

APPENDIX

A. DERIVATION OF SPATIAL EQUILIBRIUM CONDITIONS

The derivation follows Allen and Arkolakis (2014). This section is built on Allen and Arkolakis (2014) and Redding and Rossi-Hansberg (2017) with modifications that introduce location-specific natural amenities into both land and productivity.

In an economy with N locations, footloose workers in location n have Cobb-Douglas preferences defined over composite goods consumption (C_n) and amenity-adjusted residential land use (H_n).

$$U_n = \left(\frac{C_n}{\alpha} \right)^\alpha \left(\frac{H_n}{1-\alpha} \right)^{1-\alpha}, \quad 0 < \alpha < 1. \quad (1)$$

C_n is the composite good consumption at location n defined as

$$C_n = \left[\sum_{i \in N} \int_0^{M_i} c_{ni}(j)^\rho dj \right]^{\frac{1}{\rho}}, \quad (2)$$

where $c_{ni}(j)$ is a consumption good with a horizontally differentiated variety (j) produced in location i and transported to location n . The first subscript n denotes the location where final consumption occurs. The second subscript i denotes the location where production occurs. M_i is a measure of varieties in location i , endogenously determined in the model.

The composite price index P_n at location n is

$$P_n = \left[\sum_{i \in N} \int_0^{M_i} p_{ni}(j)^{1-\sigma} dj \right]^{\frac{1}{1-\sigma}}, \quad (3)$$

where $p_{ni}(j)$ is the price of variety j that is produced at location i and sold at location n . $\sigma = -\frac{1}{\rho-1}$ is the elasticity of substitution.

To introduce natural amenities into the worker's problem, we assume that the quality-adjusted residential land takes the form:

$$H_n = \bar{H}_n S_n^\xi, \quad (4)$$

where S_n denotes land attributes associated with local amenities in the community n and $\bar{H}_n$ denotes all other land attributes. To maintain positive marginal utility for consumption goods and residential land, we assume $S_n > 0$. A household derives higher utility from land in locations with higher-quality natural amenities, which implies $\xi > 0$ for all n .

A firm to produce $x_n(j)$ units of a variety j in location n use local labor $l_n(j)$ as the only inputs. It operates in a monopolistic competition market. The production incurs a fixed cost of F units of labor and a constant variable cost that depends on local productivity A_n . To introduce the quality-of-business aspect of natural amenities into the firm problem, we assume that local productivity (A_n) can be affected by natural amenities (S_n) and takes the form

$$A_n = \bar{A}_n S_n^{\eta_n}. \quad (5)$$

For amenities that increase labor productivity, $\eta_n > 0$. For amenities that decrease labor productivity, $\eta_n < 0$. The productivity impacts of amenities can vary spatially, as indicated by the subscript n . This spatial variation in the productivity impact of amenities can arise from the specialization of the local economy. For instance, cold weather and heavy snow may be unproductive for the local economies around surf resorts but productive for the local economies around ski resorts.

At a spatial equilibrium, workers should be indifferent between the locations, which implies that real wage income is equal across the locations:

$$\frac{v_n}{P_n^\alpha r_n^{1-\alpha}} = \bar{V} \quad \text{for all } n. \quad (6)$$

The spatial equilibrium of this N -location economy can be determined uniquely (to scale) by the following equation (See A for derivation details):

$$\begin{aligned} & L_n^{\tilde{\sigma}\gamma_1} \bar{A}_n^{-\frac{(\sigma-1)^2}{2\sigma-1}} \bar{H}_n^{-\frac{\sigma(\sigma-1)(1-\alpha)}{\alpha(2\sigma-1)}} S_n^{-\frac{(\sigma-1)}{2\sigma-1} \left[\eta_n(\sigma-1) + \xi\sigma\frac{(1-\alpha)}{\alpha} \right]} \\ &= C_0 \sum_{i \in N} d_{ni}^{1-\sigma} L_i^{\tilde{\sigma}\gamma_2} \bar{A}_i^{\frac{\sigma(\sigma-1)}{2\sigma-1}} \bar{H}_i^{\frac{(\sigma-1)^2(1-\alpha)}{\alpha(2\sigma-1)}} S_i^{\frac{(\sigma-1)}{2\sigma-1} \left[\eta_i\sigma + \xi_i(\sigma-1)\frac{(1-\alpha)}{\alpha} \right]}, \end{aligned} \quad (7)$$

where

$$\begin{aligned} \tilde{\sigma} &= \frac{\sigma-1}{2\sigma-1}, \quad \gamma_1 = \frac{\sigma(1-\alpha)}{\alpha}, \quad \gamma_2 = 1 + \frac{\sigma}{\sigma-1} - \frac{(\sigma-1)(1-\alpha)}{\alpha}, \\ C_0 &= \bar{W}^{1-\sigma} \frac{1}{\sigma F} \left(\frac{\sigma}{\sigma-1} \right)^{1-\sigma} \end{aligned}$$

$\bar{W}$ is a scalar determined by the labor market clearing condition $\sum_i L_i = \bar{L}$.

Substitute the equal welfare condition (6) into the spatial equilibrium condition (7), we have

$$L_n = C_1 S_n^{\xi + \eta_n \frac{\alpha(\sigma-1)}{(1-\alpha)\sigma}} \bar{A}_n^{\frac{\alpha(\sigma-1)}{(1-\alpha)\sigma}} P_n^{\frac{\alpha(1-2\sigma)}{(1-\alpha)\sigma}} \bar{H}_n, \quad (8)$$

where

$$C_1 = \frac{2\sigma-1}{(1-\alpha)\sigma} \ln(\alpha^{-\alpha}(1-\alpha)^{-(1-\alpha)}\bar{V}^{-1}) - \frac{\alpha}{(1-\alpha)\sigma} \ln(\xi).$$

To see how amenity-driven migration affects the equilibrium distribution of population, we can apply Taylor series expansion on equation (8) around observations of L , A , P and S in period t_0 , which expresses migration rate ($\Delta L_n = \frac{L_{n,t} - L_{n,t_0}}{L_{n,t_0}}$) as a function of the growth rate in local labor productivity (ΔA_n), price (ΔP_n), natural amenities (ΔS_n). It drops out of the equation because land supply $\bar{H}_n$ is assumed to be constant.

$$\Delta L_n = \left[\xi + \eta_n \frac{\alpha(\sigma-1)}{(1-\alpha)\sigma} \right] \Delta S_n + \frac{\alpha(\sigma-1)}{(1-\alpha)\sigma} \Delta A_n + \frac{\alpha(1-2\sigma)}{(1-\alpha)\sigma} \Delta P_n. \quad (9)$$

This equation will serve as the basis for empirical analysis. Compared to Equation (8), Equation (9) involves first-order differencing, which helps to control time-invariant unobserved local characteristics.

B. TECHNICAL NOTES FOR THE REGRESSIONS

B.1. Partially linear spatially varying coefficient model using BPST

The partially linear spatially varying coefficient model (PLSVCM) typically takes the form:

$$Y_n = \beta(u_n, v_n)X_n + \boldsymbol{\alpha}^\top \mathbf{Z}_n + \epsilon_n \quad (10)$$

where Y_n is the response variable and ϵ_n is the measurement error. In Equations (6) and (7), the explanatory variable X_n refers to amenity change ΔS_n and the corresponding coefficient $\beta(u_n, v_n)$ is spatially varying, whose input argument $(u, v) \in \Omega \subset \mathbb{R}^2$ typically captures the latitude and longitude coordinates of an observation i . $\mathbf{Z}_n$ is a vector of explanatory variables whose coefficient vector $\boldsymbol{\alpha}$ is not spatially varying. In Equation (6), $\mathbf{Z}_n$ includes ΔA_n and ΔP_n , while, in Equation (7), $\mathbf{Z}_n$ is empty, and $\beta(u_n, v_n) = \beta_{S,n}$ and $\beta(u_n, v_n) = \eta_n$ in Equations (6) and (7), respectively.

Due to the infinite dimensionality of $\beta(u_n, v_n)$, we consider the BPST method, which is an estimation method for bivariate function, especially useful over complex spatial domain (Lai and Wang, 2013; Mu et al., 2018; Kim and Wang, 2021). Wang et al. (2022) considers the PLSVCM using BPST. The BPST estimation starts with a partition of the domain Ω using triangulation, which is a collection of triangles $\Delta = \{\delta_1, \dots, \delta_T\}$ such that $\Omega = \bigcup_{t=1}^T \delta_t$ and the intersection of any two triangles must be either a common vertex or edge. This paper generates triangulation using “DistMesh” available at <http://persson.berkeley.edu/distmesh/>.¹

Given the triangulation Δ , we use bivariate Bernstein basis polynomials defined on each triangle δ_t that are spliced together along the common edges and vertices, to approximate the spatially varying relations $\beta(u, v)$. In particular, the spatially varying relation $\beta(u, v)$ is approximated by a linear combination of basis functions $\mathbf{B}(u, v)$: $\beta(u, v) \approx \mathbf{B}(u, v)^\top \boldsymbol{\gamma}$, for $(u, v) \in \Omega$, where $\mathbf{B}(u, v) = \{B_m(u, v)\}_{m=1}^M$ and $\boldsymbol{\gamma} = \{\gamma_m\}_{m=1}^M$. In this way, the spatially explicit functional relation $\beta(u, v)$ can be expressed as a linear combination of the basis functions with the combination coefficients $\boldsymbol{\gamma}$.

Given the spatial data $\{u_n, v_n, X_n, \mathbf{Z}_n, Y_n\}_{n=1}^N$, to estimate the spline coefficients, we follow Wang et al. (2022) and consider a regularized minimization problem

$$\min_{\beta, \boldsymbol{\alpha}} \sum_{n=1}^N \{Y_n - X_n \beta(u_n, v_n) - \boldsymbol{\alpha}^\top \mathbf{Z}_n\}^2 + \lambda \mathcal{E}(\beta), \quad (11)$$

where

$$\mathcal{E}(\beta) = \int_{\Omega} \left[(D_u^2 \beta(u, v))^2 + 2(D_u D_v \beta(u, v))^2 + (D_v^2 \beta(u, v))^2 \right] dudv,$$

and $D_{s_j}^q \beta(s_1, s_2)$ is the q th order derivative in the direction s_j , $j = 1, 2$, at a location (s_1, s_2) .

Using the BPST approximation, the minimization problem in (11) becomes

$$\min_{\boldsymbol{\gamma}, \boldsymbol{\alpha}} \sum_{n=1}^N \{Y_n - X_n \mathbf{B}(u_n, v_n)^\top \boldsymbol{\gamma} - \boldsymbol{\alpha}^\top \mathbf{Z}_n\}^2 + \lambda \boldsymbol{\gamma}^\top \mathbf{P} \boldsymbol{\gamma}, \quad (12)$$

¹Other triangulation packages include the “Delaunay” algorithm in MATLAB and “DelaunayTriangulation” in MATHEMATICA.

where $\mathbf{P}$ is the diagonal block penalty matrix satisfying $\mathcal{E}(\beta) \approx \boldsymbol{\gamma}^\top \mathbf{P} \boldsymbol{\gamma}$. The second term in the objective function (12) is a roughness penalty based on the second-order derivatives of spline functions. Tuning parameter λ controls the balance between the goodness of fit and the roughness of the function estimate. In particular, a larger value of λ entails a more parsimonious fit but a larger residual sum of squares. In addition, to ensure the smoothness of splines along the shared edges of triangles, a set of linear constraints $\mathbf{H}\boldsymbol{\gamma} = 0$ is imposed on the spline coefficients, where $\mathbf{H}$ is the linear constraints matrix. The form of $\mathbf{H}$ depends on the triangulation and the order of smoothness r . Wang et al. (2022) proposed an efficient algorithm to solve (12) subject to the constraints and also established the asymptotic consistency property of the proposed estimators.

To test whether the observed spatial variation in regression coefficients is statistically significant, we take the global stationarity test:

$$H_0 : \beta(u, v) = \bar{\beta}, \quad (13)$$

where $\bar{\beta}$ is unknown constant effect.

We construct bootstrap-based test statistics by comparing the model fitting under null and alternative hypotheses. Denote RSS_0 and RSS_1 , respectively, as the residual sums of squares (RSS) under the global null hypothesis (Equation 13) and under the alternative hypothesis (Equation 10):

$$\begin{aligned} RSS_0 &= \sum_{n=1}^N \left(Y_n - \hat{\beta} X_n - \hat{\boldsymbol{\alpha}}^\top \mathbf{Z}_n \right)^2, \\ RSS_1 &= \sum_{n=1}^N \left(Y_n - \hat{\beta}(u_n, v_n) X_n - \hat{\boldsymbol{\alpha}}^\top \mathbf{Z}_n \right)^2. \end{aligned}$$

Then, we define the test statistic for the hypothesis (Equation 13) as

$$T_0 = \frac{RSS_0 - RSS_1}{RSS_1}. \quad (14)$$

It calculates the percentage of improvement in RSS if the regression coefficients are allowed to be spatially varying under alternatives rather than fixed as constants in the null hypothesis. The null hypothesis is rejected when the improvement is large or T_0 is large. In order to obtain the p-value or critical value for the test, the bootstrapping method is used to approximate the distribution of test statistics under the null hypothesis. The details of the bootstrapping procedure can be found in (Mu et al., 2018; Kim et al., 2021).

The detailed procedure is as follows:

- Step 1. Based on the data $\{(u_n, v_n, X_n, \mathbf{Z}_n, Y_n)\}_{n=1}^N$, calculate the residuals $\hat{\varepsilon}_n = y_n - \hat{\beta}(u_n, v_n)X_n - \hat{\boldsymbol{\alpha}}^\top \mathbf{Z}_n$, $i = 1, \dots, n$, from the model (10) and obtain the centered residuals $\hat{\varepsilon}_n - \bar{\hat{\varepsilon}}$, where $\bar{\hat{\varepsilon}} = \frac{1}{n} \sum_{n=1}^N \hat{\varepsilon}_n$;
- Step 2. From the empirical distribution of the centered residuals $\hat{\varepsilon}_n - \bar{\hat{\varepsilon}}$ defined in Step 1, generate the bootstrap residuals $\{\varepsilon_n^*\}_{n=1}^N$, and set $Y_n^* = \hat{\beta}(u_n, v_n)X_n + \hat{\boldsymbol{\alpha}}^\top \mathbf{Z}_n + \varepsilon_n^*$;

- Step 3. Using the bootstrap sample $\{(u_n, v_n, X_n, \mathbf{Z}_n, Y_n^*)\}_{n=1}^N$, evaluate the bootstrap test statistic T_{nb}^* ;
- Step 4. Repeat Steps 2 and 3 B times and obtain a set of a bootstrap test statistic $\{T_{nb}^*\}_{b=1}^B$. Then, the corresponding p -value is estimated by $\hat{p} = \sum_{b=1}^B I(T_{nb}^* \geq T_0)/B$, where $I(\cdot)$ is the indicator function.

B.2. Generalized Additive Model (GAM)

Generalized Additive Models typically take the following form:

$$Y_n = a_0 + \sum_{k=1}^p f_k(X_{kn}) + \epsilon_n \quad (15)$$

where a_0 is an intercept. $f_k(\cdot)$ are unknown smooth functions capturing the effects of explanatory variables (X_{kn}) on the mean of the response variable (Y_n). ϵ_n is the independent Gaussian error with mean 0 and variance σ^2 . Our paper has a single variable ($p = 1$). Specifically, Y_n in Equation (8) is $\hat{\beta}_{S,n}$; X_{1n} is $\hat{\eta}_n$; and $f_1(X_{1n}) = f(\hat{\eta}_n)$ on $Y_n = \hat{\eta}_n$. The subscript n denotes the n -th county in the sample. The GAM model makes it possible to identify local variations in the response function.

Due to infinite dimensionality, the smooth function $f(X)$ needs to be approximated using fixed basis functions, $\mathbf{T}(X) = \{T_\ell(X)\}_{\ell=1}^L$, as follows

$$f(X) \approx \mathbf{T}^\top(X) \boldsymbol{\theta} = \sum_{\ell=1}^L \theta_\ell T_\ell(X), \quad (16)$$

where $\boldsymbol{\theta} = \{\theta_\ell\}_{\ell=1}^L$ is a vector of the corresponding coefficients and L is basis complexity.

For a large complexity of smoother, the fitted model would result in overfitting. The roughness penalty is a regularization technique useful to balance model fitting and complexity. The form of the roughness penalty is

$$J(f) = \int \left(\frac{\partial f}{\partial X} \right)^2 dX \approx \boldsymbol{\theta}^\top \mathbf{S} \boldsymbol{\theta} \quad (17)$$

where $\mathbf{S}$ is the penalty matrix.

Let $\boldsymbol{\zeta} = (\boldsymbol{\theta}^\top, \xi_0)^\top$ and $\mathcal{L}(\boldsymbol{\zeta})$ is the log-likelihood based on the regression equation (15) with basis approximation (16). By combining (16) and (17) together, we construct the penalized log-likelihood objective is as follows:

$$\hat{\boldsymbol{\zeta}} = \operatorname{argmax}_{\boldsymbol{\zeta}} \{ \mathcal{L}(\boldsymbol{\zeta}) - \boldsymbol{\zeta}^\top \mathbf{P}_{\lambda^*} \boldsymbol{\zeta} \}, \quad (18)$$

where $\mathbf{P}_{\lambda^*} = \begin{bmatrix} \mathbf{S}_{\lambda^*} & \mathbf{0} \\ \mathbf{0}^\top & 0 \end{bmatrix}$, and $\mathbf{S}_{\lambda^*} = \lambda^* \mathbf{S}$.

Then, the estimated functions become $\hat{f}(x) = \mathbf{T}^\top(x) \hat{\boldsymbol{\theta}}$. In (18), the roughness parameter, λ^* , affects the resulting estimated smoother as those control roughness of each function. They are chosen via the generalized cross-validation (GCV).

Table C1: Robustness Analyses: OLS estimates assuming
 $\hat{\beta}_{S,n} = a_0 + a_1\hat{\eta}_n + \xi_n$

Model	Measure of A	Bachelor						Highschool	
	Triangulization	Δ		Δ_-		Δ_+		Δ	
		Estimate	$Pr(> t)$	Estimate	$Pr(> t)$	Estimate	$Pr(> t)$	Estimate	$Pr(> t)$
Baseline	a_0 : Intercept	-0.0014	< 0.0001	-0.0014	< 0.0001	-0.0014	< 0.0001	-0.0018	< 0.0001
		0.0002		0.0002		0.0002		0.0002	
	a_1 : Slope	0.0050	< 0.0001	0.0053	< 0.0001	0.0053	< 0.0001	0.00630	< 0.0001
		0.0006		0.0006		0.0006		0.00031	
Extended	a_0 : Intercept	-0.0038	< 0.0001	-0.0038	< 0.0001	-0.0038	< 0.0001	-0.0044	< 0.0001
		0.0002		0.0002		0.0002		0.0002	
	a_1 : Slope	0.0074	< 0.0001	0.0078	< 0.0001	0.0080	< 0.0001	0.0087	< 0.0001
		0.0009		0.0009		0.0010		0.0004	

Note: Standard errors of the estimates are reported in parentheses.

There are many types of smoothers for different needs. In this paper, thin plate regression splines (TPRS) (Wood, 2012) are considered as a general-purpose spline basis. The TPRS has a significant benefit as it circumvents the challenges associated with determining the placement of knots in traditional regression spline modeling.

Considering a zero mean Gaussian prior distribution of ζ whose precision matrix is proportional to $\mathbf{P}_\lambda$, the Bayesian posterior distribution of ζ becomes

$$\zeta|X, Y \sim N\left(\hat{\zeta}, (\mathbf{T}_*^\top \mathbf{T}_* + \mathbf{P}_\lambda)^{-1} \sigma^2\right),$$

where $\mathbf{T}_0(X) = \{\mathbf{T}(X)^\top, 1\}^\top$ and $\mathbf{T}_* = \{\mathbf{T}_0(X_1), \dots, \mathbf{T}_0(X_n)\}^\top$.

For the implementation for estimation and Bayesian posterior credible intervals of functions and parameters, we used `gam` function in `mgcv` package (Wood, 2012) in R.

C. DETAILED REGRESSION RESULTS

The regression results for the OLS regression in stage 3 are reported in Table C2. The estimated slope coefficients are significantly positive in all cases, consistent with the model prediction.

Table C2 reports the estimation results for the spatially varying coefficients in the baseline model (Equation 4) and the extended model (including additional explanatory variables). The coefficients for the local labor productivity (ΔA) and price (ΔP) are sensitive to the different model specifications. This is because some of the additional control variables (like population density, bachelor, unemployment, Bartik measure of employment growth, crime rate in 2000) are likely correlated to the local labor productivity and price, even though they are unlikely to be correlated to the natural amenity scale, the critical aspect of the paper. The estimated coefficients for these additional control variables are robust to different model specifications.

Table C2: Estimated non-spatial varying coefficients in the baseline model of migration.

Model	Measure of A	Bachelor						Highschool	
	Triangulization	Δ		Δ_-		Δ_+		Δ	
		Estimate	$Pr(> t)$	Estimate	$Pr(> t)$	Estimate	$Pr(> t)$	Estimate	$Pr(> t)$
Baseline	Intercept	0.0112 (0.0017)	< 0.0001	0.0113 (0.0017)	< 0.0001	0.0112 (0.0017)	< 0.0001	0.0112 (0.0017)	< 0.0001
	ΔA	-0.0001 (0.0013)	0.9365	-0.0001 (0.0013)	0.9359	-0.0001 (0.0013)	0.9366	-0.0073 (0.0014)	< 0.0001
	ΔP	-0.0060 (0.0014)	< 0.0001	-0.0060 (0.0014)	< 0.0001	-0.0059 (0.0014)	< 0.0001	-0.0061 (0.0014)	< 0.0001
Extended	(Intercept)	0.0137 0.0017	< 0.0001	0.0137 0.0017	< 0.0001	0.0137 0.0017	< 0.0001	0.0137 (0.0017)	< 0.0001
	ΔA	0.0035 0.0013	0.0082	0.0035 0.0013	0.0082	0.0035 0.0013	0.0079	0.0032 (0.0017)	0.0621
	ΔP	-0.0011 0.0015	0.4717	-0.0011 0.0015	0.4660	-0.0010 0.0015	0.4825	-0.0009 (0.0015)	0.5619
	pop density	-0.0012 0.0013	0.3487	-0.0012 0.0013	0.3347	-0.0012 0.0013	0.3716	-0.0012 (0.0013)	0.3593
	white	0.0023 0.0017	0.1857	0.0023 0.0017	0.1803	0.0023 0.0017	0.1892	0.0031 (0.0017)	0.0728
	age 65+	-0.0131 0.0015	0.0000	-0.0131 0.0015	0.0000	-0.0132 0.0015	0.0000	-0.0131 (0.0015)	0.0000
	bachelor	0.0112 0.0017	0.0000	0.0112 0.0017	0.0000	0.0113 0.0017	0.0000	0.0114 (0.0018)	0.0000
	unemployment	-0.0024 0.0017	0.1453	-0.0024 0.0017	0.1525	-0.0024 0.0017	0.1514	-0.0032 (0.0017)	0.0582
	Bartik	0.0049 0.0014	0.0007	0.0048 0.0014	0.0008	0.0048 0.0014	0.0008	0.0052 (0.0014)	0.0003
	crime	-0.0038 0.0014	0.0058	-0.0038 0.0014	0.0060	-0.0038 0.0014	0.0054	-0.0038 (0.0014)	0.0057
	dist-msa	-0.0037 0.0015	0.0132	-0.0037 0.0015	0.0129	-0.0037 0.0015	0.0148	-0.0038 (0.0015)	0.0112

Note: Standard errors of the estimates are reported in parentheses.

The results from the urban and rural subsamples are similar to the full sample and are reported in the following tables. The estimated $\beta_{S,n}$ and η_n in stages 1 and 2 under different triangulation and model specifications are reported in Figure C1 and C2. The results are very similar to the estimates from the full sample.

Table C3: Rural-urban divide: OLS estimates for $\hat{\beta}_{S,n} = a_0 + a_1\hat{\eta}_n + \xi_n$

Model	Measure of A	Bachelor				Highschool			
	Region	Rural		Urban		Rural		Urban	
		Estimate	$Pr(> t)$	Estimate	$Pr(> t)$	Estimate	$Pr(> t)$	Estimate	$Pr(> t)$
Baseline	a_0 : Intercept	-0.0015	< 0.0001	-0.0031	< 0.0001	-0.0018	< 0.0001	-0.0033	< 0.0001
		0.0002		0.0003		0.0002		0.0004	
	a_1 : Slope	0.0034	< 0.0001	0.0239	< 0.0001	0.0038	< 0.0001	0.0380	< 0.0001
		0.0007		0.0019		0.0003		0.0027	
Extended	a_0 : Intercept	-0.0043	< 0.0001	-0.0043	< 0.0001	-0.0030	< 0.0001	-0.0052	< 0.0001
		0.00027		0.00027		0.0003		0.0003	
	a_1 : Slope	0.0153	< 0.0001	0.01525	< 0.0001	0.0072	< 0.0001	0.0127	< 0.0001
		0.00060		0.00060		0.0004		0.0008	

Note: Standard errors of the estimates are reported in parentheses.

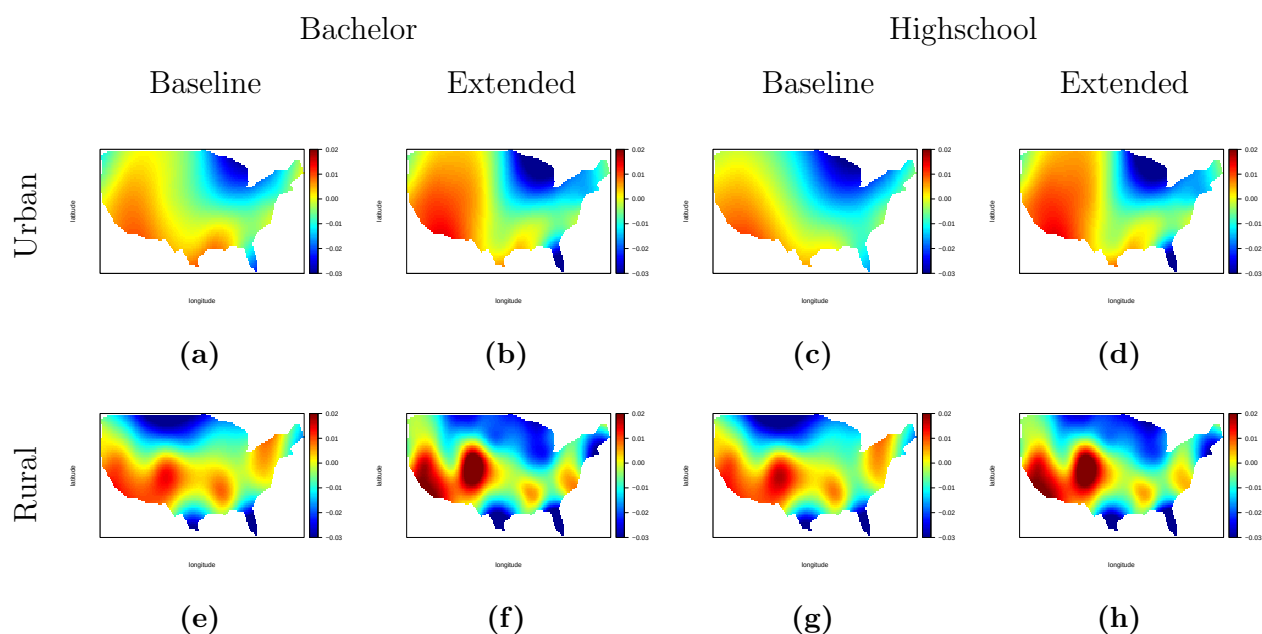
**Figure C1:** Estimates of $\hat{\beta}_{S,n}$ in the urban-rural subsamples.

Table C4: Rural-urban divide: Estimated non-spatially varying coefficients.

Model	Measure of A	Bachelor				Highschool			
	Region	Rural		Urban		Rural		Urban	
		Estimate	$Pr(> t)$	Estimate	$Pr(> t)$	Estimate	$Pr(> t)$	Estimate	$Pr(> t)$
Baseline	Intercept	0.0046	0.0307	0.0297	< 0.0001	0.0048	0.0241	0.0264	< 0.0001
		0.0021		0.0026		0.0021		0.0026	
	ΔA	0.0019	0.2231	-0.0079	0.0002	-0.0029	0.0972	-0.0150	0.0000
		0.0016		0.0021		0.0017		0.0024	
	ΔP	-0.0047	0.0104	0.0018	0.5695	-0.0046	0.0108	0.0045	0.1509
		0.0018		0.0032		0.0018		0.0031	
Extended	(Intercept)	0.0341	< 0.0001	0.0198	< 0.0001	0.0337	< 0.0001	0.0195	< 0.0001
		0.0042		0.0043		0.0042		0.0043	
	ΔA	0.0046	0.0045	0.0001	0.9749	0.0025	0.2268	0.0030	0.3633
		0.0016		0.0023		0.0021		0.0032	
	ΔP	-0.0031	0.1147	0.0007	0.8361	-0.0026	0.1742	0.0001	0.9749
		0.0019		0.0032		0.0019		0.0032	
	pop density	0.1591	0.0000	-0.0012	0.2916	0.1640	0.0000	-0.0013	0.2554
		0.0306		0.0011		0.0307		0.0011	
	white	0.0017	0.4299	0.0040	0.1151	0.0027	0.2102	0.0046	0.0793
		0.0022		0.0025		0.0022		0.0026	
	age 65+	-0.0123	0.0000	-0.0141	0.0000	-0.0118	0.0000	-0.0144	0.0000
		0.0020		0.0026		0.0020		0.0026	
	bachelor	0.0139	0.0000	0.0099	0.0000	0.0130	0.0000	0.0109	0.0000
		0.0029		0.0022		0.0030		0.0023	
	unemployment	-0.0036	0.0808	-0.0034	0.2232	-0.0042	0.0457	-0.0033	0.2272
		0.0021		0.0028		0.0021		0.0028	
	Bartik	0.0049	0.0060	0.0036	0.1599	0.0052	0.0032	0.0041	0.1170
		0.0018		0.0025		0.0018		0.0026	
	crime	-0.0075	0.0000	0.0038	0.0892	-0.0076	0.0000	0.0039	0.0825
		0.0018		0.0022		0.0018		0.0022	
	dist-msa	-0.0039	0.0439	0.0082	0.0555	-0.0038	0.0518	0.0081	0.0575
		0.0019		0.0043		0.0019		0.0043	

Note: Standard errors of the estimates are reported in parentheses.

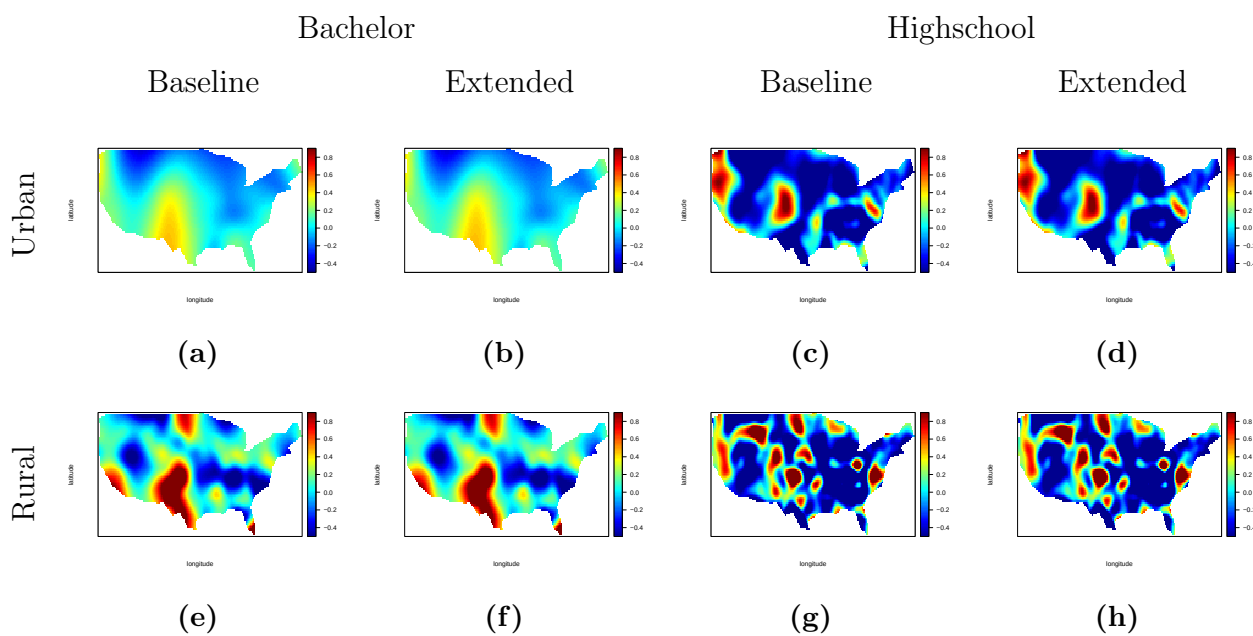


Figure C2: Estimates of $\hat{\eta}_{S,n}$ in the urban-rural subsamples.

REFERENCES

- Allen, T. and C. Arkolakis. (2014) “Trade and the Topography of the Spatial Economy,” *Quarterly Journal of Economics*, 129(3), 1085–1140.
- Kim, Myungjin and Li Wang. (2021) “Generalized spatially varying coefficient models,” *Journal of Computational and Graphical Statistics*, 30(1), 1–10.
- Kim, Myungjin, Li Wang, and Yuyu Zhou. (2021) “Spatially varying coefficient models with sign preservation of the coefficient functions,” *Journal of Agricultural, Biological and Environmental Statistics*, 26, 367–386.
- Lai, Ming-Jun and Li Wang. (2013) “Bivariate Penalized Splines for Regression,” *Statistica Sinica*, 23(3), 1399–1417. <http://doi.org/10.5705/ss.2010.278>.
- Mu, Jingru, Guannan Wang, and Li Wang. (2018) “Estimation and inference in spatially varying coefficient models,” *Environmetrics*, 29(1), e2485. <http://doi.org/10.1002/env.2485>.
- Redding, Stephen J and Esteban Rossi-Hansberg. (2017) “Quantitative Spatial Economics,” *Annual Review of Economics*, 9, 21–58.
- Wang, Yueying, Myungjin Kim, Shan Yu, Xinyi Li, Guannan Wang, and Li Wang. (2022) “Nonparametric estimation and inference for spatiotemporal epidemic models,” *Journal of Nonparametric Statistics*, 34(3), 683–705.
- Wood, Simon. (2012) mgcv: Mixed GAM Computation Vehicle with GCV/AIC/REML smoothness estimation. R package version 1.9-4. Accessed 10-27-2026.