimates are...



- Remember memory?!

- Fit the ACF of s ;



– run the regression with transformed variables over 1960-2000

$$\widehat{\Delta s_t^{\text{acf}}} = -0.363 + 2.52 \Delta y_t^{\text{acf}} - 1.41 \left(s_{t-1}^{\text{acf}} - y_{t-1}^{\text{acf}} \right) - 0.301 y_{t-1}^{\text{acf}}$$

(−0.27) (3.67) (−10.0) (−0.49)

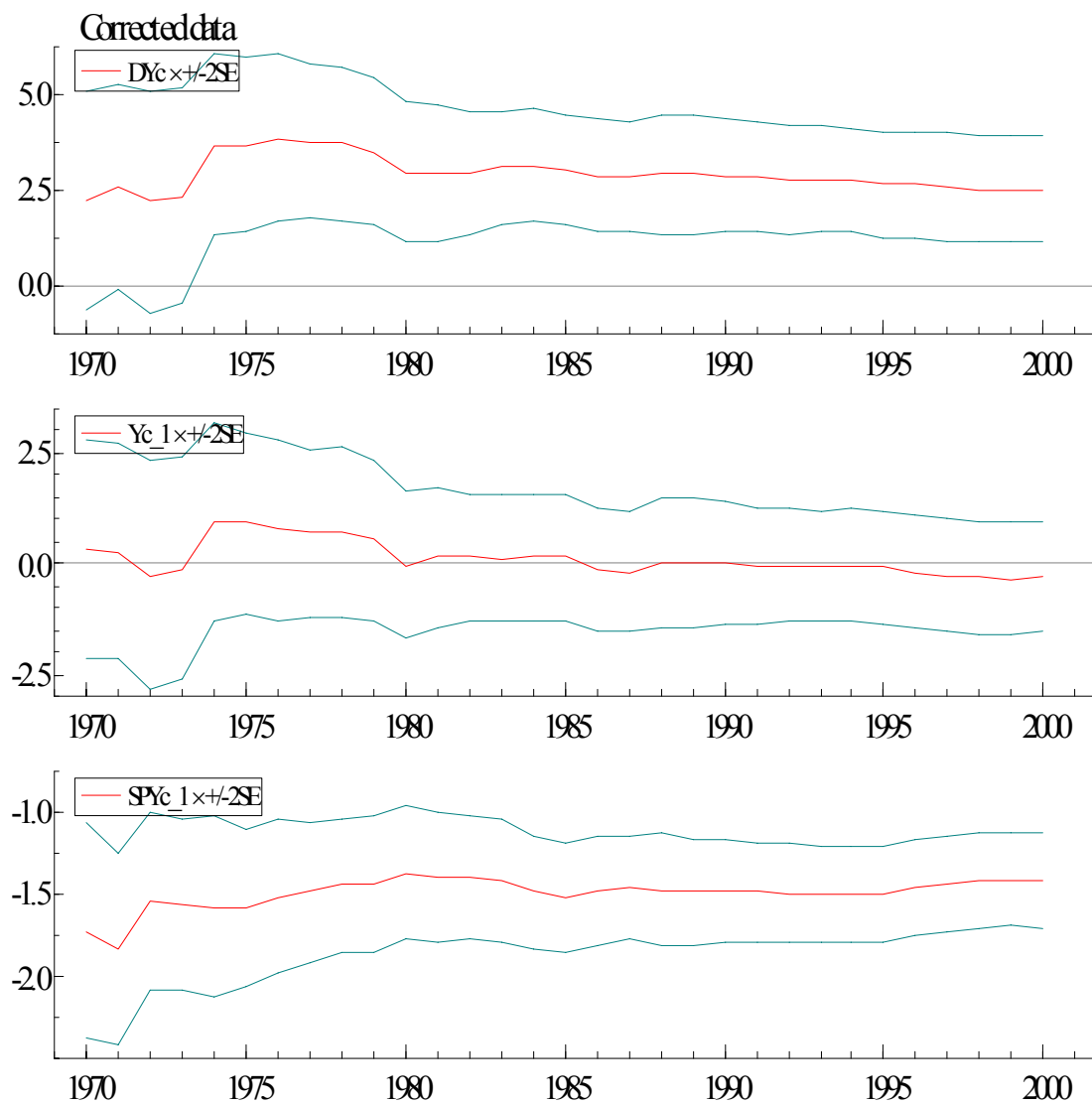
where $R^2 = 76.3\%$,

AR(2) test, $F(2, 35) : 2.47 \{9.9\%\}$

ARCH(1) test, $F(1, 35) : 1.72 \{19.9\%\}$

RESET, $F(1, 33) : 0.72 \{40.2\%\}$.

– $H_0: \delta_1 = 0$ is supported throughout the sample



- deviations from cycles around the long-run fundamental values are restored well within a year;
- good fit.

4 Extensions

- The two-step procedure is not the most efficient estimation method:
 - we provide formulae for full GLS, QMLE, etc.;
 - but qualitative results remain unchanged.

Example: QMLE case. For any given \mathbf{R} , define

$$\hat{\boldsymbol{\beta}}_{\mathbf{R}} \equiv \left(\mathbf{X}' \mathbf{R}^{-1} \mathbf{X} \right)^{-1} \mathbf{X}' \mathbf{R}^{-1} \mathbf{z}$$

as a function of \mathbf{R} . The QMLE of \mathbf{R} is obtained by maximizing

$$-\log \left| \left(\mathbf{z} - \mathbf{X} \hat{\boldsymbol{\beta}}_{\mathbf{R}} \right)' \mathbf{R}^{-1} \left(\mathbf{z} - \mathbf{X} \hat{\boldsymbol{\beta}}_{\mathbf{R}} \right) \mathbf{R} \right|$$

with respect to the parameters of the ACF: the optimization of the joint likelihood (for \mathbf{R} and $\boldsymbol{\beta}$) now depends on only 4 parameters that determine the whole autocorrelation matrix \mathbf{R} . Once the optimal value $\hat{\mathbf{R}}$ of \mathbf{R} is obtained, the MLE of $\boldsymbol{\beta}$ is $\hat{\boldsymbol{\beta}} \equiv \hat{\boldsymbol{\beta}}_{\hat{\mathbf{R}}}$.