Spatial Correlation Robust Inference

Ulrich K. Müller and Mark W. Watson
Princeton University

Presentation at Yale
April 2022

Motivation

- More and more empirical work using spatial data in development, trade, macro, etc.

- How to appropriately correct standard errors?
  - Conley (1999): Spatial analog of HAC standard errors.
  - Don't work well under moderate or high spatial dependence.
  - Many applications characterized by strong spatial correlations (Kelly (2019, 2020)).

Canonical Problem

Observe

$$y_l = \mu + u_l, \quad l = 1, \dots, n$$

- y_l (and u_l) associated with observed location $s_l \in \mathcal{S} \subset \mathbb{R}^d$.

- s_l with density g.

- With $\mathbf{s} = (s_1, \dots, s_n)$: $\mathbb{E}[u_l | \mathbf{s}] = 0$, $\mathbb{E}[u_l u_\ell | \mathbf{s}] = \sigma_u(s_l - s_\ell)$, so u_l covariance stationary.

- How to test $H_0 : \mu = \mu_0$ and construct confidence interval for μ, conditional on $\mathbf{s}$?

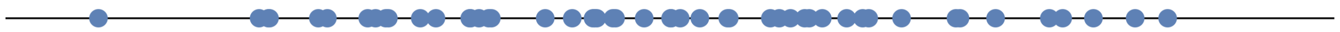
- Extensions to regression, GMM etc. follow from standard linearization arguments (see paper).

Three One-Dimensional Designs

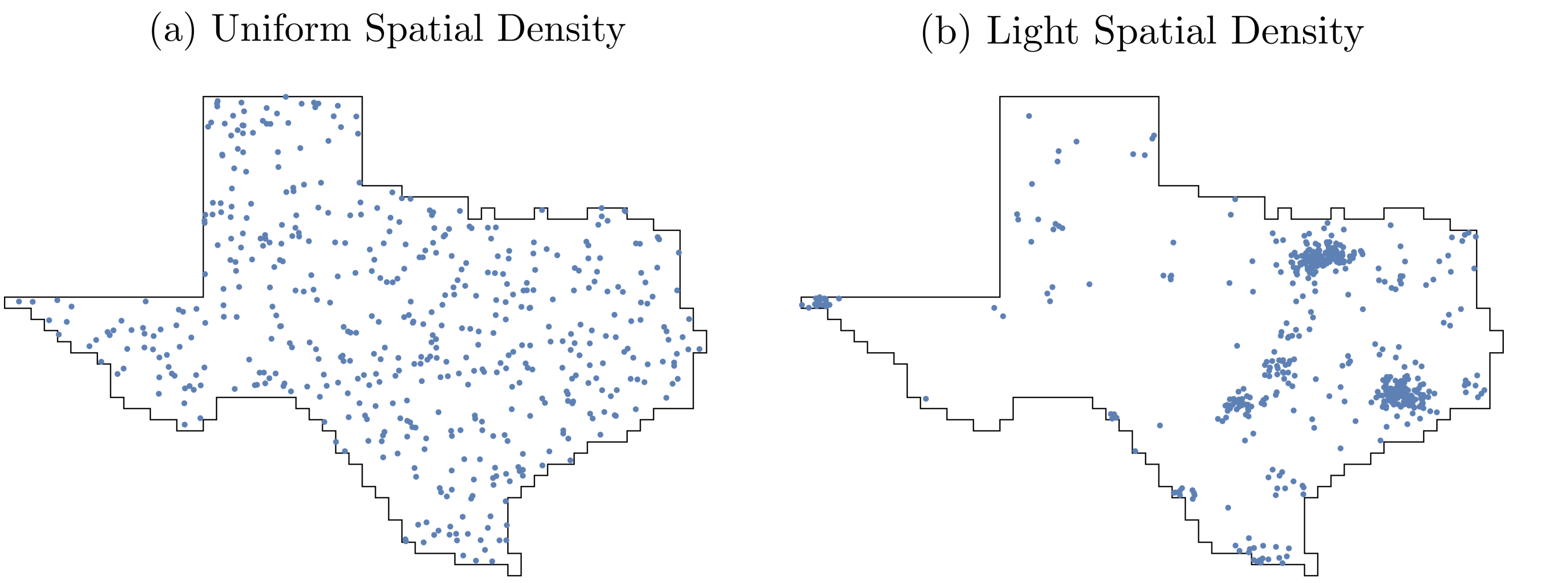
(a) Regularly spaced

(b) Uniform spatial density

(c) Triangular spatial density



Two Geographic Designs



Light data from Henderson, Squires, Storeygard and Weil (2018)

Spatial Inference

- Usual t-statistic

$$\tau = \frac{\sqrt{n}(\overline{y} - \mu_0)}{\hat{\sigma}}$$

with associated critical value cv, and confidence interval with endpoints $\bar{y} \pm \text{cv}\,\hat{\sigma}/\sqrt{n}$.

- Conley (1999): kernel-type "consistent" estimator $\hat{\sigma}^2 = n^{-1} \sum_{l,\ell} k\left(\frac{||s_l - s_\ell||}{b_n}\right) \hat{u}_l \hat{u}_\ell$ with $b_n \to \infty$ and standard normal cv .

- Bester, Conley, Hansen, and Vogelsang (2016): fixed-b version $\hat{\sigma}^2 = n^{-1} \sum_{l,\ell} k\left(\frac{||s_l - s_\ell||}{b}\right) \hat{u}_l \hat{u}_\ell$ and nonstandard cv obtained by simulating from $u_l \sim iid\mathcal{N}(0,1)$.

- Sun and Kim (2012): $\hat{\sigma}^2 = q^{-1} \sum_{i=1}^{q} \left(\sum_l n^{-1/2} w_{i,l} \hat{u}_l\right)^2$ with Fourier weights $w_{i,l}$ and student-t_q cv (analogous to Müller (2004, 2007), Phillips (2005), etc.)

This Paper

- Measure of strength of spatial dependence: average pairwise correlation

$$\bar{\rho} = \frac{1}{n(n-1)} \sum_{l=1}^{n} \sum_{\ell \neq l} \text{Cor}\left(y_l, y_\ell \left| \mathbf{s} \right.\right).$$

- Objective: construction of $\hat{\sigma}$ and cv that yield valid inference under
  - generic weak correlation $\bar{\rho} \to 0$, whether or not location distribution is uniform;
  - some strong correlation cases with $\bar{\rho} = O(1)$.

Outline of Talk

1. Definition of new "SCPC" method

2. Small sample size control under some forms of strong correlation

3. Large sample size control under generic weak correlation

4. Small sample efficiency of SCPC

Motivation for SCPC Method

- Model in vector notation: $\mathbf{y} = \mu \mathbf{l} + \mathbf{u} = \bar{y}\mathbf{l} + \hat{\mathbf{u}}$

- Numerator of t-statistic has variance $\sigma^2 = \text{Var}[\sqrt{n}\bar{y}] = n^{-1}\,\text{Var}[\mathbf{l}'\mathbf{u}]$

- Quadratic form estimators: $\hat{\sigma}^2 = \hat{\mathbf{u}}'\mathbf{Q}\hat{\mathbf{u}} = (n-1)^{-1}\sum_{i=1}^{n-1}\lambda_i(\mathbf{w}_i'\mathbf{u})^2$

- Challenge with (positive) spatial correlation: Most weighted averages $\mathbf{w}_i'\mathbf{u}$ are less variable than $\mathbf{l}'\mathbf{u}$, leading to downward bias of $\hat{\sigma}^2$.

- Solution: Select (few) weighted averages $\mathbf{w}_i'\mathbf{u}$ that are as variable as possible under plausible spatial covariance matrix, and use $\hat{\sigma}^2 = q^{-1}\sum_{i=1}^{q}(\mathbf{w}_i'\mathbf{u})^2$.

SCPC Method

- Benchmark model $\sigma_u(s_l - s_\ell) = \exp(-c||s_l - s_\ell||)$ with associated covariance matrix $\mathbf{\Sigma}(c)$ of $\mathbf{u}$.
  - Worst-case correlation has $c = c_0$, weaker correlation for $c > c_0$.

- Use q eigenvectors $\mathbf{r}_1, \ldots, \mathbf{r}_q$ of $\mathbf{M}\mathbf{\Sigma}(c_0)\mathbf{M}$ corresponding to largest eigenvalues as weights and

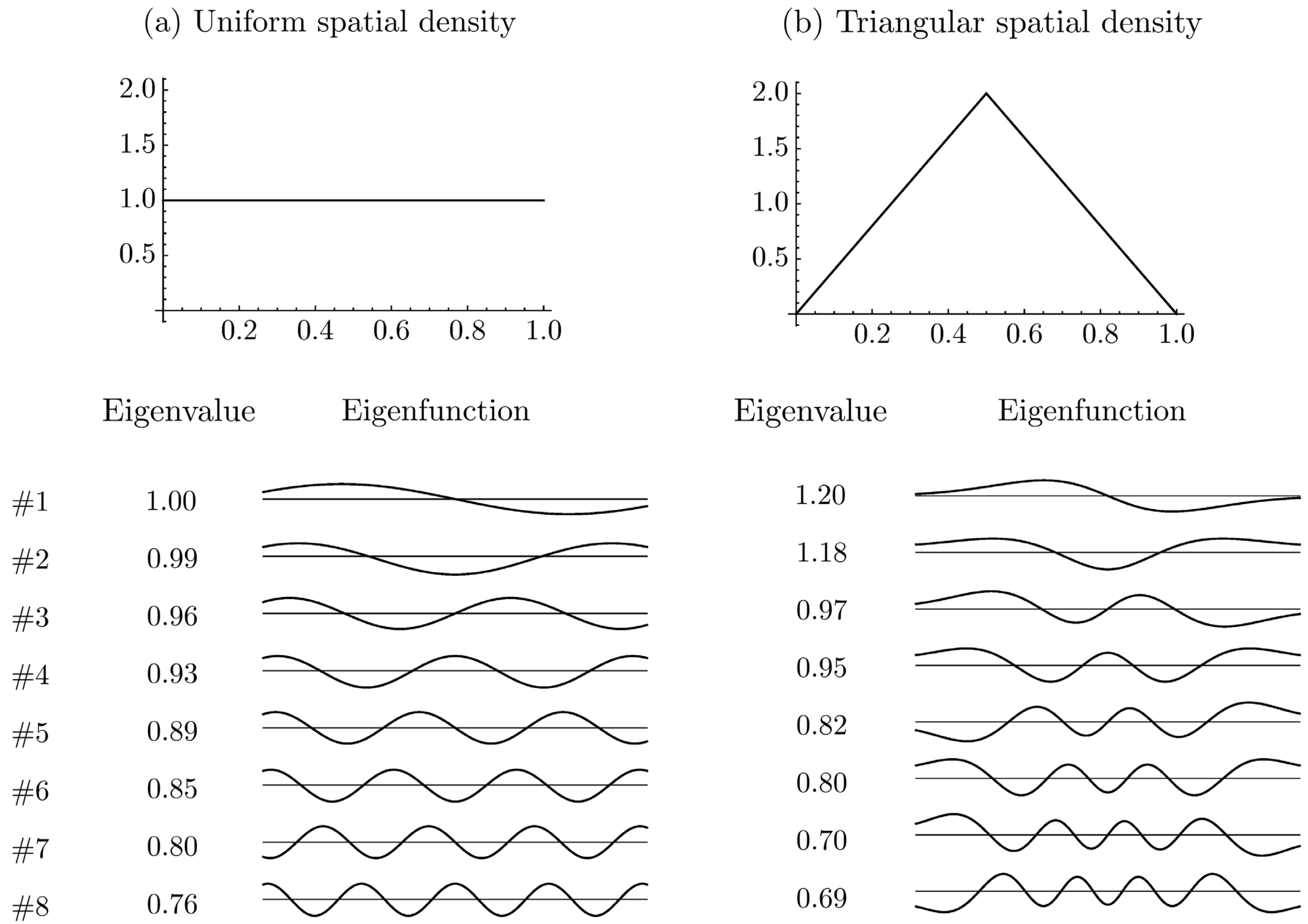
$$n\hat{\sigma}^2_{\mathsf{SCPC}}(q) = q^{-1}\sum_{i=1}^{q}(\mathbf{r}_i'\mathbf{u})^2.$$

  "Spatial Correlation Principle Components" (SCPC) of $\hat{\mathbf{u}} \sim \mathcal{N}(\mathbf{0}, \mathbf{M}\mathbf{\Sigma}(c_0)\mathbf{M})$.

- Let $\mathsf{cv}_{\mathsf{SCPC}}(q)$ solve $\sup_{c \geq c_0} \mathbb{P}^0_{\mathbf{\Sigma}(c)}(|\tau_{\mathsf{SCPC}}(q)| > \mathsf{cv}_{\mathsf{SCPC}}(q)|\mathbf{s}) = \alpha$.

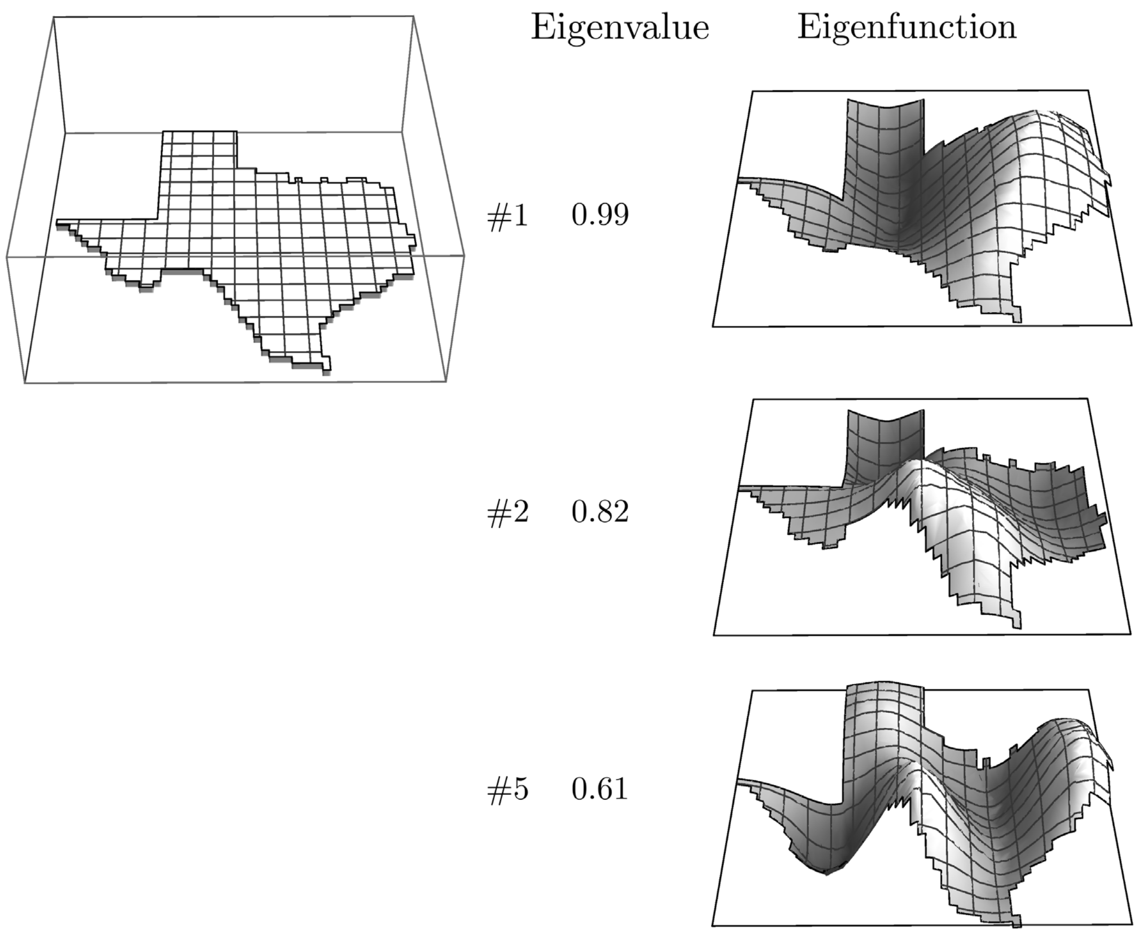
- q_{SCPC} minimizes CI length $\mathbb{E}[2\hat{\sigma}_{\mathsf{SCPC}}(q)\,\mathsf{cv}_{\mathsf{SCPC}}(q)|\mathbf{s}]$ in i.i.d. model $\mathbf{u}|\mathbf{s} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I}_n)$.

Eigenvectors in One-Dimensional Designs

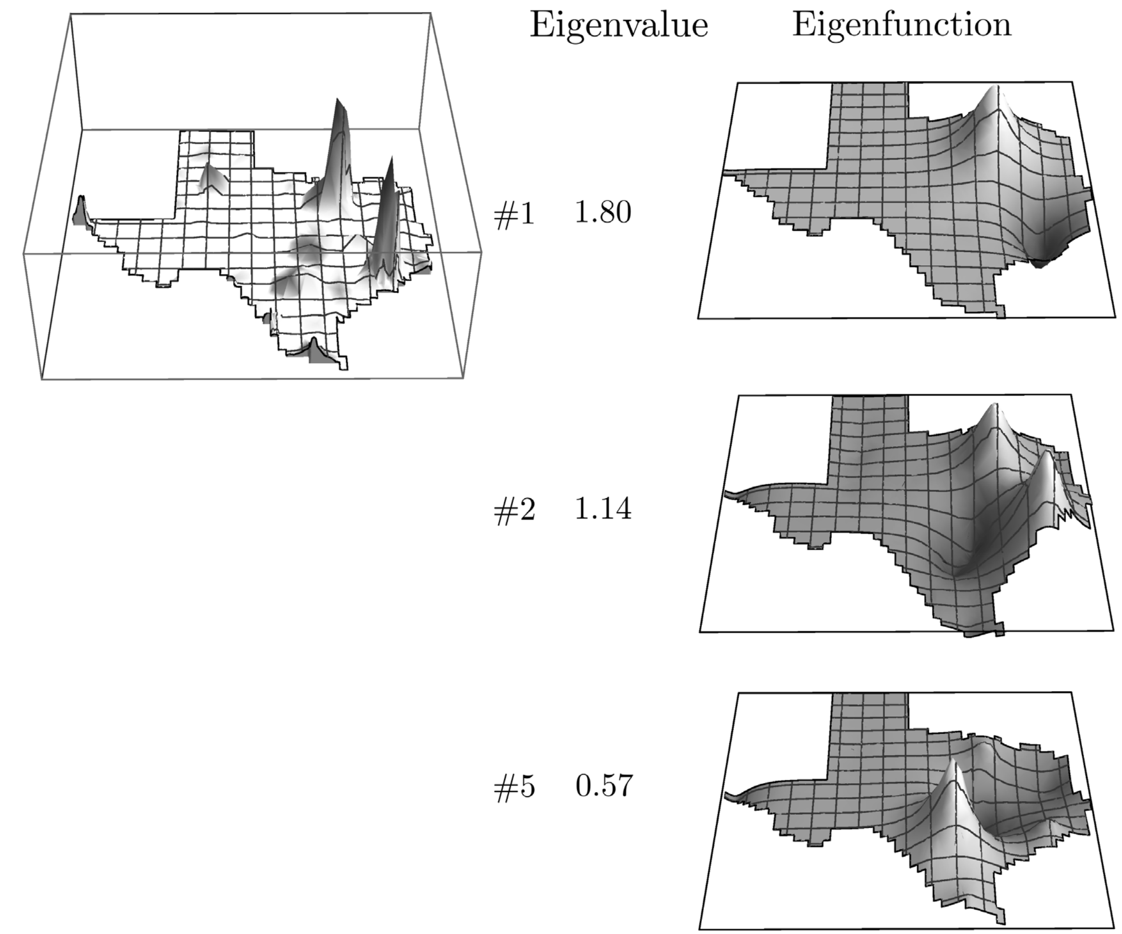


Eigenvectors in Geographic Designs

(a) Uniform spatial density



(b) Light spatial density



Properties of SCPC Inference

- Select c_0 as function of implied average correlation $\bar{\rho} = \bar{\rho}_0$.

  $\Rightarrow$ SCPC inference invariant under any linear transformation of the locations $\{s_l\}_{l=1}^n \to \{As_l\}_{l=1}^n$, such as rotations.

- For $\bar{\rho}_0 = 0.03$, $q_{\text{SCPC}} \approx 8$ (depending on $\mathbf{s}$).
  - cv$_{\text{SCPC}}$ not equal to corresponding student-t critical value: "bias aware inference".
  - Approximately 15%-30% longer confidence interval compared to oracle interval $\bar{y} \pm 1.96\sigma/\sqrt{n}$.

- SCPC easy to apply, even for (very) large n, using computational short-cuts for determination of eigenvectors if n is large (STATA and matlab code available).

- By construction, SCPC controls size in benchmark model $\mathbf{u}|\mathbf{s} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}(c))$, $c \geq c_0$.

Properties of SCPC: Roadmap

- Small sample size control in Gaussian model with "strong" correlations $\mathbf{u}|\mathbf{s} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$.
  - New inequality about null rejection probability of t-statistic with $\hat{\sigma}^2 = \hat{\mathbf{u}}'\mathbf{Q}\hat{\mathbf{u}}$ under arbitrary mixtures of covariance matrices $\boldsymbol{\Sigma}$.
  - Application to SCPC size control under mixtures of spectral densities that correspond to large class of processes that are less persistent than worst-case benchmark.
- Large sample size control under generic weak correlation.
  - Remarkably, alternative approaches to "inconsistent" spatial HAR inference do not control size when density of locations is not uniform.
- Efficiency of SCPC.

Size Control of T-statistics Under Mixtures I

- Consider Gaussian model and parametric covariance matrices $\mathbf{u}|\mathbf{s} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}^p(\theta))$, $\theta \in \Theta$.

  Seek conditions such that t-statistic is valid in *nonparametric* class of models

$$\mathbf{u}|\mathbf{s} \sim \mathcal{N}(\mathbf{0}, \int \mathbf{\Sigma}^p(\theta) d\Pi(\theta)) \quad \text{for all cdfs } \Pi.$$

- Suppose $\hat{\sigma}^2 = \hat{\mathbf{u}}'\mathbf{W}\mathbf{W}'\hat{\mathbf{u}}$, cv and $\mathbf{\Sigma}_0$ are such that

$$\mathbb{P}(|\tau| > \text{cv}) \leq \alpha \quad \text{under} \quad \mathbf{y} \sim \mathcal{N}(0, \mathbf{\Sigma}_0).$$

  Seek inequalities relating $\mathbf{\Sigma}^p(\theta)$, $\theta \in \Theta$ to $\mathbf{\Sigma}_0$ such that also

$$\mathbb{P}(|\tau| > \text{cv}) \leq \alpha \quad \text{under} \quad \mathbf{y} \sim \mathcal{N}\left(0, \int \mathbf{\Sigma}^p(\theta) d\Pi(\theta)\right)$$

  for any cdf Π on Θ.

Size Control Under Mixtures II

Theorem 1: With $\mathbf{W}^0 = (\mathbf{l}, \mathbf{W})$, let $\mathbf{\Omega}_0 = \mathbf{W}^{0\prime}\mathbf{\Sigma}_0\mathbf{W}^0$ and $\mathbf{\Omega}(\theta) = \mathbf{W}^{0\prime}\mathbf{\Sigma}^p(\theta)\mathbf{W}^0$. Suppose $\mathbf{A}_0 = \mathbf{D}(\mathrm{cv})\mathbf{\Omega}_0$ with $\mathbf{D}(\mathrm{cv}) = \mathrm{diag}(1, -\mathrm{cv}^2\,\mathbf{I}_q)$ is diagonalizable, and let $\mathbf{P}$ be its eigenvectors. Define $\mathbf{A}(\theta) = \mathbf{P}^{-1}\mathbf{D}(\mathrm{cv})\mathbf{\Omega}(\theta)\mathbf{P}$ and $\bar{\mathbf{A}}(\theta) = \frac{1}{2}(\mathbf{A}(\theta) + \mathbf{A}(\theta)')$, and suppose $\mathbf{A}_0$ and $\mathbf{A}(\theta)$, $\theta \in \Theta$ are scale normalized such that $\lambda_1(\mathbf{A}_0) = \lambda_1(\mathbf{A}(\theta)) = 1$, where $\lambda_i(\cdot)$ is the ith largest eigenvalue. Let

$$\begin{aligned} \nu_1(\theta) &= \lambda_q(-\bar{\mathbf{A}}(\theta)) - \lambda_1(\bar{\mathbf{A}}(\theta))\lambda_q(-\mathbf{A}_0) - (\lambda_1(\bar{\mathbf{A}}(\theta)) - 1) \\ \nu_i(\theta) &= \lambda_{q+1-i}(-\bar{\mathbf{A}}(\theta)) - \lambda_1(\bar{\mathbf{A}}(\theta))\lambda_{q+1-i}(-\mathbf{A}_0) \text{ for } i = 2, \ldots, q. \end{aligned}$$

If $\inf_{\theta\in\Theta}\sum_{i=1}^{j}\nu_i(\theta) \geq 0$ for all $1 \leq j \leq q$, then for any cdf Π on Θ, $\mathbb{P}(|\tau| > \mathrm{cv}) \leq \alpha$ under $\mathbf{y} \sim \mathcal{N}(0, \mathbf{\Sigma}_0)$ implies

$$\mathbb{P}(|\tau| > \mathrm{cv}) \leq \alpha \quad \text{under} \quad \mathbf{y} \sim \mathcal{N}\left(0, \int \mathbf{\Sigma}^p(\theta)d\Pi(\theta)\right).$$

Intuition

- With $\mathbf{\Omega}(\theta) = \mathbf{W}^{0\prime}\mathbf{\Sigma}^p(\theta)\mathbf{W}^0$ and $\mathbf{X} = (X_0, \mathbf{X}_{1:q}')' \sim \mathbf{\Omega}(\theta)^{1/2}\mathbf{Z} \sim \mathcal{N}(0, \mathbf{\Omega}(\theta))$

$$\begin{aligned}
\mathbb{P}\left(\tau^2 > \mathsf{cv}^2\right) &= \mathbb{P}\left(\frac{X_0^2}{\mathbf{X}_{1:q}'\mathbf{X}_{1:q}} > \mathsf{cv}^2\right) = \mathbb{P}\left(X_0^2 - \mathsf{cv}^2\,\mathbf{X}_{1:q}'\mathbf{X}_{1:q} > 0\right) \\
&= \mathbb{P}\left(\mathbf{X}'\mathbf{D}(\mathsf{cv})\mathbf{X} > 0\right) = \mathbb{P}(\mathbf{Z}'\mathbf{\Omega}(\theta)^{1/2}\mathbf{D}(\mathsf{cv})\mathbf{\Omega}(\theta)^{1/2}\mathbf{Z} > 0) \\
&= \mathbb{P}\left(\sum_{i=0}^{q}\omega_i(\theta)Z_i^2 > 0\right) = \mathbb{P}\left(Z_0^2 > -\sum_{i=1}^{q}\frac{\omega_i(\theta)}{\omega_0(\theta)}Z_i^2\right)
\end{aligned}$$

where $\mathbf{D}(\mathsf{cv}) = \mathsf{diag}(1, -\mathsf{cv}^2\,\mathbf{I}_q)$ and $\omega_i(\theta)$ are the eigenvalues of $\mathbf{\Omega}(\theta)^{1/2}\mathbf{D}(\mathsf{cv})\mathbf{\Omega}(\theta)^{1/2}$, or, equivalently, of $\mathbf{D}(\mathsf{cv})\mathbf{\Omega}(\theta)$.

- Can show: $\mathbb{P}\left(Z_0^2 > \sum_{i=1}^{q}\delta_i Z_i^2\right)$ is Schur convex in $\{\delta_i\}_{i=1}^{q}$.

- Use majorization results on eigenvalues of linear combinations of matrices, and additional calculations, to obtain result.

Robustness of SCPC Under Alternative Forms of Persistence

- If u_l is isotropic ($\sigma_u(r-s) = \tilde{\sigma}_u(||r-s||)$ for some $\tilde{\sigma}_u : \mathbb{R} \mapsto \mathbb{R}$), then spectrum $\mathfrak{f} : \mathbb{R}^d \mapsto \mathbb{R}$ satisfies $\mathfrak{f}(\boldsymbol{\omega}) = f(\omega)$ for $\omega = ||\boldsymbol{\omega}||$ and some $f : \mathbb{R} \mapsto [0, \infty)$.

- Normalize $f(0) = 1$. SCPC controls size under $f_0 = f_{c_0}^{\text{bnch}} = c_0^3/(\omega^2 + c_0^2)^{3/2}$.

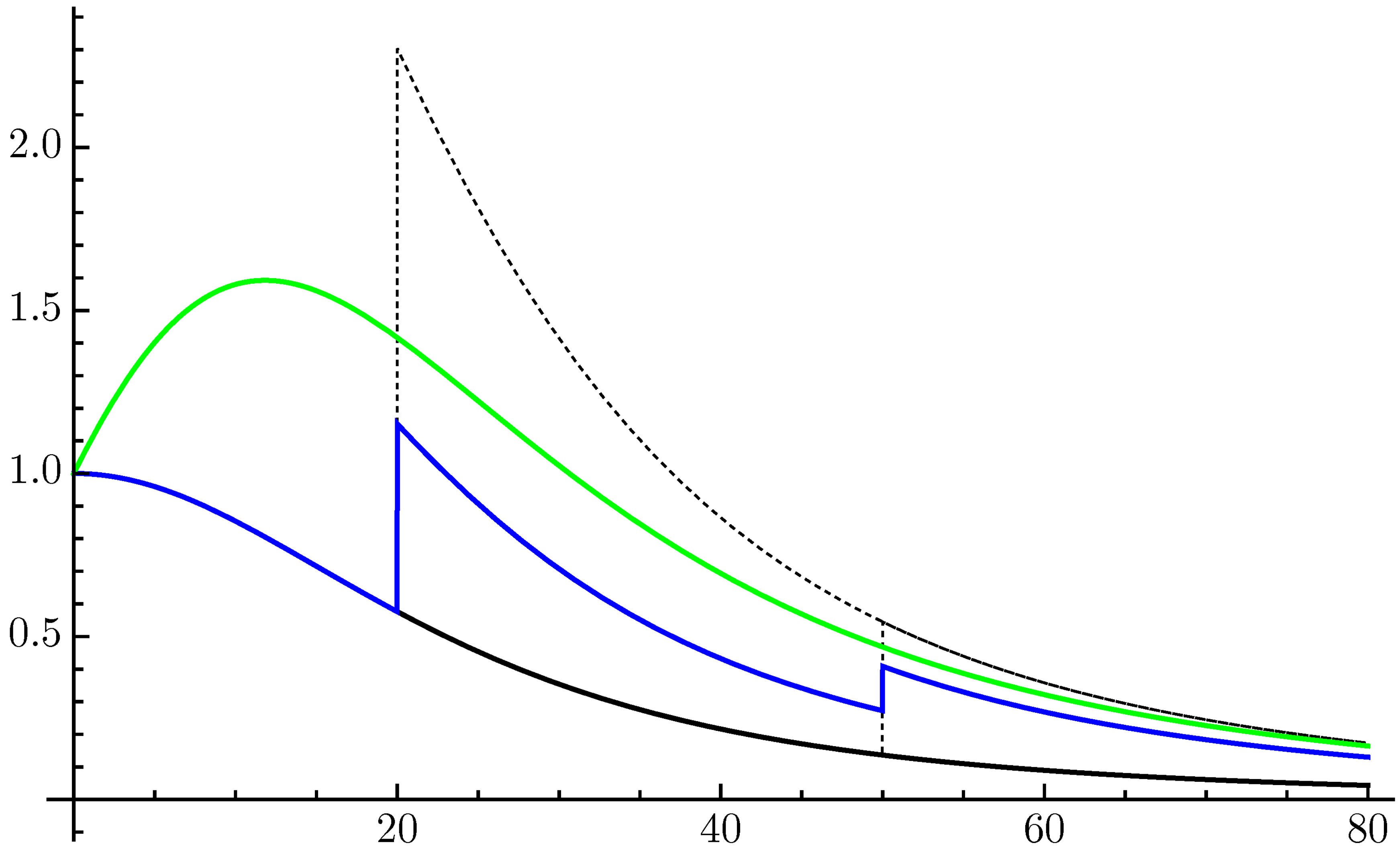
- One definition of less persistence: f is such that $r(\omega) = f(\omega)/f_0(\omega)$ (weakly) increasing in $|\omega|$.

- If $\lim_{\omega\to\infty} r(\omega) = M \geq 1$, any such r has representation $r(\omega) = 1 + (M-1)\Pi(|\omega|)$ for some CDF Π on $[0, \infty)$.

  $\Rightarrow$ Any less persistent f with $\lim_{\omega\to\infty} f(\omega)/f_0(\omega) \leq M$ has representation $f(\omega) = \int f_\theta^{\text{step}}(\omega) d\Pi(\theta)$, where $f_0^{\text{step}} = f_0$ and for $\theta > 0$

$$f_\theta^{\text{step}}(\omega) = \mathbf{1}[|\omega| \leq \theta] f_0(\omega) + \mathbf{1}[|\omega| > \theta] M \cdot f_0(\omega).$$

Robustness of SCPC Under Alternative Forms of Persistence II



Robustness of SCPC Under Alternative Forms of Persistence III

- U.S. states spatial designs:
  - Draw $n = 500$ locations from density g within each contiguous U.S. state, with $g \in \{g_{\text{uniform}}, g_{\text{light}}\}$.
  - Repeat 5 times, so obtain 240 $\mathbf{s}$ for each $g \in \{g_{\text{uniform}}, g_{\text{light}}\}$.

- Parametric covariance matrices $\mathbf{\Sigma}^p(\theta)$ implied by f_θ^{step}, $\theta \geq 0$.

- Compute $\inf_{\theta \in \Theta} \sum_{i=1}^{j} \nu_i(\theta)$ for SCPC with $\bar{\rho}_0 = 0.03$ in U.S. states spatial designs for $M = 10$.

  $\Rightarrow$ Find that SCPC inference is valid in large class of stationary Gaussian isotropic processes that are less persistent than worst-case benchmark model.

- Similar results also in classic time series case.

Weak Correlation: Set-up of Lahiri (2003)

- Location s_l are sampled i.i.d. with density g on $\mathcal{S} \subset \mathbb{R}^d$.

- Conditional on locations $\mathbf{s}_n = (s_1, \ldots, s_n)$, for some sequence $c_n > 0$,

$$u_l = B(c_n s_l)$$

with B a mean-zero stationary random field on $\mathbb{R}^d$ independent of $\mathbf{s}$ with $\mathbb{E}[B(s)B(r)] = \sigma_B(r-s)$.

- Calculation: $\bar{\rho}_n = O_p(1/c_n^d)$, so weak dependence when $c_n \to \infty$.

$\Rightarrow$ Turns out: nature of weak dependence characterized by $a_n = c_n^d/n \to a \in [0, \infty)$.

$\Rightarrow$ $a \to \infty$ corresponds to asymptotically negligible spatial correlation.

Weak Correlation: Weighted Averages

- **Lemma (Lahiri 2003):** Let $\mathbf{w}^0(s) = (1, \mathbf{w}(s)')$, $\mathbf{w}(s) = (w_1(s), \ldots, w_q(s))'$ where $w_i : \mathcal{S} \mapsto \mathbb{R}$ are continuous weight functions. Under appropriate mixing and moment assumptions on B

$$n^{-1/2} a_n^{1/2} \sum_{l=1}^{n} \mathbf{w}^0(s_l) u_l | \mathbf{s}_n \Rightarrow_p \mathcal{N}(0, \mathbf{\Omega}) \quad \text{with} \quad \mathbf{\Omega} = a\sigma_B(0)\mathbf{V}_1 + \left(\int_{\mathbb{R}^d} \sigma_B(s) ds \right) \mathbf{V}_2$$

where

$$\mathbf{V}_1 = \int_{\mathcal{S}} \mathbf{w}^0(s)\mathbf{w}^0(s)' g(s) ds \quad \text{and} \quad \mathbf{V}_2 = \int_{\mathcal{S}} \mathbf{w}^0(s)\mathbf{w}^0(s)' g(s)^2 ds.$$

$\Rightarrow$ $\mathbf{V}_1$ is what we would expect from i.i.d. data, large a corresponds to very weak correlation.
$\Rightarrow$ $\mathbf{V}_2$ proportional to $\mathbf{V}_1$ only under constant g.

- Convergence holds conditional on $\mathbf{s}_n$: Randomness in locations doesn't drive variability of weighted averages.

Weak Correlation: Projection T-statistics

- Recall: $n^{-1/2}a_n^{1/2}\sum_{l=1}^n \mathbf{w}^0(s_l)u_l|\mathbf{s}_n \Rightarrow_p \mathcal{N}(0,\mathbf{\Omega})$ with $\mathbf{\Omega} = a\sigma_B(0)\mathbf{V}_1 + \left(\int \sigma_B(s)ds\right)\mathbf{V}_2$.

- **Theorem 2:** With $\mathbf{X} = (X_0, \mathbf{X}_{1:q}')' \sim \mathcal{N}(0,\mathbf{\Omega})$, under assumptions of Lemma,

$$\mathbb{P}\left(\left|\frac{\sum_{l=1}^n u_l}{\sqrt{q^{-1}\sum_{j=1}^q \left(\sum_{l=1}^n w_j(s_l)u_l\right)^2}}\right| > \mathsf{cv}\,\middle|\,\mathbf{s}_n\right) \xrightarrow{p} \mathbb{P}\left(\frac{|X_0|}{\sqrt{q^{-1}\mathbf{X}_{1:q}'\mathbf{X}_{1:q}}} > \mathsf{cv}\right)$$

- If $\mathbf{\Omega} \propto \mathbf{V}_1 = \int \mathbf{w}^0(s)\mathbf{w}^0(s)'g(s)ds = \mathbf{I}_{q+1}$, then critical value from student-t$_q$.

  $\Rightarrow$ But $\mathbf{\Omega} \propto \mathbf{V}_1$ only if g uniform, so projection t-statistic with student-t critical value not generically valid.

  $\Rightarrow$ In light design, 5% level Sun and Kim (2012) test with 10 Fourier weights has asymptotic null rejection probability of 6.2%-30.0% under $\mathbf{\Omega} \propto \mathbf{V}_2$.

Weak Correlation: Size Control

- Since t-statistic is scale invariant, no loss of generality to normalize

$$\mathbf{\Omega} = \kappa \mathbf{V}_1 + (1-\kappa)\mathbf{V}_2, \quad \kappa \in [0,1)$$

where

$$\mathbf{V}_1 = \int_{\mathcal{S}} \mathbf{w}^0(s)\mathbf{w}^0(s)' g(s) ds \quad \text{and} \quad \mathbf{V}_2 = \int_{\mathcal{S}} \mathbf{w}^0(s)\mathbf{w}^0(s)' g(s)^2 ds.$$

$\Rightarrow$ Scalar parameter $\kappa \in [0,1)$ fully characterizes all possible limits under weak correlation.

- SCPC benchmark model $\sigma_u(s_l - s_\ell) = \exp(-c||s_l - s_\ell||)$ traces out all such limits with appropriate choices of $c = c_n \to \infty$.

Theorem 3: SCPC critical value construction $\sup_{c \geq c_0} \mathbb{P}^0_{\mathbf{\Sigma}(c)}(|\tau_{\text{SCPC}}(q)| > \text{cv}_{\text{SCPC}}(q)|\mathbf{s}) = \alpha$ implies size control under arbitrary sequences $c_n \geq c_0$, and, therefore, under generic weak correlation.

Weak Correlation: Additional Results

- Convergence of first q eigenvectors of $\mathbf{M\Sigma_0 M}$ to limiting eigenfunctions $\varphi_i : \mathcal{S} \mapsto \mathbb{R}$ in appropriate sense.

- Analogous results to those above for t-statistics with fixed-b kernel variance estimators

$$\hat{\sigma}^2 = n^{-1} \sum_{l,\ell} k\left(\frac{||s_l - s_\ell||}{b}\right) \hat{u}_l \hat{u}_\ell$$

with cv obtained from distribution in i.i.d. model, as in Bester, Conley, Hansen, and Vogelsang (2016).

$\Rightarrow$ Generic size control under weak correlation again only with g uniform.

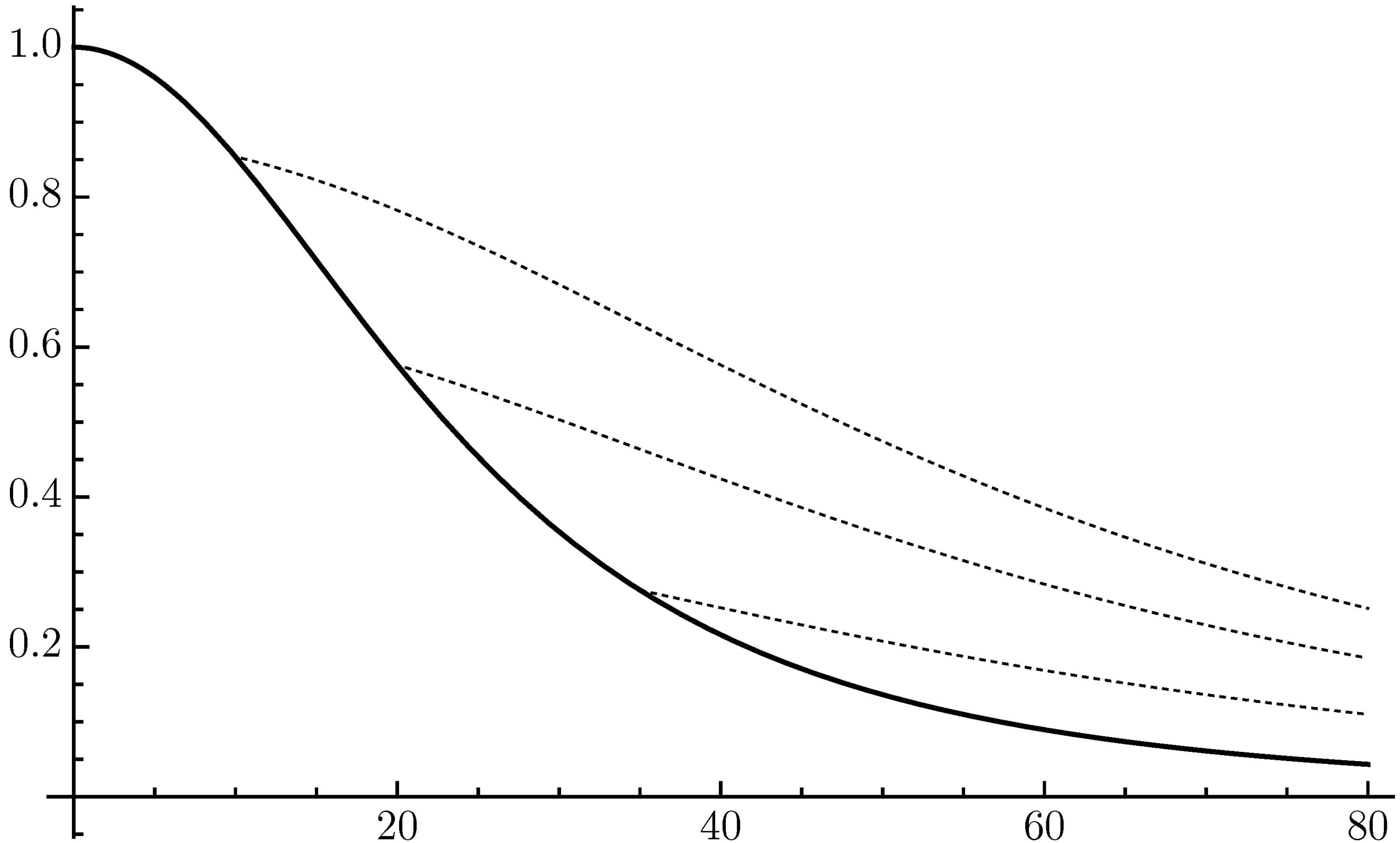
Efficiency of SCPC Confidence Intervals

- Back to small sample with $n = 500$, $\mathbf{y} \sim \mathcal{N}(\mathbf{l}\mu, \mathbf{\Sigma})$.
  For given $c_1 > c_0$, compare
  $$\mathbb{E}_{\mathbf{\Sigma}(c_1)} \left[\int \mathbf{1}[x \in \mathsf{CI}(\mathbf{y})] dx \right]$$
  under $\mathbf{y} \sim \mathcal{N}(\mathbf{l}\mu, \mathbf{\Sigma}(c_1))$, $c_1 \in \{2c_0, 5c_0, \infty\}$ ($c_1 \to \infty$ is i.i.d. model).

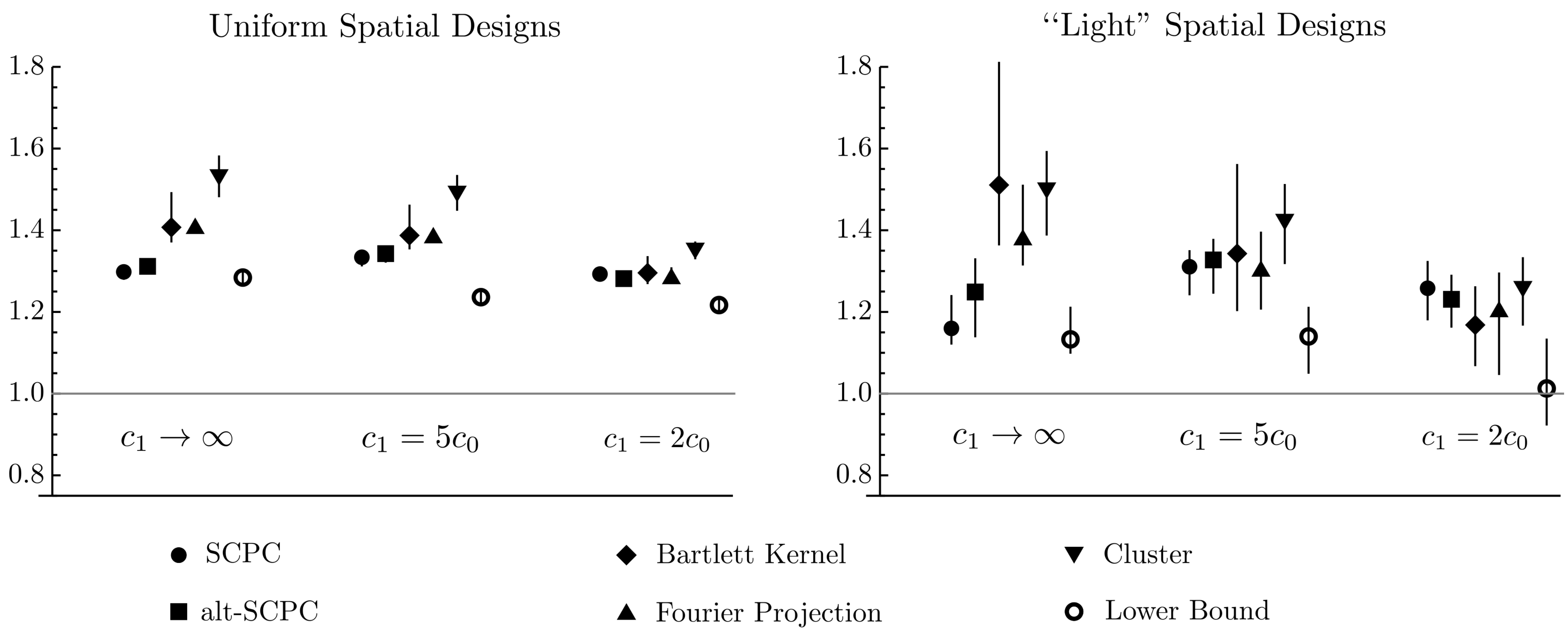
- Two comparisons: (i) with intervals from other spatial t-statistics and cv computed under $\mathbf{\Sigma}(c_0)$; (ii) with lower-bound on length of best confidence interval of the form
  $$\mathsf{CI}(\mathbf{y}) = [\bar{y} - \delta(\hat{\mathbf{u}}), \bar{y} + \delta(\hat{\mathbf{u}})]$$
  where $\delta : \mathbb{R}^n \mapsto [0, \infty)$ is scale invariant, $\delta(\lambda\hat{\mathbf{u}}) = \lambda\delta(\hat{\mathbf{u}})$, but otherwise unrestricted.

- Impose that $\mathsf{CI}(\mathbf{y})$ has no worse size control than SCPC under
  $$f_\theta^{\mathsf{kink}}(\omega) = \mathbf{1}[|\omega| \leq \theta] f_0(\omega) + \mathbf{1}[|\omega| > \theta] \frac{f_0(\theta)}{f_1(\theta)} f_1(\omega), \quad \theta > \infty$$
  for $f_0 = f_{c_0}^{\mathsf{bnch}}$ and $f_1 = f_{c_1}^{\mathsf{bnch}}$ ("approximately least favorable", as in Dou (2020)).

Efficiency of SCPC Confidence Intervals II



Efficiency of SCPC Confidence Intervals III



5th, 50th and 95th percentiles of CI length relative to oracle interval 95% confidence interval $\bar{y} \pm 1.96\sigma/\sqrt{n}$.

Lower bound computed using techniques in Müller, Elliott and Watson (2015) and Müller and Wang (2019).

Efficiency of SCPC Confidence Intervals IV

- Results imply limit on possibility of using data-dependent methods to learn about c_0:

  - Consider pre-test whether there is any spatial correlation (that is, whether one may pick $c_0 \to \infty$). If spatial correlation is detected, use very wide interval.

  - If it was possible to do this reliably, then can construct δ that (i) controls size under f_θ^{kink}; and (ii) is nearly as efficient as oracle interval in the i.i.d. case.

    $\Rightarrow$ Lower bound results imply this is impossible.

  - Same argument also for other values of c_0.

- Justifies treating c_0 (or, equivalently, $\bar{\rho}_0$) as given and construction of corresponding bias aware SCPC inference (cf. Armstrong and Kolesár (2018, 2020, 2021), etc.)

Conclusions

- New method to conduct spatial correlation robust inference that remains valid under
  - some interesting forms of strong spatial correlations
  - all forms of weak correleation
  - even when the locations are not uniformly distributed.
- Approach and results also potentially applicable in network econometrics or more generally under a given form of plausible correlation structures.