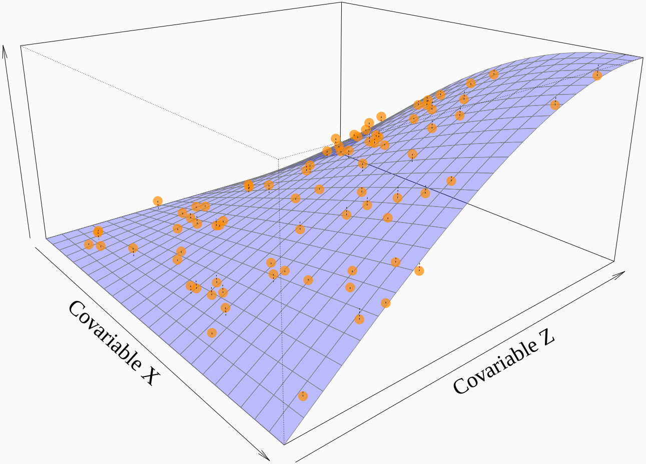
Nonparametric Inference for Functional Data

Dominik Liebl

Theoretical Subject

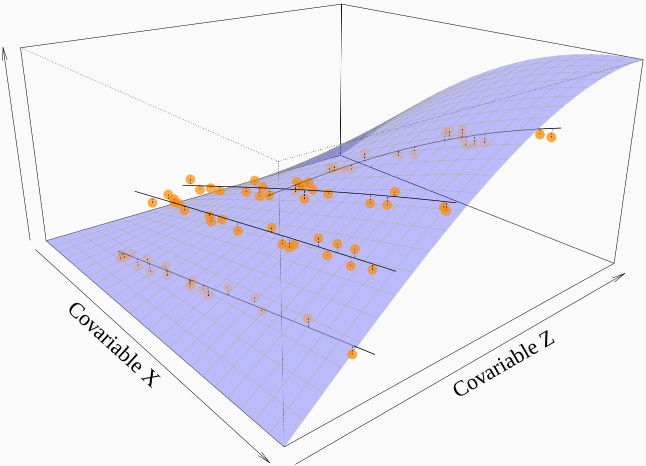
Bivariate Nonparametric Regression

Classical Data-Setup:



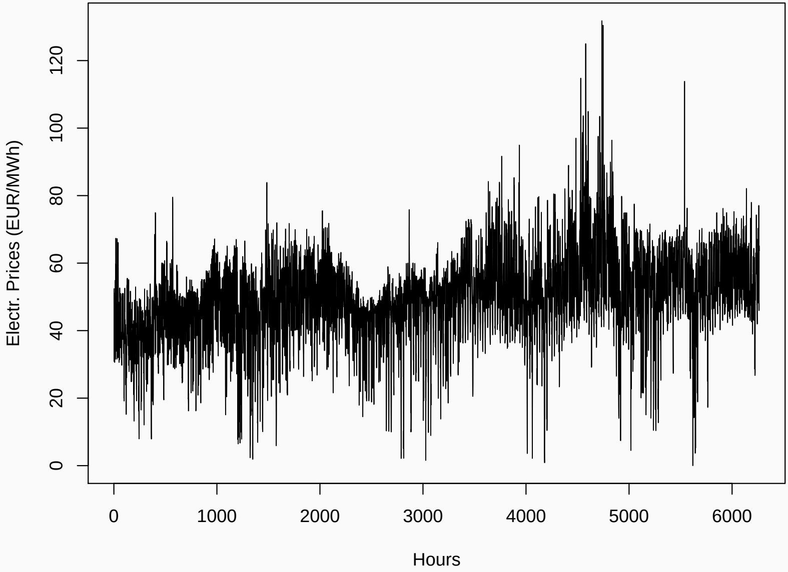
Bivariate Nonparametric Regression

Functional Data-Setup:



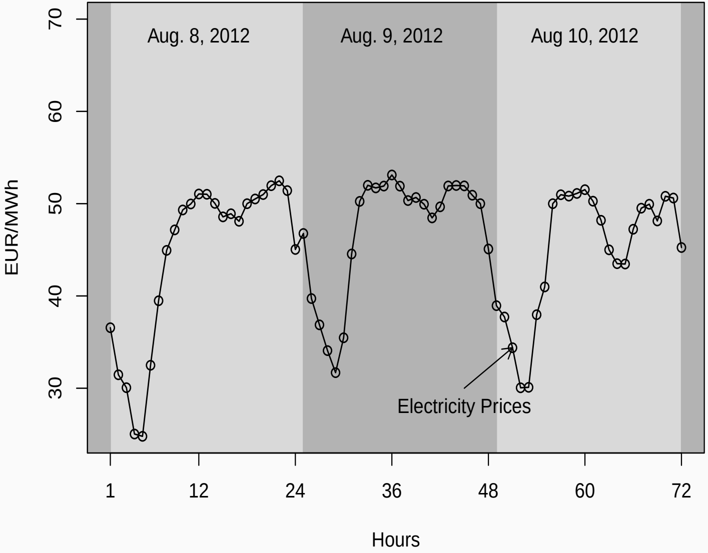
Real Data

Data: **Y**



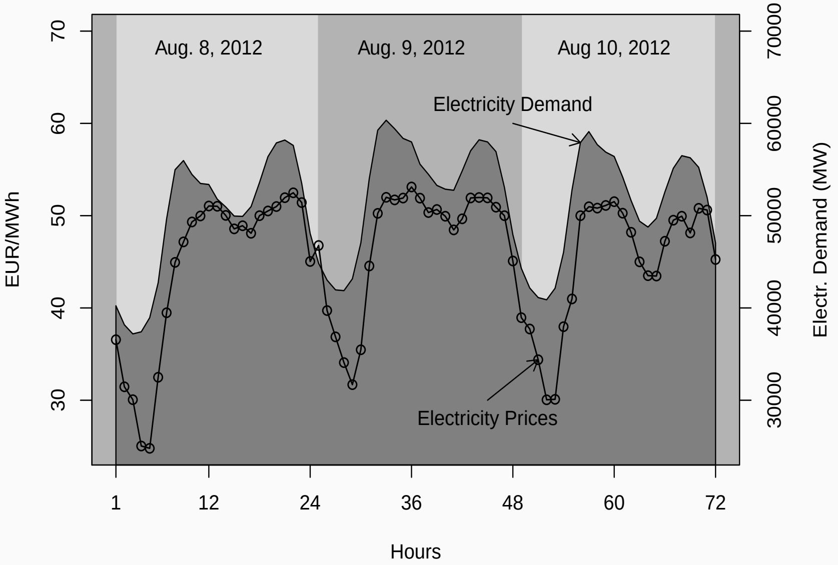
Electr. Prices (EUR/MWh)
0
20
40
60
80
100
120
0
1000
2000
3000
4000
5000
6000
Hours

Data: **Y**



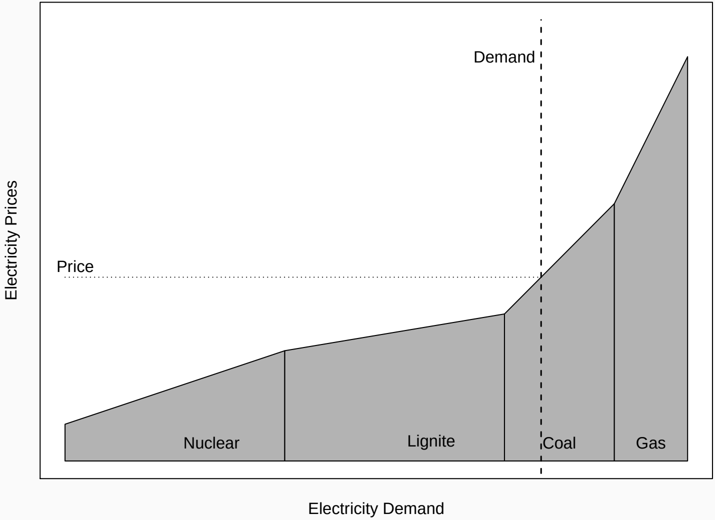
Aug. 8, 2012
Aug. 9, 2012
Aug 10, 2012
Electricity Prices
EUR/MWh
30
40
50
60
70
1
12
24
36
48
60
72
Hours

Data: **Y** and **X**

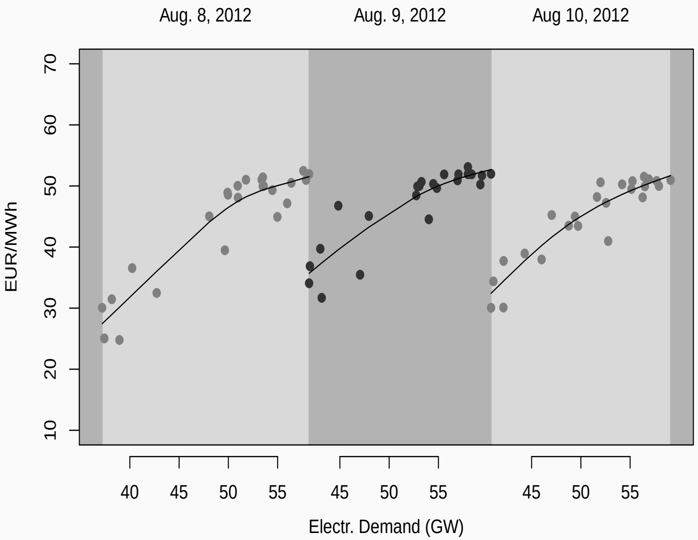


Aug. 8, 2012
Aug. 9, 2012
Aug 10, 2012
Electricity Demand
Electricity Prices
EUR/MWh
70
60
50
40
30
Electr. Demand (MW)
70000
60000
50000
40000
30000
1
12
24
36
48
60
72
Hours

Merit-Order Model

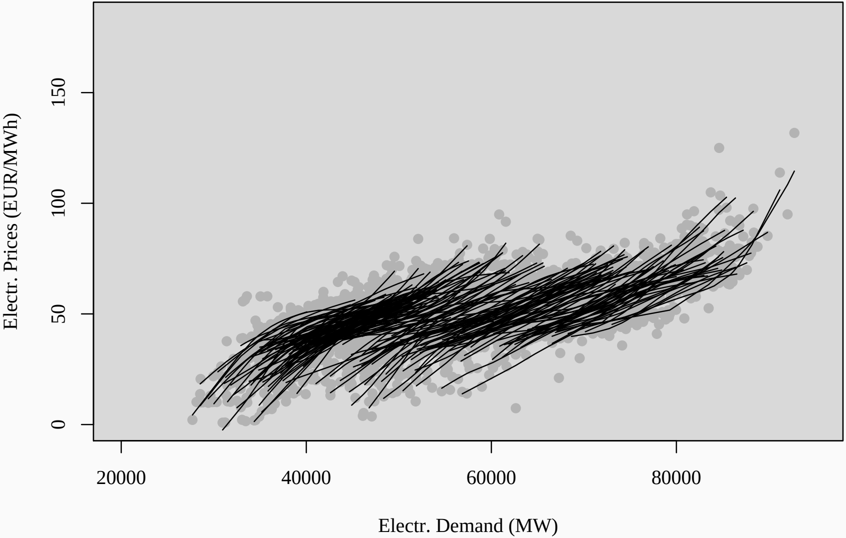


Data: $\mathcal{P}_t(.)$



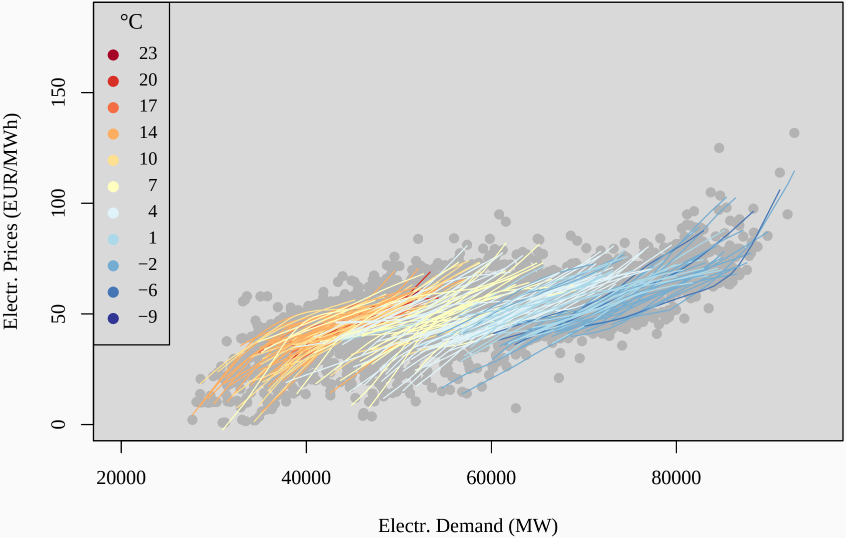
(Liebl, AOAS, 2013)

Data



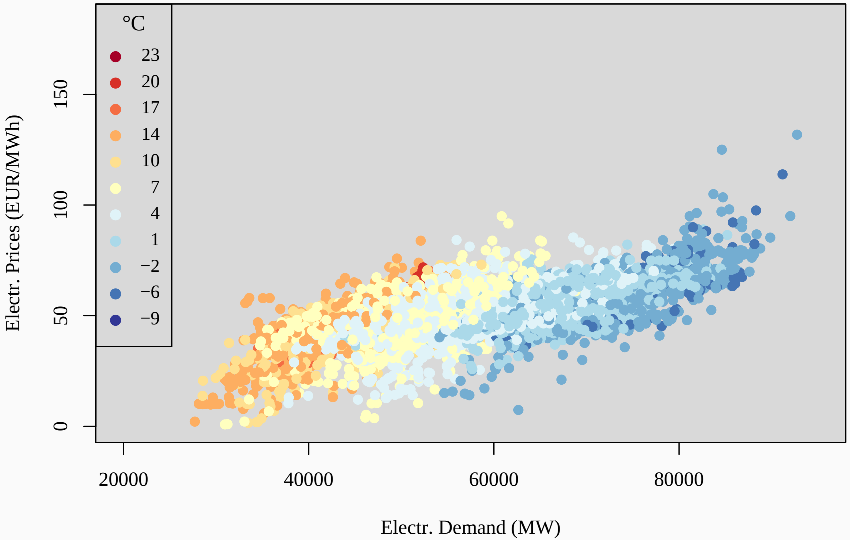
Electr. Prices (EUR/MWh)
0
50
100
150
20000
40000
60000
80000
Electr. Demand (MW)

Data



°C
23
20
17
14
10
7
4
1
−2
−6
−9
Electr. Prices (EUR/MWh)
0
50
100
150
20000
40000
60000
80000
Electr. Demand (MW)

Data


°C
23
20
17
14
10
7
4
1
−2
−6
−9
150
100
50
0
Electr. Prices (EUR/MWh)
20000
40000
60000
80000
Electr. Demand (MW)

Model & Assumptions

Model - Assumptions

$$Y_{it} = \mathcal{P}_t(X_{it}, Z_t) + \epsilon_{it}$$

- $t \in \{1, \dots, T\}$ (T Functions)
- $i \in \{1, \dots, n\}$ (n Points per Function)
- $\mathcal{P}_t(\,\cdot\,, Z_t = z) \in L^2([a(z), b(z)]) \cap C([a(z), b(z)])$
- $[a(z), b(z)] \subset \mathbb{R}$ (Compact support)
- $\mu(x, z) = \mathbb{E}(\mathcal{P}_t(x, z)) = \mathbb{E}(Y_{it} | X_{it} = x, Z_t = z)$
- $\gamma(x_1, x_2, z) = \text{Cov}(\mathcal{P}_t(x_1, z), \mathcal{P}_t(x_2, z))$

Model - Assumptions

$$Y_{it} = \mathcal{P}_t(X_{it}, Z_t) + \epsilon_{it}$$

- $t \in \{1, \ldots, T\}$ (T Functions)
- $i \in \{1, \ldots, n\}$ (n Points per Function)

- $\mathcal{P}_t(\,\cdot\,, Z_t = z) \in L^2([a(z), b(z)]) \cap C([a(z), b(z)])$
- $[a(z), b(z)] \subset \mathbb{R}$ (Compact support)

- $\mu(x, z) = \mathbb{E}(\mathcal{P}_t(x, z)) = \mathbb{E}(Y_{it} | X_{it} = x, Z_t = z)$
- $\gamma(x_1, x_2, z) = \text{Cov}(\mathcal{P}_t(x_1, z), \mathcal{P}_t(x_2, z))$

Model - Assumptions

$$Y_{it} = \mathcal{P}_t(X_{it}, Z_t) + \epsilon_{it}$$

- $t \in \{1, \dots, T\}$ (T Functions)
- $i \in \{1, \dots, n\}$ (n Points per Function)
- $\mathcal{P}_t(\,\cdot\,, Z_t = z) \in L^2([a(z), b(z)]) \cap C([a(z), b(z)])$
- $[a(z), b(z)] \subset \mathbb{R}$ (Compact support)

- $\mu(x, z) = \mathbb{E}(\mathcal{P}_t(x, z)) = \mathbb{E}(Y_{it} | X_{it} = x, Z_t = z)$
- $\gamma(x_1, x_2, z) = \text{Cov}(\mathcal{P}_t(x_1, z), \mathcal{P}_t(x_2, z))$

Model - Assumptions

$$Y_{it} = \mathcal{P}_t(X_{it}, Z_t) + \epsilon_{it}$$

- $(\mathcal{P}_t, X_{it}, Z_t, \epsilon_{it})_t$: weakly dependent and strictly stationary.
- $\mathbb{E}(\epsilon | X, Z, \mathcal{P}) = \mathbb{E}(\epsilon) = 0.$
- Y_{it} has finite fourth moments.

Model - Assumptions

$$Y_{it} = \mathcal{P}_t(X_{it}, Z_t) + \epsilon_{it}$$

- $(\mathcal{P}_t, X_{it}, Z_t, \epsilon_{it})_t$: weakly dependent and strictly stationary.

- $\mathbb{E}(\epsilon \,|\, X, Z, \mathcal{P}) = \mathbb{E}(\epsilon) = 0.$
- $\mathcal{P}_t$ and ϵ_{it} have finite fourth moments.

NP-Regression Model

Nonparametric Regression Model:

$$Y_{it} = \mathcal{P}_t(X_{it}, Z_t) + \epsilon_{it}$$

$$\Leftrightarrow Y_{it} = \boxed{\mu(X_{it}, Z_t)} + \underbrace{\mathcal{P}_t^{c}(X_{it}, Z_t)}_{=\mathcal{P}_t(X_{it}, Z_t) - \boxed{\mu(X_{it}, Z_t)}} + \epsilon_{it}$$

NP-Regression Model

Nonparametric Regression Model:

$$Y_{it} = \mu(X_{it}, Z_t) \quad + \quad \underbrace{\mathcal{P}_t^{c}(X_{it}, Z_t) + \epsilon_{it}}_{\text{Compound error structure}}$$

NP-Regression Model

Nonparametric Regression Model:

$$Y_{it} = \mu(X_{it}, Z_t) \quad + \quad \underbrace{\mathcal{P}_t^c(X_{it}, Z_t) + \epsilon_{it}}_{\text{Compound error structure}}$$

Within-function correlations:

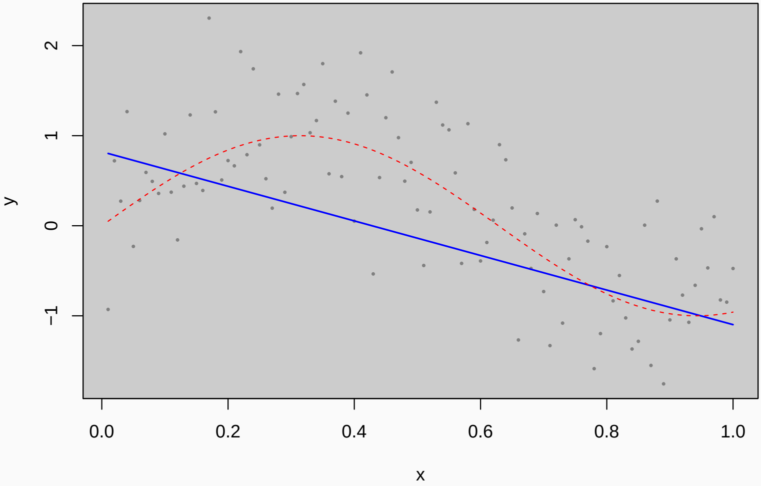
$$\text{Corr}(Y_{it}, Y_{jt} | \mathbf{X}, \mathbf{Z}) = \underbrace{\text{Corr}(\mathcal{P}_t^c(X_{it}, Z_t), \mathcal{P}_t^c(X_{jt}, Z_t) | \mathbf{X}, \mathbf{Z})}_{\approx 1}$$

if $|X_{it} - X_{jt}| \approx 0, \quad i \neq j.$

Estimation of $\mu(x,z) = \mathbb{E}(Y_{it} | X_{it} = x, Z_t = z)$

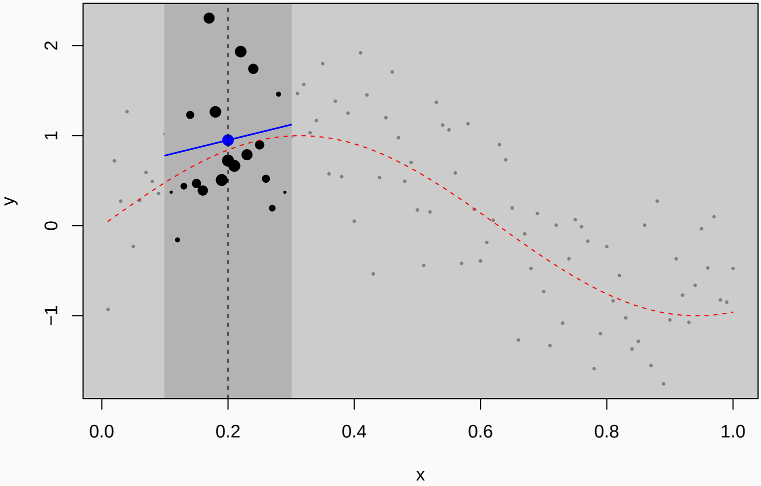
Local Linear Estimator

$$\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = \left([\mathbf{1}, \mathbf{X}]^T [\mathbf{1}, \mathbf{X}] \right)^{-1} [\mathbf{1}, \mathbf{X}]^T \mathbf{Y}$$



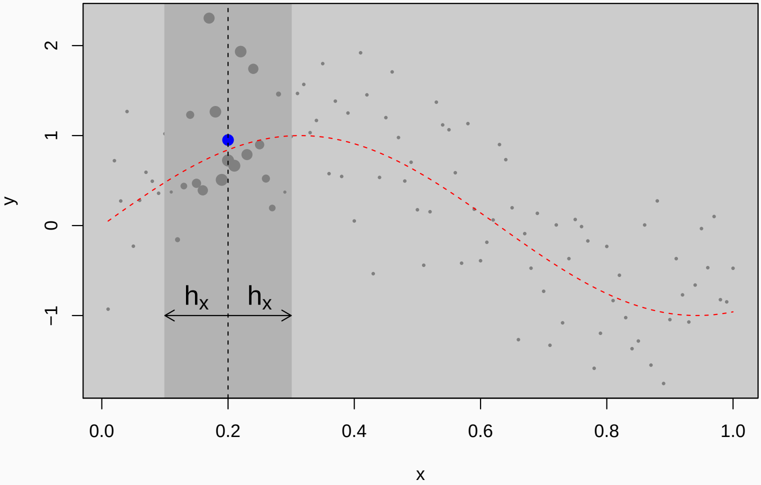
Local Linear Estimator

$$\begin{pmatrix} \hat{\alpha}(x) \\ \hat{\beta}(x) \end{pmatrix} = \left([\mathbf{1}, \mathbf{X}-x]^T W_x [\mathbf{1}, \mathbf{X}-x]\right)^{-1} [\mathbf{1}, \mathbf{X}-x]^T W_x \mathbf{Y}$$



Local Linear Estimator

$$\boxed{\hat{\alpha}(x; h_X)} = u_1^T \left([\mathbf{1}, X - x]^T W_x [\mathbf{1}, X - x]\right)^{-1} [\mathbf{1}, X - x]^T W_x Y$$



Local Linear Estimator

Estimator of the mean function:

$$\hat{\mu}(x, z; h_{\mu,X}, h_{\mu,Z}) =$$

$$u_1^\top \left([\mathbf{1}, \mathbf{X}_x, \mathbf{Z}_z]^\top \mathbf{W}_{\mu,xz} [\mathbf{1}, \mathbf{X}_x, \mathbf{Z}_z] \right)^{-1} [\mathbf{1}, \mathbf{X}_x, \mathbf{Z}_z]^\top \mathbf{W}_{\mu,xz} \mathbf{Y}$$

Problems

- How to choose the bandwidths $h_{\mu,X}$ and $h_{\mu,Z}$?

- How to do inference?

  H_0: $\mu_{\mathsf{A}}(x,z) = \mu_{\mathsf{B}}(x,z)$

  H_1: $\mu_{\mathsf{A}}(x,z) \neq \mu_{\mathsf{B}}(x,z)$

Theoretical Results

Asymptotic Setup

Double asymptotic:

- $nT \to \infty$ such that:
  - $T \to \infty$
  - $n \sim T^{\theta}$, with $0 \leq \theta < \infty$.

θ-Cases:

$n \sim T^{\theta}$ with $\theta = 0$: T large and n small

$n \sim T^{\theta}$ with $\theta > 0$: T large and n small-to-large

Further Assumptions

Bandwidths:

- $h_X, h_Z \to 0$ as $Tn \to \infty$
- $(Tn)h_X\, h_Z \to \infty$ as $Tn \to \infty$

Kernel function:

- $K(u,v) = \kappa(u)\kappa(v)$, where κ is a symmetric pdf.

Smoothness:

- Usual smoothness assumptions on $\mu(x,z)$, $f_{XZ}(x,z)$, and $f_Z(z)$.

Asymptotic Bias

Theorem (Bias of $\hat{\mu}$):

$$\overbrace{\text{Bias}\left(\hat{\mu}(x, z; h_{\mu,X}, h_{\mu,Z})\right)}^{\mathbb{E}(\hat{\mu}(x,z))-\mu(x,z)} = B_{\hat{\mu}}(x, z)\left(1 + o_p(1)\right)$$

with

$$B_{\hat{\mu}}(x, z) = \frac{\nu_2(K_\mu)}{2}\left[h^2_{\mu,X}\,\mu^{(2,0)}(x, z) + h^2_{\mu,Z}\,\mu^{(0,2)}(x, z)\right]$$

Asymptotic Variance

Theorem (Variance of $\hat{\mu}$):

$$\mathbb{V}\left(\hat{\mu}(x, z; h_{\mu,X}, h_{\mu,Z})\right) =$$

$$= \left(\boxed{V_1^{\hat{\mu}}(x,z)} + \boxed{V_2^{\hat{\mu}}(x,z)} \right)(1 + o_p(1))$$

with

$$\boxed{V_1^{\hat{\mu}}(x,z)} = (Tn)^{-1}\left[h_{\mu,X}^{-1} h_{\mu,Z}^{-1} R(K_\mu) \frac{\gamma(x,x,z) + \sigma_\epsilon^2}{f_{XZ}(x,z)} \right]$$

$$\boxed{V_2^{\hat{\mu}}(x,z)} = T^{-1}\left[h_{\mu,Z}^{-1} R(\kappa) \frac{\gamma(x,x,z)}{f_Z(z)} \right]$$

AMISE-Scenario I: $0 \le \theta \le 1/5$

$$\text{AMISE}_{\hat{\mu}}\left(h_{\mu,X}, h_{\mu,Z}\right) =$$

$$\underbrace{\overbrace{(Tn)^{-1}\, h_{\mu,X}^{-1}\, h_{\mu,Z}^{-1}\, R(K_\mu)\, Q_{\mu,1}}^{\int V_1^{\hat{\mu}}}}_{\textbf{1st Order}} + \underbrace{\overbrace{T^{-1}\, h_{\mu,Z}^{-1}\, R(\kappa)\, Q_{\mu,2}}^{\int V_2^{\hat{\mu}}}}_{\text{2nd Order}} +$$

$$\frac{\left(\nu_2(K_\mu)\right)^2}{4}\left[\underbrace{2\, h_{\mu,X}^2\, h_{\mu,Z}^2\, \mathcal{I}_{\mu,XZ}}_{\textbf{1st Order}} + \underbrace{h_{\mu,Z}^4\, \mathcal{I}_{\mu,ZZ}}_{\text{2nd Order}}\right]$$

AMISE-Scenario I: $0 \leq \theta \leq 1/5$

Classical Results in Bivariate NP-Regression:

Bandwidth Rates:

$$h_{\mu,X,\text{AMISE-I}} \sim h_{\mu,Z,\text{AMISE-I}} \sim (Tn)^{-1/6}$$

Convergence Rate for $\hat{\mu}$:

$$\text{AMISE}(\hat{\mu}) = \mathcal{O}_p\big((Tn)^{-2/3}\big)$$

AMISE-Scenario II: $1/5 < \theta < \infty$

$$\text{AMISE}_{\hat{\mu}}\left(h_{\mu,X}, h_{\mu,Z}\right) =$$

$$\underbrace{\overbrace{(Tn)^{-1}\, h_{\mu,X}^{-1}\, h_{\mu,Z}^{-1}\, R(K_\mu)\, Q_{\mu,1}}^{\int V_1^\mu}}_{\text{2nd Order}} + \underbrace{\overbrace{T^{-1}\, h_{\mu,Z}^{-1}\, R(\kappa)\, Q_{\mu,2}}^{\int V_2^\mu}}_{\textbf{1st Order}} +$$

$$\frac{\left(\nu_2(K_\mu)\right)^2}{4}\left[\underbrace{2\, h_{\mu,X}^2\, h_{\mu,Z}^2\, \mathcal{I}_{\mu,XZ}}_{\text{2nd Order}} + \underbrace{h_{\mu,Z}^4\, \mathcal{I}_{\mu,ZZ}}_{\textbf{1st Order}}\right]$$

AMISE-Scenario II: $1/5 < \theta < \infty$

Non-Classical Results in Bivariate NP-Regression:

$$h_{\mu,Z,\text{AMISE}} = \left(\frac{R(\kappa)\, Q_{\mu,2}}{(\nu_2(K_\mu))^2\, \mathcal{I}_{\mu,ZZ}} \right)^{1/5} T^{-1/5}$$

$$h_{\mu,X,\text{AMISE}} = \left(\frac{R(K_\mu)\, Q_{\mu,1}}{Tn\, (\nu_2(K_\mu))^2\, \mathcal{I}_{\mu,XZ}} \right)^{1/3} h^{-1}_{\mu,Z,\text{AMISE}}$$

Functional Data:

Uni-variate rate: $h_{\mu,Z,\text{AMISE}} \sim T^{-1/5}$

Anti-proportional: $h_{\mu,X,\text{AMISE}} \sim h^{-1}_{\mu,Z,\text{AMISE}}$

AMISE-Scenario II $1/5 < \theta < \infty$

Non-Classical Results in Bivariate NP-Regression:

$$h_{\mu,Z,\text{AMISE}} = \left(\frac{R(\kappa)\, Q_{\mu,2}}{\left(\nu_2(K_\mu)\right)^2 \mathcal{I}_{\mu,ZZ}} \right)^{1/5} T^{-1/5}$$

$$h_{\mu,X,\text{AMISE}} = \left(\frac{R(K_\mu)\, Q_{\mu,1}}{Tn\,\left(\nu_2(K_\mu)\right)^2 \mathcal{I}_{\mu,XZ}} \right)^{1/3} \left(h_{\mu,Z,\text{AMISE}}\right)^{-1}$$

Uni-variate Convergence Rate

$$\text{AMISE}(\hat{\mu}) = \mathcal{O}_p(T^{-4/5})$$

Robust Two-Sample Test Statistic

$$Z_{x,z} = \left(\frac{\left(\hat{\mu}_A(x,z) - B_{\hat{\mu}_A}(x,z)\right) - \left(\hat{\mu}_B(x,z) - B_{\hat{\mu}_B}(x,z)\right)}{\sqrt{\left(V_1^{\hat{\mu}_A}(x,z) + V_2^{\hat{\mu}_A}(x,z)\right) + \left(V_1^{\hat{\mu}_B}(x,z) + V_2^{\hat{\mu}_B}(x,z)\right)}} \right)$$

Under H_0: $\mu_A(x,z) = \mu_B(x,z)$

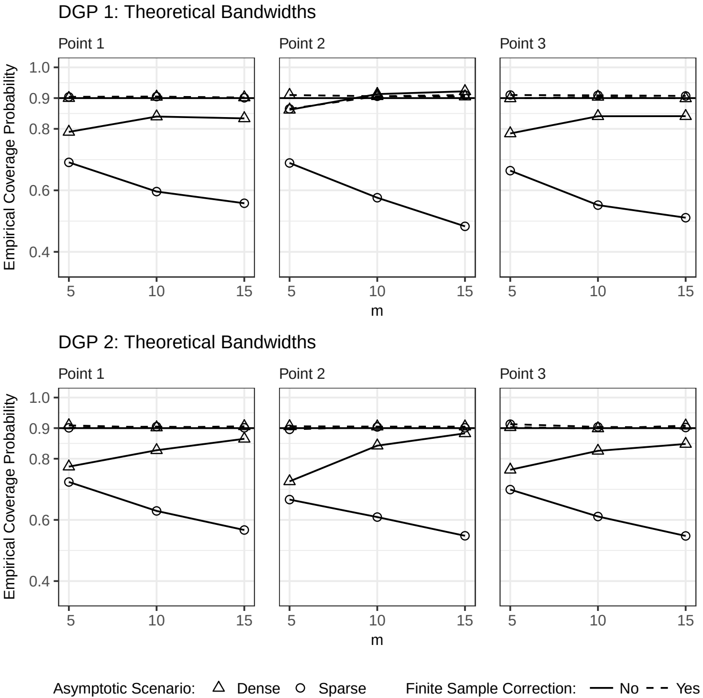
$$Z_{x,z} \to_d N(0,1) \quad \text{as} \quad nT \to \infty$$

Robust Inference under both scenarios:

"Sparse": $n \sim T^{\theta}$ with $0 \leq \theta \leq 1/5$, T large, n small

"Dense": $n \sim T^{\theta}$ with $1/5 < \theta < \infty$, T large, n small-to-large

Simulation Results


DGP 1: Theoretical Bandwidths
Point 1
Point 2
Point 3
Empirical Coverage Probability
1.0
0.9
0.8
0.6
0.4
5
10
15
m
DGP 2: Theoretical Bandwidths
Point 1
Point 2
Point 3
Empirical Coverage Probability
Asymptotic Scenario: Dense Sparse
Finite Sample Correction: No Yes

Application

March 11, 2011

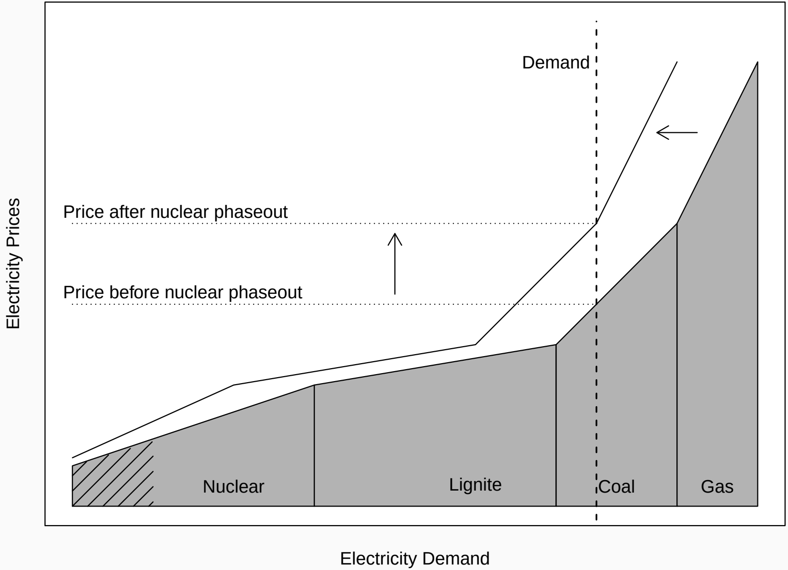

March 11, 2011

March 15, 2011

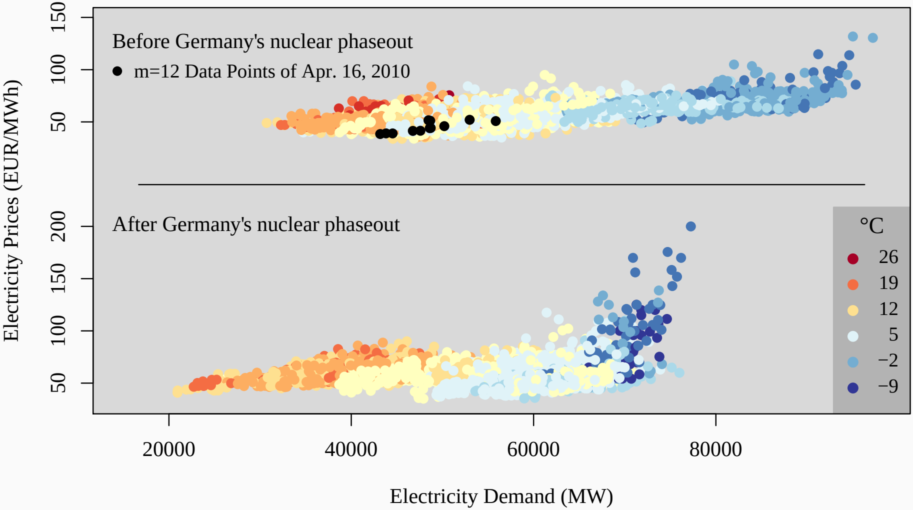

March 15, 2011


March 15, 2011

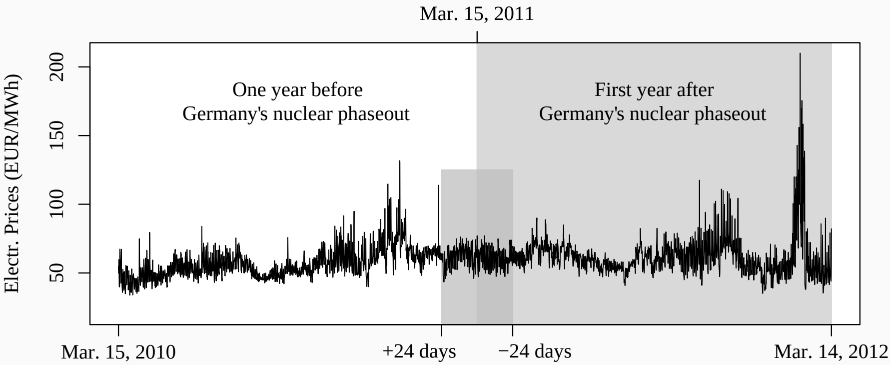
Application



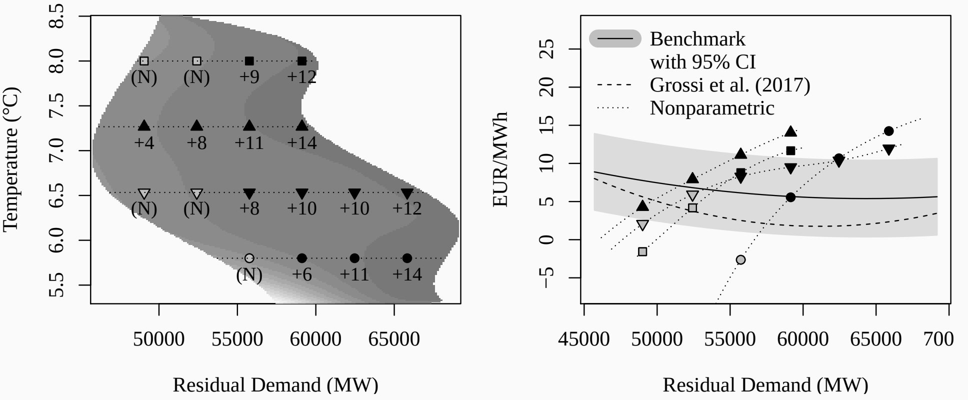
Application



Application



Application



Literature

- Grossi, L., Heim, S. and Waterson, M. (2017). The impact of the German response to the Fukushima earthquake. Energy Economics, 66, 450–465
- Zhang, X., & Wang, J. L. (2016). From sparse to dense functional data and beyond. The Annals of Statistics, 44(5), 2281–2321
- Jiang, C.-R. and Wang, J.-L. (2010). Covariate adjusted functional principal components analysis for longitudinal data. The Annals of Statistics, 38(2), 1194–1226
- **Liebl, D.** (2013). Modeling and forecasting electricity prices: A functional data perspective. The Annals of Applied Statistics, 7(3), 1562-1592
- **Liebl, D.** (2019). Inference for sparse and dense functional data with covariate adjustments. Journal of Multivariate Analysis, 170, 315-335
- **Liebl, D.** (2019). Nonparametric testing for differences in electricity prices: The case of the Fukushima nuclear accident. The Annals of Applied Statistics, 13(2), 1128-1146

Thank you!

www.dliebl.com

@domliebl