USDA ERS Broadband: *Comparison of boundary analysis methods for spatial models.*

Abstract

This a weekly updated memo documenting progress in statistical modeling for the USDA ERS Broadband Project. Currently we aim to perform exploratory analysis, synthetic experiments comparing methods that detect geographical boundaries separating response behavior within a fixed spatial reference domain. We aim to also decide on a modeling framework suitable for the CoreLogic data.

Meeting: 10/21/2020

We aim to perform synthetic experiments showcasing the performance of spatial regression discontinuity design (see Lee and Lemieux (2010), Hidano et al. (2015), Keele and Titiunik (2015)) under a variety of spatial patterns. All the experiments showcased compare regression discontinuity designs under (a) ordinary least squares, (b) two stage least squares with spatially lagged errors (Hidano et al. (2015), Kelejian and Prucha (1998)) (c) *parametric* fully Bayesian hierarchical spatial regression (Finley et al. (2007)) modeling frameworks. The reason behind the choice of a parametric spatial regression is motivated from a desired setup for comparing regression discontinuity designs with methods for wombling (Womble (1951)). We consider the synthetic data generating process (outlined in Hidano et al. (2015))

$$Y = \beta_0 + \alpha X + \delta \left(sin(\overline{lat}) + cos(\overline{long}) \right) + \epsilon,$$

where $\beta_0 = 60$, $\alpha = -1.5$, $\delta = 2.5$, $\epsilon \sim N(0,1)$, $\overline{lat}$ and $\overline{long}$ are the standardized latitude and longitude respectively. This introduces location specific effects to within the response. X is the covariate determining the spatial location based discontinuity, i.e. it is 0 or 1 based on the location of the response. We select $\mathbf{s} = (s_x, s_y)^\top \in (0,20) \times (0,20) \subset \mathbb{R}^2$ as the spatial reference domain. The synthetic patterns considered are shown in figure 1. For the patterns shown the X variable is selected as follows,

1. Sharp Regression Discontinuity:

   (a) $X = 0 \cdot I(s_x < 10) + 1 \cdot I(s_x \geq 10)$, (Open Curve)

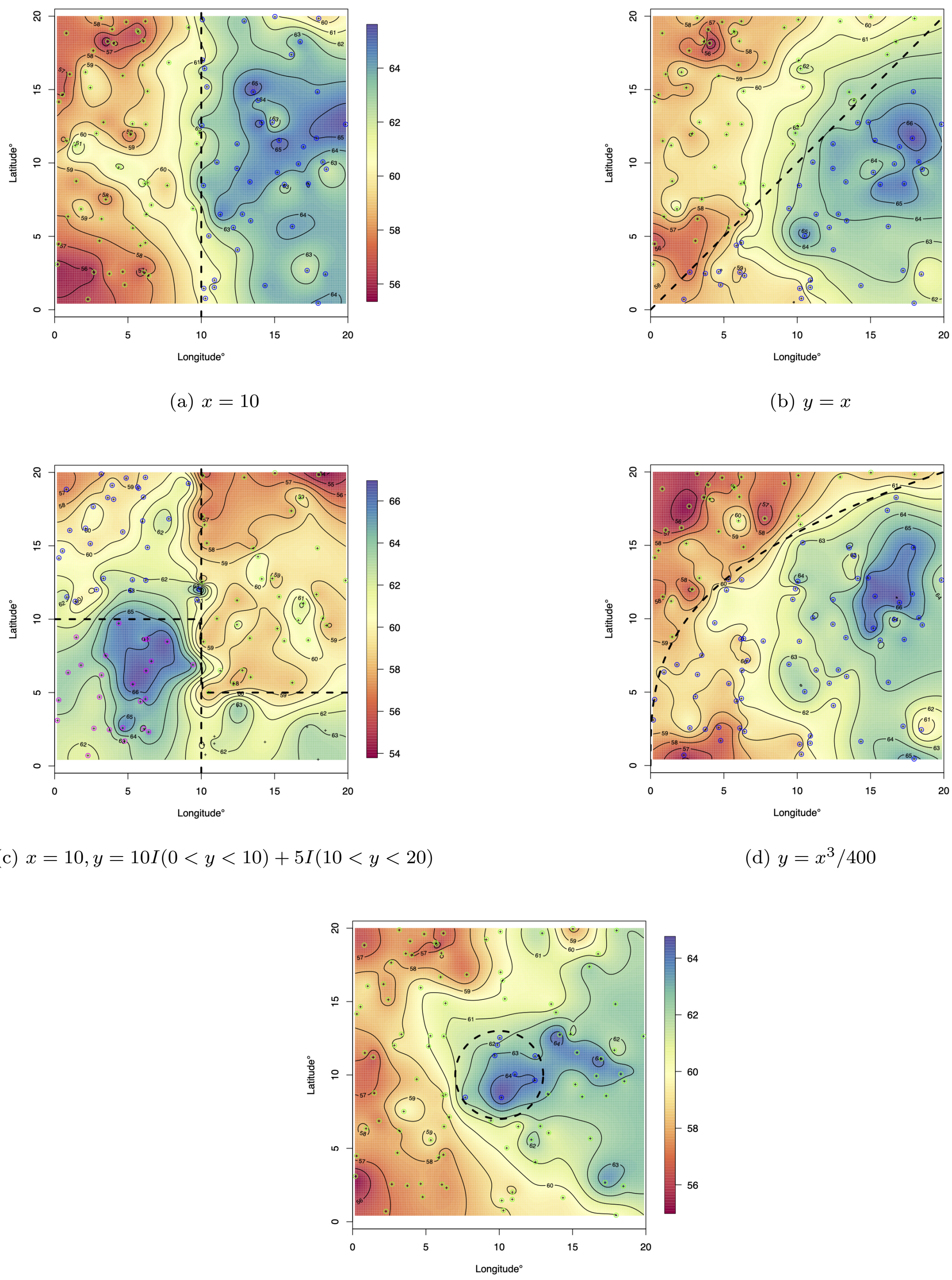
   (b) $X = 1 \cdot I(s_y < s_x) + 0 \cdot I(s_y \geq s_x)$, (Open Curve)

   (d) $X = 1 \cdot I(s_y < s_x^3/400) + 0 \cdot I(s_y \geq s_x^3/400)$, (Open Curve)

   (e) $X = 1 \cdot I\left((s_y - 10)^2 + (s_x - 10)^2 < 36\right) + 0 \cdot I\left((s_y - 10)^2 + (s_x - 10)^2 \geq 36\right)$ (Closed Curve)

2. Fuzzy Regression Discontinuity:

- $X = -5 \cdot I(s_y > 10, s_x \geq 5) + 0.5 \cdot I(s_y < 10, s_x \geq 10) + 5 \cdot I(s_y < 10, s_x < 10) + 0 \cdot I(s_y > 10, s_x < 5)$, (Open Curve)



(a) $x = 10$

(b) $y = x$

(c) $x = 10, y = 10I(0 < y < 10) + 5I(10 < y < 20)$

(d) $y = x^3/400$

(e) $(y - 10)^2 + (x - 10)^2 = 6^2$

Figure 1: Synthetic spatial patterns.

The models being compared are:

1. OLS Model: $Y = \beta_0 + \alpha X + \epsilon$, $\epsilon \sim N(0, \sigma^2)$
2. OLS Model: $Y = \beta_0 + \alpha X + \lambda WY + u$, $u = \rho Mu + \epsilon$, $\epsilon \sim N(0, \sigma^2)$, $|\lambda| < 1$,$|\rho| < 1$, where M and W are spatial loading matrices (Kelejian and Prucha (1998)) eq. 1.
3. Spatial Model: $Y(\mathbf{s}) = \beta_0 + \alpha X + \mathbf{Z}(\mathbf{s}) + \epsilon$, $\mathbf{Z}(\mathbf{s}) \sim N(0, \Sigma)$, $\epsilon \sim N(0, \tau^2)$, $\Sigma = \sigma^2 \exp(-\phi_s ||\mathbf{s} - \mathbf{s}'||)$, where $\mathbf{s}, \mathbf{s}'$ are locations within the random field. Σ corresponds to the exponential covariance kernel.

It is important to note that the accuracy in estimating the regression discontinuity (if present) is highly dependent on the predictive power of the model. To that end, we use the root mean square error (RMSE) to measure model performance under the various synthetic patterns. Under each pattern we performed 10 replications and the number of locations ranged between $\{100, 200, 500\}$. Tables 1 and 2 shows the superiority in performance for the parametric spatial model based on RMSE.

Table 1: Table showing RMSE and associated standard deviation (in brackets) for estimated parameters under the three different modeling frameworks for synthetic patterns pertaining to sharp regression discontinuity.

Pattern (Sharp Reg. Discontinuity)	N	OLS				2SLS				Spatial			
		β_0	α	Fitted	Cont. Prob.	β_0	α	Fitted	Cont. Prob.	β_0	α	Fitted	Cont. Prob.
$x = 10$	100	3.050 (0.262)	3.326 (0.222)	1.563 (0.074)	0.000	6.975 (0.877)	3.182 (0.403)	1.420 (0.066)	0.000	1.032 (0.341)	1.036 (0.424)	0.800 (0.071)	0.800
	200	3.136 (0.112)	3.390 (0.168)	1.548 (0.058)	0.000	7.156 (0.765)	3.322 (0.215)	1.395 (0.055)	0.000	0.459 (0.202)	0.693 (0.337)	0.798 (0.031)	0.700
	500	3.107 (0.077)	3.399 (0.131)	1.543 (0.036)	0.000	6.998 (0.186)	3.269 (0.109)	1.398 (0.034)	0.000	0.333 (0.166)	0.514 (0.296)	0.882 (0.025)	0.600
$y = x$	100	2.579 (0.248)	2.249 (0.361)	1.941 (0.149)	0.000	6.254 (1.292)	2.198 (0.406)	1.836 (0.151)	0.000	0.670 (0.418)	0.530 (0.349)	0.734 (0.092)	1.000
	200	2.511 (0.125)	2.111 (0.168)	2.047 (0.091)	0.000	6.908 (0.626)	2.108 (0.223)	1.919 (0.097)	0.000	0.272 (0.251)	0.332 (0.236)	0.823 (0.051)	1.000
	500	2.493 (0.096)	2.115 (0.170)	2.019 (0.041)	0.000	6.883 (0.472)	2.002 (0.138)	1.877 (0.047)	0.000	0.236 (0.151)	0.150 (0.103)	0.861 (0.029)	1.000
$y = x^3/400$	100	2.034 (0.134)	2.571 (0.255)	1.960 (0.090)	0.000	4.480 (1.696)	2.263 (0.400)	1.905 (0.099)	0.000	0.745 (0.249)	1.092 (0.377)	0.759 (0.055)	0.700
	200	1.968 (0.126)	2.281 (0.302)	2.077 (0.065)	0.000	5.390 (1.194)	1.800 (0.229)	1.995 (0.082)	0.000	0.244 (0.211)	0.277 (0.213)	0.803 (0.055)	1.000
	500	2.022 (0.042)	2.444 (0.114)	2.059 (0.061)	0.000	4.844 (0.718)	2.125 (0.09)	2.001 (0.067)	0.000	0.275 (0.2)	0.2700 (0.216)	0.886 (0.025)	0.900
$(y-10)^2 + (x-10)^2 = 3^2$	100	2.159 (0.224)	1.040 (0.383)	2.193 (0.221)	0.600	5.697 (3.582)	1.259 (0.754)	2.045 (0.145)	0.300	0.574 (0.510)	0.400 (0.338)	0.768 (0.065)	0.900
	200	2.199 (0.263)	1.092 (0.315)	2.214 (0.077)	0.700	5.824 (2.228)	0.604 (0.620)	2.131 (0.102)	0.100	0.387 (0.361)	0.299 (0.271)	0.817 (0.052)	0.900
	500	2.209 (0.119)	1.101 (0.176)	2.232 (0.065)	0.700	5.928 (1.175)	0.317 (0.287)	2.164 (0.058)	0.000	0.265 (0.140)	0.209 (0.145)	0.873 (0.028)	0.900

Table 2: Table showing RMSE and associated standard deviation (in brackets) for estimated parameters under the three different modeling frameworks for a synthetic pattern pertaining to fuzzy regression discontinuity.

Pattern (Fuzzy Reg. Discontinuity)	N	OLS						2SLS						Spatial					
		β_0	α_1	α_2	α_3	Fitted	Cont. Prob.	β_0	α_1	α_2	α_3	Fitted	Cont. Prob.	β_0	α_1	α_2	α_3	Fitted	Cont. Prob.
$x = 10, y = 10I(0 < y < 10) + 5I(10 < y < 20)$	100	2.139 (0.326)	1.219 (0.469)	2.437 (0.403)	2.339 (0.330)	1.489 (0.115)	0.200	6.339 (1.558)	0.822 (0.556)	2.716 (0.409)	2.705 (0.384)	1.355 (0.108)	0.100	0.660 (0.429)	0.840 (0.557)	0.763 (0.480)	0.715 (0.422)	0.738 (0.072)	0.966
	200	2.445 (0.291)	0.891 (0.363)	2.661 (0.508)	2.630 (0.193)	1.517 (0.050)	0.233	6.724 (0.788)	0.523 (0.270)	2.955 (0.442)	3.023 (0.193)	1.375 (0.070)	0.066	0.421 (0.346)	0.343 (0.205)	0.618 (0.465)	0.660 (0.365)	0.822 (0.034)	0.966
	500	2.218 (0.226)	1.130 (0.213)	2.386 (0.222)	2.508 (0.310)	1.505 (0.031)	0.033	5.970 (0.686)	0.726 (0.223)	2.744 (0.252)	2.870 (0.339)	1.403 (0.040)	0.000	0.352 (0.199)	0.298 (0.237)	0.257 (0.260)	0.277 (0.192)	0.852 (0.033)	1.000

For the Bayesian hierarchical spatial model, we have the following parameters, $\boldsymbol{\theta} = \{\boldsymbol{\beta}^\top, \phi_s, \sigma^2, \tau^2\}$, where $\boldsymbol{\beta} = (\beta_0, \alpha)^\top$ with the posterior likelihood being specified by

$$p(\boldsymbol{\theta}|y) \propto U(\phi_s|a_\phi, b_\phi) \times IG(\sigma^2|a_\sigma, b_\sigma) \times IG(\tau^2|a_\tau, b_\tau) \times N(\boldsymbol{\beta}|\mu_\beta, \Sigma_\beta) \times N(\mathbf{Z}|0, \Sigma(\sigma^2, \phi_s)) \\ \times \prod N(y|\beta_0 + \alpha X, \tau^2) \tag{1}$$

We use the following hyperparameter specifications: $a_\phi = 0$, $b_\phi = 30$, $a_\sigma = 2, b_\sigma = 1$, $a_\tau = 2, b_\tau = 0.1$, $\mu_\beta = (0,0)^\top$, $\Sigma_\beta = diag\{1 \times 10^{-3}, 1 \times 10^{-3}\}$. The posterior for ϕ_s is sampled using an *adaptive Metropolis within Gibbs sampler*. It was observed that for each simulation the true value $\tau^2 = 1$ was contained within the highest posterior density (HPD) intervals for the Bayesian hierarchical spatial model.

Scaleable Spatial Processes

We use two approximate spatial processes to provide scaleability to the spatial models viz. Nearest Neighbor Gaussian Processes (NNGP) and Integrated Nested Laplace Approximation (INLA). A synthetic comparison for a sharp regression discontinuity design $x = 10$ is shown below.

Table 3: Table showing RMSE for estimates from NNGP and INLA for a shrap regression discontinuity pattern, $x = 10$.

N	INLA		NNGP		Prediction	
	β	α	β	α	INLA	NNGP
1000	0.353	0.111	0.209	0.089	2.258	2.258
	(0.070)	(0.075)	(0.075)	(0.071)	0.029	0.029
5000	0.247	0.038	0.139	0.036	2.248	2.250
	(0.033)	(0.036)	(0.078)	(0.076)	0.013	0.014
10000	0.113	0.022	0.061	0.016	**4.498**	2.251
	(0.044)	(0.050)	(0.108)	(0.068)	**7.114**	0.005

Wombling

For the purpose of synthetic illustrations for wombling we require differentiability of the underlying random field (as opposed to just continuity in the RD scenario). This is established by using differentiable kernels while constructing Σ for the spatial models. In terms of kernel choices we have,

$$\Sigma = \begin{cases} \tilde{K}(||\Delta||) = \sigma^2 \exp(-\phi_s ||\Delta||^\nu), & \nu \in [0,2], \text{ (Power exponential Class)} \\ \tilde{K}(||\Delta||) = \sigma^2 (\phi_s ||\Delta||)^\nu K_\nu(\phi_s ||\Delta||). & \text{(Matérn Class).} \end{cases}$$

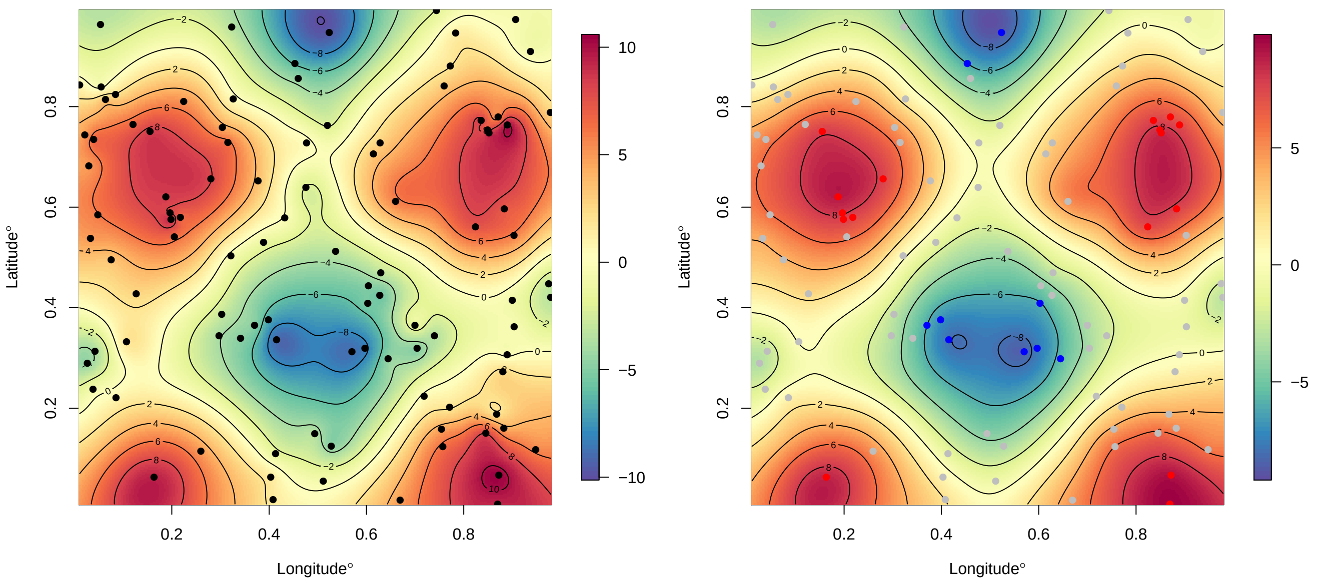
ν controls the smoothness for process realizations. For the power exponential class $\nu = 1$ is the exponential kernel which is just continuous admitting no derivatives, $\nu = 2$ is the squared exponential or gaussian covariance, which is infinitely differentiable. Within the Matérn class we have $\nu = 3/2$ admitting gradients and $\nu = 5/2$ admitting both first and second order gradients. If $\nu \to \infty$ we obtain the Gaussian covariance.

Furthermore, to check the validity of our estimation of gradients it is beneficial to have an analytic function that has a closed form analytic expression of the gradient. We select the following form within the spatial reference domain $\mathbf{s} = (s_x, s_y)^\top \in (0,1) \times (0,1)$ in $\mathbb{R}^2$.

$$Y \sim N(5(\sin(3\pi s_x) + \cos(3\pi s_y)), \tau^2), \qquad \tau^2 = 1$$

Fig. (2) (a) shows the synthetic spatial pattern. Before we demonstrate gradient estimation, a spatial model with a Matérn kernel with $\nu = 5/2$ is fit to the data. The advantage of fitting such a kernel is that both first and second order gradients exist for process realizations. Given the choice of our synthetic data generating process we should have,

$$\nabla_x Y = 15\pi \cos(3\pi s_x), \qquad \nabla_y Y = -15\pi \sin(3\pi s_y)$$
$$\nabla_{xx} Y = -45\pi^2 \sin(3\pi s_x), \quad \nabla_{yy} Y = -45\pi^2 \cos(3\pi s_y), \quad \nabla_{xy} Y = 0.$$

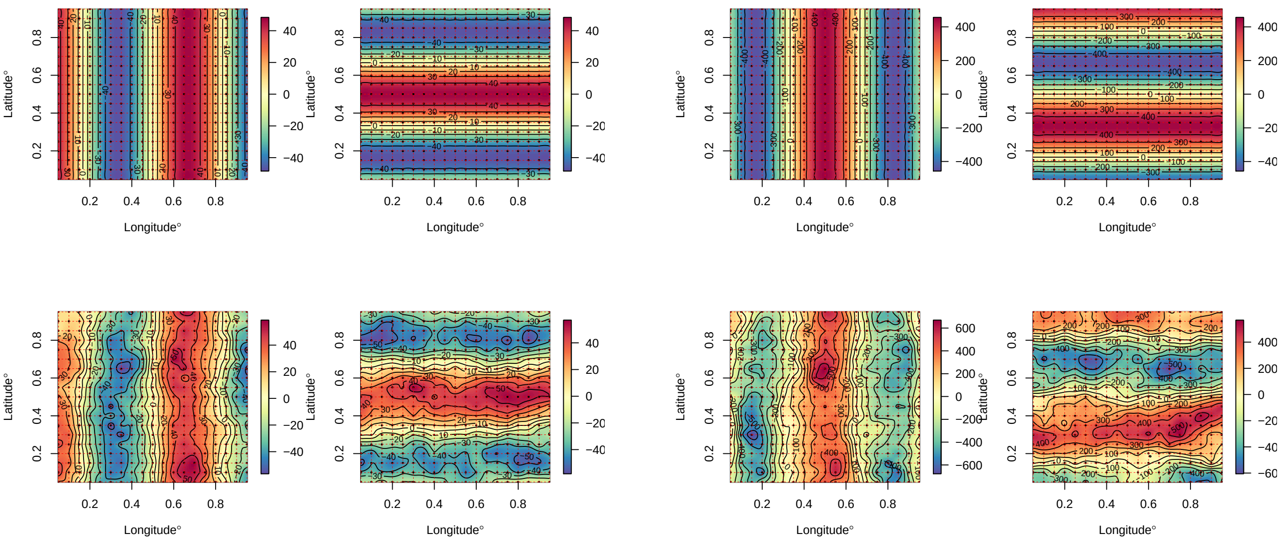


(a) (b)

Figure 2: Interpolated spatial plots showing (a) the true pattern (b) the estimated fit from a spatial model with a Matérn($\nu = 5/2$) kernel. Significant random effects in (b) are color coded accordingly based on their algebraic sign

From the fitted spatial process, we extract posterior samples of process parameters $\{\sigma^2, \phi_s, \tau^2, \mathbf{Z}(\mathbf{s})\}$ to estimate the gradients at regularly spaced points on a grid post-MCMC.

In a realistic setup wombling amounts to locating significant gradients and curvature along points on a curve, once such a curve is available we could replace the points on the grid with points on the curve to



(a) (b)

Figure 3: Interpolated spatial plots showing (a) the true and estimated first order gradients, ∇_x and ∇_y (b) the true and estimated pure second order gradients, ∇_{xx} and ∇_{yy} over 360 points on a regular grid (marked using a dashed red line).

compute gradients/curvature along its normal direction.

Validity of the Continuity assumption: a model-based approach

Along with the conventional distance to boundary approach, using model based information criteria to inform about continuity/differentiability of the random field. For instance, for the first sharp discontinuity pattern we have

(a) Exponential Kernel (Continuous only) **AIC=499.32, BIC = 772.87, DIC= 289.32**

(b) Matérn($\nu = 3/2$) (Derivative exists) AIC=547.94, BIC = 821.48, DIC= 337.94

(c) Matérn($\nu = 5/2$) (Curvature exists) AIC=566.64, BIC = 840.18, DIC= 356.64

(d) Gaussian (Infinitely differentiable) AIC=576.89, BIC =890.44, DIC=376.89

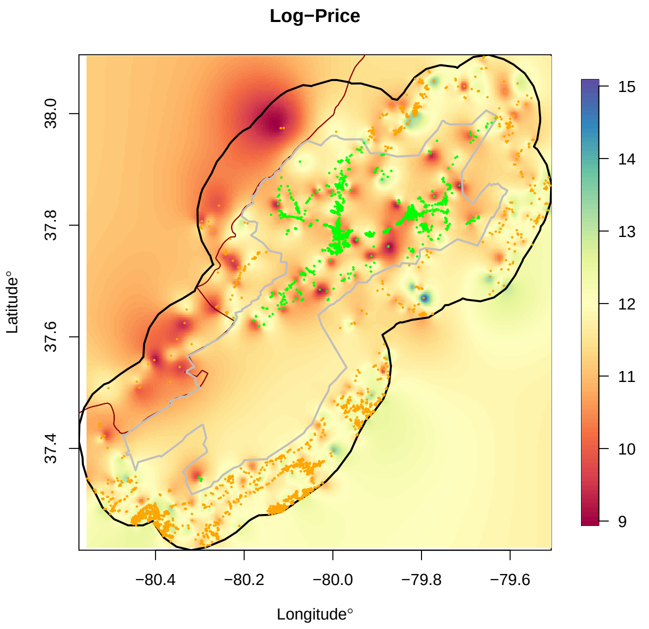
The lower information criteria indicates a better model. The chosen exponential kernel points to only continuity and not differentiability of the random field thereby making it a better fit for RD analysis. As opposed to smooth synthetic patterns like the one considered for wombling presented with Materń($\nu = 5/2$) showing the least information criteria indicating differentiability of the random field.

Table 4: Project Details for the BIPs compared.

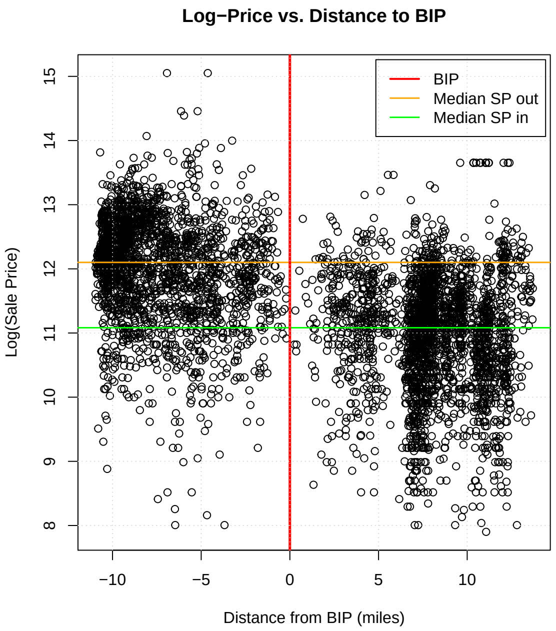
RUSID	Project ID	Name	Fisc.Yr.	Obligation Date	Original Obligation	Total Net Obligation (CY 2020)	Adv. Amt.	Resc. Amt.	Year5_Subsc.	Net Grant	Net Loan
VA 1108	VA1108-A39	LUMOS Telephone Inc.	2010	08/10/10	$8,062,088	$7,692,552	$7,692,552	$369,536	Yes	$7,692,552	$-
WV 1103	WV1103-A40	Hardy Telecommunications, Inc.	2010	08/06/10	$31,648,274	$31,622,782	$31,622,782	$25,492	Yes	$22,135,944	$9,486,838

CoreLogic Data: Spatial modeling, regression discontinuity design and wombling

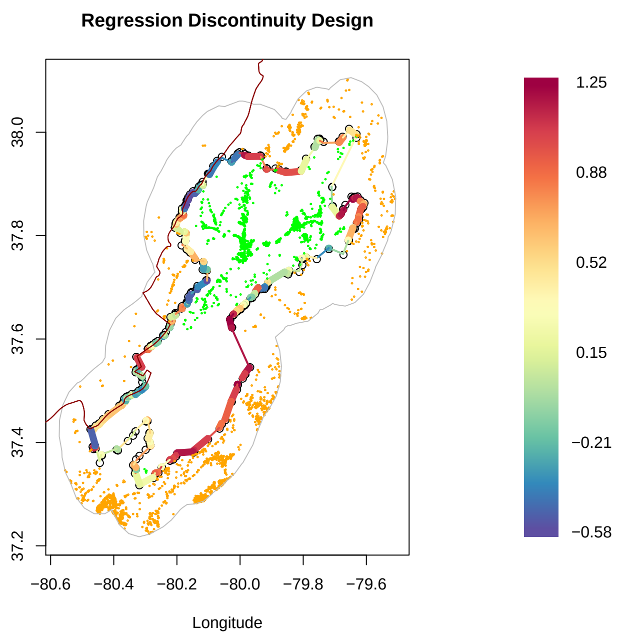
BIP: VA1108



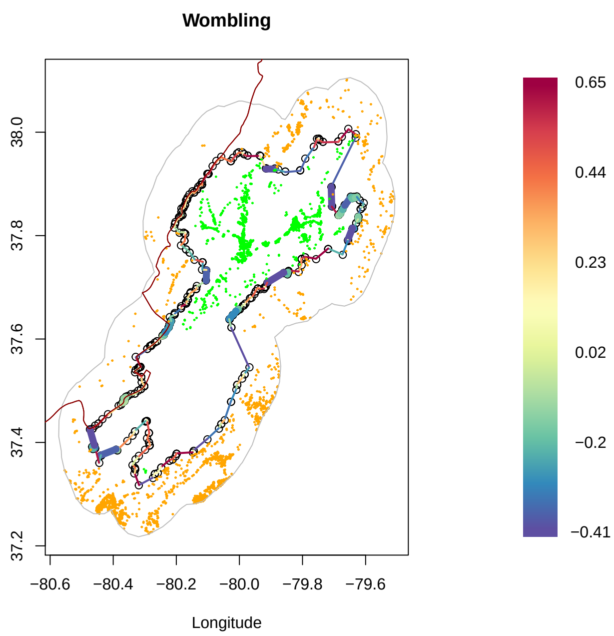
(a)



(b)



(c)



(d)

Figure 4: (a) Spatially interpolated surface of log-house prices in and around the BIP (15 mile radius) (b) log-house prices plotted against distance to BIP (c) RDD applied to fitted model, the difference of predictions of log-house prices, $\lim_{d\downarrow 0}\log(\text{Price})_{BIP}(\mathbf{s}+d) - -\lim_{d\uparrow 0}\log(\text{Price})_{BIP}(\mathbf{s}-d)$ (scale in log(US-Dollars)) significant discontinuities are marked in bold (b) Gradient analysis (Wombling) performed for the BIP (scale in log(US-Dollars)). *Note: Raw estimated gradients are first scaled by length of line segment. Direction (of the associated normal to the line segment) of gradients are inside → outside of BIP, i.e. positive gradients ⇒ inside > outside log-price and vice-versa.* Significant gradients are marked in bold. Gradient analysis took 22.17 mins.

BIP: WV1103

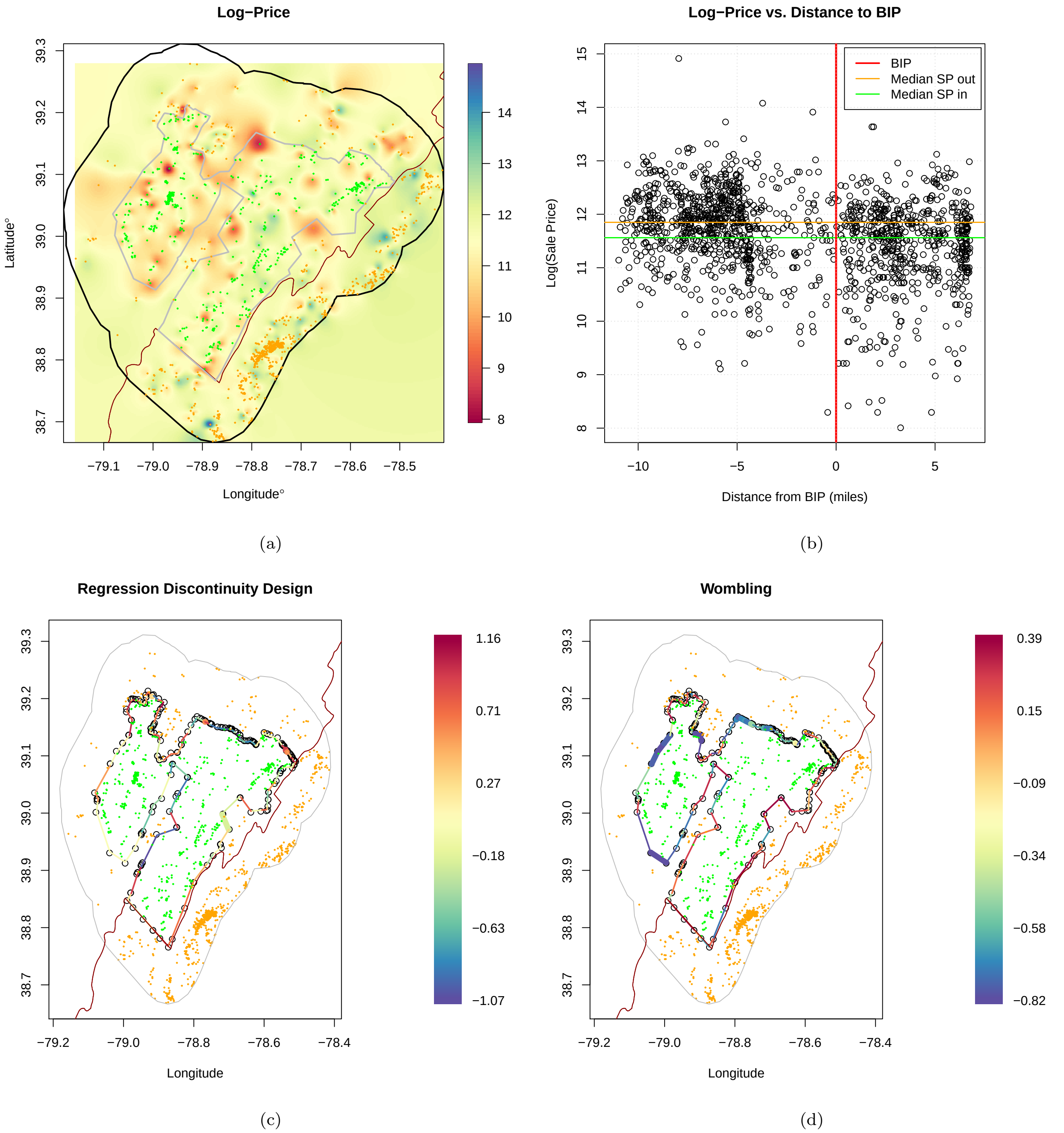


Figure 5: (a) Spatially interpolated surface of log-house prices in and around the BIP (15 mile radius) (b) log-house prices plotted against distance to BIP (c) RDD applied to fitted model, the difference of predictions of log-house prices, $\lim_{d\downarrow 0} \log(\text{Price})_{BIP}(\mathbf{s}+d) - -\lim_{d\uparrow 0} \log(\text{Price})_{BIP}(\mathbf{s}-d)$ (scale in log(US-Dollars)), significant discontinuities are marked in bold (b) Gradient analysis (Wombling) performed for the BIP (scale in log(US-Dollars)). *Note: Raw estimated gradients are first scaled by length of line segment. Direction (of the associated normal to the line segment) of gradients are inside $\rightarrow$ outside of BIP, i.e. positive gradients $\Rightarrow$ inside $>$ outside log-price and vice-versa.* Significant gradients are marked in bold. Gradient analysis took 4.18 mins.

Appendix: Tables

Table 5: Fitted model coefficients from the spatial model for BIP VA1108.

Effect	mean	sd	0.025quant	0.5quant	0.975quant	mode
(intercept)	12.4275	1.4713	9.5387	12.4274	15.3140	12.4273
BIP	0.4344	0.0956	0.2473	0.4340	0.6230	0.4334
age	-0.0007	0.0004	-0.0016	-0.0007	0.0001	-0.0007
nbaths	0.3367	0.0161	0.3051	0.3367	0.3682	0.3367
sqft_ratio	0.2083	0.0305	0.1484	0.2083	0.2680	0.2083
acres	0.3163	0.1378	0.0459	0.3163	0.5866	0.3163
land_square_footage	-0.2515	0.1368	-0.5202	-0.2515	0.0169	-0.2515
bedrooms	0.0127	0.0067	-0.0003	0.0127	0.0259	0.0127
transaction_type:3	0.1972	0.0775	0.0451	0.1973	0.3492	0.1973
transaction_type:6	-0.1041	0.2175	-0.5311	-0.1041	0.3226	-0.1042
transaction_type:7	-0.0977	0.1508	-0.3938	-0.0977	0.1981	-0.0977
bldg_code:A0G	0.5311	0.3483	-0.1513	0.5305	1.2159	0.5294
bldg_code:C00	0.2151	0.3138	-0.3996	0.2146	0.8319	0.2137
bldg_code:MA0	-0.2234	0.6718	-1.5422	-0.2235	1.0949	-0.2237
bldg_code:R00	-0.0649	0.1294	-0.3157	-0.0662	0.1931	-0.0686
bldg_code:RM1	-0.4514	0.2715	-0.9828	-0.4521	0.0829	-0.4533
bldg_code:RM2	-0.0222	0.1349	-0.2794	-0.0249	0.2509	-0.0306
bldg_code:RS0	0.2945	0.1193	0.0694	0.2911	0.5390	0.2842
pri_cat_code:B	-0.2246	0.1905	-0.5986	-0.2246	0.1490	-0.2246
pri_cat_code:D	0.1345	0.3893	-0.6299	0.1345	0.8983	0.1345
zoning:1	-0.0712	0.2919	-0.6447	-0.0711	0.5012	-0.0709
zoning:2	0.1349	0.3535	-0.5590	0.1348	0.8287	0.1346
zoning:3	-0.5959	0.2257	-1.0391	-0.5960	-0.1529	-0.5960
zoning:A-1	-0.1928	0.3751	-0.9324	-0.1918	0.5404	-0.1896
zoning:A1	0.2625	0.0825	0.0999	0.2626	0.4240	0.2629
zoning:A2	0.4099	0.1471	0.1191	0.4106	0.6970	0.4119
zoning:AG1	0.2509	0.1276	0.0007	0.2508	0.5015	0.2506
zoning:AG1S	1.1493	0.6905	-0.2064	1.1493	2.5039	1.1493
zoning:AG3	0.2700	0.1365	0.0045	0.2690	0.5403	0.2672
zoning:AL	0.4828	0.5061	-0.5112	0.4829	1.4756	0.4830
zoning:AR	0.0518	0.0794	-0.1042	0.0518	0.2074	0.0519
zoning:ARS	-0.1711	0.6760	-1.4984	-0.1711	1.1552	-0.1711
zoning:AV	0.1982	0.3300	-0.4496	0.1982	0.8455	0.1981
zoning:B-1	-0.1537	0.5257	-1.1862	-0.1535	0.8773	-0.1533
zoning:B-2	-0.0911	0.5225	-1.1179	-0.0909	0.9332	-0.0903
zoning:B1	-0.3057	0.2627	-0.8214	-0.3058	0.2096	-0.3058
zoning:C-1	-0.0387	0.5390	-1.0973	-0.0386	1.0187	-0.0384
zoning:C1	-0.4109	0.2570	-0.9156	-0.4109	0.0932	-0.4109
zoning:C2	0.0417	0.2052	-0.3612	0.0417	0.4444	0.0417
zoning:C3	1.1550	0.9419	-0.6945	1.1550	3.0028	1.1552
zoning:CFB2	0.0952	0.7075	-1.2934	0.0950	1.4835	0.0947
zoning:CFBD	0.2815	0.4371	-0.5767	0.2815	1.1389	0.2816
zoning:CFBG	-0.2434	0.2403	-0.7148	-0.2436	0.2287	-0.2439
zoning:CFR1	0.1179	0.1988	-0.2722	0.1177	0.5086	0.1174
zoning:CFR2	-0.2428	0.2026	-0.6401	-0.2430	0.1554	-0.2435
zoning:CFR3	0.0334	0.2711	-0.4986	0.0333	0.5657	0.0330
zoning:CFUK	0.4527	0.3136	-0.1629	0.4527	1.0679	0.4526
zoning:CN	0.5815	0.6681	-0.7305	0.5816	1.8921	0.5818
zoning:FC	0.7314	0.2293	0.2822	0.7310	1.1822	0.7302
zoning:IG	-0.1992	0.2182	-0.6276	-0.1993	0.2294	-0.1995
zoning:M1	-0.9778	0.2620	-1.4924	-0.9778	-0.4638	-0.9777

	mean	sd	0.025quant	0.5quant	0.975quant	mode
zoning:PR	0.5510	0.2569	0.0469	0.5508	1.0559	0.5503
zoning:R-1	-0.1497	0.5488	-1.2284	-0.1494	0.9261	-0.1487
zoning:R-2	-0.4660	0.6172	-1.6783	-0.4659	0.7445	-0.4656
zoning:R-4	0.5890	0.1895	0.2174	0.5888	0.9614	0.5884
zoning:R1	0.4615	0.0800	0.3043	0.4615	0.6184	0.4616
zoning:R1S	1.1004	0.6685	-0.2120	1.1004	2.4118	1.1004
zoning:R2	0.1168	0.0859	-0.0516	0.1168	0.2854	0.1167
zoning:R3	-0.0713	0.1063	-0.2797	-0.0714	0.1374	-0.0715
zoning:R4	0.1573	0.6748	-1.1676	0.1572	1.4811	0.1572
zoning:RR	-0.2018	0.2662	-0.7226	-0.2025	0.3225	-0.2039
zoning:RR-1	0.2995	0.3205	-0.3284	0.2989	0.9300	0.2977
zoning:RRA1	-0.2894	0.3612	-0.9998	-0.2891	0.4182	-0.2883
zoning:RSF	1.1787	0.1723	0.8403	1.1787	1.5167	1.1788
after-2010	*0.0110*	0.0336	-0.0549	0.0110	0.0771	0.0110
after-2010 x BIP	*0.0674*	0.0451	-0.0212	0.0674	0.1559	0.0674
hs_or_less	-0.1843	0.2176	-0.6126	-0.1840	0.2419	-0.1833
renters	0.1785	0.1606	-0.1411	0.1799	0.4900	0.1827
poverty	-0.0552	0.3088	-0.6616	-0.0552	0.5507	-0.0551
age_65_older	-0.0604	0.4090	-0.8703	-0.0581	0.7359	-0.0533
hispanic	0.1078	0.2451	-0.3730	0.1075	0.5896	0.1069
black	0.1360	0.1619	-0.1821	0.1360	0.4537	0.1360
family	0.4691	0.2070	0.0630	0.4689	0.8758	0.4686
foreign	-0.3589	0.4082	-1.1631	-0.3579	0.4396	-0.3561

Table 6: Fitted model coefficients from the spatial model for BIP WV1103.

Effects	mean	sd	0.025quant	0.5quant	0.975quant	mode
(intercept)	4.0544	6.4612	-8.6459	4.0585	16.7204	4.0672
BIP	0.0969	0.1400	-0.1754	0.0960	0.3738	0.0942
age	0.0001	0.0007	-0.0013	0.0001	0.0014	0.0001
nbaths	0.2211	0.0300	0.1621	0.2211	0.2800	0.2211
sqft_ratio	0.1141	0.1509	-0.1822	0.1141	0.4101	0.1141
acres	-0.4761	0.5856	-1.6279	-0.4754	0.6713	-0.4741
land_square_footage	0.6173	0.5839	-0.5274	0.6167	1.7648	0.6153
bedrooms	0.0427	0.0267	-0.0098	0.0427	0.0952	0.0427
transaction_type:1	0.3582	0.6459	-0.9098	0.3580	1.6259	0.3578
transaction_type:3	0.4559	0.6743	-0.8679	0.4557	1.7793	0.4554
transaction_type:6	0.0106	0.6758	-1.3159	0.0104	1.3371	0.0101
transaction_type:7	0.1679	0.6632	-1.1339	0.1677	1.4695	0.1674
bldg_code:C00	0.3039	1.8017	-3.2372	0.3044	3.8388	0.3057
bldg_code:R00	-0.0142	1.6776	-3.3135	-0.0132	3.2761	-0.0110
bldg_code:RCA	0.0313	1.7582	-3.4259	0.0323	3.4796	0.0345
bldg_code:RM0	-0.1446	1.7466	-3.5796	-0.1434	3.2803	-0.1408
bldg_code:RM1	-0.2206	1.7433	-3.6485	-0.2196	3.1985	-0.2175
bldg_code:RM2	-0.4926	1.7024	-3.8408	-0.4915	2.8459	-0.4890
bldg_code:RS0	0.0165	22.3767	-43.9165	0.0158	43.9127	0.0165
bldg_code:RT0	-0.1787	1.6817	-3.4865	-0.1775	3.1190	-0.1749
bldg_code:X0M	-0.6285	1.6995	-3.9703	-0.6276	2.7050	-0.6256
pri_cat_code:B	-0.8341	0.3817	-1.5835	-0.8342	-0.0852	-0.8342
pri_cat_code:C	-0.8581	0.6471	-2.1284	-0.8582	0.4119	-0.8585
pri_cat_code:D	0.1525	0.3378	-0.5108	0.1525	0.8154	0.1525
zoning:0000	-0.0587	0.2420	-0.5340	-0.0587	0.4159	-0.0586
zoning:A2	0.0482	0.1996	-0.3369	0.0453	0.4502	0.0396
zoning:R2	-0.0769	0.4575	-0.9741	-0.0774	0.8221	-0.0783
zoning:RA	0.0165	22.3767	-43.9165	0.0158	43.9127	0.0165
zoning:RR1	-0.0599	0.2484	-0.5458	-0.0610	0.4315	-0.0630

Table 7: Fitted model coefficients from the spatial model for BIP WV1103.

Effects	mean	sd	0.025quant	0.5quant	0.975quant	mode
after-2010	*0.0018*	0.0934	-0.1817	0.0018	0.1849	0.0019
after-2010 x BIP	*-0.0278*	0.1049	-0.2337	-0.0279	0.1780	-0.0279
hs_or_less	-0.0636	0.8432	-1.7245	-0.0619	1.5866	-0.0585
renters	-0.2416	0.6205	-1.4619	-0.2411	0.9746	-0.2400
poverty	-1.3835	1.4376	-4.2029	-1.3860	1.4464	-1.3908
age_65_older	-0.1044	1.7163	-3.4538	-0.1125	3.2863	-0.1289
hispanic	6.4113	4.8210	-3.0833	6.4194	15.8515	6.4362
black	0.5787	0.6605	-0.7225	0.5797	1.8732	0.5818
family	-0.3155	1.7246	-3.7030	-0.3157	3.0703	-0.3161
foreign	-11.7145	9.6347	-30.5914	-11.7311	7.2356	-11.7635

References

Finley, A.O., Banerjee, S., Carlin, B.P., 2007. spbayes: an R package for univariate and multivariate hierarchical point-referenced spatial models. Journal of statistical software 19, 1.

Hidano, N., Hoshino, T., Sugiura, A., 2015. The effect of seismic hazard risk information on property prices: Evidence from a spatial regression discontinuity design. Regional Science and Urban Economics 53, 113–122.

Keele, L.J., Titiunik, R., 2015. Geographic boundaries as regression discontinuities. Political Analysis 23, 127–155.

Kelejian, H.H., Prucha, I.R., 1998. A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. The Journal of Real Estate Finance and Economics 17, 99–121.

Lee, D.S., Lemieux, T., 2010. Regression discontinuity designs in economics. Journal of economic literature 48, 281–355.

Womble, W.H., 1951. Differential systematics. Science 114, 315–322.