Low-Frequency Econometrics

Ulrich K. Müller and Mark W. Watson
Princeton University

November 2015

Motivation

Some econometric questions abut time series concern low-frequency behavior

1. Is a time series mean reverting, so that there is a stable long-run mean?

2. What is the variability of the sample average of a mean reverting series?

   ($\Rightarrow$ confidence interval for population mean)

3. What kind of values do we expect many periods hence?

4. What is the degree of long-run covariability?

Extracting Low-Frequency Variability

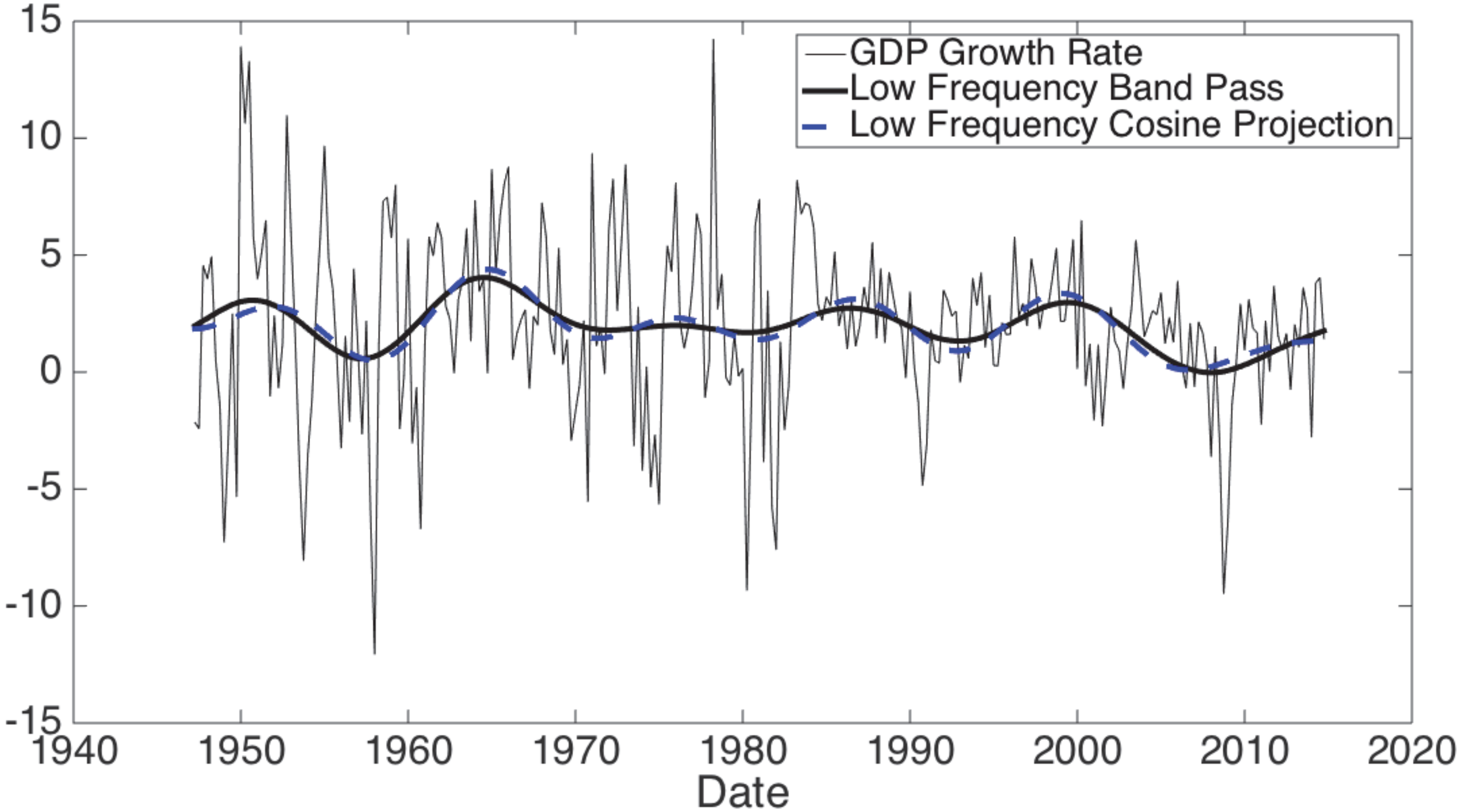
- Focus on, say, cycles of period greater than 10 years

- Could use band-pass filter

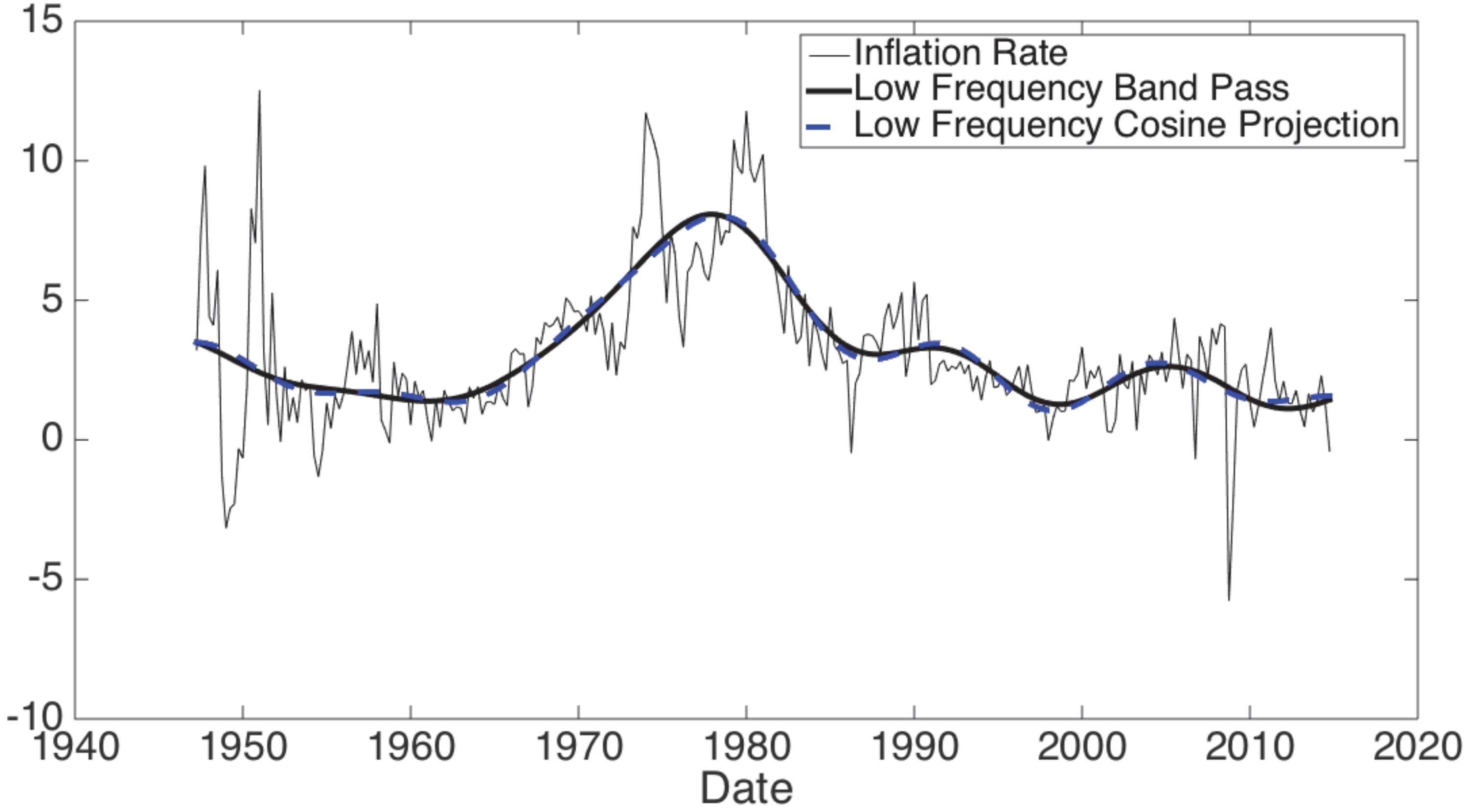
- Alternatively, project series on small number low-frequency trigonometric series

  $\Rightarrow$ Convenient for inference, as we shall see

Band-Pass Filter vs Low-Frequency Projection



Band-Pass Filter vs Low-Frequency Projection



Low-Frequency Transformation

- Let $\Psi_j(s) = \sqrt{2}\cos(j\pi s)$

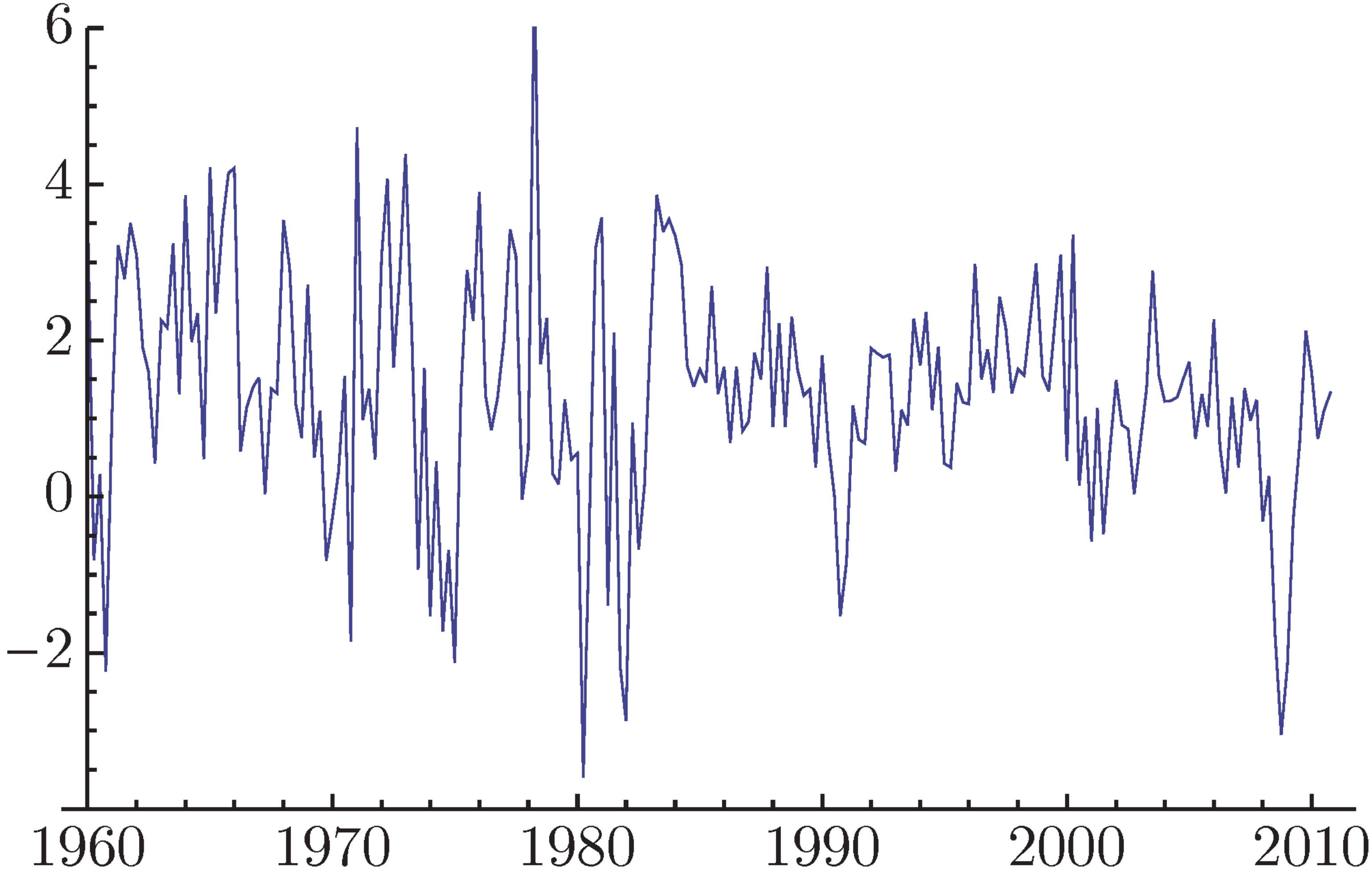
- For a scalar sequence $\{a_t\}_{t=1}^{T}$, define

$$\begin{aligned} A_j &= \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \Psi_j(t/T) a_t \\ &= \frac{\sqrt{2}}{\sqrt{T}} \sum_{t=1}^{T} \cos(\pi j t/T) a_t, \quad j = 1, \cdots, q \end{aligned}$$

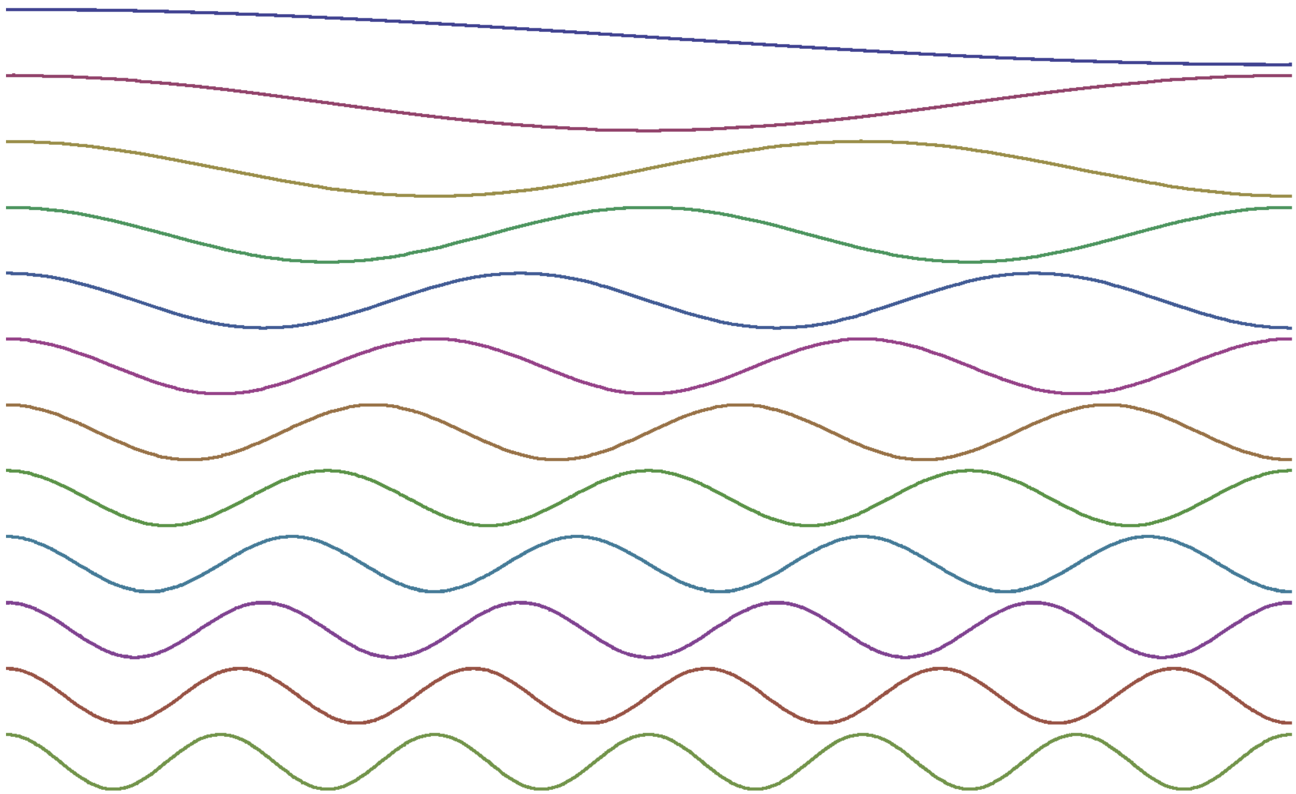
- Number q determines which frequencies are extracted

  $\Rightarrow q = 12$ extracts periods greater than 10 years from 60 years of data

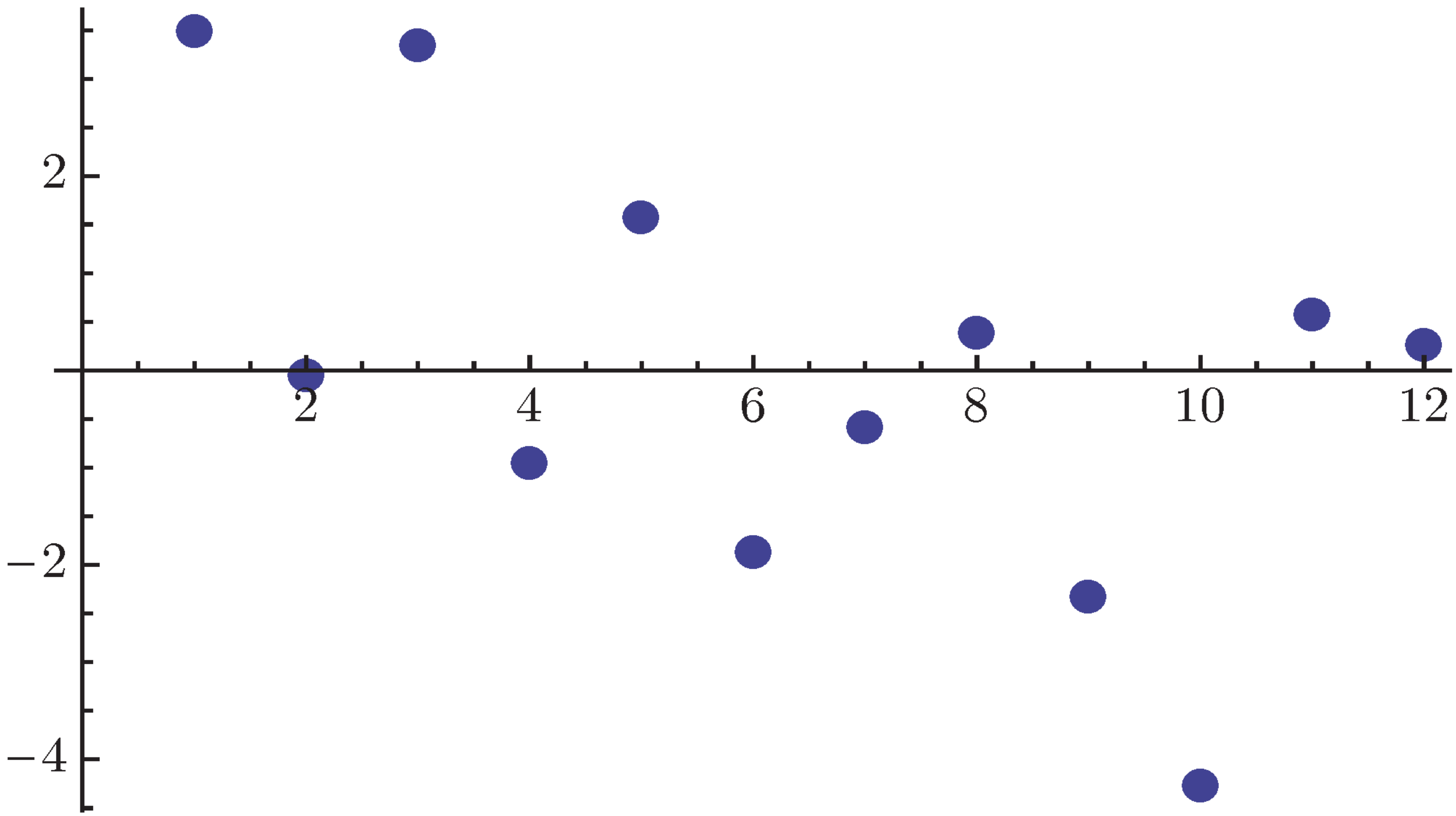
Example: U.S. Postwar GDP Growth



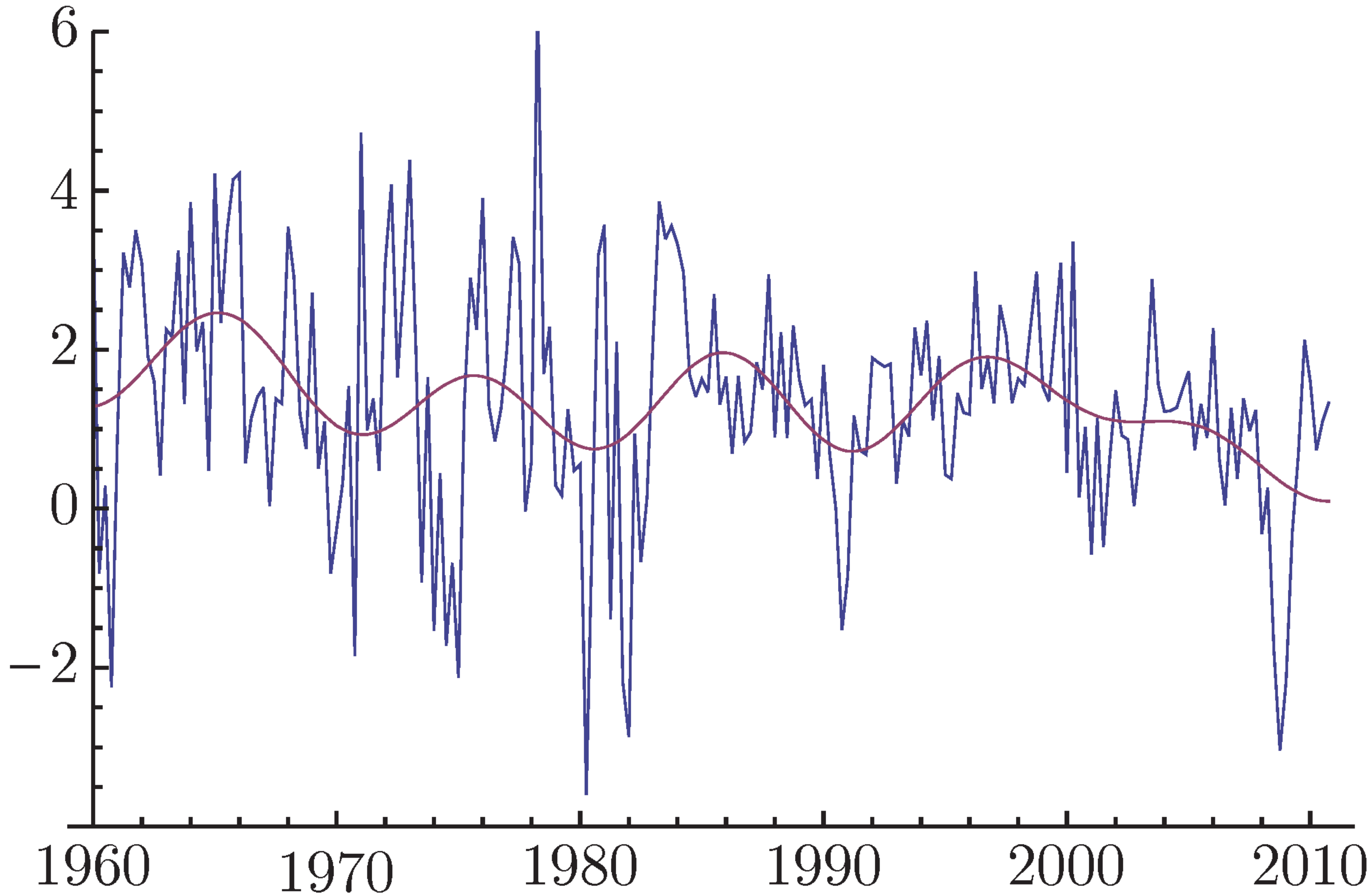
$q = 12$ Cosine Weights



LF Transforms of GDP Growth



Projection on $q = 12$ Cosines



Low Frequency Transformation

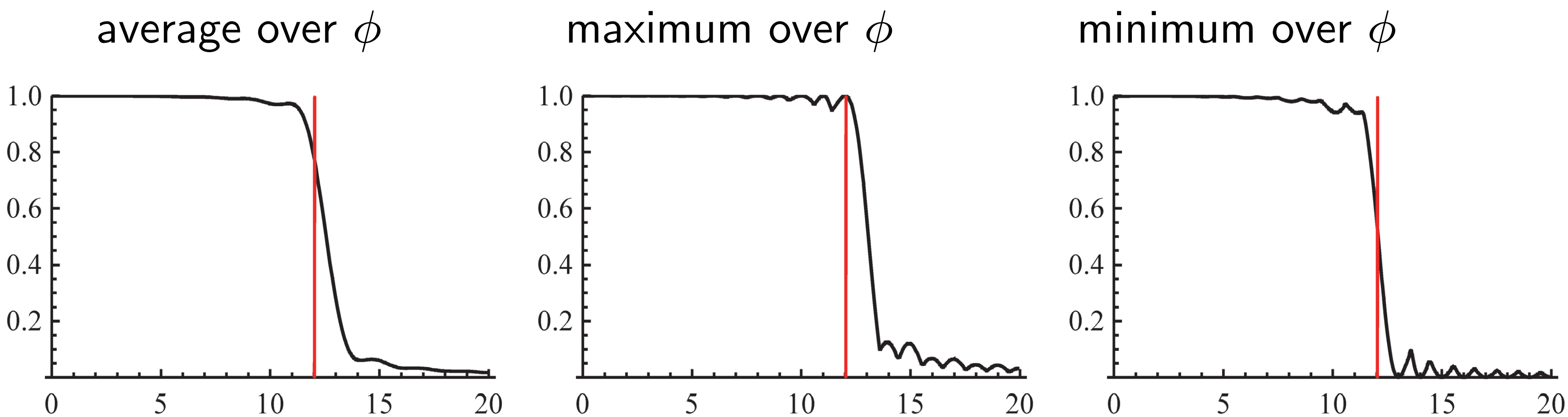
- Claim: The q numbers $A_1, \cdots, A_q$ summarize variability of $\{a_t\}_{t=1}^T$ for frequencies lower than $q\pi/T$.

- Consider R^2 from regression of generic periodic series

$$a_t = \sin(\pi r t/T + \phi)$$

on $\Psi_j(t/T) = \sqrt{2}\cos(j\pi t/T)$, $j = 1, \cdots, q$.

R^2 as Function of r for $q = 12$



- Ideally, $R^2 = 1$ for $r \leq 12$ and $R^2 = 0$ for $r > 12$ for all ϕ.

Low-Frequency Econometrics

- For data spanning 60 years, low-frequency sampling information (=periods greater than 10 years) summarized by $q = 12$ weighted averages of the original data.

- Central idea: Answer all questions about low-frequency properties with these q data points
  - Captures the notion that information about low-frequency behavior is scarce
  - Avoids modelling and potential misspecification of higher frequency aspects

Standard Asymptotics for Time Series

- Under a wide range of primitive conditions on the dependent and heterogeneous mean-zero process $\{u_t\}$, a Central Limit Theorem holds for all fractions of the sample, i.e. for all $0 \leq r_1 < r_2 \leq s_1 < s_2 \leq 1$,

$$\begin{pmatrix} \frac{1}{\sqrt{T}} \sum_{t=\lfloor r_1 T \rfloor+1}^{\lfloor r_2 T \rfloor} u_t \\ \frac{1}{\sqrt{T}} \sum_{t=\lfloor s_1 T \rfloor+1}^{\lfloor s_2 T \rfloor} u_t \end{pmatrix} \Rightarrow \mathcal{N}\left(0, \begin{pmatrix} \omega^2(r_2 - r_1) & 0 \\ 0 & \omega^2(s_2 - s_1) \end{pmatrix}\right)$$

- This (almost) implies the "Functional" Central Limit Theorem for nicely behaved I(0) processes

$$T^{-1/2} \sum_{t=1}^{\lfloor \cdot T \rfloor} u_t \Rightarrow \omega W(\cdot)$$

Implication for Low Frequency Transformations

- Suppose $y_t = \mu + u_t$, where $T^{-1/2} \sum_{t=1}^{\lfloor \cdot T \rfloor} u_t \Rightarrow \sigma W(\cdot)$.

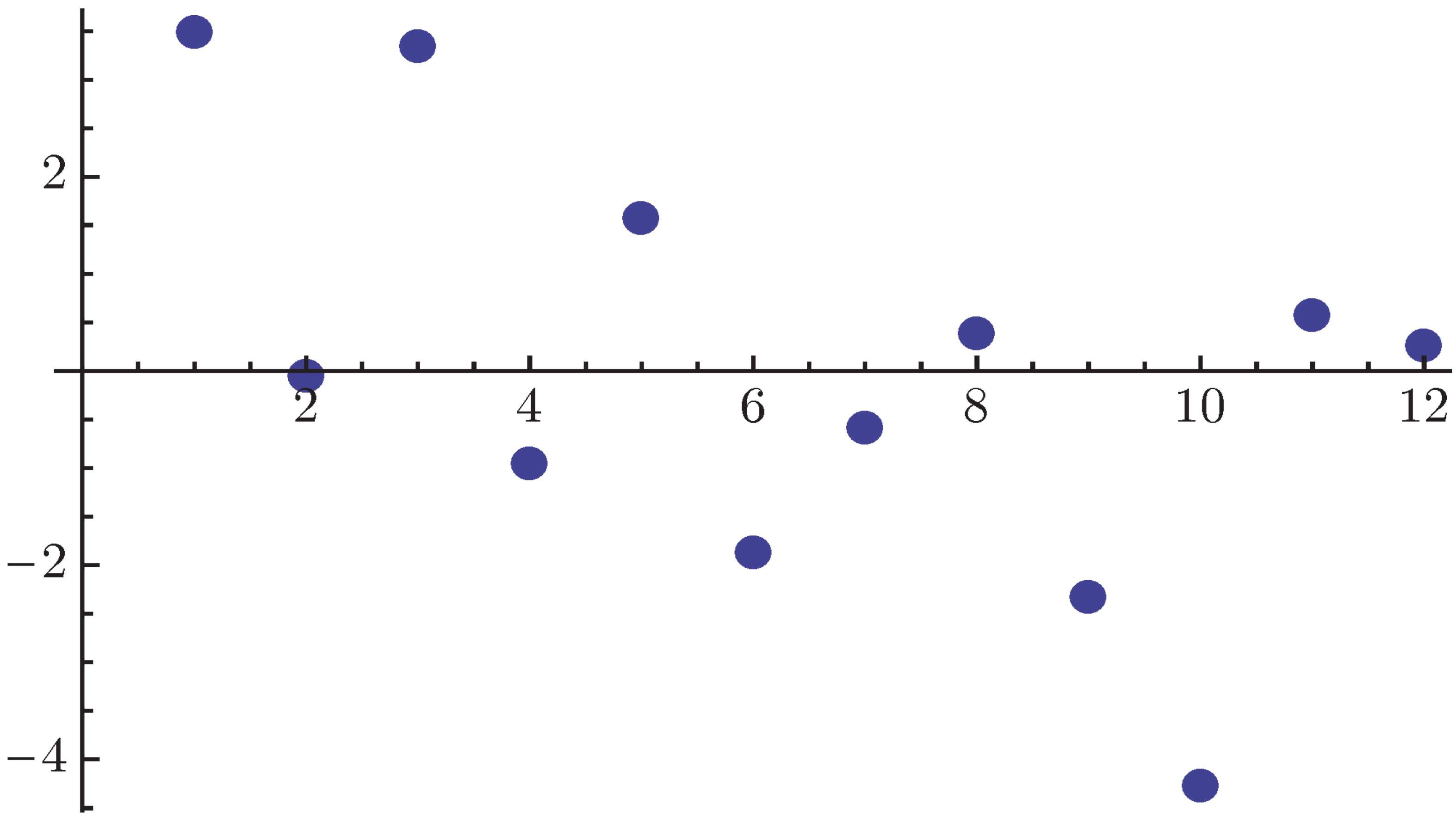
- Then weighted averages of u_t also become approximately Gaussian:

  With $Y_j = T^{-1/2} \sum_{t=1}^{T} \Psi_j(t/T) y_t$

$$\left\{Y_j\right\}_{j=1}^{q} \Rightarrow \left\{\omega \int_0^1 \Psi_j(s) dW(s)\right\}_{j=1}^{q} \sim iid \mathcal{N}(0, \omega^2)$$

  where Z_j $\sim$i.i.d.$\mathcal{N}(0,1)$, since $\int_0^1 \Psi_i(s)\Psi_j(s)ds = \mathbf{1}[i = j]$ (and $\int_0^1 \Psi_i(s)ds = 0$).

Low Frequency Transformation GDP Growth



Beyond the I(0) Model

- Standard forms of persistence modelling in time series (I(1), fractional, local-to-unity, etc.) have noise process u_t that satisfy

$$T^{-\alpha} \sum_{t=1}^{\lfloor \cdot T \rfloor} u_t \Rightarrow \omega G(\cdot)$$

for some α and Gaussian process G.

- Example: I(1) model $u_t = \sum_{s=1}^{t} \nu_s$, $\nu_t \sim$I(0) satisfies this with $\alpha = \frac{3}{2}$ and $G(s) = \int_0^s W(r)dr$.

- With $y_t = \mu + u_t$, this still implies that suitably scaled LF transforms $\left\{Y_j\right\}_{j=1}^{q}$ become asymptotically mean-zero Gaussian, with $q \times q$ covariance matrix $\mathbf{\Sigma}$ that depends on covariance kernel of G.

Applications

1. Confidence intervals for population mean

   (Kiefer and Vogelsang (2002, 2005), Müller (2004, 2007, 2014), Phillips (2006))

2. Testing I(0) property

   $\Rightarrow$ if applied to putative error correction term, test null hypothesis of cointegration

   (Bierens (1997), Phillips (1998), Wright (2000), Müller and Watson (2008, 2013))

3. Inference about degree of persistence

   (Müller and Watson (2008))

Applications, ctd.

4. Long-run forecasting

   (Müller and Watson (2014))

5. Multiple time series

   ⇒ Low-frequency covariability, low-frequency regression

   (New in survey paper: "Low-frequency Econometrics"

   Available under www.princeton.edu/∼umueller)

Application 1: Confidence Intervals

- We observe

$$y_t = \mu + u_t, \quad u_t \sim I(0)$$

and want to construct a CI for μ.

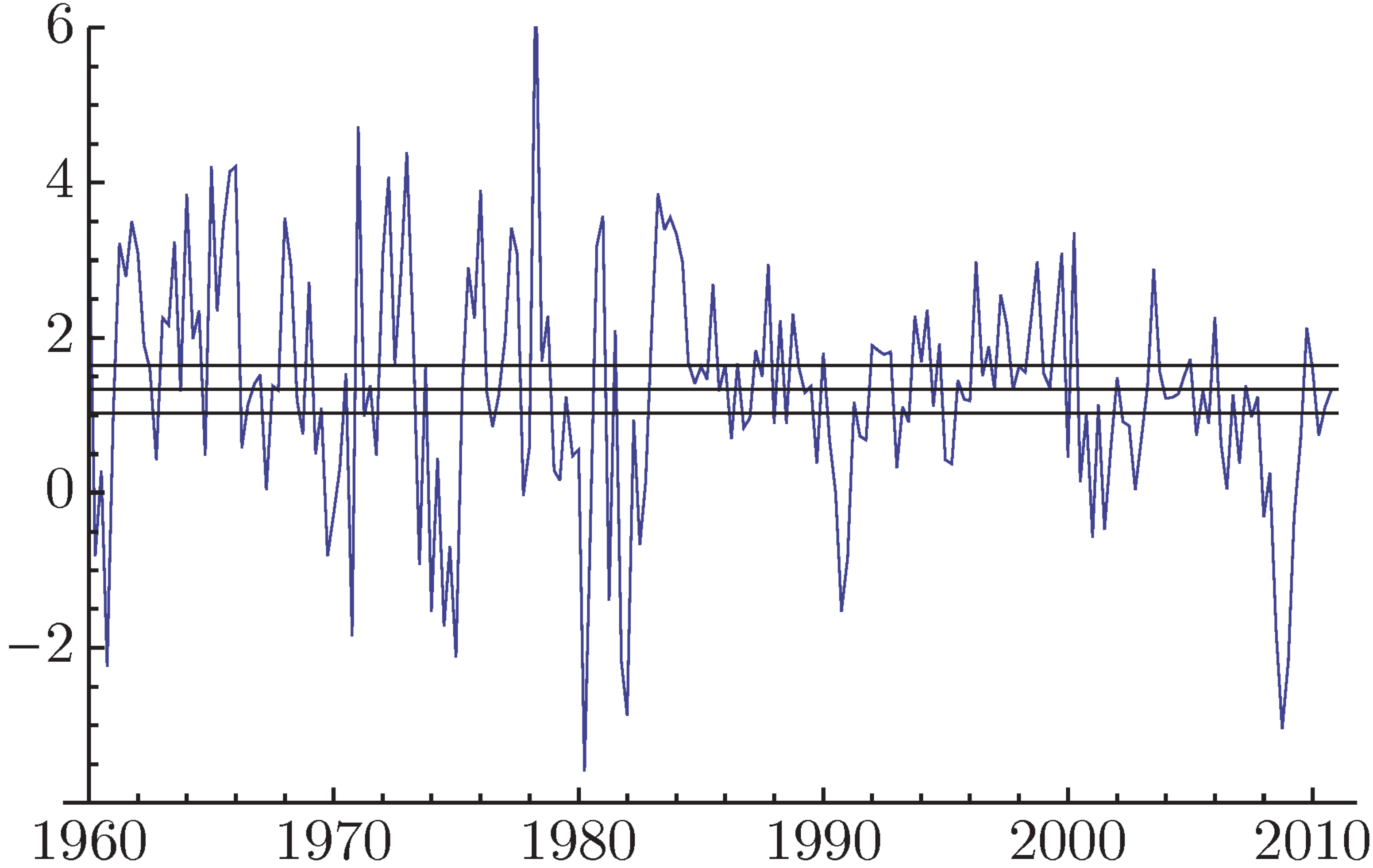
- With $Y_0 = \sqrt{T}(\bar{y} - \mu)$, it follows similar to before that

$$\left\{Y_j\right\}_{j=0}^{q} \Rightarrow iid\mathcal{N}(0, \omega^2).$$

Thus

$$\sqrt{T}\frac{\bar{y} - \mu}{\sqrt{\sum_{j=1}^{q} Y_j^2}} \Rightarrow \text{Student-t}_q$$

95% Confidence Interval for Mean Growth



Application 2: Testing I(0) Property

- Consider local-level model (sum of I(0) and I(1))

$$y_t = \mu + u_t + \frac{g}{T} \sum_{s=1}^{t} \eta_s$$

with u_t and η_t independent I(0) and of common long-run variance ω^2, so that

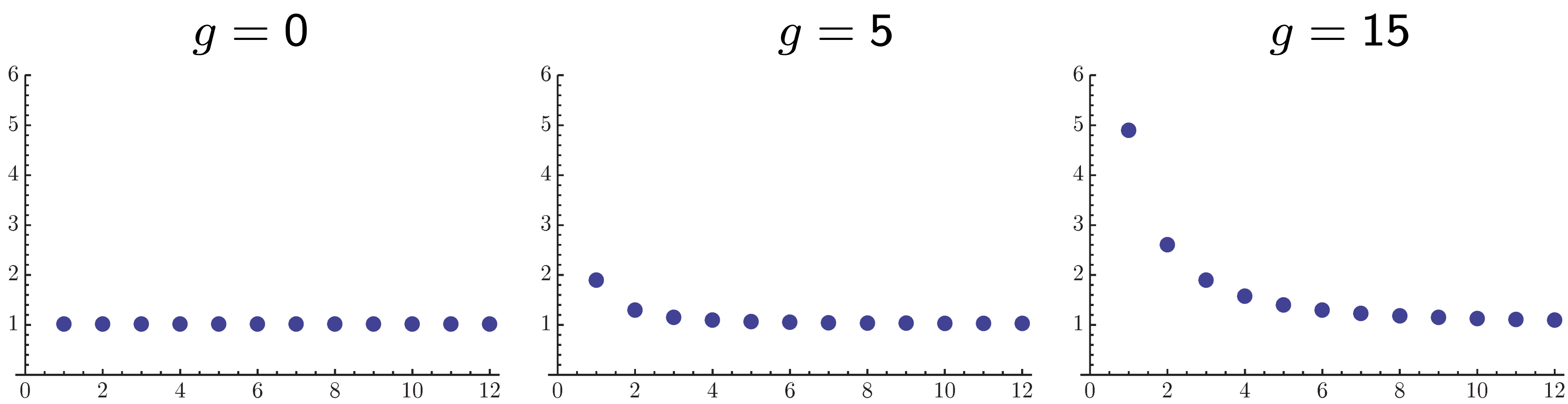
$$T^{-1/2} \sum_{t=1}^{\lfloor \cdot T \rfloor} (y_t - \mu) \Rightarrow \omega W_u(\cdot) + \omega g \int_0^{\cdot} W_\eta(s) ds$$

- With $Y_j = T^{-1/2} \sum_{t=1}^{T} \Psi_j(t/T) y_t$, it follows

$$\begin{aligned} \left\{Y_j\right\}_{j=1}^{q} &\Rightarrow \left\{\omega \int_0^1 \Psi_j(s) dW_u(s) + \omega g \int_0^1 \Psi_j(s) W_\eta(s) ds\right\}_{j=1}^{q} \\ &\sim i\mathcal{N}\left(0, \omega^2 \left(1 + \frac{g^2}{(\pi j)^2}\right)\right) \end{aligned}$$

Standard Deviations of LF Transforms, $\omega = 1$

$$\left\{Y_j\right\}_{j=1}^{q} \Rightarrow i\mathcal{N}\left(0, \omega^2\left(1 + \frac{g^2}{(\pi j)^2}\right)\right)$$



Test Statistic

- Impose scale invariance on test φ: $\varphi(Y) = \varphi(cY)$ for $c > 0$

  $\Rightarrow$ Test is function of maximal invariant $U = Y/\sqrt{Y'Y}$

- Best scale invariant test of

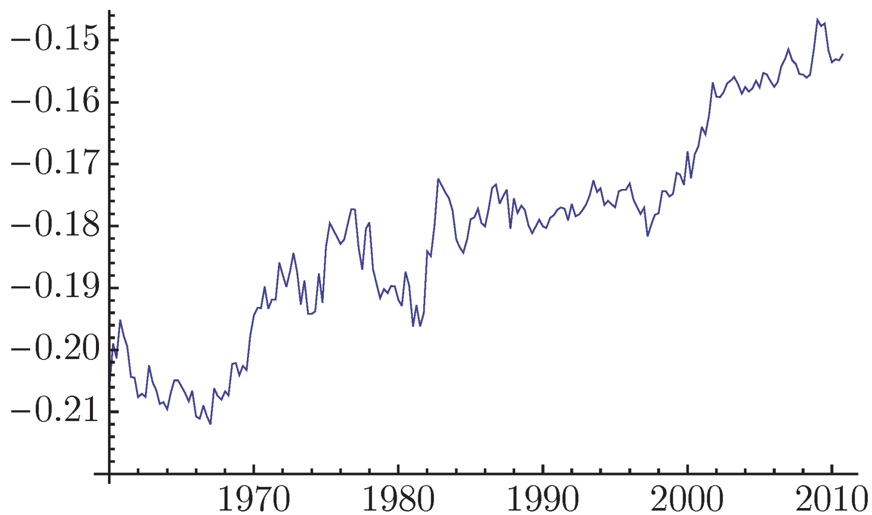
$$
\begin{aligned}
H_0 &: \left\{Y_j\right\}_{j=1}^{q} \sim iid\mathcal{N}(0,\omega^2) \text{ vs.} \\
H_1 &: \left\{Y_j\right\}_{j=1}^{q} \sim i\mathcal{N}\left(0,\omega^2\left(1+\frac{g^2}{(\pi j)^2}\right)\right)
\end{aligned}
$$

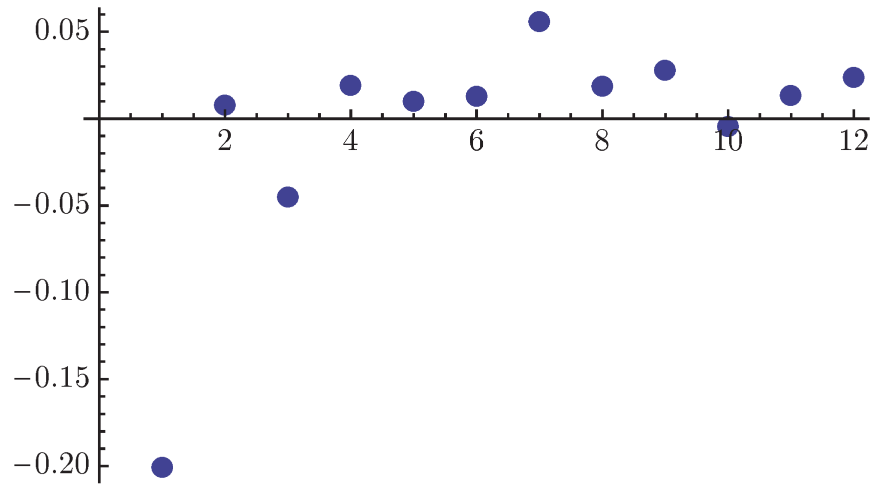
  for a given value of g rejects for large values of LR statistic based on U, or

$$
\frac{\sum_{j=1}^{q} Y_j^2}{\sum_{j=1}^{q}(1+\frac{g^2}{(\pi j)^2})^{-1} Y_j^2}
$$

- Choice of $g = 10$ in test statistic leads to powerful test for all values of g

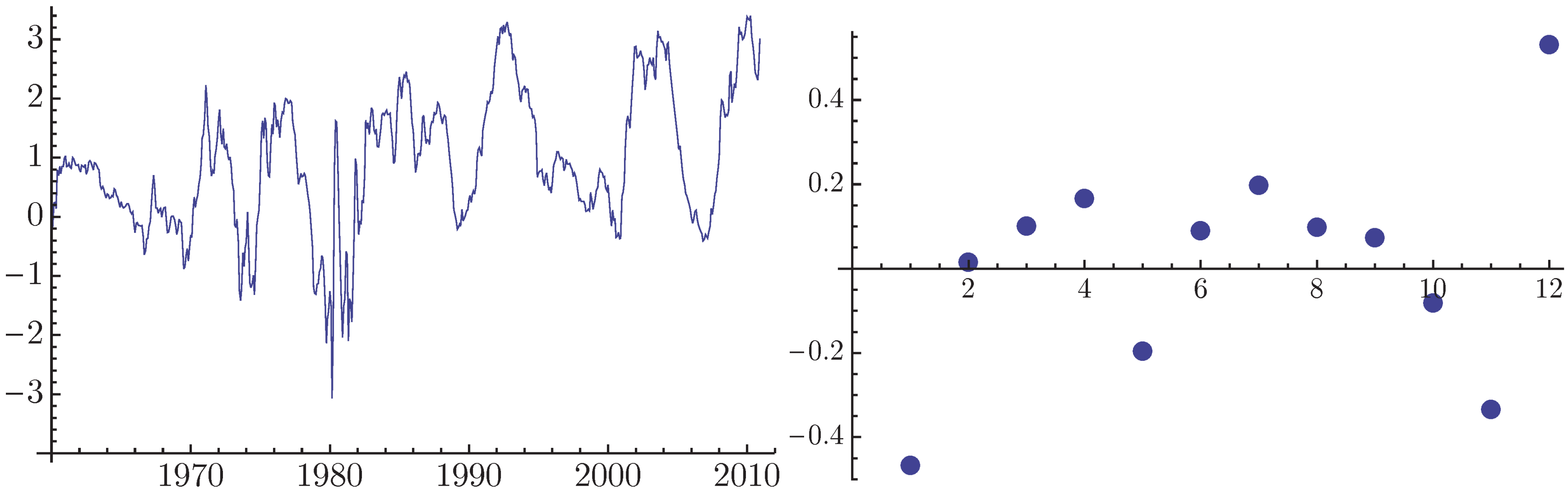
Postwar Consumption/Income Ratio ($q = 12$)





$\Rightarrow$ reject cointegration

Postwar 10 year – 1 year Interest Spread ($q = 12$)



⇒ do not reject cointegration

Application 3: Inference about Degree of Persistence

- Suppose $u_t \sim I(d)$ with $d \in (-1/2, 3/2)$, that is $(1-L)^d u_t = \varepsilon_t$.

- Then $T^{-1/2-d} \sum_{t=1}^{\lfloor \cdot T \rfloor} u_t \Rightarrow \omega W_d(\cdot)$, where W_d is fractional Brownian motion $(-1/2 < d < 1/2)$ or integrated fractional Brownian motion $(1/2 < d < 3/2)$.

  Thus
  $$T^{-d} Y \Rightarrow Z \sim \mathcal{N}(0, \Sigma(d))$$
  and $U = Y/\sqrt{Y'Y} \Rightarrow Z/\sqrt{Z'Z}$.

- Under scale invariance, inference about d becomes inference about covariance matrix $\Sigma(d)$ of Z.

Confidence Set for d

- For any d_0, derive best scale invariant test of

$$H_0 : d = d_0 \quad \text{vs} \quad H_1 : d = D \sim \text{Uniform}[-1/2; 3/2]$$

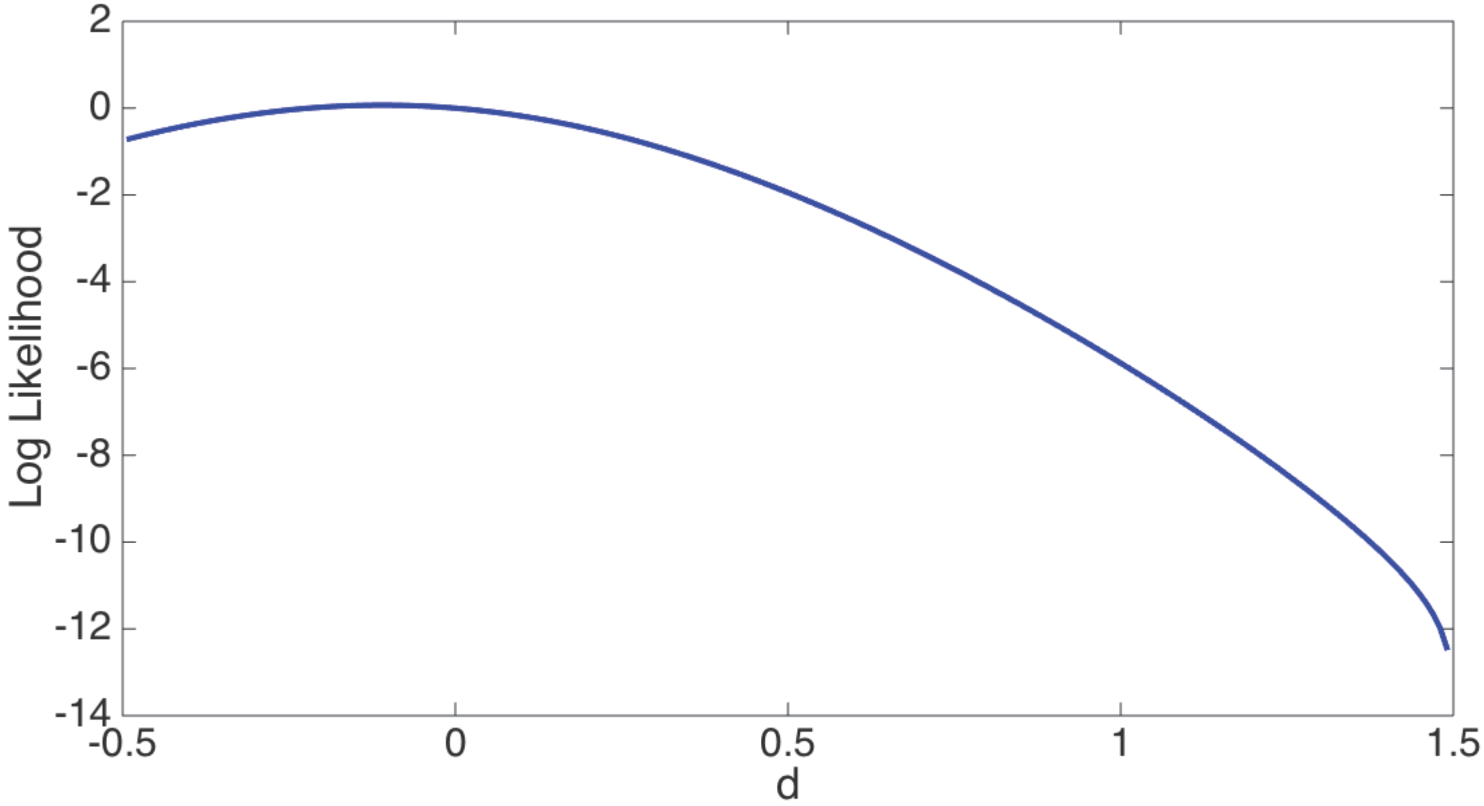
$\Rightarrow$ reject for large values of

$$\frac{E_D[f_U(U; D)]}{f_U(U; d_0)}$$

- Collect values of d_0 for which test does not reject

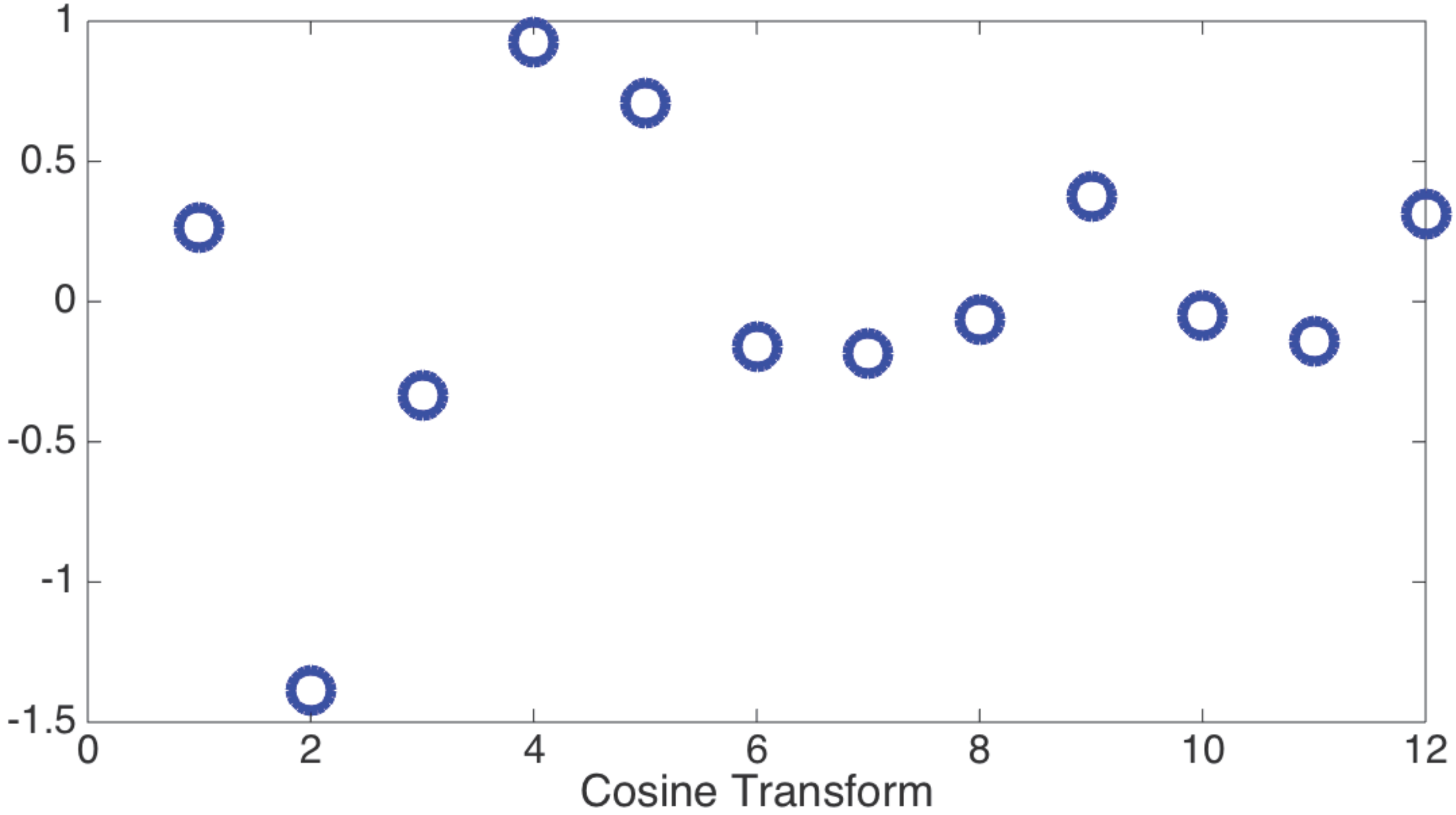
- Pratt (1961): Resulting confidence set minimizes average expected length

GDP Growth: Log-Likelihood of U and CI

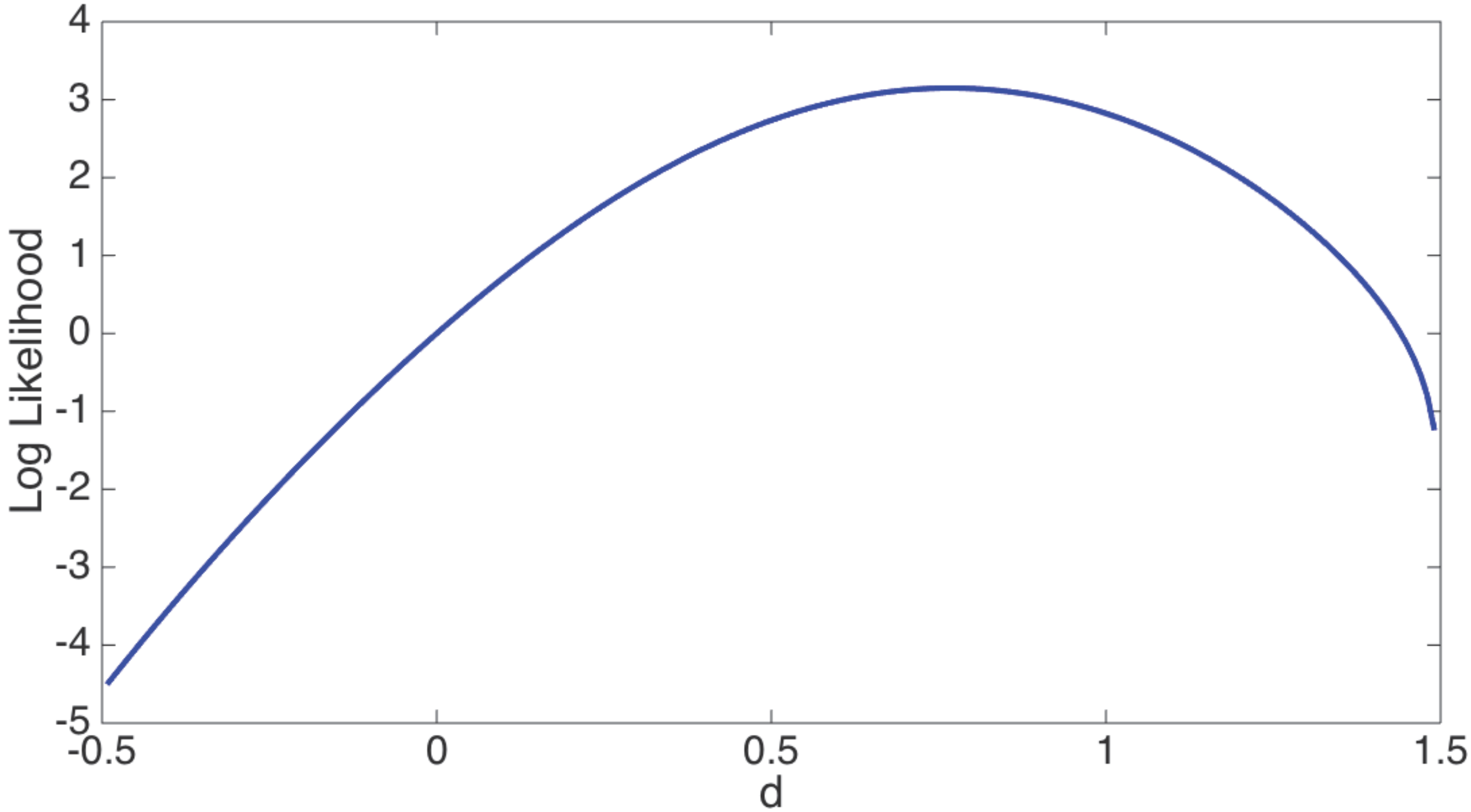


90% CI for d: $[-0.49; 0.39]$

Inflation: LF Transforms



Inflation: Log-Likelihood of U and CI



90% CI for d: $[0.29; 1.18]$

Application 4: Long-Run Forecasts

- Objective: Predictive interval that contains average of future values

$$\bar{y}_{T:T+h} = h^{-1} \sum_{t=1}^{h} y_{T+t}$$

with prespecified probability, for $h = \lfloor rT \rfloor$ a fraction of the sample size

- Allow rich low-frequency dynamics for $y_t = \mu + u_t$:

$$\begin{aligned} u_t &= \varepsilon_{1t} + (bT)^{-d} \eta_t \\ (1 - \rho T)^d \eta_t &= \varepsilon_{2t} \end{aligned}$$

where $\rho = \rho_T = 1 - c/T$, $d \in [-1/2, 3/2]$ and $(\varepsilon_{1T}, \varepsilon_{2T})$ uncorrelated I(0) with long-run variance ω^2

$\Rightarrow$ "bcd-model": Nests local-level model, fractional model and local-to-unity AR(1) model as special cases

Application 4: Long-Run Forecasts

- Under weak primitive conditions on $(\varepsilon_{1t}, \varepsilon_{2t})$, can establish central limit theorem

$$\begin{pmatrix} T^{1/2}(\bar{y}_{T:T+h} - \mu) \\ T^{1/2}(\bar{y} - \mu) \\ Y_1 \\ \vdots \\ Y_q \end{pmatrix} \Rightarrow \mathcal{N}(0, \omega\mathbf{\Sigma}(\theta))$$

where $\theta = (b, c, d)$

- With μ, ω and θ known, appropriate predictive interval follows readily from formula for conditional normal distribution

Application 4: Long-Run Forecasts

- Problem of unknown parameters

  ⇒ Mean parameter μ and scale parameter ω can be handled by invariance

- Bayesian approach: Employ prior on θ and use posterior average over θ conditional forecasts

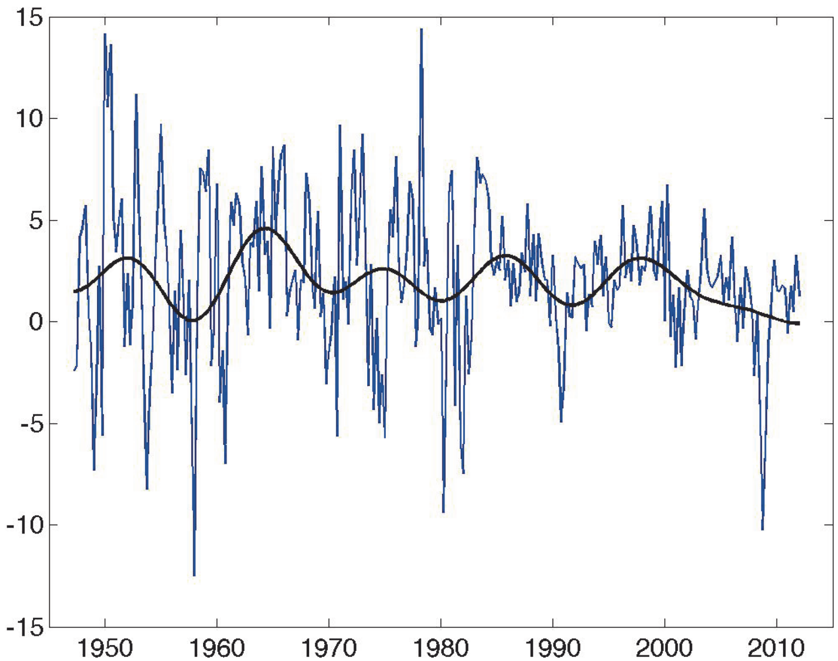
  ⇒ But no coverage of future value under repated sampling under all θ by construction

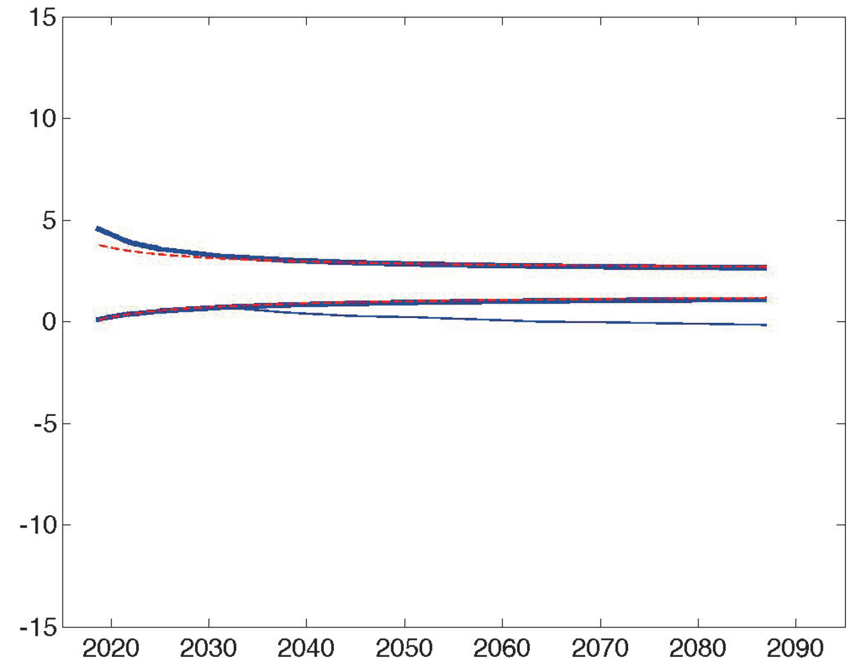
- Enlarge Bayes set in a average expected length manner to ensure coverage

  ⇒ Weighted length subject to coverage constraint can be cast as Lagrangian problem, only challenge is to numerically determine Lagrange multipliers

  ⇒ Use numerical methods similar to Elliott, Müller and Watson (2015)

Average GDP Growth 90% Forecast Intervals



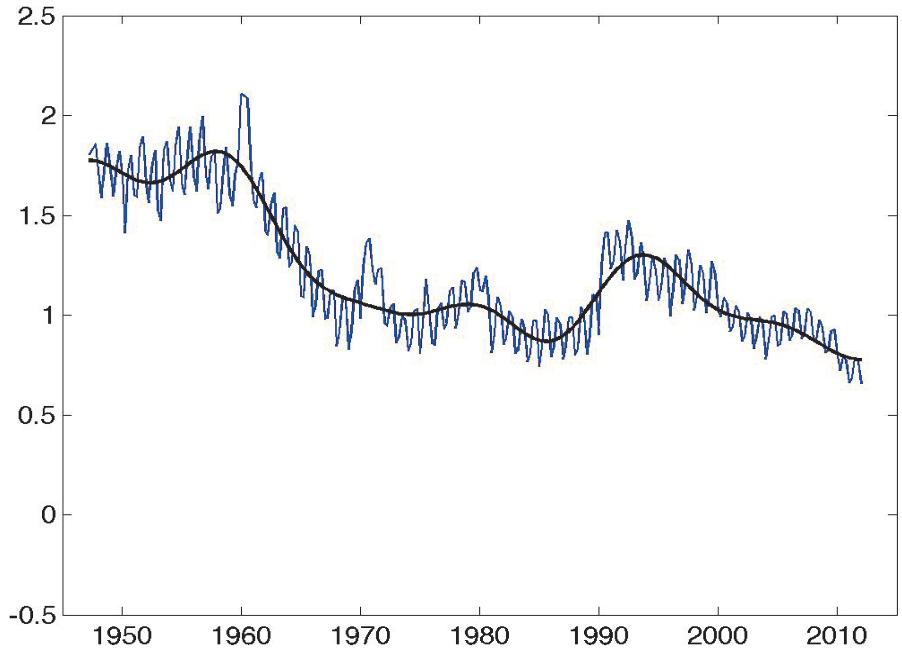


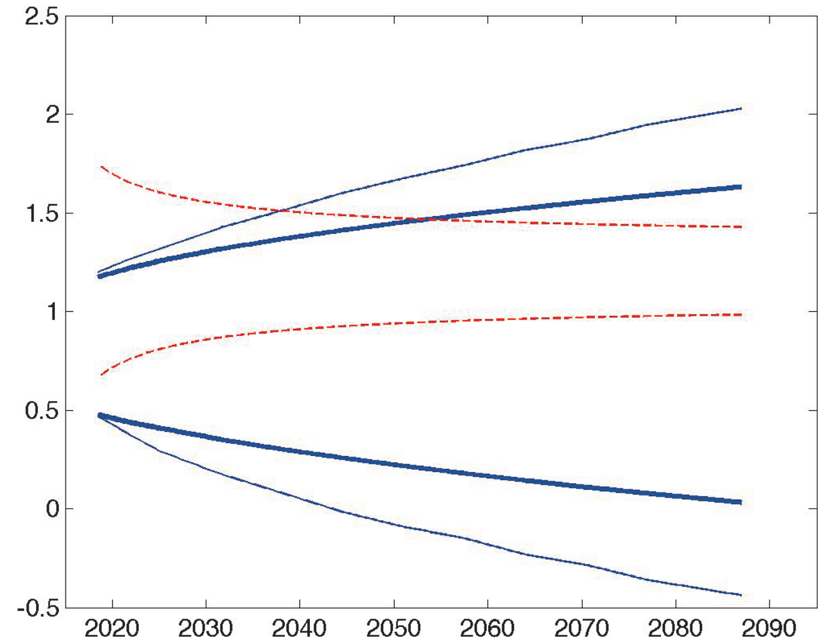
Dashed: I(0)

Thick: Bayes

Thin: 90% Coverage

Average Population Growth 90% Forecast Intervals



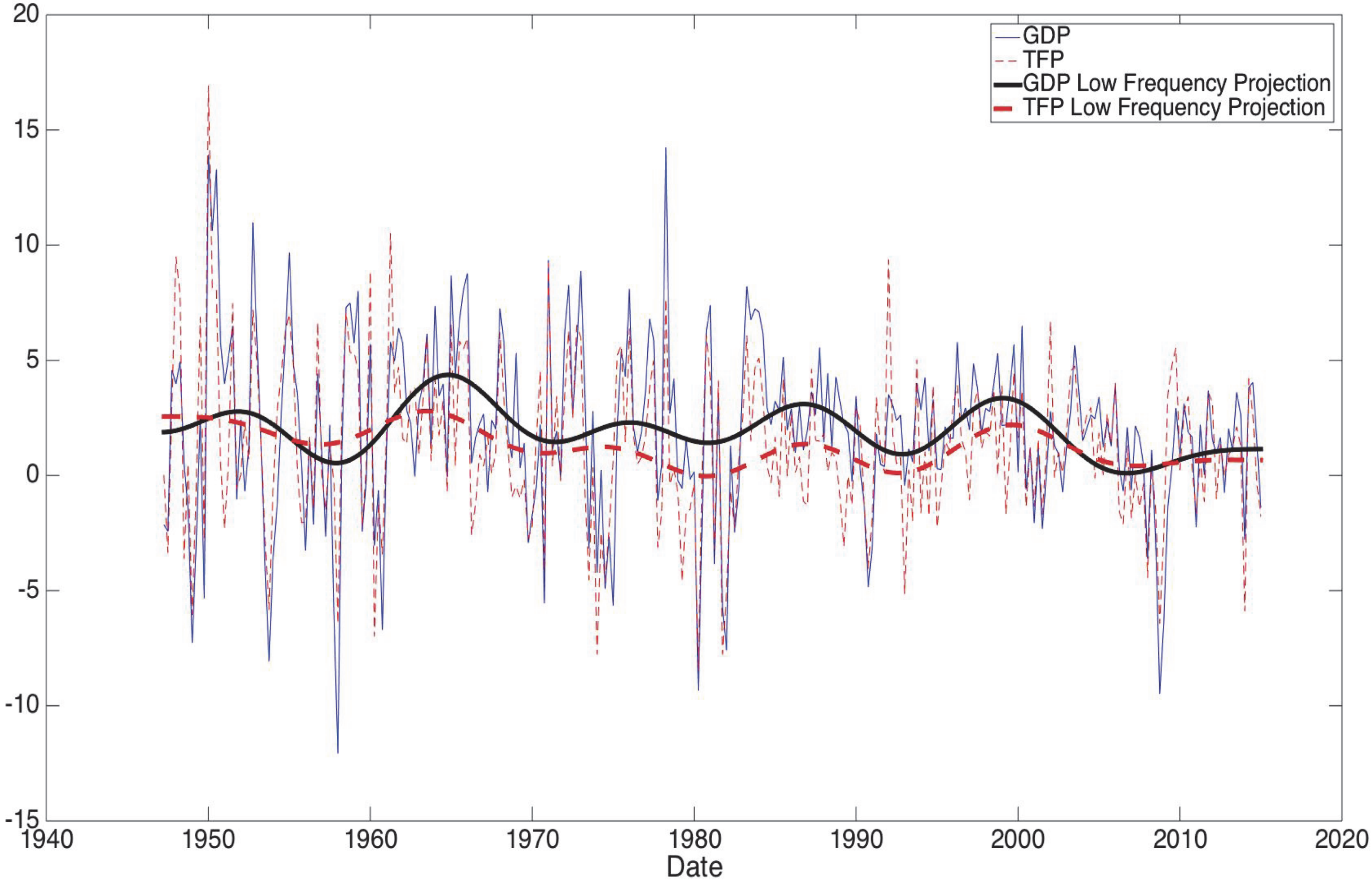


Dashed: I(0)

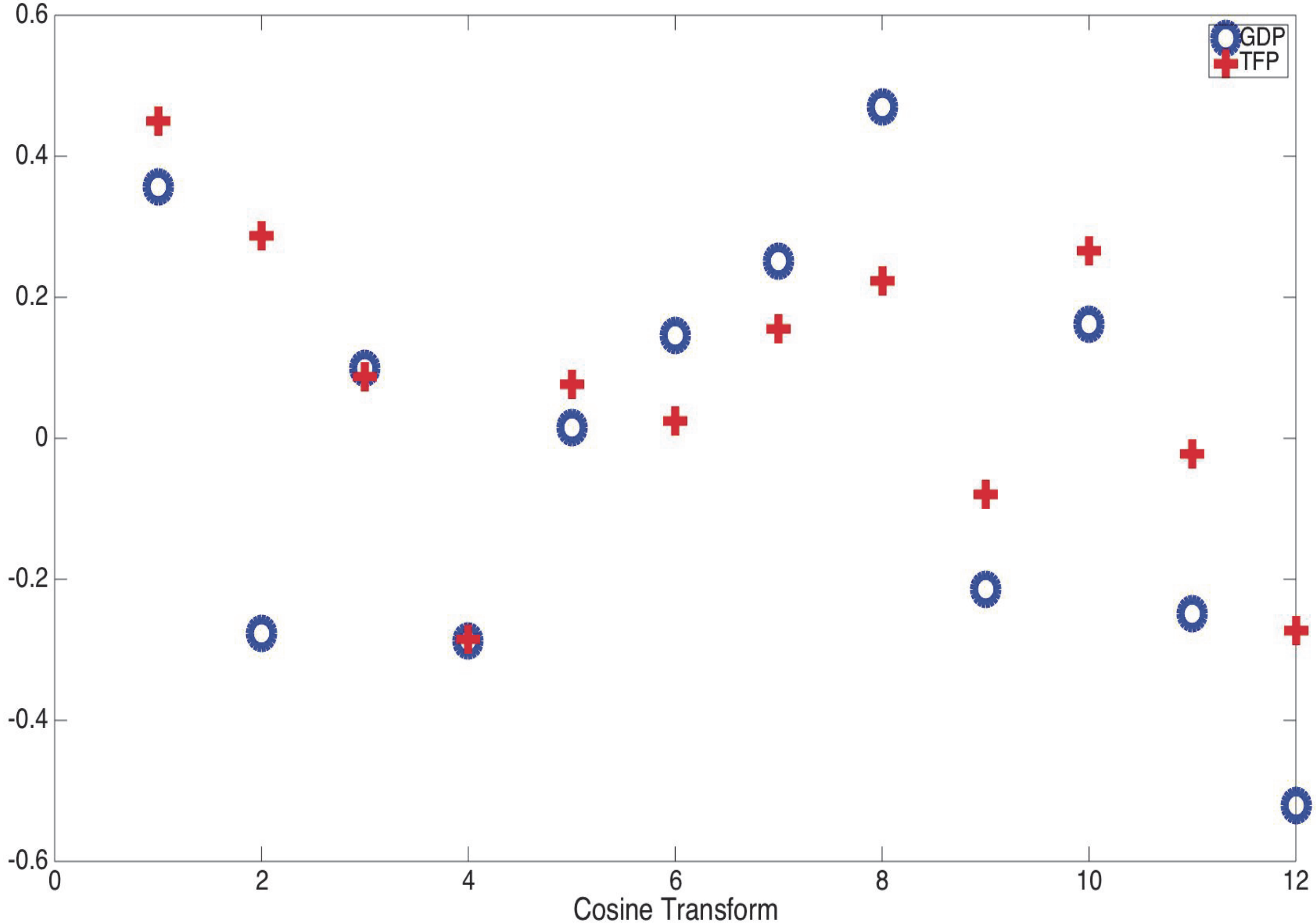
Thick: Bayes

Thin: 90% Coverage

Application 5: Multiple Series



LF Transforms



Multiple Series Questions

- Confidence intervals for joint mean

- Degree of covariability

- Conditional low-frequency variability, low-frequency regression

Bivariate $I(0)$ Model

- Suppose $\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix} + \begin{pmatrix} u_{xt} \\ u_{yt} \end{pmatrix}$, and

$$T^{-1/2} \sum_{t=1}^{\lfloor \cdot T \rfloor} \begin{pmatrix} u_{xt} \\ u_{yt} \end{pmatrix} \Rightarrow \Omega^{1/2} W(\cdot)$$

- Then, with $X_0 = \sqrt{T}(\bar{x} - \mu_x)$ and $Y_0 = \sqrt{T}(\bar{y} - \mu_y)$

$$\begin{pmatrix} X_j \\ Y_j \end{pmatrix} \Rightarrow iid\mathcal{N}\left(0, \Omega\right), j = 0, \ldots, q$$

- Parameters of interest: 2×1 vector $(\mu_x, \mu_y)'$ and 2×2 matrix Ω

Inference about (μ_x, μ_y)

- Natural estimator for Ω:

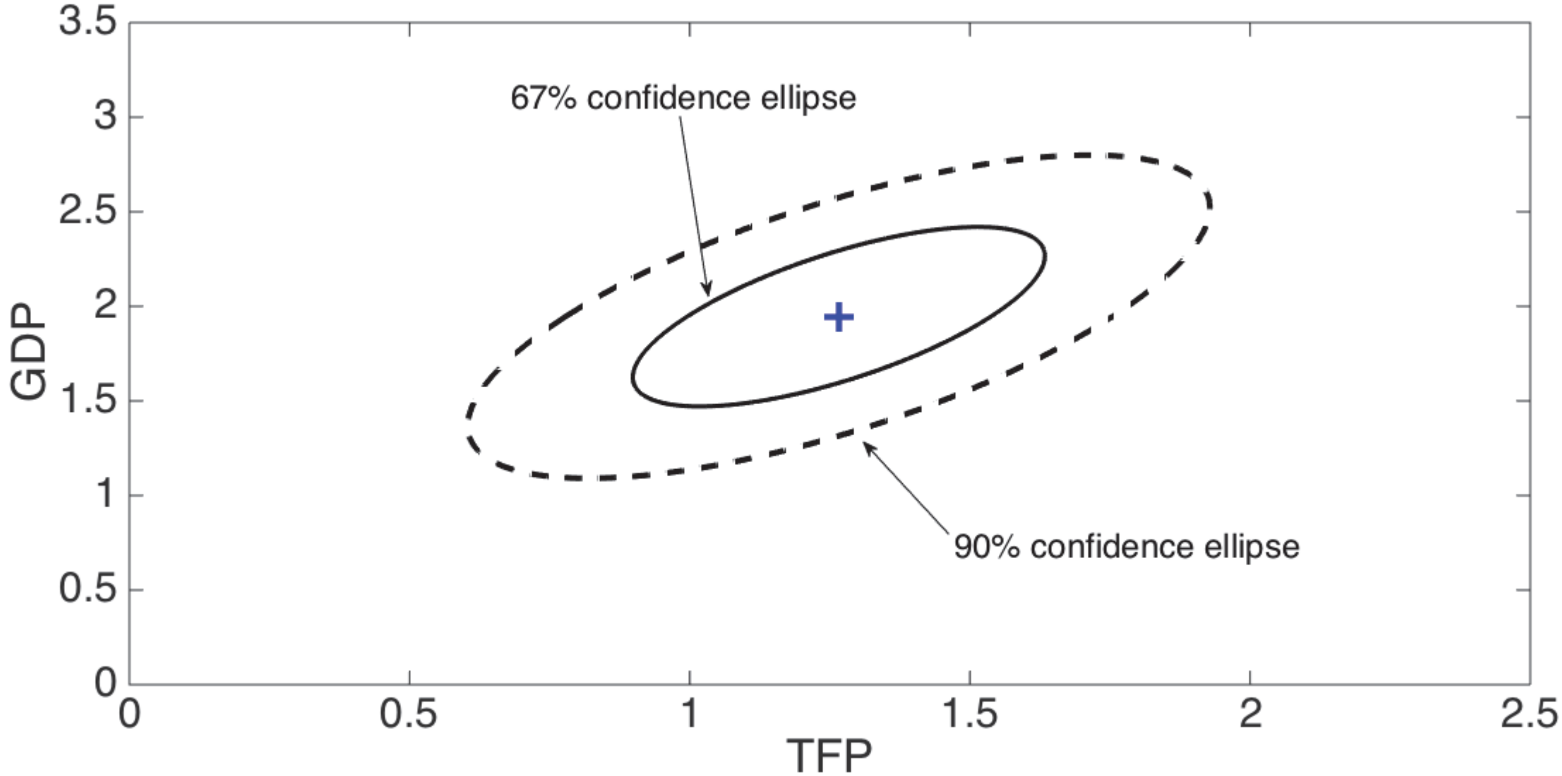
$$\hat{\Omega} = q^{-1} \sum_{j=1}^{q} \begin{pmatrix} X_j \\ Y_j \end{pmatrix} \begin{pmatrix} X_j \\ Y_j \end{pmatrix}' \Rightarrow q^{-1} \text{Wishart}(\Omega, q)$$

- Hotelling (1932):

$$T \begin{pmatrix} \bar{x} - \mu_x \\ \bar{y} - \mu_y \end{pmatrix}' \hat{\Omega}^{-1} \begin{pmatrix} \bar{x} - \mu_x \\ \bar{y} - \mu_y \end{pmatrix} \Rightarrow \frac{2q}{q-1} F_{2,q-1}$$

$\Rightarrow$ Large sample confidence sets for (μ_x, μ_y) are ellipses with radius determined by critical value of F distribution

Confidence Set for GDP and TFP Growth Means



Inference about (Functions of) Ω

- $\hat{\Omega} \Rightarrow q^{-1}\text{Wishart}(\Omega, q)$

  $\Rightarrow$ Could be used to do inference about 2×2 matrix $\Omega = \begin{pmatrix} \Omega_{xx} & \Omega_{xy} \\ \Omega_{yx} & \Omega_{yy} \end{pmatrix}$

- Correlation $\rho = \Omega_{xy}/\sqrt{\Omega_{xx}\Omega_{yy}}$ is naturally estimated by

$$\hat{\rho} = \hat{\Omega}_{xy}/\sqrt{\hat{\Omega}_{xx}\hat{\Omega}_{yy}}$$

  $\Rightarrow$ Distribution of $\hat{\rho}$ is known and depends on Ω only through ρ

  $\Rightarrow$ Easy to obtain confidence intervals for ρ based on $\hat{\rho}$

  $\Rightarrow$ In example: 90% CI for ρ is $[0.29; 0.86]$

Low-Frequency Regression

- Alternative function of Ω of potential interest: $\beta = \Omega_{xx}^{-1}\Omega_{xy}$

  $\Rightarrow$ In large sample approximation $(X_j, Y_j) \Rightarrow iid\mathcal{N}(0, \Omega)$, conditional mean of Y_j given $X = (X_1, \ldots, X_q)$ is βX_j

  $\Rightarrow$ Parameter β describes how low-frequency variability of x_t predicts low frequency variability of y_t

- In large sample approximation, small sample Gaussian linear regression:

$$Y_j = \beta X_j + \varepsilon_j, \quad \varepsilon_j | X \sim iid\mathcal{N}(0, \sigma^2)$$

  for $j = 1, \ldots, q$ with $\sigma^2 = \Omega_{yy} - \Omega_{yx}\Omega_{xx}^{-1}\Omega_{xy}$

Low-Frequency Regression, Ctd.

- Inference about β via usual small sample results:

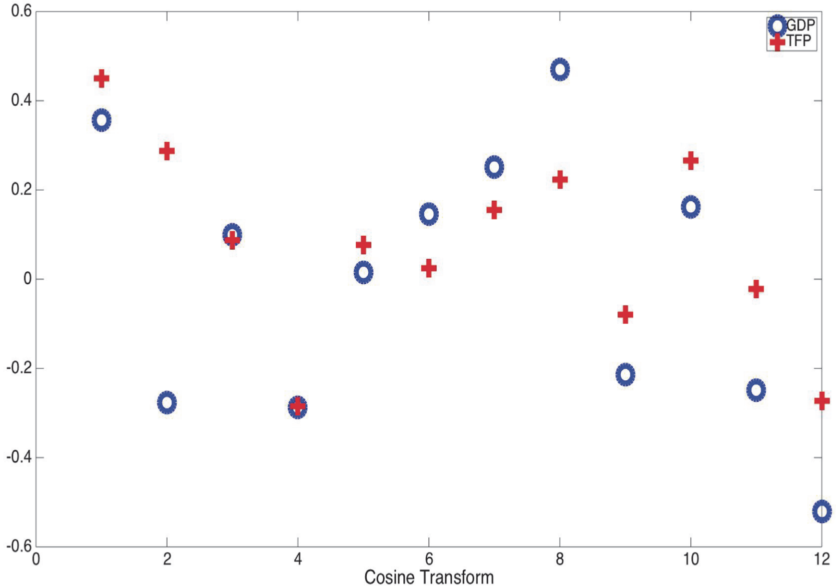
$$\hat{\beta} = (\sum_{j=1}^{q} X_j^2)^{-1} \sum_{j=1}^{q} X_j Y_j, \quad \hat{\beta}|X \overset{a}{\sim} \mathcal{N}\left(\beta, \sigma^2 (\sum_{j=1}^{q} X_j^2)^{-1}\right)$$

$$\hat{\sigma}^2 = (q-1)^{-1} \sum_{j=1}^{q} (Y_j - \hat{\beta} X_j)^2, \; \hat{\sigma}^2/\sigma^2 \overset{a}{\sim} \chi^2_{q-1}$$

so that with $se(\hat{\beta}) = \sqrt{\hat{\sigma}^2 (\sum_{j=1}^{q} X_j^2)^{-1}}$

$$\text{t-stat} = \frac{\hat{\beta} - \beta}{se(\hat{\beta})} \Rightarrow \text{Student-t}_{q-1}$$

LF Regression GDP Growth on TFP Growth



Statistic	
Sample size (q)	12
$\hat{\beta}$	0.87
$se(\hat{\beta})$	0.29
t-statistic	3.04
90% CI for β	[0.36;1.39]
Standard error ($\hat{\sigma}$)	0.22
R^2	0.46
90% CI for ρ^2	[0.08;0.74]

Note: In basic model of balanced growth, permanent TFP shock leads to long-run permanent increase in GDP of $1/(1-\alpha) = 1.5$ with elasticity of output relative to capital $\alpha = 2/3$

Alternative Interpretation of Ω

- Could alternatively measure low-frequency covariability by covariance of LF-projections $(\hat{x}_t, \hat{y}_t)'$, averaged over time t, that is by

$$\tilde{\Omega} = T^{-1} \sum_{t=1}^{T} E\left[\begin{pmatrix} \hat{x}_t - \mu_x \\ \hat{y}_t - \mu_y \end{pmatrix} \begin{pmatrix} \hat{x}_t - \mu_x \\ \hat{y}_t - \mu_y \end{pmatrix}' \right]$$

- Projection coefficients are X_j and Y_j, and cosines are orthonormal:

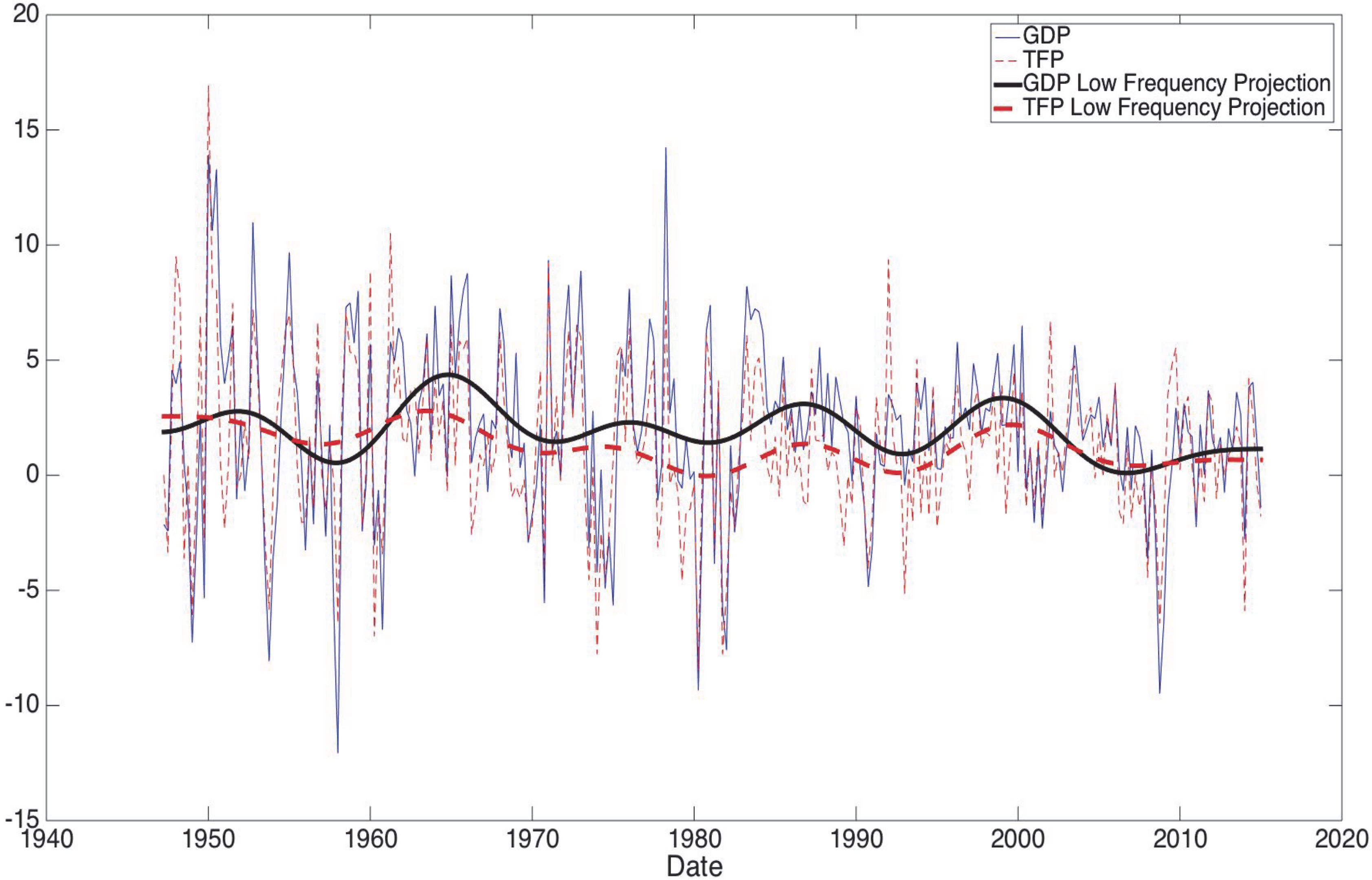
$$\frac{T}{q+1} \tilde{\Omega} = \Omega = E\left[\begin{pmatrix} X_j \\ Y_j \end{pmatrix} \begin{pmatrix} X_j \\ Y_j \end{pmatrix}' \right]$$

- Equivalence also holds for estimators

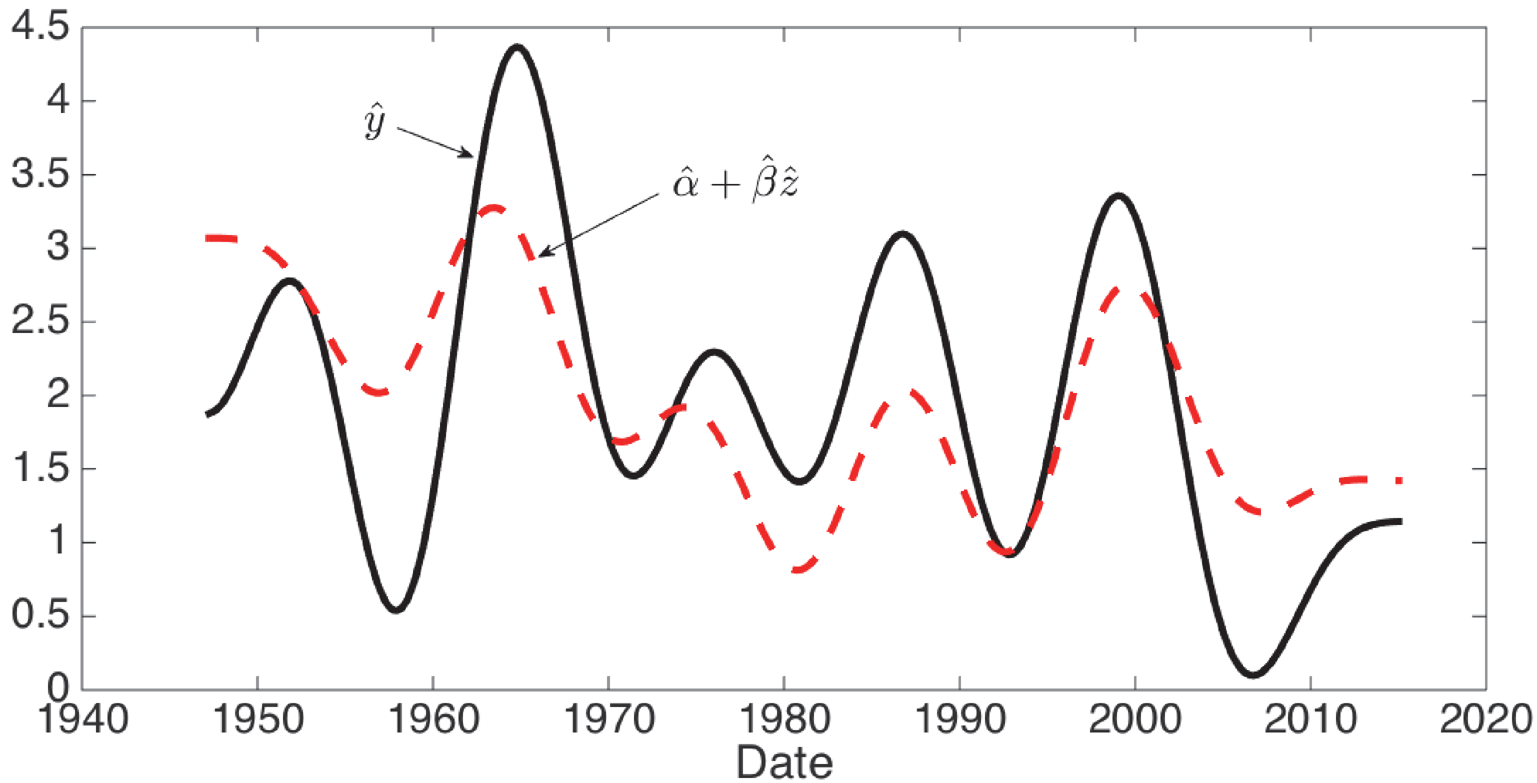
$\Rightarrow$ $\hat{\rho}$ is time series sample correlation between $\hat{x}_t$ and $\hat{y}_t$

$\Rightarrow$ $\hat{\beta}$ is coefficient in time series regression of $\hat{y}_t$ on $\hat{x}_t$ and a constant

LF Projections GDP and TFP Growth



GDP Projection $\hat{y}$ and Predicted Values of LF Regression $\hat{\alpha} + \hat{\beta}\hat{x}$



Conclusions

Ultra low-frequency econometrics:

- Explicitly addresses scarcity of relevant information
- Avoids modelling and thus misspecification at higher frequencies
- Leads to tractable small sample Gaussian inference problems
- More topics: Non-Gaussian limits of X_j (stochastic volatility), breaks, measures of covariability that allow for persistence, etc.