Sieve bootstrap inference for time-varying coefficient models

Marina Friedrich, Yicong Lin

EcoDep Seminary, Cergy, 18.05.2022

Vrije Universiteit Amsterdam, Tinbergen Institute

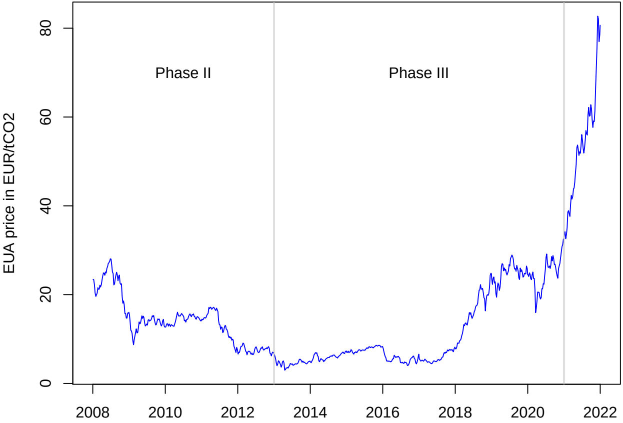
Outline

- Empirical application
- (Econometric) Motivation
- The model and estimation technique
- The sieve bootstrap and confidence bands
- Asymptotic theory
- Simulation results

Empirical application: The EU Emission Trading System (ETS)

- one of the largest and long lived cap-and-trade systems
- 31 countries (EU plus Iceland, Norway and Liechtenstein), covers around 40% of EU's GHG emissions
- from power sector and energy intensive industries (10,000 firms) as well as civil aviation (500 operators)
- started operating in 2005
  - ✓ Phase I: 2005-2007
  - ✓ Phase II: 2008-2012
  - ✓ Phase III: 2013-2020
  - ✓ Phase IV: 2021-2030
- current goal: 43% emission reduction by 2030 (vs. 2005)
- linear reduction factor increased from 1.74% to 2.2%

Development of allowance prices

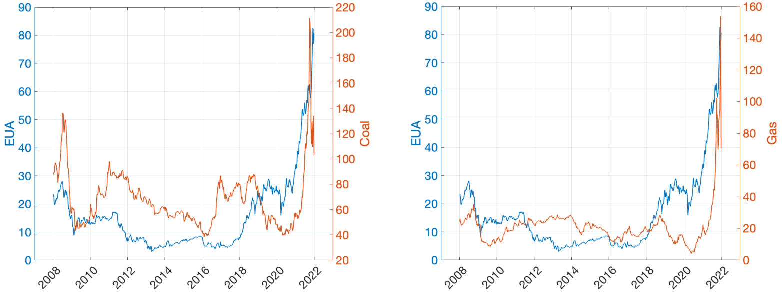


EU ETS empirical literature

- **Theory**: among most important price drivers are the *coal price (-)* and the *gas price (+)*. Also, economic activity *(+)*, weather variables *(+)* and renewables *(-)*
- **Practice**: hard to find empirical evidence
- Previous studies faced challenges:
  - ✓ insignificant and/or wrong sign of coefficients (for coal)
  - ✓ they split the sample in parts, take out outliers related to certain events or allow for different pricing regimes
  - ✓ high sensitivity to choice of data series
  - ✓ cointegration analysis shows little to no evidence of a stable long-run relationship

This paper

- A possible explanation for previous findings might be an unstable relationship between the allowance price and its fundamental drivers.
- We look at the relationship in a time-varying regression approach.
- We find evidence of time variation in the coefficients.



(a) EUA and coal

(b) EUA and gas

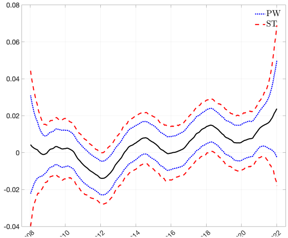
Figure 1: Joint plot of (a) EUA Dec futures and coal prices and (b) EUA Dec futures and gas prices (2008-2018). Source: ICE via Nasdaq Data Link and FactSet

- Weekly data from Jan. 2008 to Dec. 2021
- EUA prices, coal, gas and oil prices, STOXX 50, temperature
- Most series contain a unit root $\rightarrow$ log returns

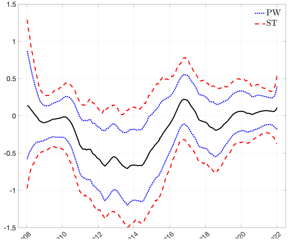
	(a)			(b)		
	$\hat{\beta}$	se_{HAC}	p-value	$\hat{\beta}$	se_{HAC}	p-value
Intercept	0.004	0.002	0.041	0.004	0.002	0.041
Coal	-0.072	0.047	0.126	-0.038	0.047	0.412
Gas	0.113	0.036	0.002	0.115	0.035	0.001
Oil	0.176	0.049	0.000			
STOXX 50				0.277	0.079	0.000
Temperature	0.000	0.001	0.687	0.000	0.001	0.830
Constancy test		reject***			reject***	

The results of linear time-constant models by OLS regressions. Dependent variable: EUA log returns. HAC-robust standard error estimates se_{HAC} (with Bartlett kernel and the bandwidth selected by Andrews (1991) using AR(1) model, see p.835)

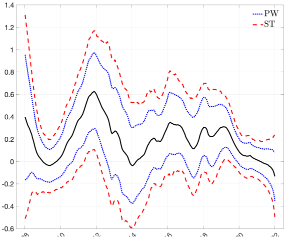
Time-varying coefficient estimates



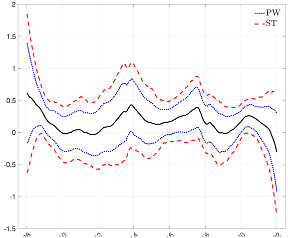
(a) Trend



(b) Coal



(c) Gas



(d) Oil

(Econometric) Motivation and contribution

Bootstrap is a frequently used tool in the context of time-varying coefficient models. It is used to study e.g.

- price elasticities for energy demand (Liddle et al., 2020),
- income inequality and CO_2 emissions (Uddin et al., 2020),
- the environmental Kuznets curve (Churchill et al., 2020).

Advantages of bootstrapping:

- bootstrap is easy to implement
- it has good small sample properties (simulation study)
- estimation of nuisance parameters is not necessary

(Econometric) Motivation and contribution

Bootstrap validity has not been provided for these types of models.

Our theoretical contribution includes:

- We establish validity of the sieve bootstrap to construct confidence intervals in the context of linear time-varying coefficient regression models.
- We provide both pointwise and simultaneous results.
- (We develop a bootstrap-based test for parameter constancy.)

Literature on simultaneous confidence bands (SCBs)

- Bühlmann (1998):
  - ✓ sieve bootstrap method for deterministic trend model
- Neumann and Polzehl (1998):
  - ✓ wild bootstrap method for deterministic trend model
- Wu and Zhao (2007):
  - ✓ derive SCBs for trend model using asymptotic theory
- Friedrich, Smeekes and Urbain (2020):
  - ✓ autoregressive wild bootstrap for deterministic trend model

- Zhou and Wu (2010):
  - ✓ derive SCBs based on asymptotic theory for the more general model (with regressors)

Model and estimation

Consider the following linear time-varying coefficient model:

$$y_t = \boldsymbol{\beta}_t' \mathbf{x}_t + z_t = \beta_{0,t} + \sum_{j=1}^{d} \beta_{j,t} x_{j,t} + z_t, \qquad t = 1 \ldots, n,$$

where:

- $\boldsymbol{\beta}_t$ is a $(d+1)$-dimensional vector of time-varying coefficients, with $\boldsymbol{\beta}_t := \boldsymbol{\beta}(t/n)$ being a vector of smooth functions,
- $\mathbf{x}_t$ a vector of covariates; $\{z_t, \mathbf{x}_t\}$ a stationary process
- $z_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$ with $\psi_0 = 1$, where $\{\varepsilon_t\}$ is i.i.d.

We use a local linear Kernel estimator:

- originally proposed for this model by Cai (2007)
- does not suffer from boundary effects

The local linear estimator

The estimator is based on the following approximation:

$$y_t \approx \boldsymbol{\beta}(\tau)'\boldsymbol{x}_t + \boldsymbol{\beta}^{(1)}(\tau)'\boldsymbol{x}_t(t/n - \tau) + z_t =: \boldsymbol{\theta}(\tau)'\tilde{\boldsymbol{x}}_t(\tau) + z_t.$$

The local linear estimator minimizes the following weighted sum of squares:

$$\widehat{\boldsymbol{\theta}}(\tau) = \begin{pmatrix} \widehat{\boldsymbol{\beta}}(\tau) \\ \widehat{\boldsymbol{\beta}}^{(1)}(\tau) \end{pmatrix} = \arg\min_{\boldsymbol{\theta}} \sum_{t=1}^{n} \left(y_t - \tilde{\boldsymbol{x}}_t(\tau)'\boldsymbol{\theta}\right)^2 K\left(\frac{t/n - \tau}{h}\right),$$

where $K(\cdot)$ is a kernel function and $h > 0$ is a bandwidth.

The local linear estimator

Local linear estimator $\widehat{\boldsymbol{\theta}}(\tau) = \left(\widehat{\boldsymbol{\beta}}(\tau) \quad \widehat{\boldsymbol{\beta}}^{(1)}(\tau)\right)'$

$$\widehat{\boldsymbol{\theta}}(\tau) = \begin{pmatrix} \mathbf{S}_{n,0}(\tau) & \mathbf{S'}_{n,1}(\tau) \\ \mathbf{S}_{n,1}(\tau) & \mathbf{S}_{n,2}(\tau) \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{T}_{n,0}(\tau) \\ \mathbf{T}_{n,1}(\tau) \end{pmatrix} \equiv \mathbf{S}_n^{-1}(\tau)\mathbf{T}_n(\tau),$$

where for $k = 0, ..., 2$:

$$\mathbf{S}_{n,k}(\tau) = \frac{1}{n}\sum_{t=1}^{T} \mathbf{x}_t\mathbf{x}_t' \, (t/n - \tau)^k \, K\left(\frac{t/n - \tau}{h}\right)$$

$$\mathbf{T}_{n,k}(\tau) = \frac{1}{n}\sum_{t=1}^{T} \mathbf{x}_t \, (t/n - \tau)^k \, K\left(\frac{t/n - \tau}{h}\right) y_t$$

Bandwidth selection

- The bandwidth parameter h plays an important role as it determines the smoothness of the estimator.
- We investigate four different data driven selection methods.
- The criteria make use of the hat-matrix Q_h which satisfies $\hat{\mathbf{y}} = Q_h \mathbf{y}$, where $\mathbf{y} = (y_1, \ldots, y_n)'$ and $\hat{\mathbf{y}} = (\hat{y}_1, \ldots, \hat{y}_n)'$.

(1) Generalized Cross Validation (Zhou and Wu, 2010):

$$GCV(h) = \frac{n^{-1} \left| \mathbf{y} - \hat{\mathbf{y}} \right|^2}{\left(1 - \mathrm{tr}\left\{Q_h\right\}/n\right)^2}$$

Bandwidth selection (cont.)

(2) Akaike Information Criterion based bandwidth selection (Cai, 2007):

$$AIC(h) = \log \hat{\sigma}^2 + 2\frac{\operatorname{tr}\{Q_h\} + 1}{n - \operatorname{tr}\{Q_h\} - 2}$$

(3) Local Modified Cross Validation (adapted from Vieu, 1991):

$$\hat{h}_\tau = \arg\min_h \mathrm{LMCV}_\tau(h)$$

$$\mathrm{LMCV}_\tau(h) = \frac{1}{n}\sum_{t=1}^{n}\left[y_t - \boldsymbol{x}_t'\widehat{\boldsymbol{\beta}}^{l,h}(t/n)\right]^2 w_\tau\left(t/n\right),$$

where $w_\tau\left(\cdot\right)$ is a weight function and $\widehat{\boldsymbol{\beta}}^{l,h}(\tau)$ is the leave-$(2l+1)$-out estimator constructed by omitting the observations $\lfloor \tau n \rfloor + i$, $-l \leq i \leq l$.

Sieve bootstrap algorithm (1/2)

1. Calculate the residual series:

$$\hat{z}_t = y_t - \mathbf{x}_t' \widetilde{\boldsymbol{\beta}}(t/n), \qquad t = 1, \ldots, n,$$

   where $\widetilde{\boldsymbol{\beta}}(t/n)$ uses oversmoothing bandwidth $\tilde{h} > h$.

2. Fit an autoregressive model of order p to residuals:

$$\widehat{\varepsilon}_{t,p} = \hat{z}_t - \sum_{j=1}^{p} \widehat{\phi}_j \hat{z}_{t-j}, \qquad t = p+1, ..., n,$$

   Obtain recentered residuals $\{\tilde{\varepsilon}_{t,p}\}$.

3. Draw randomly with replacement from $\{\tilde{\varepsilon}_{t,p}\}$ to obtain $\{\varepsilon_t^*\}$.
4. Calculate the bootstrap errors $z_t^* = \sum_{j=1}^{p} \widehat{\phi}_j z_{t-j}^* + \varepsilon_t^*$ and generate

$$y_t^* = \mathbf{x}_t' \widetilde{\boldsymbol{\beta}}(t/n) + z_t^*, \qquad t = 1, \ldots, n.$$

Sieve bootstrap algorithm (2/2)

5. Obtain the bootstrap estimator $\hat{\boldsymbol{\beta}}^*(\cdot)$ using the same bandwidth h as used for the original estimate $\hat{\boldsymbol{\beta}}(\cdot)$.

6. Repeat Steps 3 to 5 B times, and let

$$\hat{q}_{j,\alpha}(\tau) = \inf\left\{u \in \mathbb{R} : \mathbb{P}^*\left(\hat{\beta}_j^*(\tau) - \tilde{\beta}_j(\tau) \leq u\right) \geq \alpha\right\}$$

   denote, for $j = 0, \ldots, d$, the bootstrap quantiles.

The bootstrap quantiles can directly be used to construct pointwise confidence intervals.

We additionally propose a procedure to construct simultaneous confidence bands.

Confidence intervals

- Using these quantiles, pointwise two-sided confidence intervals for a confidence level of $(1-\alpha)$ can be constructed as

$$I^*_{j,n,\alpha}(\tau) = \left[\hat{\beta}_j(\tau) - \hat{q}_{j,1-\alpha/2}(\tau), \hat{\beta}_j(\tau) - \hat{q}_{j,\alpha/2}(\tau)\right].$$

- These confidence intervals are only valid for a certain point in time. They should satisfy, for $j = 0, \ldots, d$:

$$\liminf_{n\to\infty} \mathbb{P}\left[\left(\beta_j(\tau) \in I^{(p)}_{j,n,\alpha}(\tau)\right)\right] \geq 1-\alpha \qquad \tau \in G.$$

- For an arbitrary set of time points G, simultaneous bands should satisfy, for $j = 0, \ldots, d$:

$$\liminf_{n\to\infty} \left[\mathbb{P}\left(\beta_j(\tau) \in I^{G}_{j,n,\alpha}(\tau) \quad \forall \tau \in G\right)\right] \geq 1-\alpha.$$

Construction of SCBs over a set G

For all $j = 0, \ldots, d$:

Step 1 For varying $\alpha_p \in [1/B, \alpha]$, obtain pointwise quantiles $\hat{q}_{j,\alpha_p/2}(\tau)$ and $\hat{q}_{j,1-\alpha_p/2}(\tau)$ for all $\tau \in G$.

Step 2 Choose α_s as

$$\alpha_s = \underset{\alpha_p \in [1/B, \alpha]}{\arg\min} \left| \mathbb{P}^* \left[\hat{q}_{j,\alpha_p/2}(\tau) \leq \hat{\beta}_j^*(\tau) - \tilde{\beta}_j(\tau) \leq \hat{q}_{j,1-\alpha_p/2}(\tau) \right. \right. \\ \left. \left. \forall \tau \in G \right] - (1-\alpha) \right|.$$

Step 3 Given α_s from Step 2, construct the simultaneous confidence bands as

$$I_{j,n,\alpha}^G(\tau) = \left[\hat{\beta}_j(\tau) - \hat{q}_{j,1-\alpha_s/2}(\tau), \hat{\beta}_j(\tau) - \hat{q}_{j,\alpha_s/2}(\tau) \right] \qquad \tau \in G.$$

Assumption A1

Suppose $\{(z_t, \boldsymbol{x}_t)\}_{t\in\mathbb{Z}}$ *is a strictly stationary, mixing process satisfying the following:*

(a) $\{\boldsymbol{x}_t\}_{t\in\mathbb{Z}}$ *is a strictly stationary* α*-mixing process with* $\alpha(m) = O(m^{-\varphi})$, $\varphi = \max\{(2+\delta)(1+\delta)/\delta, 3(1+\delta)/\delta\}$ *and* $\mathrm{E}\,\|\boldsymbol{x}_t\|^{2(2+\delta)} < \infty$, $\delta > 0$.

(b) *Assume* $z_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$ *with* $\psi_0 = 1$, *where* $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ *is an* $i.i.d.$ *sequence of continuous variables with* $\mathrm{E}(\varepsilon_t) = 0$ *and* $\mathrm{E}\,|\varepsilon_t|^{2(2+\delta)} < \infty$.

(c) *The density function* f_ε *of* ε_t *satisfies* $\int_{x\in\mathbb{R}} |f_\varepsilon(x+a) - f_\varepsilon(x)|\, dx \le M|a|$, $M < \infty$, *whenever* $|a| \le \tau$ *for some* $\tau > 0$.

(d) *The lag polynomial function* $\Psi : z \mapsto \sum_{j=0}^{\infty} \psi_j z^j$, $z \in \mathbb{C}$, *is bounded away from zero for* $|z| \le 1$. *Moreover, for some* $\lambda > (5\varphi + 7)/4$, $|\psi_j| \ll C_1 j^{-\lambda}$ *as* $j \to \infty$.

(e) *The following moment conditions hold almost surely:* (i) $\mathrm{E}\,(z_t|\boldsymbol{x}_t) = 0$; (ii) $\mathrm{E}\,(z_t z_s|\boldsymbol{x}_t \boldsymbol{x}_s') = \mathrm{E}\,(z_t z_s)$, $s, t \in \mathbb{Z}$.

Assumptions

A2 *The coefficient functions* $\beta_j(\cdot) \in \mathcal{C}^3[0,1]$, $j = 0, 1, \ldots, d$.

A3 *The kernel function* $K(\cdot)$ *is positive, symmetric, Lipschitz continuous, and has compact support* $[-1, 1]$ *with* $\mu_0 \equiv \int_{-1}^{1} K(u)\, du = 1$.

A4 *The bandwidth* $h \equiv h(n)$ *satisfies* $\max\left\{h, \frac{\ln n}{nh}, \frac{1}{nh^2}, nh^7, h^4 \ln n\right\} \overset{n\to\infty}{\to} 0$.

B1 *The oversmoothing bandwidth* $\tilde{h} = \tilde{h}(n)$ *satisfies* $\max\left\{\tilde{h}, nh\tilde{h}^4, h \ln n / \tilde{h}\right\} \to 0$ *as* $n \to \infty$.

B2 *The lag order* $p = p(n) \to \infty$ *with* $p \max\left\{\tilde{h}, \left(\ln n/(n\tilde{h})\right)^{1/4}\right\} \to 0$ *as* $n \to \infty$.

Theorem 1 - Bootstrap validity

Theorem

Under Assumptions A1, A2, A3, A4, B1 and B2, we have for any fixed $\tau \in (0,1)$,

$$\sqrt{nh}\boldsymbol{H}\left(\widehat{\boldsymbol{\theta}}^*(\tau) - \widetilde{\boldsymbol{\theta}}(\tau) - h^2\boldsymbol{b}(\tau)\right)$$
$$\xrightarrow{d^*}_p \mathcal{N}\left(\boldsymbol{0}, \begin{pmatrix} \nu_0 & \\ & \nu_2/\mu_2^2 \end{pmatrix} \otimes \left(\boldsymbol{\Omega}_0^{-1}\boldsymbol{\Lambda}\boldsymbol{\Omega}_0^{-1}\right)\right),$$

and $\sqrt{nh}\boldsymbol{H}\left(\widehat{\boldsymbol{\theta}}(\tau) - \boldsymbol{\theta}(\tau) - h^2\boldsymbol{b}(\tau)\right)$ *converges in distribution to the same limit.*

$$\mu_k = \int_{-1}^{1} u^k K(u)du, \qquad \nu_k = \int_{-1}^{1} u^k K^2(u)du$$

$$\boldsymbol{b}(\tau) = \frac{1}{2}\begin{pmatrix} \mu_2\boldsymbol{\beta}^{(2)}(\tau) \\ \boldsymbol{0} \end{pmatrix} \qquad \boldsymbol{H} = \operatorname{diag}\left(\boldsymbol{I}_{d+1}, h\boldsymbol{I}_{d+1}\right)$$

$$\boldsymbol{\Omega}_0 = \mathrm{E}\left(\boldsymbol{x}_t\boldsymbol{x}_t'\right) \qquad \boldsymbol{\Lambda} = \sum_{j=-\infty}^{\infty} \operatorname{cov}\left(\boldsymbol{x}_t z_t, \boldsymbol{x}_{t+j} z_{t+j}\right)$$

Theorem 2: Uniform Validity

Theorem

Under the assumptions in Theorem 1, for any fixed $\tau_0 \in (0, 1)$*,*

$$\left\{\sqrt{nh}\boldsymbol{H}\left(\widehat{\boldsymbol{\theta}}(\tau_0+\tau h)-\boldsymbol{\theta}(\tau_0+\tau h)-h^2\boldsymbol{b}(\tau_0)\right)\right\}_{\tau\in[-1,1]} \xrightarrow{w} \left\{\boldsymbol{W}(\tau)\right\}_{\tau\in[-1,1]},$$

$$\left\{\sqrt{nh}\boldsymbol{H}\left(\widehat{\boldsymbol{\theta}}^*(\tau_0+\tau h)-\widetilde{\boldsymbol{\theta}}(\tau_0+\tau h)-h^2\boldsymbol{b}(\tau_0)\right)\right\}_{\tau\in[-1,1]} \xrightarrow{w} \left\{\boldsymbol{W}(\tau)\right\}_{\tau\in[-1,1]} \text{ in probability,}$$

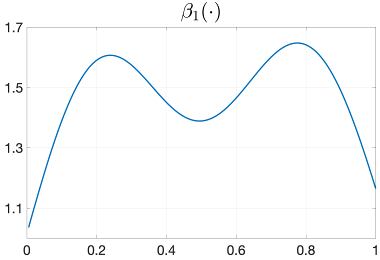
where $\{\boldsymbol{W}(\tau)\}_{\tau\in[-1,1]}$ *is a multivariate Gaussian process with* $\mathrm{E}\,\boldsymbol{W}(\tau)=\boldsymbol{0}$ *and*

$$\operatorname{cov}\left(\boldsymbol{W}(\tau_1),\boldsymbol{W}(\tau_2)\right) = \left[\operatorname{diag}\left(1,\mu_2^{-1}\right)\boldsymbol{\kappa}(\tau_1,\tau_2)\operatorname{diag}\left(1,\mu_2^{-1}\right)\right]\otimes\left(\boldsymbol{\Omega}_0^{-1}\boldsymbol{\Lambda}\boldsymbol{\Omega}_0^{-1}\right).$$

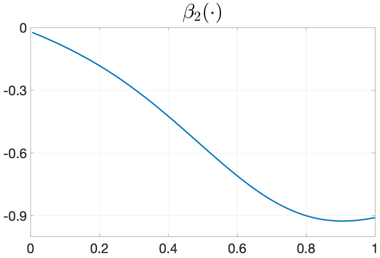
Simulation study - Data generating process

Inspired by the empirical application, consider the following DGP:

$$y_t = \beta_1(t/n)x_{1,t} + \beta_2(t/n)x_{2,t} + u_t$$



(a) $\beta_1(\cdot)$



(b) $\beta_2(\cdot)$

Data generating process (cont.)

- The two coefficient functions are given by

$$\beta_1(t) = 1.5\exp(-10(t-0.2)^2) + 1.6\exp(-8(t-0.8)^2)$$
$$\beta_2(t) = -0.5t - 0.5\exp(-5(t-0.8)^2).$$

- We allow for a dynamic specifications between the regressors $x_{1,t}$ and $x_{2,t}$:

$$\begin{pmatrix} x_{1,t} \\ x_{2,t} \end{pmatrix} = \begin{pmatrix} 0.3 & 0.1 \\ 0.1 & 0.2 \end{pmatrix} \begin{pmatrix} x_{1,t-1} \\ x_{2,t-1} \end{pmatrix} + \begin{pmatrix} \xi_{1,t} \\ \xi_{2,t} \end{pmatrix},$$

where $(\xi_{1,t}, \xi_{2,t})'$ are i.i.d. bivariate standard normal.

Data generating process (cont.)

- The error term follows an ARMA(1,1) given by

$$u_t = \phi u_{t-1} + \psi \epsilon_{t-1} + \epsilon_t,$$

  with ϵ_t generated as

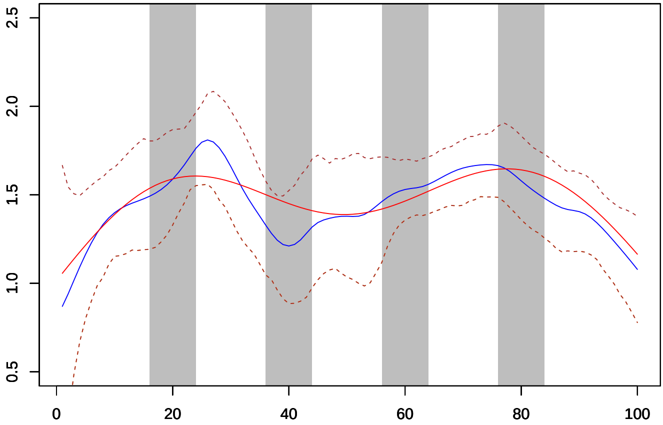
$$\epsilon_t \sim N\left(0, \frac{1-\phi^2}{1+\psi^2+2\phi\psi}\right).$$

- The variance of ϵ_t is such that the signal to noise ratio does not change when ϕ and ψ are varied.

Specifications

- 2000 samples of length $n = 200$
- 1299 bootstrap replications
- Epanechnikov kernel given by $K(x) = \frac{3}{4}(1 - x^2)\mathbb{1}_{\{|x| \leq 1\}}$
- Bandwidths: $h = 0.06$ to $h = 0.21$ in steps of 0.03
- $\tilde{h} = Ch^{5/9}$ for $C = 2$
- $G_{sub} = U_1(h) \cup U_4(h)$ and $G = \bigcup_{i=1}^{4} U_i(h)$, with $U_i(h) = \{(i/5) - h + j/100;\ j = 0, \ldots, [200h]\}$
- In addition: simultaneity over the whole sample

An example



β_1 with estimate and confidence bands

AR(1) with $\phi = 0.3$, shaded area is G with $h = 0.06$

Simulation results: empirical coverage for $\beta_1(\cdot)$

Coverage		Pointwise		G_{sub}		G		Full Sample	
(ϕ, ψ)	h	SB	SimBased	SB	SimBased	SB	SimBased	SB	SimBased
(0,0)	0.06	0.947	0.883	0.941	0.409	0.930	0.176	0.861	0.017
	0.09	0.947	0.925	0.927	0.643	0.924	0.448	0.895	0.121
	0.12	0.945	0.933	0.927	0.704	0.931	0.606	0.922	0.180
	0.15	0.939	0.931	0.916	0.560	0.919	0.517	0.911	0.224
	0.18	0.927	0.918	0.898	0.291	0.896	0.278	0.899	0.241
	0.21	0.908	0.894	0.860	0.226	0.854	0.220	0.854	0.221
(0.5,0)	0.06	0.943	0.865	0.925	0.345	0.914	0.130	0.841	0.009
	0.09	0.943	0.914	0.926	0.580	0.920	0.374	0.897	0.093
	0.12	0.939	0.924	0.917	0.628	0.918	0.515	0.909	0.168
	0.15	0.933	0.921	0.915	0.515	0.919	0.467	0.910	0.215
	0.18	0.924	0.908	0.892	0.272	0.894	0.260	0.895	0.203
	0.21	0.910	0.888	0.864	0.211	0.859	0.203	0.857	0.206

Simulation results: empirical coverage for $\beta_2(\cdot)$

Coverage		Pointwise		G_{sub}		G		Full Sample	
(ϕ, ψ)	h	SB	SimBased	SB	SimBased	SB	SimBased	SB	SimBased
(0,0)	0.06	0.946	0.886	0.925	0.408	0.917	0.175	0.853	0.021
	0.09	0.946	0.927	0.933	0.647	0.931	0.461	0.909	0.126
	0.12	0.946	0.935	0.945	0.705	0.939	0.593	0.930	0.193
	0.15	0.946	0.933	0.938	0.566	0.940	0.520	0.930	0.212
	0.18	0.946	0.929	0.936	0.315	0.937	0.303	0.937	0.250
	0.21	0.946	0.922	0.931	0.263	0.932	0.258	0.932	0.262
(0.5,0)	0.06	0.945	0.877	0.923	0.393	0.920	0.159	0.834	0.015
	0.09	0.944	0.919	0.932	0.594	0.932	0.410	0.904	0.109
	0.12	0.943	0.928	0.929	0.654	0.936	0.556	0.926	0.172
	0.15	0.944	0.928	0.939	0.538	0.931	0.485	0.920	0.215
	0.18	0.944	0.924	0.930	0.298	0.928	0.285	0.926	0.235
	0.21	0.942	0.917	0.926	0.254	0.925	0.252	0.927	0.253

Simulation results: empirical length for $\beta_1(\cdot)$

Length		Pointwise		G_{sub}		G		Full Sample	
(ϕ, ψ)	h	SB	SimBased	SB	SimBased	SB	SimBased	SB	SimBased
(0,0)	0.06	0.663	0.603	0.948	0.603	1.021	0.603	1.024	0.603
	0.09	0.526	0.564	0.754	0.564	0.807	0.564	0.813	0.564
	0.12	0.450	0.514	0.649	0.514	0.681	0.514	0.696	0.514
	0.15	0.401	0.469	0.582	0.469	0.604	0.469	0.607	0.469
	0.18	0.366	0.432	0.536	0.432	0.547	0.432	0.551	0.432
	0.21	0.338	0.402	0.494	0.402	0.499	0.402	0.502	0.402
(0.5,0)	0.06	0.717	0.638	1.024	0.638	1.105	0.638	1.109	0.638
	0.09	0.580	0.611	0.830	0.611	0.889	0.611	0.897	0.611
	0.12	0.502	0.563	0.722	0.563	0.758	0.563	0.775	0.563
	0.15	0.450	0.516	0.652	0.516	0.678	0.516	0.681	0.516
	0.18	0.412	0.477	0.603	0.477	0.617	0.477	0.620	0.477
	0.21	0.382	0.445	0.558	0.445	0.564	0.445	0.566	0.445

Simulation results: empirical length for $\beta_2(\cdot)$

Length		Pointwise		G_{sub}		G		Full Sample	
(ϕ, ψ)	h	SB	SimBased	SB	SimBased	SB	SimBased	SB	SimBased
(0,0)	0.06	0.678	0.621	0.969	0.621	1.044	0.621	1.047	0.621
	0.09	0.538	0.582	0.771	0.582	0.826	0.582	0.832	0.582
	0.12	0.461	0.529	0.664	0.529	0.697	0.529	0.713	0.529
	0.15	0.411	0.480	0.596	0.480	0.619	0.480	0.622	0.480
	0.18	0.375	0.438	0.548	0.438	0.560	0.438	0.564	0.438
	0.21	0.346	0.403	0.505	0.403	0.511	0.403	0.513	0.403
(0.5,0)	0.06	0.703	0.642	1.004	0.642	1.083	0.642	1.087	0.642
	0.09	0.568	0.609	0.814	0.609	0.871	0.609	0.879	0.609
	0.12	0.492	0.557	0.707	0.557	0.742	0.557	0.760	0.557
	0.15	0.441	0.507	0.639	0.507	0.664	0.507	0.667	0.507
	0.18	0.403	0.466	0.591	0.466	0.604	0.466	0.607	0.466
	0.21	0.375	0.430	0.546	0.430	0.553	0.430	0.555	0.430

Concluding remarks and future research

- The contribution of this paper:
  - ✓ Sieve bootstrap validity for time-varying coefficient models
  - ✓ Simultaneous confidence bands
  - ✓ Extensive simulation study
  - ✓ Application to EU Emissions Trading System
- Future research:
  - ✓ Methods: Extension to a panel setting
  - ✓ Application: Study the effect of sea surface temperature and soil moisture on U.S. soy yield

References i

[1] Bühlmann, P. (1998). Sieve bootstrap for smoothing in nonstationary time series. Annals of Statistics 26, 48-83.

[2] Cai, Z. (2007). Trending time-varying coefficient time series models with serially correlated errors. Journal of Econometrics 136, 163-188.

[3] Friedrich, M., Smeekes, S. and J.-P. Urbain (2020). Autoregressive wild bootstrap inference for nonparametric trends. Journal of Econometrics 214(1), 81-109.

[4] Neumann, M.H. and J. Polzehl (1998). Simultaneous bootstrap confidence bands in nonparametric regression. Journal of Nonparametric Statistics 9, 307-333.

[5] Karmakar, S., Richter, S. and Wu, W. B. (2021). Simultaneous inference for time-varying models. Journal of Econometrics, 227(2), 408–428.

[6] Wu, W.B. and Z. Zhao(2007). Inference of trends in time series. Journal of the Royal Statistical Society Series B 69, 391-410.

[7] Zhou, Z. and W.B. Wu (2010). Simultaneous inference of linear models with time varying coefficients. Journal of the Royal Statistical Society Series B 72, 513-531.