Peter C. B. Phillips' Contributions and Influence in Financial Econometrics

Jun Yu[1]

[1]Singapore Management University

March 2008

Outline

Outline

Overview 1: Indirect influence

1. Unit root and cointegration in financial time series.

Overview 1: Indirect influence

1. Unit root and cointegration in financial time series.
2. Fractional integration in volatility. References: Phillips and Shimostu (2005, Annals). Financial Applications: Lieberman and Phillips (2007), Huang, et al (2007)

Overview 1: Indirect influence

1. Unit root and cointegration in financial time series.
2. Fractional integration in volatility. References: Phillips and Shimostu (2005, Annals). Financial Applications: Lieberman and Phillips (2007), Huang, et al (2007)
3. Local to unity. References: Phillips (1988). Financial Applications: Valkanov (1999)

Overview 1: Indirect influence

1. Unit root and cointegration in financial time series.
2. Fractional integration in volatility. References: Phillips and Shimostu (2005, Annals). Financial Applications: Lieberman and Phillips (2007), Huang, et al (2007)
3. Local to unity. References: Phillips (1988). Financial Applications: Valkanov (1999)
4. Edgeworth and saddlepoint approximations. References: Phillips (1977, 1978), Holy and Phillips (1979). Financial Applications: Aït-Sahalia and Yu (2007)

Overview 1: Indirect influence

1. Unit root and cointegration in financial time series.
2. Fractional integration in volatility. References: Phillips and Shimostu (2005, Annals). Financial Applications: Lieberman and Phillips (2007), Huang, et al (2007)
3. Local to unity. References: Phillips (1988). Financial Applications: Valkanov (1999)
4. Edgeworth and saddlepoint approximations. References: Phillips (1977, 1978), Holy and Phillips (1979). Financial Applications: Aït-Sahalia and Yu (2007)
5. Nonlinear cointegration. References: Phillips and Park (2002). Financial Applications: Hu and Phillips (2004), Huang and Yu (2008).

Overview 2: Direct influence

1. Identification.

Overview 2: Direct influence

1. Identification.
2. Parametric methods – discretization bias.

Overview 2: Direct influence

1. Identification.
2. Parametric methods – discretization bias.
3. Nonparametric methods – specification bias.

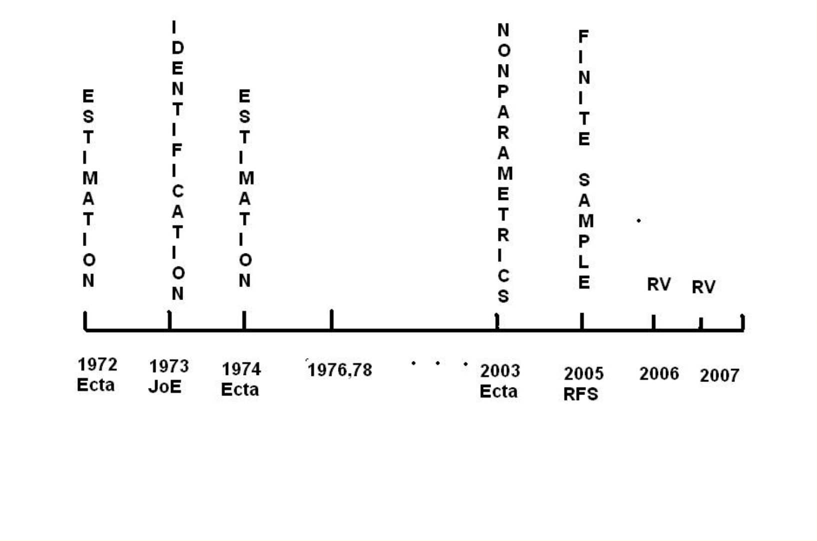
Overview 2: Direct influence

1. Identification.
2. Parametric methods – discretization bias.
3. Nonparametric methods – specification bias.
4. Finite sample issues – finite sample bias.

Overview 2: Direct influence

1. Identification.
2. Parametric methods – discretization bias.
3. Nonparametric methods – specification bias.
4. Finite sample issues – finite sample bias.
5. Others

Overview 3: Time Series Plot



Outline

The Model: Phillips (1973, JoE)

- First study of identification of continuous time econometric models.

The Model: Phillips (1973, JoE)

- First study of identification of continuous time econometric models.
- The model

$$DX(t) = AX(t) + \xi(t), \tag{1}$$

The Model: Phillips (1973, JoE)

- First study of identification of continuous time econometric models.
- The model

$$DX(t) = AX(t) + \xi(t), \tag{1}$$

- $X(t)$ is a n-dimensional vector of random functions observed at discrete points in time. $\xi(t)$ is the pure noise term with covariance $\Sigma\Sigma'$.

The Model: Phillips (1973, JoE)

- First study of identification of continuous time econometric models.
- The model

$$DX(t) = AX(t) + \xi(t), \tag{1}$$

- $X(t)$ is a n-dimensional vector of random functions observed at discrete points in time. $\xi(t)$ is the pure noise term with covariance $\Sigma\Sigma'$.
- It nests the $A_0(m)$ model of Dai and Singleton (2001, JF):

$$dX(t) = AX(t)dt + \Sigma dB(t) \tag{2}$$

The Model: Phillips (1973, JoE)

- First study of identification of continuous time econometric models.
- The model

$$DX(t) = AX(t) + \xi(t), \tag{1}$$

- $X(t)$ is a n-dimensional vector of random functions observed at discrete points in time. $\xi(t)$ is the pure noise term with covariance $\Sigma\Sigma'$.
- It nests the $A_0(m)$ model of Dai and Singleton (2001, JF):

$$dX(t) = AX(t)dt + \Sigma dB(t) \tag{2}$$

- The exact discrete time model is:

$$X_t = BX_{t-1} + \varepsilon_t, \tag{3}$$

where $B = \exp(Ah)$, $\varepsilon_t = \int_0^h \exp(sA)\xi(t-s)ds$.

The Identification Problem

- Aliasing problem: different continuous time models may be indistinguishable when sampled at discrete points in time.

The Identification Problem

- Aliasing problem: different continuous time models may be indistinguishable when sampled at discrete points in time.
- While B can be consistently estimated from (3), A cannot be identified from B.

The Identification Problem

- Aliasing problem: different continuous time models may be indistinguishable when sampled at discrete points in time.
- While B can be consistently estimated from (3), A cannot be identified from B.
- Solution: Prior restrictions on A.

The Identification Problem

- Aliasing problem: different continuous time models may be indistinguishable when sampled at discrete points in time.
- While B can be consistently estimated from (3), A cannot be identified from B.
- Solution: Prior restrictions on A.
- Extensions: Hansen and Sargent (1983, Ecta).

Outline

The Basic Model: Phillips (1972, Ecta)

- First time in the literature addressing discretization bias.

The Basic Model: Phillips (1972, Ecta)

- First time in the literature addressing discretization bias.
- Model

$$DX(t) = A(\theta)X(t) + \xi(t).$$

The Basic Model: Phillips (1972, Ecta)

- First time in the literature addressing discretization bias.
- Model

$$DX(t) = A(\theta)X(t) + \xi(t).$$

- The exact discrete time model

$$X_t = B(\theta)X_{t-1} + \varepsilon_t = \exp(A(\theta)h)X_{t-1} + \varepsilon_t$$

The Basic Model: Phillips (1972, Ecta)

- First time in the literature addressing discretization bias.
- Model

$$DX(t) = A(\theta)X(t) + \xi(t).$$

- The exact discrete time model

$$X_t = B(\theta)X_{t-1} + \varepsilon_t = \exp(A(\theta)h)X_{t-1} + \varepsilon_t$$

- Why exact discrete time model? Not subject to discretization bias!

The Basic Model: Phillips (1972, Ecta)

- First time in the literature addressing discretization bias.
- Model
$$DX(t) = A(\theta)X(t) + \xi(t).$$
- The exact discrete time model
$$X_t = B(\theta)X_{t-1} + \varepsilon_t = \exp(A(\theta)h)X_{t-1} + \varepsilon_t$$
- Why exact discrete time model? Not subject to discretization bias!
- Euler discretization:
$$X_t = (1 - Ah)X_{t-1} + \epsilon_t.$$

Continue

- When $A = -1$, $h = 1/12$, e^{Ah} is 0.92 whereas $1 - Ah$ is 0.9167. If $A = 1$, $h = 1$, then e^{Ah} is 0.3679 whereas $1 - Ah$ is 0.

Continue

- When $A=-1$, $h=1/12$, e^{Ah} is 0.92 whereas $1-Ah$ is 0.9167. If $A=1$, $h=1$, then e^{Ah} is 0.3679 whereas $1-Ah$ is 0.
- GLS procedure for the exact discrete time model:

$$\min(X_t - \exp(A(\theta)h)X_{t-1})'S(X_t - \exp(A(\theta)h)$$

Continue

- When $A = -1$, $h = 1/12$, e^{Ah} is 0.92 whereas $1 - Ah$ is 0.9167. If $A = 1$, $h = 1$, then e^{Ah} is 0.3679 whereas $1 - Ah$ is 0.
- GLS procedure for the exact discrete time model:

$$\min(X_t - \exp(A(\theta)h)X_{t-1})'S(X_t - \exp(A(\theta)h)$$

- Gaussian estimation (QMLE).

The Extensions: Phillips (1974, Ecta)

- Extension 1 – *model with identities*

$$DX(t) = A(\theta)X(t) + H\xi(t),$$

where A is $n \times n$ but H is $n \times r$ with $r < n$.

The Extensions: Phillips (1974, Ecta)

- Extension 1 – *model with identities*

$$DX(t) = A(\theta)X(t) + H\xi(t),$$

where A is $n \times n$ but H is $n \times r$ with $r < n$.

- The exact discrete time model

$$X_t = \exp(A(\theta))X_{t-1} + \varepsilon_t,$$

where $\varepsilon_t = \int_0^1 \exp(sA)H\xi(t-s)ds$.

The Extensions: Phillips (1974, Ecta)

- Extension 1 – *model with identities*

$$DX(t) = A(\theta)X(t) + H\xi(t),$$

  where A is $n \times n$ but H is $n \times r$ with $r < n$.

- The exact discrete time model

$$X_t = \exp(A(\theta))X_{t-1} + \varepsilon_t,$$

  where $\varepsilon_t = \int_0^1 \exp(sA)H\xi(t-s)ds$.

- Conditions are needed to ensure the positivity of the covariance of ε_t.

- Extension 2 – *model with exogenous variables*

$$DX(t) = A(\theta h)X(t) + BZ(t) + H\xi(t),$$

where B is $n \times m$ and $Z(t)$ is a vector of exogenous variables.

- Extension 2 – *model with exogenous variables*

$$DX(t) = A(\theta h)X(t) + BZ(t) + H\xi(t),$$

where B is $n \times m$ and $Z(t)$ is a vector of exogenous variables.

- The exact discrete time model

$$X_t = \exp(A(\theta)h)X_{t-1} + \int_0^h \exp(sA)BZ(t-s)ds + \varepsilon_t.$$

- Extension 2 – *model with exogenous variables*

$$DX(t) = A(\theta h)X(t) + BZ(t) + H\xi(t),$$

where B is $n \times m$ and $Z(t)$ is a vector of exogenous variables.

- The exact discrete time model

$$X_t = \exp(A(\theta)h)X_{t-1} + \int_0^h \exp(sA)BZ(t-s)ds + \varepsilon_t.$$

- This is not estimable unless $Z(t-s)$ is analytically integrable.

- Extension 2 – *model with exogenous variables*

$$DX(t) = A(\theta h)X(t) + BZ(t) + H\xi(t),$$

where B is $n \times m$ and $Z(t)$ is a vector of exogenous variables.

- The exact discrete time model

$$X_t = \exp(A(\theta)h)X_{t-1} + \int_0^h \exp(sA)BZ(t-s)ds + \varepsilon_t.$$

- This is not estimable unless $Z(t-s)$ is analytically integrable.
- Phillips (1974) proposed techniques to approximate $\int_0^h \exp(sA)BZ(t-s)ds$.

- Extension 2 – *model with exogenous variables*

$$DX(t) = A(\theta h)X(t) + BZ(t) + H\xi(t),$$

where B is $n \times m$ and $Z(t)$ is a vector of exogenous variables.

- The exact discrete time model

$$X_t = \exp(A(\theta)h)X_{t-1} + \int_0^h \exp(sA)BZ(t-s)ds + \varepsilon_t.$$

- This is not estimable unless $Z(t-s)$ is analytically integrable.
- Phillips (1974) proposed techniques to approximate $\int_0^h \exp(sA)BZ(t-s)ds$.
- Further extensions: mixed stock and flow variables, higher order models.

Interest Rate Models: Yu and Phillips (2001, Ects J)

- Models

$$dr(t) = \kappa(\mu - r(t))dt + \sigma(r_t)dB(t), \tag{4}$$

Interest Rate Models: Yu and Phillips (2001, Ects J)

- Models

$$dr(t) = \kappa(\mu - r(t))dt + \sigma(r_t)dB(t), \tag{4}$$

- Exact discrete time model

$$r(t) = e^{-\kappa h} r(t-1) + \mu(1 - e^{-\kappa h}) + \int_0^h e^{-\kappa(h-\tau)} \sigma(r_{t+\tau}) dB(\tau), \tag{5}$$

Interest Rate Models: Yu and Phillips (2001, Ects J)

- Models

$$dr(t) = \kappa(\mu - r(t))dt + \sigma(r_t)dB(t), \tag{4}$$

- Exact discrete time model

$$r(t) = e^{-\kappa h} r(t-1) + \mu(1 - e^{-\kappa h}) + \int_0^h e^{-\kappa(h-\tau)} \sigma(r_{t+\tau}) dB(\tau), \tag{5}$$

- The error is not normal.

Continue

- To get an exact discrete model with Gaussian errors, Yu and Phillips (2001) used

$$r(t_{j+1}) = \mu(1 - e^{-\kappa h_{j+1}}) + e^{-\kappa h_{j+1}} r(t_j) + M(h_{j+1}), \qquad (6)$$

where $M(h)$ is a continuous martingale with quadratic variation

$$[M]_h = \sigma^2 \int_0^h e^{-2\kappa(h-\tau)} \sigma^2(t+\tau) d\tau, \qquad (7)$$

$$h_{j+1} = \inf\{s | [M_j]_s \geq a\} = \inf\{s | \int_0^s e^{-2\kappa(s-\tau)} \sigma^2(t_j+\tau) d\tau \geq a\}, \qquad (8)$$

with a predetermined value for a.

Continue

- To get an exact discrete model with Gaussian errors, Yu and Phillips (2001) used

$$r(t_{j+1}) = \mu(1 - e^{-\kappa h_{j+1}}) + e^{-\kappa h_{j+1}} r(t_j) + M(h_{j+1}), \qquad (6)$$

where $M(h)$ is a continuous martingale with quadratic variation

$$[M]_h = \sigma^2 \int_0^h e^{-2\kappa(h-\tau)} \sigma^2(t+\tau) d\tau, \qquad (7)$$

$$h_{j+1} = \inf\{s | [M_j]_s \geq a\} = \inf\{s | \int_0^s e^{-2\kappa(s-\tau)} \sigma^2(t_j+\tau) d\tau \geq a\}, \qquad (8)$$

with a predetermined value for a.

- According to Dambis, Dambus-Schwarz theorem, $M(h_{j+1}) \sim N(0, a)$.

Outline

Method 1: Bandi and Phillips (2003, Ecta)

- Why nonparametric models: not subject to specification bias!
- First fully nonparametric estimation of SDF without stationarity

Method 1: Bandi and Phillips (2003, Ecta)

- Why nonparametric models: not subject to specification bias!
- First fully nonparametric estimation of SDF without stationarity
- Time homogenous diffusion nonparametric model

$$dX(t) = \mu(X(t))dt + \sigma(X(t))dB(t), \tag{9}$$

Method 1: Bandi and Phillips (2003, Ecta)

- Why nonparametric models: not subject to specification bias!
- First fully nonparametric estimation of SDF without stationarity
- Time homogenous diffusion nonparametric model

$$dX(t) = \mu(X(t))dt + \sigma(X(t))dB(t), \tag{9}$$

- Note that

$$E\left[dX(t) \mid X(t)\right] = \mu(X(t))dt, \tag{10}$$

$$E\left[(dX(t))^2 \mid X(t)\right] = \sigma^2(X(t))dt, \tag{11}$$

Method 1: Bandi and Phillips (2003, Ecta)

- Why nonparametric models: not subject to specification bias!
- First fully nonparametric estimation of SDF without stationarity
- Time homogenous diffusion nonparametric model

$$dX(t) = \mu(X(t))dt + \sigma(X(t))dB(t), \tag{9}$$

- Note that

$$E\left[dX(t) \mid X(t)\right] = \mu(X(t))dt, \tag{10}$$

$$E\left[(dX(t))^2 \mid X(t)\right] = \sigma^2(X(t))dt, \tag{11}$$

- Assume X_t is observed at equispaced points $(\Delta_{n,T}, 2\Delta_{n,T}, \cdots, n\Delta_{n,T})$ over the period $[0, T]$. So $n = T\Delta_{n,T}$. It is assumed $n \to \infty$, $T \to \infty$ and $\Delta_{n,T} = T/n \to 0$.

Continue

- The nonparametric estimator of $\mu(X(t))$ is

$$\frac{\sum_{i=1}^{n} K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,t}}\right)\tilde{\mu}_{n,T}(x)}{\sum_{i=1}^{n} K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,t}}\right)}$$

- The nonparametric estimator of $\sigma^2(X(t))$ is

$$\frac{\sum_{i=1}^{n} K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,t}}\right)\tilde{\sigma}^2_{n,T}(x)}{\sum_{i=1}^{n} K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,t}}\right)}$$

Continue

- The nonparametric estimator of $\mu(X(t))$ is

$$\frac{\sum_{i=1}^{n} K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,t}}\right)\tilde{\mu}_{n,T}(x)}{\sum_{i=1}^{n} K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,t}}\right)}$$

- The nonparametric estimator of $\sigma^2(X(t))$ is

$$\frac{\sum_{i=1}^{n} K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,t}}\right)\tilde{\sigma}^2_{n,T}(x)}{\sum_{i=1}^{n} K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,t}}\right)}$$

$$\tilde{\mu}_{n,T}(x) = \frac{1}{m_{n,T}(i\Delta_{n,T})\Delta_{n,T}} \sum_{j=0}^{m_{n,T}(i\Delta_{n,T})-1} \left[X_{t(i\Delta_{n,T})_j+\Delta_{n,T}} - X_{t(i\Delta_{n,T})_j}\right]$$

Continue

- The nonparametric estimator of $\mu(X(t))$ is

$$\frac{\sum_{i=1}^{n} K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,t}}\right)\tilde{\mu}_{n,T}(x)}{\sum_{i=1}^{n} K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,t}}\right)}$$

- The nonparametric estimator of $\sigma^2(X(t))$ is

$$\frac{\sum_{i=1}^{n} K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,t}}\right)\tilde{\sigma}^2_{n,T}(x)}{\sum_{i=1}^{n} K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,t}}\right)}$$

$$\tilde{\mu}_{n,T}(x) = \frac{1}{m_{n,T}(i\Delta_{n,T})\Delta_{n,T}} \sum_{j=0}^{m_{n,T}(i\Delta_{n,T})-1} [X_{t(i\Delta_{n,T})_j+\Delta_{n,T}} - X_{t(i\Delta_{n,T})_j}]$$

$$\tilde{\sigma}^2_{n,T}(x) = \frac{1}{m_{n,T}(i\Delta_{n,T})\Delta_{n,T}} \sum_{j=0}^{m_{n,T}(i\Delta_{n,T})-1} [X_{t(i\Delta_{n,T})_j+\Delta_{n,T}} - X_{t(i\Delta_{n,T})_j}]^2$$

Continue

- The nonparametric estimator of $\mu(X(t))$ is

$$\frac{\sum_{i=1}^{n} K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,t}}\right)\tilde{\mu}_{n,T}(x)}{\sum_{i=1}^{n} K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,t}}\right)}$$

- The nonparametric estimator of $\sigma^2(X(t))$ is

$$\frac{\sum_{i=1}^{n} K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,t}}\right)\tilde{\sigma}^2_{n,T}(x)}{\sum_{i=1}^{n} K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,t}}\right)}$$

$$\tilde{\mu}_{n,T}(x) = \frac{1}{m_{n,T}(i\Delta_{n,T})\Delta_{n,T}} \sum_{j=0}^{m_{n,T}(i\Delta_{n,T})-1} [X_{t(i\Delta_{n,T})_j+\Delta_{n,T}} - X_{t(i\Delta_{n,T})_j}]$$

$$\tilde{\sigma}^2_{n,T}(x) = \frac{1}{m_{n,T}(i\Delta_{n,T})\Delta_{n,T}} \sum_{j=0}^{m_{n,T}(i\Delta_{n,T})-1} [X_{t(i\Delta_{n,T})_j+\Delta_{n,T}} - X_{t(i\Delta_{n,T})_j}]^2$$

Continue

- Identification 1: aliasing is avoided by the double limit

Continue

- Identification 1: aliasing is avoided by the double limit
- Identification 2: to separately identify the drift and diffusion functions, the condition of recurrence is assumed.

Continue

- Identification 1: aliasing is avoided by the double limit
- Identification 2: to separately identify the drift and diffusion functions, the condition of recurrence is assumed.
- Application 1: Bandi (2004, JFE)

Continue

- Identification 1: aliasing is avoided by the double limit
- Identification 2: to separately identify the drift and diffusion functions, the condition of recurrence is assumed.
- Application 1: Bandi (2004, JFE)
- Application 2: Corradi and Disto (2007)

Continue

- Identification 1: aliasing is avoided by the double limit
- Identification 2: to separately identify the drift and diffusion functions, the condition of recurrence is assumed.
- Application 1: Bandi (2004, JFE)
- Application 2: Corradi and Disto (2007)
- Application 3: Jeffrey, Linton, Nguyen, Phillips (2004)

Method 2: Bandi and Phillips (2006, JoE)

- Estimate parametric models via nonparametric methods

$$dX(t) = \mu(X(t);\theta_1)dt + \sigma(X(t);\theta_2)dB(t), \tag{12}$$

Method 2: Bandi and Phillips (2006, JoE)

- Estimate parametric models via nonparametric methods

$$dX(t) = \mu(X(t);\theta_1)dt + \sigma(X(t);\theta_2)dB(t), \tag{12}$$

- The estimator of θ_1:

$$\min ||\tilde{\mu}_{n,T} - \mu(X(t);\theta_1)||$$

Method 2: Bandi and Phillips (2006, JoE)

- Estimate parametric models via nonparametric methods

$$dX(t) = \mu(X(t);\theta_1)dt + \sigma(X(t);\theta_2)dB(t), \qquad (12)$$

- The estimator of θ_1:

$$\min ||\tilde{\mu}_{n,T} - \mu(X(t);\theta_1)||$$

- The estimator of θ_2:

$$\min ||\tilde{\sigma}^2_{n,T} - \sigma^2(X(t);\theta_2)||$$

Method 3: Phillips and Yu, (JoE, forthcoming)

- Estimate parametric models via nonparametric and parametric methods

$$dX(t) = \mu(X(t);\theta_1)dt + \sigma(X(t);\theta_2)dB(t),$$

Method 3: Phillips and Yu, (JoE, forthcoming)

- Estimate parametric models via nonparametric and parametric methods

$$dX(t) = \mu(X(t);\theta_1)dt + \sigma(X(t);\theta_2)dB(t),$$

- If the diffusion term is known (i.e. $\sigma(X_t;\theta_2) = \sigma(X_t)$), the exact continuous record log-likelihood is

$$\ell_{IF}(\theta_1) = \int_0^T \frac{\mu(X_t;\theta_1)}{\sigma^2(X_t)} dX_t - \frac{1}{2}\int_0^T \frac{\mu^2(X_t;\theta_1)}{\sigma^2(X_t)} dt.$$

Method 3: Phillips and Yu, (JoE, forthcoming)

- Estimate parametric models via nonparametric and parametric methods

$$dX(t) = \mu(X(t);\theta_1)dt + \sigma(X(t);\theta_2)dB(t),$$

- If the diffusion term is known (i.e. $\sigma(X_t;\theta_2) = \sigma(X_t)$), the exact continuous record log-likelihood is

$$\ell_{IF}(\theta_1) = \int_0^T \frac{\mu(X_t;\theta_1)}{\sigma^2(X_t)} dX_t - \frac{1}{2}\int_0^T \frac{\mu^2(X_t;\theta_1)}{\sigma^2(X_t)} dt.$$

- On the other hand

$$[X]_T = \int_0^T (dX_t)^2 = \int_0^T \sigma^2(X_t;\theta_2)dt,$$

where $[X]_T$ is the quadratic variation of X, which can be consistently estimated by $[X_\Delta]_T$ as $\Delta \to 0$.

Method 4: Phillips and Yu (2006)

- Estimate integrated volatility using nonparametric methods

$$dp^*(t) = \sigma(t)dB(t), \tag{13}$$

where $p^*(t)$ is the equilibrium log-price and $\sigma(t)$ is a càdlàg volatility process. The quantity of interest is $IV = \int_0^1 \sigma^2(t)dt$.

Method 4: Phillips and Yu (2006)

- Estimate integrated volatility using nonparametric methods

$$dp^*(t) = \sigma(t)dB(t), \tag{13}$$

where $p^*(t)$ is the equilibrium log-price and $\sigma(t)$ is a càdlàg volatility process. The quantity of interest is $IV = \int_0^1 \sigma^2(t)dt$.

- Assume that $p^*(t)$ is observed at $\{t_{i,m} = \frac{i}{m} : i = 0, ..., m\}$ with $h = \frac{1}{m}$

Method 4: Phillips and Yu (2006)

- Estimate integrated volatility using nonparametric methods

$$dp^*(t) = \sigma(t)dB(t), \tag{13}$$

where $p^*(t)$ is the equilibrium log-price and $\sigma(t)$ is a càdlàg volatility process. The quantity of interest is $IV = \int_0^1 \sigma^2(t)dt$.

- Assume that $p^*(t)$ is observed at $\{t_{i,m} = \frac{i}{m} : i = 0, ..., m\}$ with $h = \frac{1}{m}$
- A nonparametric estimate is RV

$$\sum_{i=1}^{m} [p^*_{i,m} - p^*_{i-1,m}]^2 := RV^{(m)}(p^*).$$

Method 4: Phillips and Yu (2006)

- Estimate integrated volatility using nonparametric methods

$$dp^*(t) = \sigma(t)dB(t), \tag{13}$$

where $p^*(t)$ is the equilibrium log-price and $\sigma(t)$ is a càdlàg volatility process. The quantity of interest is $IV = \int_0^1 \sigma^2(t)dt$.

- Assume that $p^*(t)$ is observed at $\{t_{i,m} = \frac{i}{m} : i = 0, ..., m\}$ with $h = \frac{1}{m}$
- A nonparametric estimate is RV

$$\sum_{i=1}^{m} [p^*_{i,m} - p^*_{i-1,m}]^2 := RV^{(m)}(p^*).$$

- Barndorff-Nielsen and Shephard (2002) showed that

$$\sqrt{m}\left[RV^{(m)}(p^*) - IV\right] |\sigma^2(t) \xrightarrow{d} MN\left(0, 2\int_0^1 \sigma^4(t)dt\right). \tag{14}$$

Continue

- When m is large, Model (13) cannot describe the data well. A particularly important stylized fact in actual stock data ultra-high frequencies is flat trading.

Continue

- When m is large, Model (13) cannot describe the data well. A particularly important stylized fact in actual stock data ultra-high frequencies is flat trading.
- New model

$$p_{i,m} = \begin{cases} p^*_{i,m} & \text{if } \xi_i = 1 \\ p_{i-1,m} & \text{if } \xi_i = 0 \end{cases}, \tag{15}$$

where ξ_i is a Bernoulli sequence with $E(\xi_i = 1) = \pi$.

Continue

- When m is large, Model (13) cannot describe the data well. A particularly important stylized fact in actual stock data ultra-high frequencies is flat trading.
- New model

$$p_{i,m} = \begin{cases} p^*_{i,m} & \text{if } \xi_i = 1 \\ p_{i-1,m} & \text{if } \xi_i = 0 \end{cases}, \qquad (15)$$

where ξ_i is a Bernoulli sequence with $E(\xi_i = 1) = \pi$.

- New CLT

$$\sqrt{m}\left[RV^{(m)}(p) - IV\right] \xrightarrow{d} MN\left(0, \frac{4-2\pi}{\pi}\int_0^1 \sigma^4(t)dt\right), \qquad (16)$$

Outline

ML – General Results

- Parametric models with a linear drift

$$dX(t) = \kappa(\mu - X(t))dt + \sigma(X(t);\theta)dB(t), \tag{17}$$

ML – General Results

- Parametric models with a linear drift

$$dX(t) = \kappa(\mu - X(t))dt + \sigma(X(t);\theta)dB(t), \tag{17}$$

- Suppose $p(X_{ih}|X_{(i-1)h},\theta)$ is the transition probability density. The Markov property of model (12) implies

$$\ell_{TD}(\theta) = \ln(p(X_{ih}|X_{(i-1)h},\theta)).$$

ML – General Results

- Parametric models with a linear drift

$$dX(t) = \kappa(\mu - X(t))dt + \sigma(X(t);\theta)dB(t), \qquad (17)$$

- Suppose $p(X_{ih}|X_{(i-1)h},\theta)$ is the transition probability density. The Markov property of model (12) implies

$$\ell_{TD}(\theta) = \ln(p(X_{ih}|X_{(i-1)h},\theta)).$$

- MLE has good asymptotic properties.

ML – General Results

- Parametric models with a linear drift

$$dX(t) = \kappa(\mu - X(t))dt + \sigma(X(t);\theta)dB(t), \qquad (17)$$

- Suppose $p(X_{ih}|X_{(i-1)h},\theta)$ is the transition probability density. The Markov property of model (12) implies

$$\ell_{TD}(\theta) = \ln(p(X_{ih}|X_{(i-1)h},\theta)).$$

- MLE has good asymptotic properties.
- Plug-in estimates are themselves MLEs

Continue

- Only in rare cases (BM, Vasicek, Square root), do the transition density $\ln(p(X_{ih}|X_{(i-1)h},\theta))$ have analytical expressions.

Continue

- Only in rare cases (BM, Vasicek, Square root), do the transition density $\ln(p(X_{ih}|X_{(i-1)h},\theta))$ have analytical expressions.
- Euler scheme

$$X_{ih} = (1-\kappa h)X_{(i-1)h} + \kappa\mu + \sigma(X_{(i-1)h},\theta)\sqrt{h}\epsilon_i. \tag{18}$$

Continue

- Only in rare cases (BM, Vasicek, Square root), do the transition density $\ln(p(X_{ih}|X_{(i-1)h},\theta))$ have analytical expressions.
- Euler scheme

$$X_{ih} = (1-\kappa h)X_{(i-1)h} + \kappa\mu + \sigma(X_{(i-1)h},\theta)\sqrt{h}\epsilon_i. \tag{18}$$

- Discretization bias in Euler: approximate $e^{-\kappa h}$ by $1-\kappa h$.

Continue

- Only in rare cases (BM, Vasicek, Square root), do the transition density $\ln(p(X_{ih}|X_{(i-1)h},\theta))$ have analytical expressions.
- Euler scheme

$$X_{ih} = (1-\kappa h)X_{(i-1)h} + \kappa\mu + \sigma(X_{(i-1)h},\theta)\sqrt{h}\epsilon_i. \tag{18}$$

- Discretization bias in Euler: approximate $e^{-\kappa h}$ by $1-\kappa h$.
- Refinements of the Euler scheme are possible (Bergstrom, Milstein, Kessler, Nowman, Durham and Gallant).

Continue

- Only in rare cases (BM, Vasicek, Square root), do the transition density $\ln(p(X_{ih}|X_{(i-1)h},\theta))$ have analytical expressions.
- Euler scheme

$$X_{ih} = (1-\kappa h)X_{(i-1)h} + \kappa\mu + \sigma(X_{(i-1)h},\theta)\sqrt{h}\epsilon_i. \tag{18}$$

- Discretization bias in Euler: approximate $e^{-\kappa h}$ by $1-\kappa h$.
- Refinements of the Euler scheme are possible (Bergstrom, Milstein, Kessler, Nowman, Durham and Gallant).
- Closed-form approximations of transition density (Aït-Sahalia, 2002, Aït-Sahalia and Yu, 2006).

Finite sample problems: Phillips and Yu, 2005, RFS

- Finite sample estimation bias is substantial, often much more serious than the discretization bias or specification bias.

Finite sample problems: Phillips and Yu, 2005, RFS

- Finite sample estimation bias is substantial, often much more serious than the discretization bias or specification bias.
- Square root model

$$dX(t) = \kappa(\mu - X(t))dt + \sigma\sqrt{X(t)}dB(t)$$

Finite sample problems: Phillips and Yu, 2005, RFS

- Finite sample estimation bias is substantial, often much more serious than the discretization bias or specification bias.
- Square root model

$$dX(t) = \kappa(\mu - X(t))dt + \sigma\sqrt{X(t)}dB(t)$$

- Assume $\kappa = 0.1, \mu = 0.1, \sigma = 0.1$. Simulate 1000 samples, each with 120 monthly observations (ie $h = 1/12$).

Finite sample problems: Phillips and Yu, 2005, RFS

- Finite sample estimation bias is substantial, often much more serious than the discretization bias or specification bias.
- Square root model

$$dX(t) = \kappa(\mu - X(t))dt + \sigma\sqrt{X(t)}dB(t)$$

- Assume $\kappa = 0.1, \mu = 0.1, \sigma = 0.1$. Simulate 1000 samples, each with 120 monthly observations (ie $h = 1/12$).

Exact and Approximate ML Estimation						
True Value $\kappa = 0.1$						
	Exact	Euler	Milst	Nowman	Infill	Herm
Mean	.2403	.2419	.2444	.2386	.2419	.2400
Std	.2777	.2867	.2867	.2771	.2867	.2762
RMSE	.3112	.3199	.3210	.3098	.3199	.3096
Mean of diff	NA	.0016	.0041	-.0017	.0016	**-.0003**
Std of diff	NA	.0500	.0453	.0162	.0500	**.0043**

Continue

- Finite sample estimation bias translates into bias in pricing

Continue

- Finite sample estimation bias translates into bias in pricing
- 1000 samples, each with 600 monthly observations (ie $\Delta = 1/12$), are simulated from the square root model with $(\kappa, \mu, \sigma) = (0.1, 0.08, 0.02)$.

Continue

- Finite sample estimation bias translates into bias in pricing
- 1000 samples, each with 600 monthly observations (ie $\Delta = 1/12$), are simulated from the square root model with $(\kappa, \mu, \sigma) = (0.1, 0.08, 0.02)$.

Parameter		κ	Bond Price	Option Price
True Value		**0.1**	**0.8503**	**2.3920**
Exact ML of CIR	Mean	0.1845	0.8438	1.8085
	RMSE	0.1319	0.0103	0.9052
Euler ML of CIR	Mean	0.1905	0.8433	1.7693
	RMSE	0.1397	0.0111	0.9668
ML of Vasicek (misspecified)	Mean	0.1746	0.8444	1.8837
	RMSE	0.1175	0.0088	0.7637

Continue

- Finite sample estimation bias translates into bias in pricing
- 1000 samples, each with 600 monthly observations (ie $\Delta = 1/12$), are simulated from the square root model with $(\kappa, \mu, \sigma) = (0.1, 0.08, 0.02)$.

Parameter		κ	Bond Price	Option Price
True Value		**0.1**	**0.8503**	**2.3920**
Exact ML of CIR	Mean	0.1845	0.8438	1.8085
	RMSE	0.1319	0.0103	0.9052
Euler ML of CIR	Mean	0.1905	0.8433	1.7693
	RMSE	0.1397	0.0111	0.9668
ML of Vasicek (misspecified)	Mean	0.1746	0.8444	1.8837
	RMSE	0.1175	0.0088	0.7637

- Bias from nonlinear plug-in

Continue

- First source of bias – estimation bias from dynamic models.

Continue

- First source of bias – estimation bias from dynamic models.
- Let $\rho = e^{-\kappa h} \approx 1 - \kappa h$.

$$\hat{\rho} - \rho \approx \frac{1 + 3\rho}{n}.$$

so

$$\hat{\kappa} - \kappa \approx -\frac{1 + 3\rho}{nh} = -\frac{1 + 3\rho}{T}.$$

Continue

- First source of bias – estimation bias from dynamic models.
- Let $\rho = e^{-\kappa h} \approx 1 - \kappa h$.

$$\hat{\rho} - \rho \approx \frac{1 + 3\rho}{n}.$$

so

$$\hat{\kappa} - \kappa \approx -\frac{1 + 3\rho}{nh} = -\frac{1 + 3\rho}{T}.$$

- Second source of bias – nonlinearity in pricing

Jackknife, Phillips and Yu, 2005, RFS

- Let N be the number of observations in the whole sample and decompose the sample into m consecutive subsamples each with ℓ observations, so that $N = m \times \ell$. Phillips and Yu (2005a) proposed the following jackknife estimator of κ, estimator

$$\hat{\kappa}_{jack} = \frac{m}{m-1}\hat{\kappa}_N - \frac{\sum_{i=1}^{m} \hat{\kappa}_{li}}{m^2 - m}, \tag{19}$$

Continue

- To reduce the bias in pricing, jackknife prices directly

$$\hat{P}_{jack} = 2\hat{P}_T - \frac{1}{2}(\hat{P}_{1,T/2} + \hat{P}_{2,T/2}).$$

Continue

- To reduce the bias in pricing, jackknife prices directly

$$\hat{P}_{jack} = 2\hat{P}_T - \frac{1}{2}(\hat{P}_{1,T/2} + \hat{P}_{2,T/2}).$$

- The bias in the jackknife estimate $\hat{P}_{jack}$ is of order $O(T^{-2})$ rather than $O(T^{-1})$.

Continue

- To reduce the bias in pricing, jackknife prices directly

$$\hat{P}_{jack} = 2\hat{P}_T - \frac{1}{2}(\hat{P}_{1,T/2} + \hat{P}_{2,T/2}).$$

- The bias in the jackknife estimate $\hat{P}_{jack}$ is of order $O(T^{-2})$ rather than $O(T^{-1})$.
- The jackknife almost completely removed the bias in $\hat{\theta}_T^{ML}$.

Continue

- To reduce the bias in pricing, jackknife prices directly

$$\hat{P}_{jack} = 2\hat{P}_T - \frac{1}{2}(\hat{P}_{1,T/2} + \hat{P}_{2,T/2}).$$

- The bias in the jackknife estimate $\hat{P}_{jack}$ is of order $O(T^{-2})$ rather than $O(T^{-1})$.
- The jackknife almost completely removed the bias in $\hat{\theta}_T^{ML}$.
- The jackknife reduced the bias in option prices in ML down to 10-15%.

Simulation Methods, Phillips and Yu, 2007

- Any bias that occurs in parameter estimation with actual data, will also be present with simulated data.

Simulation Methods, Phillips and Yu, 2007

- Any bias that occurs in parameter estimation with actual data, will also be present with simulated data.
- Similarly, any bias that occurs in contingent-claim price estimation with actual data, will also be present with simulated data.

Simulation Methods, Phillips and Yu, 2007

- Any bias that occurs in parameter estimation with actual data, will also be present with simulated data.
- Similarly, any bias that occurs in contingent-claim price estimation with actual data, will also be present with simulated data.
- Bias can arise from estimation of dynamic model, nonlinear transformation and discretization. Hence the form of the bias can be complicated.

Simulation Methods, Phillips and Yu, 2007

- Any bias that occurs in parameter estimation with actual data, will also be present with simulated data.
- Similarly, any bias that occurs in contingent-claim price estimation with actual data, will also be present with simulated data.
- Bias can arise from estimation of dynamic model, nonlinear transformation and discretization. Hence the form of the bias can be complicated.
- Bias is computed by simulation.

Outline

Other Contributions

1. Model with time-varying variances. References: Phillips and Xu (2006, 2007).

Other Contributions

1. Model with time-varying variances. References: Phillips and Xu (2006, 2007).
2. Tail behavior in stock return distributions. Reference: McFarland, McMahon and Phillips (1996).

Other Contributions

1. Model with time-varying variances. References: Phillips and Xu (2006, 2007).
2. Tail behavior in stock return distributions. Reference: McFarland, McMahon and Phillips (1996).
3. Efficiency in ForEx markets. Reference: McFarland and Phillips (1997).

Other Contributions

1. Model with time-varying variances. References: Phillips and Xu (2006, 2007).
2. Tail behavior in stock return distributions. Reference: McFarland, McMahon and Phillips (1996).
3. Efficiency in ForEx markets. Reference: McFarland and Phillips (1997).
4. Federal funds target rates. Reference: Hu and Phillips (2004).

Outline

References

- Bandi, F.M. and P.C.B. Phillips, 2003, Fully Nonparametric Estimation of Scalar Diffusion Models, *Econometrica*, 71, 241-283.
- Bandi, F.M. and P.C.B. Phillips, 2007, A simple approach to the parametric estimation of potentially nonstationary diffusions. *Journal of Econometrics*, 137, 354-395.
- Holly, A. and P.C.B. Phillips, 1979, A Saddlepoint Approximation to the Distribution of the k-Class Estimator of a Coefficient in a Simultaneous System, *Econometrica*, 47, 1527-1548.
- Hu, L and P.C.B. Phillips,, 2004, Dynamics of the Federal Funds Target Rate: A Nonstationary Discrete Choice Approach, *Journal of Applied Econometrics*, 19, 851-867.

References: Continue

- McFarland, J., P. C. McMahon, and P.C.B. Phillips, 1996, Robust Tests of Forward Exchange Market Efficiency with Empirical Evidence from the 1920's, *Journal of Applied Econometrics*, 11, 1 - 23.
- McFarland, J., and P.C.B. Phillips, 1997, Forward Exchange Market Unbiasedness: the Case of the Australian Dollar since 1984, *Journal of International Money and Finance*, 16, 885-907.
- Phillips, P.C.B., 1972, The Structural Estimation of a Stochastic Differential Equation System, *Econometrica*, 40, 1021-1041.
- Phillips, P.C.B., 1973, The Problem of Identification in Finite Parameter Continuous Time Models, *Journal of Econometrics*, 1, 351-262.
- Phillips, P.C.B., 1974, The Estimation of Some Continuous Time Models, *Econometrica*, 42, 803-823.

References: Continue

- Phillips, P.C.B., 1977, Approximations to Some Finite Sample Distributions Associated with a First Order Stochastic Difference Equation, *Econometrica*, 45, 463-485.
- Phillips, P.C.B., 1978, Edgeworth and Saddlepoint Approximations in a First Order Non-Circular Autoregression, *Biometrika*, 65, 91-98.
- Phillips, P.C.B., 1988, Regression Theory for Near-Integrated Time Series, *Econometrica*, 56, 1021-1044.
- Phillips, P.C.B., 1991, Error Correction and Long Run Equilibria in Continuous Time, *Econometrica*, 967-980.
- Phillips, P.C.B. and J. Yu, 2005, Jackknifing bond option prices. *Review of Financial Studies*.
- Phillips, P.C.B. and J. Yu, 2005b, Comments: A selective overview of nonparametric methods in financial econometrics. *Statistical Science*, 20, 338–343.

References: Continue

- Phillips, P.C.B. and J. Yu, 2007, Maximum Likelihood and Gaussian Estimation of Continuous Time Models in Finance. *Handbook of Financial Time Series*, forthcoming.
- Phillips, P.C.B. and J. Yu, 2007, A Two-Stage Realized Volatility Approach to Estimation of Diffusion Processes with Discrete Data. *Journal of Econometrics*, forthcoming.
- Yu, J., and P. C. B. Phillips, 2001, A Gaussion Approach for Estimating Continuous Time Models of Short Term Interest Rates, *Econometrics Journal*, 4, 211-225.