Online Supplemental Appendix to: Uniform Nonparametric Inference for Spatially Dependent Panel Data

Abstract

This supplemental appendix contains the simulation results for the proposed uniform inference procedure. Section S.1 describes the setting and Section S.2 reports the results.

S.1 The setting

We consider the following data generating process (DGP). The underlying nonparametric regression model satisfies

$$Y_{it} = g\left(X_{it}\right) + \epsilon_{it}, \quad \text{where} \quad g\left(x\right) = \frac{\exp\left(x\right)}{1+\exp\left(x\right)}.$$

Following Driscoll and Kraay (1998), we simulate ϵ_{it} from a one-factor model given by

$$\begin{cases} \epsilon_{it} = \sigma_{\epsilon}(\lambda_i f_t^{\epsilon} + v_{it}^{\epsilon}), & v_{it}^{\epsilon} \overset{i.i.d.}{\sim} \mathcal{N}\left(0, 1-\lambda_i^2\right), \\ f_t^{\epsilon} = \rho f_{t-1}^{\epsilon} + u_t^{\epsilon}, & u_t^{\epsilon} \overset{i.i.d.}{\sim} \mathcal{N}\left(0, 1-\rho^2\right), \quad \rho \in \{0, 0.5\}, \end{cases}$$

where the v_{it}^{ϵ} and u_t^{ϵ} variables are mutually independent in both time-series and cross-sectional dimensions. Under this construction, the factor f_t^{ϵ} has a standard normal distribution with autocorrelation ρ, and $\epsilon_{it} \sim \mathcal{N}\left(0, \sigma_{\epsilon}^2\right)$. The contemporaneous correlation between ϵ_{it} and ϵ_{jt} is $\lambda_i\lambda_j$ for $i \neq j$. We draw the λ_i loadings independently from the Uniform$[0,1]$ distribution so that the average cross-sectional correlation $\mathbb{E}[\lambda_i\lambda_j]$ is 0.25.

The X_{it} variables are simulated in a similar fashion according to

$$\begin{cases} X_{it} = \sigma_X(\lambda_i f_t^X + v_{it}^X), & v_{it}^X \overset{i.i.d.}{\sim} \mathcal{N}\left(0, 1-\lambda_i^2\right), \\ f_t^X = \rho f_{t-1}^X + u_t^X, & u_t^X \overset{i.i.d.}{\sim} \mathcal{N}\left(0, 1-\rho^2\right), \end{cases}$$

where the $\left(v_{it}^X, u_t^X\right)$ variables are independent of $(v_{it}^{\epsilon}, u_t^{\epsilon})$. We set the standard deviation of X_{it} to $\sigma_X = 2.5$ so that $g\left(\cdot\right)$ is visibly nonlinear on the $[-2\sigma_X, 2\sigma_X]$ interval. The resulting variance of $g\left(X_{it}\right)$ is approximately 0.12. We further set $\sigma_{\epsilon}^2 = 0.04$ so that roughly 75% of Y_{it}'s variance is contributed by $g\left(X_{it}\right)$.

We need a collection of approximating functions to implement the nonparametric series regression. Following Li et al. (2020), we first rescale X_{it} to the $[-1,1]$ interval using the transformation $x \mapsto 2\Phi\left((x-\mu_X)/s_X\right)-1$, where $\Phi\left(\cdot\right)$ is the standard normal distribution function and μ_X and s_X are the sample mean and standard deviation of X_{it}, respectively. The approximating functions are then set to be Legendre polynomials of the transformed X_{it}; recall that the kth-order Legendre polynomial is given by $\mathscr{L}_k\left(x\right) \equiv \frac{1}{2^k k!}\frac{d^k}{dx^k}\left(x^2-1\right)^k$, and $\mathscr{L}_j\left(\cdot\right)$ is orthogonal to $\mathscr{L}_k\left(\cdot\right)$ under the Lebesgue measure on $[-1,1]$ for $j \neq k$. Evidently, this type of monotone smooth transformation on X_{it} is innocuous in the nonparametric estimation context. The main purpose of doing so is to make the regressors approximately orthogonal, which generally improves the numerical stability of the series regression, especially when a large number of series terms are included.

In each simulation, we compute the proposed uniform confidence band for $g\left(\cdot\right)$ at the 95% confidence level as described in Algorithm 1 and report its uniform coverage probability calculated using 10,000 Monte Carlo replications. For comparison, we also report similar results for the confidence bands proposed by Belloni et al. (2015) and Li and Liao (2020). Recall that Belloni et al.'s method is valid for independent data, and Li and Liao's method allows for serial dependence but rules out spatial dependence. For ease of reference, we label these two benchmark methods as BCCK and LL, respectively, and refer to our proposal as the *robust* method.

We conduct Monte Carlo experiments for a variety of scenarios. Specifically, we consider $T \in \{200, 500, 2000\}$. The small-sample case with $T = 200$ may be relevant for macroeconomic settings (e.g., 50 years of quarterly data), whereas the large sample size with $T = 2000$ is easily obtainable on the daily frequency. Since our asymptotic theory relies on $T \to \infty$, we expect the proposed method to have better performance in the latter setting. To illustrate how the size of the cross section may affect inference in the presence of spatial dependence, we also consider $N \in \{5, 25, 100\}$. Intuitively, since the BCCK and LL methods ignore the (positive) spatial dependence, they tend to underestimate the sampling variability of the functional estimator, which in turn leads to under-coverage, especially when N is large. Finally, we check the robustness of the proposed procedure with respect to the number of series terms by setting $m \in \{6, 8, 10, 12\}$; for example, when $m = 12$, the approximating functions include the constant term and Legendre polynomials up to the eleventh power.

S.2 Results

Table S.1 reports the uniform coverage rates for the case with $\rho = 0$. In this setting, there is no serial correlation in the data, but spatial dependence is present. Looking at the $N = 5$ case displayed on the top panel, we see that the proposed robust method delivers adequate finite-sample coverage. In the small-sample case with $T = 200$, we observe 3–5% under-coverage; but we can confirm that this is mainly a small-sample issue, because the size distortion essentially disappears when $T = 2000$. These findings also appear to be robust with respect to the number of series terms m. On the other hand, the BCCK and LL methods both exhibit more severe size distortions than the robust method. It is important to note that increasing T does little for reducing the under-coverage of these benchmark methods, which suggests that their lack of coverage is not a small-sample issue, but rather because they fail to account for spatial dependence in the data.

The contrast between the robust method and the two benchmarks become even more clear for larger cross sections (i.e., $N = 25$ or 100). We only focus on the $N = 100$ case for brevity. From the bottom panel of Table S.1, we see that the coverage rates of the proposed robust method are quite close to the 95% nominal level. In sharp contrast, the coverage

Table S.1: Coverage Rates of Uniform Confidence Bands ($\rho = 0$)

	Robust			BCCK			LL		
T	200	500	2000	200	500	2000	200	500	2000
				Case I: $N = 5$					
$m = 6$	0.923	0.937	0.943	0.883	0.893	0.892	0.885	0.895	0.892
$m = 8$	0.910	0.936	0.944	0.885	0.898	0.906	0.893	0.904	0.905
$m = 10$	0.903	0.930	0.941	0.887	0.903	0.910	0.900	0.906	0.910
$m = 12$	0.894	0.922	0.942	0.885	0.899	0.914	0.903	0.905	0.916
				Case II: $N = 25$					
$m = 6$	0.923	0.934	0.938	0.617	0.614	0.589	0.630	0.620	0.589
$m = 8$	0.920	0.935	0.946	0.672	0.658	0.666	0.691	0.665	0.670
$m = 10$	0.925	0.937	0.947	0.703	0.707	0.705	0.733	0.718	0.706
$m = 12$	0.917	0.934	0.944	0.730	0.731	0.730	0.763	0.748	0.734
				Case III: $N = 100$					
$m = 6$	0.927	0.933	0.930	0.209	0.205	0.151	0.220	0.210	0.152
$m = 8$	0.926	0.934	0.947	0.261	0.263	0.259	0.279	0.269	0.260
$m = 10$	0.922	0.940	0.947	0.293	0.298	0.291	0.322	0.308	0.293
$m = 12$	0.921	0.934	0.943	0.327	0.324	0.327	0.362	0.338	0.330

Note: This table reports the coverage rates of 95% uniform confidence bands for the proposed (spatio-temporal) robust method, as well as the methods of Belloni et al. (2015) and Li and Liao (2020), labeled as robust, BCCK, and LL, respectively. The autoregressive coefficient is set to $\rho = 0$. The HAC estimation is performed using the Newey–West estimator with bandwidth $\lfloor 0.75T^{1/3} \rfloor$. The coverage rates are computed based on $10,000$ Monte Carlo replications.

rates of the BCCK and LL methods range from 15% to 36%, exhibiting extremely severe under-coverage in this large-N scenario.

We next turn to the case with $\rho = 0.5$, so that both serial and spatial dependence are present in the data. The results are reported in Table S.2. Here, we observe that the robust method has more under-coverage than it does in the $\rho = 0$ case, which might be attributed to the well-known difficulty in HAC estimation in the presence of serial dependence. Nevertheless, we still find that the proposed robust method has far better coverage properties than the BCCK and LL methods.

Overall, these simulation results show that the proposed uniform confidence band has adequate finite-sample coverage properties. Its performance appears to be robust with respect to the size of the cross section and the number of approximating functions used in the nonparametric series regression. In contrast, in the presence of spatial dependence, the benchmark BCCK and LL methods suffer from rather severe size distortions, especially when N is large. In view of these findings, we recommend using the proposed method in panel-data applications. This message extends the lesson of Driscoll and Kraay (1998) to the present new context concerning uniform nonparametric inference.

Table S.2: Coverage Rates of Uniform Confidence Bands ($\rho = 0.5$)

	Robust			BCCK			LL		
T	200	500	2000	200	500	2000	200	500	2000
				Case I: $N = 5$					
$m = 6$	0.893	0.916	0.935	0.774	0.774	0.780	0.813	0.814	0.826
$m = 8$	0.887	0.914	0.933	0.796	0.807	0.810	0.832	0.839	0.847
$m = 10$	0.875	0.908	0.933	0.807	0.820	0.828	0.842	0.847	0.856
$m = 12$	0.860	0.906	0.930	0.804	0.831	0.837	0.843	0.861	0.860
				Case II: $N = 25$					
$m = 6$	0.872	0.901	0.917	0.397	0.389	0.370	0.442	0.435	0.412
$m = 8$	0.875	0.907	0.929	0.454	0.447	0.456	0.504	0.489	0.493
$m = 10$	0.869	0.907	0.929	0.481	0.495	0.494	0.535	0.535	0.531
$m = 12$	0.868	0.907	0.930	0.524	0.531	0.526	0.582	0.570	0.559
				Case III: $N = 100$					
$m = 6$	0.869	0.896	0.918	0.100	0.097	0.067	0.121	0.116	0.084
$m = 8$	0.867	0.902	0.922	0.136	0.136	0.129	0.158	0.155	0.146
$m = 10$	0.862	0.894	0.920	0.154	0.148	0.141	0.181	0.168	0.157
$m = 12$	0.866	0.899	0.923	0.180	0.166	0.175	0.215	0.187	0.193

Note: This table reports the coverage rates of 95% uniform confidence bands for the proposed (spatio-temporal) robust method, as well as the methods of Belloni et al. (2015) and Li and Liao (2020), labeled as robust, BCCK, and LL, respectively. The autoregressive coefficient is set to $\rho = 0.5$. The HAC estimation is performed using the Newey–West estimator with bandwidth $\lfloor 0.75T^{1/3} \rfloor$. The coverage rates are computed based on 10,000 Monte Carlo replications.

References

Belloni, A., V. Chernozhukov, D. Chetverikov, and K. Kato (2015). Some new asymptotic theory for least squares series: Pointwise and uniform results. *Journal of Econometrics 186*(2), 345 – 366.

Driscoll, J. C. and A. C. Kraay (1998). Consistent covariance matrix estimation with spatially dependent panel data. *The Review of Economics and Statistics 80*(4), 549–560.

Li, J. and Z. Liao (2020). Uniform nonparametric inference for time series. *Journal of Econometrics 219*(1), 28–51.

Li, J., Z. Liao, and R. Quaedvlieg (2020). Conditional superior predictive ability. *Review of Economic Studies, forthcoming*.