A network regression model with an estimated interaction matrix

Jose Olmo
Univ. Zaragoza & Southampton

Marcos Sanso-Navarro
Univ. Zaragoza

EEA-ESEM 2022. Boconi University, Milan
22 August, 2022

Table of Contents

Background literature

The spatial econometrics literature has developed models that treat three different types of interactions across units:

- Endogenous interactions affecting the dependent variable (endogenous peer effects).
- Exogenous interactions in the explanatory variables (contextual effects)
- Interaction effects among the error terms (cross-correlation between unobserved components)

A standard formulation of a network regression is the linear-in-means model of peer effects introduced by Manski (1993, RES). In this model agents' outcomes depend on their own characteristics, their peers' characteristics, and their peers' outcomes.

Contributions

- This paper proposes a network regression model that incorporates exogenous and endogenous neighboring effects.
- The network coefficients are modelled as a functional coefficient $w(d)$, with d the network variable, see Fan and Gijbels (1996) for local polynomial approximations, and Park et al (2009, JASA) for B-splines.
- Use local Taylor expansions of the function $w(d)$ over a partition of the support of the network variable from a set of disjoint intervals.
- Derive the asymptotic properties of the estimator of the functional network coefficient and propose pointwise and uniform tests for network effects.
- Study Environmental Engel curves and find strong neighboring effects between households' income and the amount of pollution embodied in their consumption.

The baseline model

A Nonparametric NLX model:

$$y_i = \lambda x_i + \sum_{\substack{j=1 \\ j \neq i}}^{N} w(d_{ij}) x_j + \varepsilon_i, \; i = 1, \ldots, N, \tag{1}$$

with $w(d_{ij})$ the functional coefficient with the network effects and $d_{ij} = f(z_i, z_j)$, for $\{z_i, z_j\}$ realizations of a random variable Z evaluated at different units.

- Use a partition of the support of the network variable $d \in \chi$ into K disjoint intervals characterized by a grid of K points $\{z_1, \ldots, z_K\}$.
- Let $[z_k - h, z_k + h)$ be a generic interval of the partition and let $p_k = P\{d \in [z_k - h, z_k + h)\}$ be the probability of belonging to a given interval and such that $\sum_{k=1}^{K} p_k = 1$.

Assumptions I

(A1) $\{(x_i, z_i, \varepsilon_i)\}$ is an *iid* sequence across index i and y_i is generated from model (1). The regressor $E[x_i^4] < \infty$, for $i = 1, \ldots, N$.

(A2) The functional coefficient $w(d)$ is $(q+1)$-times continuously differentiable on (and extension of) the compact set $\chi \subset \mathbb{R}^+$, with $q \geq 0$ fixed.

(A3) The network variable $d \in \chi$ is continuously distributed with Lebesgue density that is bounded, and bounded away from zero on χ.

(A4) $E[\varepsilon_i \mid X_i = x, D_i = \overline{d}_i] = 0$ for $\overline{d}_i = \{d_{i1}, \ldots, d_{i,i-1}, d_{i,i+1}, \ldots, d_{iN}\}$; $\sigma^2(x, \overline{d}_i) = E(\varepsilon_i^2 \mid X_i = x, D_i = \overline{d}_i)$ is continuous and bounded away from zero, and $E[\varepsilon_i^4 \mid X_i = x, D_i = \overline{d}_i] < \infty$, for all i and any $(x, \overline{d}_i) \in R \times \chi^{N-1}$.

Assumptions II

(A5) Let K denote the number of disjoint intervals covering the compact set χ. Then, we require $K/N \to 0$ and $N/K^{q+1} \to 0$ as $K, N \to \infty$; q is the order of the Taylor expansion (a fixed parameter).

(A6) The number of intervals K depends on the tuning parameter h such that $h \asymp K^{-1}$, where for scalars a and b, $a \asymp b$ denotes that $C_* \leq a \leq C^* b$ for positive constants C_* and C^*. Similarly, we assume $p_k \asymp K^{-1}$. By construction, $p_k = K^{-1}$ if p_k is exactly the same across intervals.

The functional coefficient capturing network effects

$$w(d) = \sum_{k=1}^{K} \sum_{m=0}^{q} \frac{1}{m!} w^{(m)}(z_k)(d - z_k)^m 1_k(d) + R(d), \quad (2)$$

with $w^{(m)}(z_k)$ the m^{th}−derivative of $w(\cdot)$ evaluated at z_k and

$R(d) = \sum_{k=1}^{K} w^{(q+1)}(c_k)(d - z_k)^{q+1} 1_k(d)$, with $c_k \in (z_k - h, z_k + h)$.

Plugging-in in the above expression, we obtain

$$y_i = \lambda x_i + \sum_{k=1}^{K} \sum_{m=0}^{q} \gamma_{km} \widetilde{x}_i^{(km)} + \overline{R}_i + \varepsilon_i, \quad (3)$$

with $\widetilde{x}_i^{(km)} = \sum_{\substack{j=1 \\ j \neq i}}^{N} x_j (d_{ij} - z_k)^m 1_k(d_{ij})$ be regression variables;

$\gamma_{km} = \frac{1}{m!} w^{(m)}(z_k)$ and $\overline{R}_i = \sum_{\substack{j=1 \\ j \neq i}}^{N} R(d_{ij}) x_j$.

Partitioning estimator for NLX model

The NLX model in matrix form (without $\overline{R}$) is $Y = X\lambda + \mathbb{X}\,\Gamma + \varepsilon$. Using the partitioned inverse, a suitable estimator of λ is

$$\widehat{\lambda} = \left(\widehat{X}_u'\widehat{X}_u\right)^{-1}\widehat{X}_u'(Y - \widetilde{Y}), \quad (4)$$

with $\widehat{X}_u = M_{\mathbb{X}}X$, where $M_{\mathbb{X}} = I_N - \mathbb{P}_{\mathbb{X}}$ and $\mathbb{P}_{\mathbb{X}} = \mathbb{X}(\mathbb{X}'\mathbb{X})^{-1}\mathbb{X}'$; $\widetilde{Y} = \mathbb{P}_{\mathbb{X}}Y$ is the projection of Y on $\mathbb{X} = [\mathbb{X}_1, \ldots, \mathbb{X}_K]$. Similarly,

$$\widehat{\Gamma}_k = \left(\sum_{i=1}^{N}\mathbb{X}_{ki}'\mathbb{X}_{ki}\right)^{-1}\sum_{i=1}^{N}\mathbb{X}_{ki}'(y_i - x_i\widehat{\lambda}), \quad (5)$$

for each $\widehat{\Gamma}_k = (\widehat{\gamma}_{k0}, \ldots, \widehat{\gamma}_{kq})'$. Then,

$$\widehat{w}(d) = \sum_{k=1}^{K}\sum_{m=0}^{q}\widehat{\gamma}_{km}(d - z_k)^m 1_k(d) = \sum_{k=1}^{K}\widehat{\Gamma}_k' v_k(d) 1_k(d), \quad (6)$$

with $v_k(d) = [1, (d - z_k), (d - z_k)^2, \ldots, (d - z_k)^q]'$.

Variance estimation

Let $\Phi_0 = E[X' M_{\mathbb{X}} X]$ and $\Psi_0 = E[X' M_{\mathbb{X}} X \varepsilon^2]$. The sample counterparts are $\widehat{\Phi} = \frac{1}{N}\sum_{i=1}^{N} \widehat{X}'_{ui}\widehat{X}_{ui}$ and $\widehat{\Psi} = \frac{1}{N}\sum_{i=1}^{N} \widehat{X}'_{ui}\widehat{X}_{ui} e_i^2$, where $e_i = y_i - \widehat{\lambda} x_i - \sum_{k=1}^{K} \mathbb{X}_{ki}\, \widehat{\Gamma}_k$. Then,

$$\widehat{V}\left(\widehat{\lambda}\right) = \frac{1}{N}\widehat{\Phi}^{-1}\widehat{\Psi}\widehat{\Phi}^{-1}. \tag{7}$$

Similarly,

$$\widehat{V}\left(\widehat{\Gamma}_k\right) = \frac{1}{\alpha_N p_k}\widehat{Q}_k^{-1}\widehat{A}_k\widehat{Q}_k^{-1}, \tag{8}$$

with $\alpha_N = N(N-1)$ a standardizing constant for the network coefficient.

Note that $Q_k = \frac{1}{(N-1)p_k} E[\mathbb{X}'_{ki}\mathbb{X}_{ki}] = E[\overline{X}'_{k,ij}\overline{X}_{k,ij}]/p_k$, with $\overline{X}_{k,ij} = (x_j 1_k(d_{ij}), x_j(d_{ij} - z_k)1_k(d_{ij}), \ldots, x_j(d_{ij} - z_k)^q 1_k(d_{ij}))$. Then,

$$\widehat{Q}_k = \frac{1}{\alpha_N p_k} \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} \overline{X}'_{k,ij}\overline{X}_{k,ij}.$$

Furthermore,

$$\widehat{A}_k = \frac{1}{\alpha_N p_k} \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} \overline{X}'_{k,ij}\overline{X}_{k,ij} e_i^2$$

is a suitable estimator of $A_k = \frac{1}{(N-1)p_k} E[\mathbb{X}'_{ki}\mathbb{X}_{ki}\varepsilon_i^2] = E[\overline{X}'_{k,ij}\overline{X}_{k,ij}\varepsilon_i^2]/p_k$.

A natural estimator of the variance of $\widehat{w}(d)$ in (6) is

$$\widehat{V}(\widehat{w}(d)) = \sum_{k=1}^{K} v'_k(d)\widehat{V}(\widehat{\Gamma}_k)v_k(d)1_k(d). \tag{9}$$

Asymptotic convergence of the NLX estimator

Assumption B. The matrices Φ_0 and Ψ_0 are positive definite, such that $\|\Phi_0\| < \infty$, $\|\Psi_0\| < \infty$, $\|\Phi_0^{-1}\| < \infty$ and $\|\Psi_0^{-1}\| < \infty$. Similarly, we impose $E[\|X' M_{\mathbb{X}} X\|^2] < \infty$. We also assume $\|Q_k\| < \infty$ and $\|A_k\| < \infty$, for $k = 1, \ldots, K$.

Lemma 1.- *Under assumptions A and B, for $k = 1, \ldots, K$, it follows that $\|\widehat{Q}_k - Q_k\| = O_p\left(\frac{K^{\nu}}{\sqrt{N}}\right)$, with $\nu = 0$, if $d_{ij} \neq d_{ji}$, and $\nu = 1/2$ if the network variable is symmetric ($d_{ij} = d_{ji}$, for all $i, j = 1, \ldots N$) as $K, N \to \infty$.*

Proposition 1.- *Under assumptions A and B, it follows that $\|\widehat{\lambda} - \lambda\| = O_p(1/\sqrt{N})$ and $\|\widehat{\Gamma}_k - \Gamma_k\| = O_p(\sqrt{K}/N)$, for $k = 1, \ldots, K$, as $K, N \to \infty$.*

Proposition 2.- *Under assumptions A and B,*

$$\sqrt{N}(\widehat{\lambda} - \lambda) \xrightarrow{d} N\left(0, \Phi_0^{-1}\Psi_0\Phi_0^{-1}\right). \tag{10}$$

These results allow us to derive the uniform convergence of the estimator of the functional coefficient.

Theorem 1.- *Under assumptions A and B,*

$$\sup_{d \in \chi} |\widehat{w}(d) - w(d)| = O_p\left(\sqrt{K}/N + K^{-(q+1)}\right). \tag{11}$$

The uniform convergence is determined by a variance term $\sqrt{K}/N$ given by the estimation of the network parameters and a bias term $K^{-(q+1)}$ driven by the approximation error of the remainder terms of the Taylor expansions.

Asymptotic convergence of variance estimator

Let $V_K(d) \equiv \alpha_N \sum_{k=1}^{K} v_k'(d) V(\widehat{\Gamma}_k) v_k(d) 1_k(d)$ and
$\widehat{V}_K(d) = \sum_{k=1}^{K} v_k'(d) \widehat{Q}_k^{-1} \widehat{A}_k \widehat{Q}_k^{-1} v_k(d) 1_k(d)/p_k$.

Proposition 3.- *Under assumptions A and B, for $k = 1, \ldots, K$ and any $d \in \chi$ fixed, it holds that*

(i) $|\widehat{V}_K(d) - V_K(d)| = O_p(K/N)$.

(ii) $V\left(\sqrt{\alpha_N}(\widehat{w}(d) - w(d))\right) = V_K(d) + O(N/K^{2(q+1)})$.

Theorem 2.- *Under assumptions A and B, for any $d \in \chi$ fixed, it follows that*

$$\sqrt{\alpha_N}\, \frac{\widehat{w}(d) - w(d)}{\widehat{V}_K^{1/2}(d)} \xrightarrow{d} N(0,1). \qquad (12)$$

Weak convergence and uniform tests

The stochastic process $\widehat{w}(d)$ is not asymptotically tight.

The weak convergence of $\widehat{w}(d)$ can be obtained adapting the strong approximation results in Cattaneo et al. (2020, Annals of Statistics).

Proposition 4.- *Under assumptions A and B, the estimator $\widehat{w}(d)$, for $d \in \chi$, satisfies that*

$$\sqrt{\alpha_N}\,\frac{\widehat{w}(d) - w(d)}{\widehat{V}_K^{1/2}(d)} \xrightarrow{w} \mathbb{G}(d), \tag{13}$$

with $\xrightarrow{w}$ denoting weak convergence and $\mathbb{G}(d)$ a zero-mean Gaussian process defined on $d \in \chi$.

The asymptotic distribution of the supremum functional is

$$\sqrt{\alpha_N}\sup_{d\in\chi}\left|\frac{\widehat{w}(d) - w(d))}{\widehat{V}_K^{1/2}(d)}\right| \xrightarrow{d} \sup_{d\in\chi}\left|\mathbb{G}(d)\right|, \text{ as } N \to \infty. \tag{14}$$

Hypothesis testing

Testing for the presence of network effects can be formulated as $H_0 : \sup_{d \in \chi} |w(d)| = 0$, against the alternative $H_A : \sup_{d \in \chi} |w(d)| > 0$.

Let $f(\cdot)$ be some known functional specification. Then, the null hypothesis can be modified such that $H_{0f} : \sup_{d \in \chi} |w(d) - f(d)| = 0$, with of $d \in \chi$, against the alternative $H_{Af} : \sup_{d \in \chi} |w(d) - f(d)| > 0$.

We propose the test statistic

$$T_N = \sqrt{\alpha_N} \sup_{d \in \chi} \left| \frac{\widehat{w}(d) - f(d)}{\widehat{V}_K^{1/2}(d)} \right|, \tag{15}$$

where the functional form $f(d)$ depends on the null hypothesis under study.

Approximating the critical values

Obtaining asymptotic critical values for these tests is difficult because the asymptotic distribution is non-standard and cannot be tabulated.

Apply a p-value transformation method for testing the null hypothesis of interest. Operate conditionally on a realization of $\{(x_i, y_i)\}_{i=1}^N$, denoted as ω_N. Then,

$$\sqrt{\alpha_N}\frac{\widehat{w}(d) - w(d)}{V_k^{1/2}(d)} = \frac{\frac{1}{\sqrt{\alpha_N} p_k}\sum_{i=1}^{N}\sum_{k=1}^{K} v_k(d)' \widehat{Q}_k^{-1} \mathbb{X}_{ki}' e_{0i} 1_k(d)}{V_k^{1/2}(d)} + o_p(1),$$

where e_{0i} are the residuals of the data generating process obtained under the null hypothesis, e.g. $e_{0i} = y_i - X_i\widehat{\lambda}$, with $\widehat{\lambda}$ the OLS estimator of the regression model without network effects obtained under the null hypothesis $H_0 : \sup_{d\in\chi} |w(d)| = 0$.

Construct independent replicas of T_N under H_0

Let $\mathbb{G}_N^\star$ be a conditional zero-mean Gaussian process with the same covariance kernel as $\mathbb{G}(d)$. This process can be simulated by generating a vector of *iid* random variables $\epsilon = (\epsilon_1, \ldots, \epsilon_N)'$ to construct the simulated residuals $e_0^\star = e_0 \otimes \epsilon$. Then,

$$\mathbb{G}_N^\star(d) = \frac{\frac{1}{\sqrt{\alpha_N} p_k} \sum_{i=1}^{N} \sum_{k=1}^{K} v_k(d)' \widehat{Q}_k^{-1} \mathbb{X}_{ki}' e_{0i}^\star 1_k(d)}{V_k^{1/2}(d)}, \tag{16}$$

and $T_N^\star = \sup_{d \in \chi} |\mathbb{G}_N^\star(d)|$. Then,

$$P_{\omega_N}\{T_N^\star > T_N\} \to P_{H_0}\left\{T_N > \sup_{d \in \chi} |\mathbb{G}(d)|\right\}, \quad \text{as } N \to \infty, \tag{17}$$

with P_{ω_N} the probability distribution conditional on the sample ω_N, and P_{H_0} the probability distribution of $\sup_{d \in \chi} |\mathbb{G}(d)|$ under the null hypothesis.

Model selection methods for the tuning parameters

$$\text{Mallows (1973):} \widehat{h}_M = \underset{\{h\}}{\arg\min} \left\{ \widehat{\sigma}_e^2 \left(1 + \frac{C}{Nh} \right) \right\}, \tag{18}$$

with $\widehat{\sigma}_e^2 = \frac{1}{N} \sum_{i=1}^{N} e_i^2$ obtained under conditional homoscedasticity.

$$\text{Craven and Wahba (1978):} \widehat{h}_{GCV} = \underset{\{h\}}{\arg\min} \left\{ \frac{\widehat{\sigma}_e^2}{\left(1 - \frac{C}{Nh}\right)^2} \right\}. \tag{19}$$

Methods accounting for the order of the Taylor expansion (q):

$$\textit{Akaike} : \widehat{h}_{AIC} = \underset{\{h,q\}}{\arg\min} \left\{ \ln \widehat{\sigma}_e^2 + 2 \frac{(q+1)\lceil C/2h \rceil + 1}{N} \right\}, \tag{20}$$

$$\textit{Bayesian} : \widehat{h}_{BIC} = \underset{\{h,q\}}{\arg\min} \left\{ \ln \widehat{\sigma}_e^2 + \frac{((q+1)\lceil C/2h \rceil + 1) \ln N}{N} \right\}. \tag{21}$$

Monte-Carlo simulations

$$DGP: y_i = x_i\gamma + \sum_{\substack{j=1 \\ j\neq i}}^{N} w(d_{ij})x_j + \varepsilon_i, \text{ for } i = 1, \ldots, N, \tag{22}$$

with $x_i \sim N(0,1)$; $d_{ij} = |x_i - x_j|$ and $\varepsilon_i \sim \text{iid}N(0, \log^2|1 + x_i|)$.

Two specifications of the functional network effects for the simulation exercise:

- Exponential function $w(d_{ij}) = \beta \exp(-\theta d_{ij})$.
- Gaussian kernel $w(d_{ij}) = \beta \exp\left(-\frac{1}{2}(\theta d_{ij})^2\right)$, with $\beta, \theta > 0$.

We choose $\lambda = 1, 0.25$; $\beta = 0.1$ and $\theta = 5, 9$, for $N = 100, 200, 500$ and $h = 0.05, 0.075, 0.10$.

Bias of estimators of λ and $w(d)$.

				Model 1: Exponential function					Model 2: Gaussian kernel function				
λ	h	θ	N	λ	$w(h/2)$	$w(h)$	$w(3h/2)$	$w(2h)$	λ	$w(h/2)$	$w(h)$	$w(3h/2)$	$w(2h)$
1	0.05	5	100	0.144	-0.006	0.011	0.018	0.017	0.004	0.002	0.001	0.003	0.007
			250	0.095	0.010	0.014	0.015	0.014	0.046	0.005	0.003	0.003	0.005
			500	0.064	0.007	0.010	0.011	0.012	0.078	0.005	0.004	0.005	0.006
		9	100	0.011	0.000	-0.001	-0.005	-0.005	0.010	0.010	-0.002	-0.001	0.005
			250	0.030	-0.005	0.004	0.003	0.002	-0.002	0.009	-0.002	-0.001	0.003
			500	0.033	-0.003	0.003	0.001	-0.001	-0.002	0.005	-0.003	-0.001	0.003
	0.075	5	100	0.055	-0.001	-0.001	-0.002	0.000	0.006	0.013	0.008	0.006	0.003
			250	0.115	-0.005	0.000	-0.001	-0.002	0.001	0.003	-0.002	-0.001	0.001
			500	0.016	-0.006	-0.002	-0.002	-0.003	0.004	0.003	-0.002	-0.001	0.002
		9	100	0.008	-0.016	0.004	0.002	0.000	-0.019	0.024	-0.010	0.000	0.008
			250	-0.001	-0.017	0.005	0.001	-0.005	-0.007	0.030	-0.009	-0.002	0.009
			500	0.009	-0.012	0.005	-0.001	-0.005	0.000	0.023	-0.011	0.000	0.012
	0.1	5	100	0.015	-0.010	0.003	0.002	-0.003	0.003	0.012	-0.004	-0.001	0.003
			250	0.032	-0.007	0.001	-0.001	-0.004	0.008	0.009	-0.004	-0.001	0.004
			500	0.058	-0.008	0.000	-0.002	-0.003	-0.001	0.009	-0.004	0.000	0.005
		9	100	-0.020	-0.020	0.013	0.001	-0.008	0.000	0.058	-0.024	0.003	0.026
			250	0.002	-0.021	0.011	-0.001	-0.010	-0.003	0.055	-0.026	0.001	0.024
			500	0.002	-0.022	0.010	-0.001	-0.009	-0.008	0.054	-0.026	0.002	0.025
0.25	0.05	5	100	0.030	0.002	0.002	0.002	0.004	0.006	-0.005	0.001	0.004	0.003
			250	0.075	0.003	0.003	0.003	0.004	0.013	0.004	0.001	-0.001	0.000
			500	0.044	0.001	0.002	0.003	0.002	0.020	0.001	0.001	0.001	0.001
		9	100	-0.006	0.000	0.002	0.001	0.002	-0.011	0.003	-0.001	-0.001	0.003
			250	0.007	-0.003	0.000	0.001	0.001	0.000	0.003	0.001	0.000	0.000
			500	0.009	0.002	0.001	-0.001	-0.001	-0.001	0.002	-0.001	-0.001	0.000
	0.075	5	100	0.004	0.007	0.002	-0.001	0.000	0.006	0.006	0.002	0.000	-0.002
			250	0.032	0.002	0.002	0.001	0.000	0.000	-0.002	-0.001	0.000	0.001
			500	0.058	0.000	-0.001	-0.001	-0.001	-0.005	0.000	0.000	0.001	0.000
		9	100	-0.001	-0.005	0.001	-0.001	-0.004	-0.002	-0.002	-0.003	0.004	0.006
			250	0.002	-0.005	0.000	-0.001	-0.002	-0.002	0.007	-0.003	-0.001	0.001
			500	0.002	-0.002	0.001	-0.001	-0.002	0.001	0.006	-0.003	-0.001	0.003
	0.1	5	100	-0.013	-0.008	0.002	0.004	0.003	-0.003	0.002	0.000	0.001	-0.001
			250	0.015	-0.006	0.000	0.001	0.001	-0.005	0.006	-0.001	-0.002	-0.001
			500	0.013	0.000	0.000	-0.001	-0.001	-0.002	0.003	0.000	0.001	0.001
		9	100	0.011	0.003	0.004	-0.003	-0.006	0.005	0.007	-0.007	0.004	0.012
			250	0.001	-0.005	0.002	-0.001	-0.002	0.000	0.013	-0.006	0.002	0.006
			500	0.005	-0.004	0.003	-0.001	-0.002	-0.008	0.015	-0.007	0.000	0.006

Root mean square error of estimators of λ and $w(d)$.

				Model 1: Exponential function					Model 2: Gaussian kernel function				
λ	h	θ	N	λ	$w(h/2)$	$w(h)$	$w(3h/2)$	$w(2h)$	λ	$w(h/2)$	$w(h)$	$w(3h/2)$	$w(2h)$
1	0.05	5	100	0.251	0.242	0.109	0.117	0.102	0.205	0.239	0.100	0.108	0.100
			250	0.128	0.091	0.046	0.048	0.043	0.125	0.086	0.037	0.041	0.036
			500	0.070	0.052	0.030	0.030	0.028	0.116	0.043	0.019	0.021	0.019
		9	100	0.201	0.242	0.101	0.108	0.095	0.210	0.232	0.101	0.107	0.097
			250	0.122	0.086	0.037	0.039	0.036	0.120	0.086	0.037	0.041	0.036
			500	0.093	0.039	0.017	0.019	0.018	0.084	0.042	0.018	0.020	0.018
	0.075	5	100	0.234	0.205	0.084	0.091	0.083	0.220	0.201	0.087	0.090	0.080
			250	0.176	0.072	0.030	0.033	0.029	0.135	0.071	0.030	0.033	0.029
			500	0.142	0.037	0.016	0.017	0.016	0.099	0.034	0.014	0.016	0.015
		9	100	0.220	0.200	0.083	0.092	0.081	0.224	0.206	0.085	0.091	0.080
			250	0.134	0.069	0.029	0.032	0.030	0.137	0.077	0.033	0.034	0.032
			500	0.099	0.036	0.016	0.016	0.015	0.101	0.042	0.019	0.016	0.018
	0.1	5	100	0.244	0.179	0.072	0.079	0.069	0.253	0.183	0.074	0.078	0.071
			250	0.164	0.064	0.025	0.030	0.027	0.162	0.064	0.026	0.029	0.026
			500	0.143	0.033	0.013	0.014	0.013	0.128	0.032	0.013	0.014	0.014
		9	100	0.259	0.180	0.077	0.083	0.071	0.248	0.189	0.078	0.079	0.074
			250	0.157	0.067	0.028	0.029	0.027	0.162	0.086	0.037	0.029	0.035
			500	0.124	0.039	0.017	0.014	0.016	0.128	0.062	0.029	0.014	0.028
0.25	0.05	5	100	0.219	0.244	0.099	0.109	0.099	0.204	0.242	0.101	0.110	0.096
			250	0.136	0.085	0.037	0.041	0.037	0.121	0.086	0.037	0.041	0.035
			500	0.066	0.041	0.019	0.020	0.017	0.086	0.041	0.018	0.020	0.017
		9	100	0.199	0.247	0.101	0.108	0.098	0.215	0.250	0.105	0.115	0.100
			250	0.125	0.086	0.037	0.039	0.035	0.122	0.085	0.038	0.041	0.037
			500	0.084	0.041	0.017	0.019	0.017	0.088	0.041	0.018	0.019	0.017
	0.075	5	100	0.207	0.199	0.084	0.091	0.082	0.219	0.202	0.081	0.087	0.078
			250	0.133	0.072	0.029	0.033	0.030	0.141	0.072	0.031	0.034	0.031
			500	0.113	0.035	0.014	0.016	0.014	0.103	0.034	0.014	0.016	0.014
		9	100	0.216	0.203	0.082	0.089	0.080	0.215	0.208	0.085	0.094	0.082
			250	0.135	0.072	0.031	0.033	0.030	0.133	0.070	0.031	0.032	0.029
			500	0.097	0.035	0.014	0.016	0.014	0.100	0.034	0.015	0.016	0.014
	0.1	5	100	0.262	0.179	0.075	0.078	0.069	0.252	0.171	0.075	0.077	0.069
			250	0.161	0.065	0.026	0.028	0.026	0.167	0.064	0.027	0.028	0.025
			500	0.126	0.031	0.013	0.014	0.013	0.125	0.031	0.013	0.014	0.012
		9	100	0.247	0.183	0.072	0.078	0.070	0.248	0.180	0.075	0.082	0.072
			250	0.170	0.068	0.025	0.030	0.027	0.171	0.067	0.027	0.029	0.027
			500	0.128	0.031	0.013	0.014	0.012	0.127	0.034	0.015	0.014	0.014

Empirical coverage rates of confidence interval at $\alpha = 0.05$

			Model 1: Exponential function					Model 2: Gaussian kernel function				
h	θ	N	λ	$w(h/2)$	$w(h)$	$w(3h/2)$	$w(2h)$	λ	$w(h/2)$	$w(h)$	$w(3h/2)$	$w(2h)$
0.050	5	100	0.178	0.096	0.118	0.126	0.118	0.112	0.092	0.106	0.120	0.106
		250	0.232	0.080	0.092	0.090	0.068	0.066	0.058	0.084	0.084	0.084
		500	0.324	0.050	0.064	0.056	0.082	0.084	0.078	0.046	0.052	0.058
	7	100	0.120	0.100	0.090	0.116	0.124	0.112	0.098	0.126	0.112	0.098
		250	0.082	0.070	0.078	0.086	0.070	0.088	0.060	0.078	0.092	0.060
		500	0.066	0.044	0.058	0.060	0.048	0.042	0.058	0.064	0.074	0.074
	9	100	0.116	0.114	0.132	0.122	0.118	0.092	0.110	0.108	0.112	0.120
		250	0.082	0.080	0.082	0.060	0.054	0.074	0.068	0.058	0.066	0.076
		500	0.052	0.050	0.036	0.044	0.048	0.080	0.050	0.064	0.064	0.060
0.075	5	100	0.102	0.114	0.094	0.102	0.110	0.104	0.100	0.100	0.104	0.112
		250	0.088	0.060	0.064	0.066	0.070	0.074	0.068	0.062	0.072	0.068
		500	0.058	0.082	0.080	0.054	0.064	0.064	0.048	0.050	0.044	0.068
	7	100	0.106	0.086	0.098	0.108	0.114	0.092	0.124	0.102	0.116	0.094
		250	0.072	0.056	0.066	0.080	0.068	0.076	0.072	0.072	0.060	0.072
		500	0.068	0.052	0.054	0.046	0.056	0.072	0.056	0.086	0.068	0.054
	9	100	0.098	0.108	0.096	0.102	0.100	0.110	0.122	0.088	0.102	0.090
		250	0.078	0.064	0.060	0.064	0.052	0.044	0.066	0.076	0.072	0.068
		500	0.068	0.056	0.050	0.044	0.082	0.066	0.066	0.062	0.074	0.080
0.100	5	100	0.118	0.092	0.106	0.072	0.080	0.138	0.098	0.116	0.106	0.092
		250	0.096	0.058	0.060	0.076	0.060	0.102	0.092	0.094	0.072	0.082
		500	0.056	0.054	0.050	0.052	0.074	0.084	0.046	0.048	0.062	0.054
	7	100	0.138	0.102	0.122	0.102	0.106	0.112	0.110	0.092	0.088	0.138
		250	0.074	0.062	0.072	0.056	0.062	0.102	0.062	0.076	0.072	0.100
		500	0.056	0.054	0.044	0.054	0.046	0.074	0.044	0.048	0.048	0.066
	9	100	0.080	0.094	0.104	0.096	0.090	0.132	0.110	0.158	0.092	0.130
		250	0.080	0.054	0.068	0.076	0.050	0.078	0.094	0.094	0.062	0.056
		500	0.066	0.052	0.070	0.058	0.064	0.064	0.082	0.084	0.066	0.078

Empirical coverage rates of confidence interval at $\alpha = 0.05$

			Model 1: Exponential function					Model 2: Gaussian kernel function				
h	θ	N	λ	$w(h/2)$	$w(h)$	$w(3h/2)$	$w(2h)$	λ	$w(h/2)$	$w(h)$	$w(3h/2)$	$w(2h)$
0.050	5	100	0.186	0.100	0.090	0.096	0.110	0.094	0.098	0.112	0.112	0.128
		250	0.260	0.078	0.062	0.068	0.094	0.090	0.068	0.088	0.072	0.062
		500	0.328	0.054	0.076	0.070	0.056	0.052	0.068	0.052	0.052	0.070
	7	100	0.104	0.108	0.094	0.098	0.112	0.104	0.122	0.118	0.120	0.114
		250	0.062	0.056	0.060	0.054	0.066	0.058	0.050	0.062	0.064	0.062
		500	0.068	0.060	0.052	0.046	0.060	0.056	0.056	0.062	0.042	0.052
	9	100	0.116	0.128	0.084	0.130	0.130	0.096	0.128	0.114	0.100	0.108
		250	0.068	0.060	0.066	0.078	0.086	0.056	0.066	0.074	0.062	0.064
		500	0.062	0.058	0.070	0.056	0.064	0.052	0.048	0.054	0.048	0.060
0.075	5	100	0.106	0.092	0.082	0.112	0.090	0.122	0.106	0.104	0.108	0.144
		250	0.082	0.078	0.072	0.060	0.062	0.080	0.070	0.054	0.048	0.060
		500	0.092	0.058	0.064	0.066	0.054	0.048	0.060	0.084	0.072	0.062
	7	100	0.088	0.100	0.118	0.110	0.118	0.122	0.104	0.098	0.088	0.118
		250	0.070	0.068	0.072	0.064	0.056	0.072	0.074	0.068	0.068	0.072
		500	0.076	0.054	0.060	0.056	0.066	0.068	0.068	0.072	0.056	0.064
	9	100	0.112	0.108	0.096	0.114	0.082	0.122	0.106	0.124	0.148	0.120
		250	0.072	0.068	0.072	0.074	0.064	0.084	0.064	0.066	0.066	0.066
		500	0.056	0.060	0.062	0.084	0.070	0.042	0.072	0.064	0.060	0.038
0.100	5	100	0.124	0.092	0.090	0.086	0.112	0.110	0.096	0.110	0.100	0.088
		250	0.076	0.078	0.072	0.080	0.070	0.084	0.068	0.072	0.072	0.068
		500	0.074	0.046	0.052	0.040	0.050	0.098	0.066	0.068	0.072	0.070
	7	100	0.104	0.100	0.114	0.154	0.112	0.102	0.104	0.112	0.102	0.100
		250	0.080	0.068	0.068	0.066	0.058	0.078	0.054	0.060	0.070	0.080
		500	0.082	0.048	0.068	0.060	0.080	0.046	0.064	0.072	0.052	0.056
	9	100	0.112	0.104	0.114	0.108	0.110	0.096	0.102	0.128	0.062	0.092
		250	0.118	0.090	0.066	0.070	0.066	0.078	0.080	0.074	0.082	0.082
		500	0.068	0.054	0.062	0.064	0.066	0.052	0.074	0.086	0.054	0.074

Empirical power of marginal t-test for $H_0 : w(d) = 0$.

h	θ	N	Model 1: Exponential function λ	$w(h/2)$	$w(h)$	$w(3h/2)$	$w(2h)$	Model 2: Gaussian kernel function λ	$w(h/2)$	$w(h)$	$w(3h/2)$	$w(2h)$
0.05	5	100	0.990	0.128	0.234	0.180	0.200	0.988	0.152	0.286	0.210	0.214
		250	1.000	0.250	0.624	0.444	0.442	1.000	0.216	0.682	0.588	0.628
		500	1.000	0.554	0.958	0.882	0.886	1.000	0.576	0.982	0.968	0.982
	7	100	0.992	0.146	0.216	0.180	0.170	0.998	0.138	0.238	0.196	0.232
		250	1.000	0.196	0.566	0.372	0.332	1.000	0.218	0.676	0.562	0.614
		500	1.000	0.558	0.952	0.856	0.808	1.000	0.552	0.988	0.950	0.956
	9	100	0.994	0.154	0.210	0.144	0.138	0.994	0.108	0.252	0.210	0.196
		250	1.000	0.206	0.560	0.372	0.310	1.000	0.234	0.646	0.560	0.536
		500	1.000	0.554	0.942	0.794	0.690	1.000	0.542	0.976	0.930	0.940
0.075	5	100	0.988	0.138	0.288	0.186	0.182	0.980	0.134	0.314	0.286	0.260
		250	1.000	0.298	0.716	0.474	0.470	1.000	0.302	0.846	0.712	0.692
		500	1.000	0.682	0.986	0.946	0.902	1.000	0.700	0.992	0.988	0.990
	7	100	0.982	0.132	0.222	0.170	0.148	0.988	0.120	0.320	0.250	0.266
		250	0.998	0.304	0.676	0.382	0.288	1.000	0.252	0.788	0.668	0.640
		500	1.000	0.648	0.968	0.864	0.816	1.000	0.742	0.996	0.982	0.970
	9	100	0.994	0.134	0.218	0.160	0.152	0.986	0.136	0.318	0.244	0.216
		250	1.000	0.278	0.626	0.352	0.272	1.000	0.302	0.764	0.590	0.512
		500	1.000	0.684	0.962	0.768	0.628	1.000	0.706	0.986	0.956	0.912
0.10	5	100	0.966	0.154	0.278	0.186	0.168	0.960	0.134	0.402	0.308	0.274
		250	0.996	0.350	0.752	0.490	0.432	0.998	0.372	0.878	0.724	0.720
		500	1.000	0.762	0.984	0.922	0.876	1.000	0.788	0.998	0.998	0.992
	7	100	0.972	0.150	0.248	0.150	0.144	0.948	0.148	0.336	0.282	0.230
		250	1.000	0.320	0.722	0.392	0.296	0.998	0.318	0.852	0.680	0.586
		500	1.000	0.766	0.982	0.874	0.686	1.000	0.804	0.998	0.990	0.950
	9	100	0.972	0.188	0.240	0.118	0.132	0.948	0.144	0.320	0.182	0.164
		250	0.998	0.328	0.690	0.344	0.204	0.996	0.378	0.818	0.556	0.368
		500	1.000	0.764	0.966	0.722	0.474	1.000	0.816	0.994	0.936	0.822

Empirical power of marginal t-test for $H_0 : w(d) = 0$.

			Model 1: Exponential function					Model 2: Gaussian kernel function				
h	θ	N	λ	$w(h/2)$	$w(h)$	$w(3h/2)$	$w(2h)$	λ	$w(h/2)$	$w(h)$	$w(3h/2)$	$w(2h)$
0.05	5	100	0.420	0.110	0.236	0.202	0.188	0.402	0.108	0.306	0.276	0.226
		250	0.704	0.224	0.614	0.476	0.430	0.636	0.190	0.710	0.588	0.602
		500	0.960	0.578	0.960	0.856	0.866	0.902	0.552	0.984	0.964	0.984
	7	100	0.378	0.116	0.206	0.176	0.182	0.390	0.152	0.302	0.214	0.214
		250	0.646	0.208	0.576	0.390	0.370	0.652	0.214	0.696	0.578	0.590
		500	0.918	0.578	0.950	0.828	0.788	0.882	0.526	0.978	0.958	0.956
	9	100	0.382	0.138	0.196	0.134	0.132	0.404	0.144	0.278	0.240	0.196
		250	0.630	0.208	0.580	0.368	0.282	0.652	0.262	0.656	0.532	0.538
		500	0.902	0.578	0.944	0.768	0.660	0.850	0.586	0.974	0.944	0.928
0.075	5	100	0.372	0.154	0.260	0.192	0.180	0.354	0.140	0.312	0.274	0.284
		250	0.548	0.264	0.720	0.524	0.470	0.510	0.262	0.812	0.706	0.734
		500	0.788	0.664	0.980	0.904	0.874	0.792	0.752	0.990	0.982	0.984
	7	100	0.342	0.130	0.238	0.160	0.144	0.372	0.162	0.358	0.252	0.226
		250	0.520	0.304	0.658	0.412	0.338	0.536	0.274	0.800	0.658	0.632
		500	0.712	0.692	0.982	0.854	0.782	0.708	0.704	0.996	0.978	0.966
	9	100	0.326	0.166	0.230	0.164	0.174	0.372	0.174	0.310	0.226	0.218
		250	0.566	0.276	0.612	0.348	0.226	0.554	0.312	0.760	0.596	0.526
		500	0.726	0.666	0.964	0.738	0.572	0.728	0.722	0.988	0.972	0.924
0.10	5	100	0.308	0.152	0.276	0.186	0.190	0.354	0.174	0.372	0.268	0.290
		250	0.452	0.372	0.782	0.492	0.426	0.450	0.366	0.878	0.750	0.752
		500	0.634	0.768	0.986	0.934	0.862	0.568	0.754	0.998	0.992	0.986
	7	100	0.306	0.134	0.250	0.122	0.126	0.324	0.144	0.318	0.218	0.210
		250	0.442	0.356	0.720	0.410	0.294	0.446	0.374	0.842	0.674	0.586
		500	0.562	0.732	0.978	0.856	0.716	0.564	0.768	0.994	0.978	0.944
	9	100	0.312	0.138	0.228	0.132	0.096	0.316	0.174	0.376	0.226	0.152
		250	0.416	0.338	0.672	0.310	0.186	0.456	0.360	0.808	0.548	0.378
		500	0.566	0.786	0.978	0.756	0.436	0.606	0.834	0.992	0.958	0.808

Empirical size and power of the uniform test

		H_0: No network structure					
		Size			Power		
θ	N	h=0.05	0.075	0.1	h=0.05	0.075	0.1
5	100	0.071	0.056	0.065	0.998	0.998	1.000
	250	0.049	0.052	0.036	1.000	1.000	1.000
	500	0.073	0.032	0.040	1.000	1.000	1.000
9	100	0.065	0.062	0.071	0.985	0.997	0.998
	250	0.039	0.042	0.054	1.000	1.000	1.000
	500	0.035	0.036	0.052	1.000	1.000	1.000
		H_{0f}: Exponential function					
		Size			Power		
θ	N	h=0.05	0.075	0.1	h=0.05	0.075	0.1
5	100	0.056	0.058	0.046	0.994	1.000	0.999
	250	0.024	0.042	0.041	1.000	1.000	1.000
	500	0.030	0.030	0.034	1.000	1.000	1.000
9	100	0.055	0.057	0.071	0.980	0.994	0.998
	250	0.037	0.042	0.028	1.000	1.000	1.000
	500	0.038	0.033	0.034	1.000	1.000	1.000

Empirical application

Extend Levinson and OBrien (2019, REStat) by incorporating neighboring effects in the relationship between households' income and pollution.

Levinson and OBrien (2019, REStat) has two main objectives:

- Find the shape of the relationship between income and pollution, the magnitude of the slope, and the study of its curvature.
- Analyze shifts in the EEC in terms of income increases (movements along the curve), or in terms of regulation-induced price increases (movements of the curve).

These authors find that the EECs are upward sloping, reflecting that richer households are more pollutant, and that the rate at which pollution increases with income is less than one-for-one.

The model with network effects:

We estimate the following specification:

$$p_{it} = \lambda_{1t} y_{it} + \sum_{\substack{j=1 \\ j \neq i}}^{N} w(d_{ij,t}) y_{jt} + \lambda_{2t} y_{it}^2 + \varepsilon_{it}, \quad (23)$$

where p_{it} and y_{it} are pollution and after-tax income, respectively. The network structure is given by $d_{ij,t} = |y_{it} - y_{jt}|$ such that

$$p_{it} = \lambda_{1t} y_{it} + \sum_{k=1}^{K} \sum_{m=0}^{q} \gamma_{km,t} y_{it}^{(km)} + \lambda_{2t} y_{it}^2 + \varepsilon_{it}, \quad (24)$$

with $y_{it}^{(km)} = \sum_{\substack{j=1 \\ j \neq i}}^{N} y_{jt} (d_{ij,t} - z_k)^m 1_k(d_{ij,t})$, and $1_k(d_{ij,t}) = 1(|d_{ij,t} - z_k| \leq h)$;

$\gamma_{km,t} = \frac{1}{m!} w_t^{(m)}(z_k)$ is the Taylor expansion for $m = 0, 1, 2$.

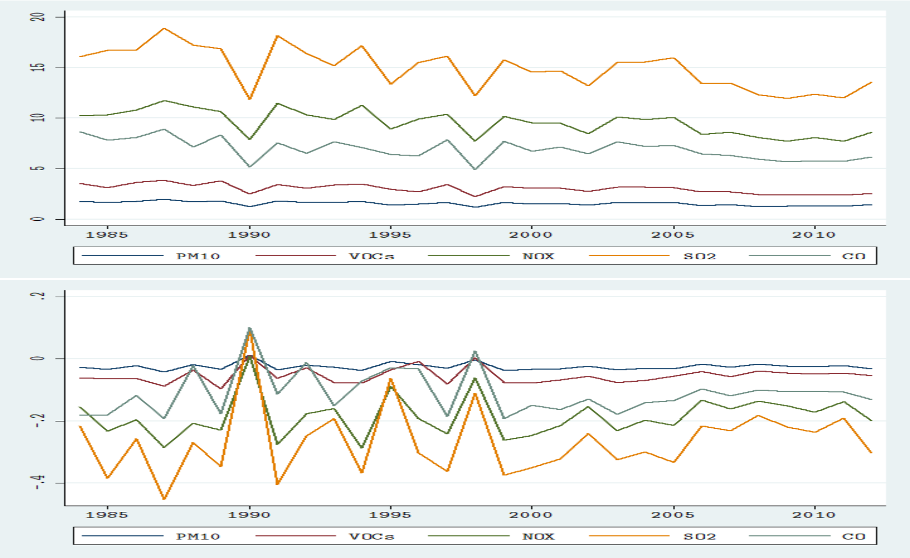

Figure: Network regression estimation of EECs in the U.S., 1984-2012.

Figure: Network regression estimation of EECs, 1984-2012. Functional coefficient.

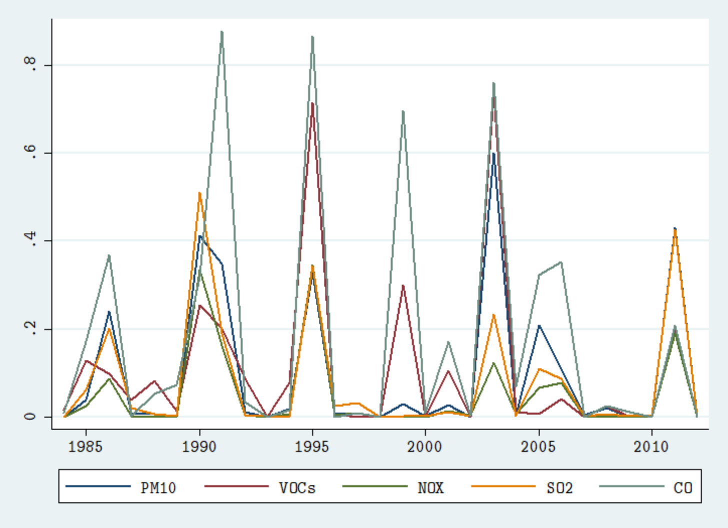

Figure: Uniform test statistic, p-values.

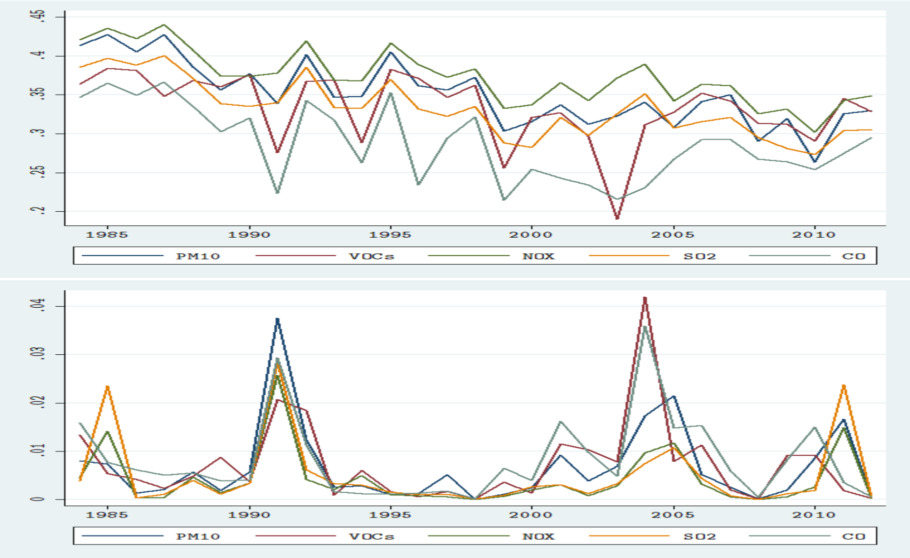

Figure: Network regression estimation of EECs in the U.S., 1984-2012.

Conclusions

- This paper proposes a network regression model to account for the influence of peer contextual effects on the outcome variable.
- The parameters capturing spillover effects are modelled as a functional coefficient that depends on a network variable.
- The estimator of the slope and network effects in our network regression model is consistent and asymptotically normal.
- The slope estimator converges to a standard normal distribution at a parametric $\sqrt{N}$ rate and the network estimator converges at a rate N.
- We find strong empirical evidence of neighboring effects on the relationship between different forms of environmental pollution and after-tax household income.