Modeling Spatial Dependence: A Simulation-Based Comparison of Parametric and Semi-Parametric Approaches

Master in Statistics Mathematics

Alejandro M. Ouslan

Supervisors:
Dr. Raul E. Macchiavelli
Dra. Damaris Santana
Dr. Julio C. Hernandez
Dr. Roberto Rivera Santiago

University of Puerto Rico, Mayagüez

November 17, 2025

Abstract

This research aims to compare the performance of spatial regression models that rely on predefined weight matrices with that of semi-parametric regression models using spatial smoothers.

1 Proposal Keywords

Spatial simulation, Spatial Regressions, GAMs, Tensor Products

2 Introduction

Economic models are commonly used to simplify and represent complex real-world relationships. Since all models are approximations, additional components are often introduced to better capture underlying patterns. Especially in micro economices where the assumption of independent and identically distributed (i.i.d) can not be reasonably assumed due to its time factor or close proximity making them influence each outer. Spatial models are not a novel concept. Through simulation, this research seeks to understand the limitations of various spatial models and identify contexts in which specific models are preferable.

More specifically in this research we seek to find reasonable strategies to best specify spatial models, where are it can be reasonably be presumed that there is some spatial effect but there is no clear answer on how spatial agents relate to each other.

3 Background and Motivation

Starting with out, economic theory tells that you can model the productivity of a given economy as a function of labor (L) and capital (K) this can be model in the following equation.

$$Y = AK^{\alpha}L^{\beta}$$

Where:

- Y is the total output (real GDP) produced.
- A is total factor productivity (TFP), capturing the efficiency with which inputs are used.
- K is the quantity of physical capital used in production.
- L is the quantity of labor employed.
- α is the output elasticity of capital, representing the percentage change in output resulting from a 1% change in capital, holding labor constant.
- β is the output elasticity of labor, representing the percentage change in output resulting from a 1% change in labor, holding capital constant.

In practice, though varios data sources L, K and Y are known and α and β are unknown. The unknown parameters are of special interest to economist and for public policy as they highlight areas where local government could invest to increase productivity. However as this models are representations of more complex models is reasonable to asume there is some unseen components in our models. An example of this could be the pretense of underground economies, this referring to transactions that could be reflected in the productivity but can not be reasonably measured through the official means and this could introduce uncertainty into our model.

Turning the Cobb-Douglas equation into a stochastic model can be done in the following format:

$$\begin{aligned} Y &= AK^{\alpha}L^{\beta} \\ \log(Y) &= \log(AK^{\alpha}L^{\beta}) \\ \log(Y) &= \log(A) + \alpha\log(K) + \beta\log(L) + \epsilon; \epsilon \sim N(0,\sigma) \end{aligned}$$

Returning to the original form would give:

$$Y = AK^{\alpha}L^{\beta}e^{\epsilon} \tag{1}$$

3.1 Spatial Modeling

Turning the Cobb-Douglas into a stochastic model that follows the convential linear regression gives access to the ability to introduce space into the model. However it is importent to devine space in this context. Spatial regression models are relatively simple yet expressive tools. The Spatial Durbin Model (SDM) can be expressed as:

$$y = X\beta + \rho W X + \epsilon \tag{2}$$

In this formulation, $X\beta$ represents the standard linear component, and $\rho W X$ incorporates spatial influence through a known weight matrix W. However, in real-world applications, W is not truly known and must be inferred or chosen by the researcher. For example, W can be represented in different forms, such as:

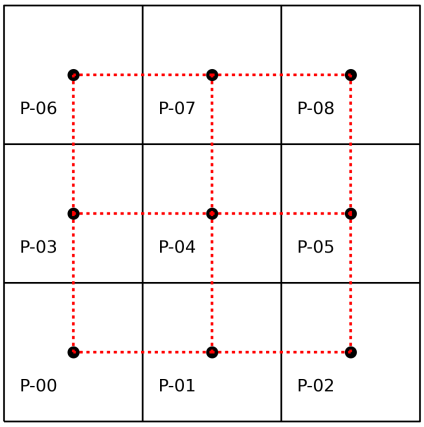

Figure 1: Rook Contiguity Matrix

Mathematically, the rook model can be represented as:

$$\begin{bmatrix}
0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0
\end{bmatrix}$$

Another popular variant is the Queen's contiguity model, which includes all bordering neighbors:

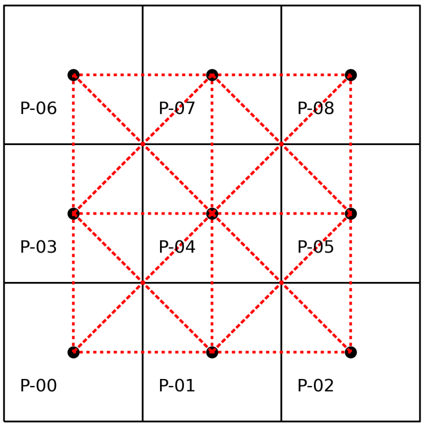

Figure 2: Queen Contiguity Matrix

Its corresponding matrix:

$$\begin{bmatrix} 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \end{bmatrix}$$

Another approach is the k-nearest neighbors (KNN) model, where influence decreases with distance:

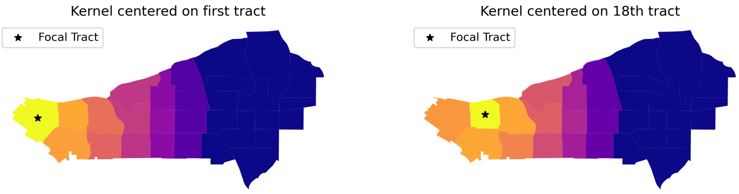

Figure 3: K-Nearest Neighbors (KNN) Model

Clearly, there are numerous ways to define the W matrix, and no universally accepted method exists for choosing the most appropriate one. In practice, W is unknown and selecting it remains a major modeling challenge.

3.2 spatial Cobb-Douglas

Now Turning the Cobb-Douglas into a panel form is given by the follwoing model:

$$y_{it} = \mu_i + \gamma_t + \alpha k_{it} + \beta l_{it} + \epsilon_{it} \tag{3}$$

where:

(a) $\log(A)_{it} \equiv \mu_i + \gamma_t + \epsilon_{it}$

(b) $\log(L)_{it} \equiv l_{it}$

(c) $\log(K)_{it} \equiv k_{it}$

Using the notation form the previos section we can add the spatial componant to the Cobb-Douglas function and returning as the orinal function is as follwos:

$$
\begin{aligned}
y_{it} &= \mu_i + \gamma_t + \alpha k_{it} + \beta l_{it} + \rho \sum_{j=1}^{N} w_{ij} y_{jt} + \epsilon_{it} \\
\log(Y_{it}) &= \log(A_{it}) + \alpha \log(K_{it}) + \beta \log(L_{it}) + \rho \sum_{j=1}^{N} w_{ij} \log(Y_{jt}) \\
Y_{it} &= A_{it} K_{it}^{\alpha} L_{it}^{\beta} \exp \left\{ \rho \sum_{j=1}^{N} w_{ij} \log(Y_{jt}) \right\} \\
Y_{it} &= A_{it} K_{it}^{\alpha} L_{it}^{\beta} e^{\rho \sum_{j=1}^{N} w_{ij} \log(Y_{jt})}
\end{aligned}
$$

$$Y_{it} = A_{it} K_{it}^{\alpha} L_{it}^{\beta} \prod_{j=1}^{N} Y_{it}^{\rho w_{ij}} \tag{4}$$

4 Aims and Objectives

The primary objective of this research is to understand and compare the Spatial Durbin Model (SDM) and a semi-parametric model using tensor product smoothers. We will investigate under what circumstances each model performs better in terms of prediction accuracy and reliability. Furthermore, we aim to evaluate trade-offs such as ease of implementation, assumptions, computational cost, and interpretability.

We restate the SDM as:

$$y = X\beta + \rho W X + \epsilon \tag{5}$$

Where W is predefined by the researcher, though in practice the actual spatial process remains unknown. Choosing an appropriate W is often left to domain experts.

5 Research Plan and Methodology

5.1 Model Comparison

We begin with the classical Ordinary Least Squares (OLS) regression:

$$y_i = \alpha + \sum_{i=1}^{p} x_i \beta_i + \epsilon \tag{6}$$

To account for spatial correlation, we introduce a spatial weight matrix W into the regression:

$$y = X\beta + \rho W X + \epsilon \tag{7}$$

Which expands to:

$$y_{it} = \alpha + \sum_{i=1}^{p} x_{it} \beta_i + \rho \sum_{j=1}^{N} w_{ij} x_{jt} + \epsilon \tag{8}$$

Alternatively, spatial terms may be incorporated into the dependent variable:

$$y = \alpha + X\beta + \rho W Y + \epsilon \tag{9}$$

Or into the error structure:

$$
\begin{aligned}
y &= \alpha + X\beta + u \\
u &= \gamma W u + \epsilon
\end{aligned} \tag{10}
$$

The semi-parametric model with a spatial smoother can be expressed as:

$$y = \alpha + \sum_{i=1}^{p} x_{it}\beta_i + f(C_i) + \epsilon \tag{11}$$

Where C_i is the centroid of each observation and $f(C_i)$ is a spatial smoothing function.

We hypothesize that semi-parametric models may outperform spatial regression models, especially when the true W matrix is unknown or misspecified.

5.2 Spatial Smoother

Previously it was motioned that $f(C_i)$ would be the penalized spatial smoother used as an alternative to using the weight to remove the spatial correlation. Stratign from the basis function we can express any X as a sum of cubic function. This is can be showen in the following equation:

$$f(x) = \sum_{j=1}^{k} b_j(x)\beta_j \tag{12}$$

This would generate the following graph.

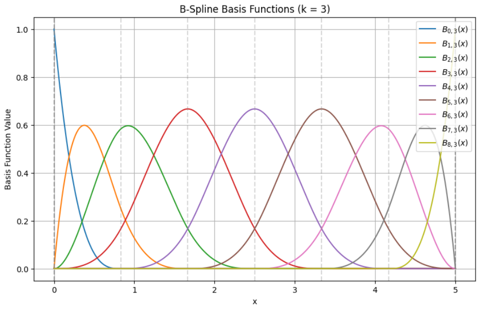

Figure 4: Basis spline functions

Addapting to include other variables can be done by first selecting the basis function for each factor.

$$f_x(x) = \sum_{j=1}^{k} b_j(x)\beta_j \quad f_z(z) = \sum_{j=1}^{k} d_j(z)\delta_j$$

we can then let $b_j(x)$ vary smoothly accoriding to z this give us the following:

$$\beta_i(z) = \sum_{l=1}^{L} \delta_{jl} d_l(z)$$

which gives:

$$f_{xz}(x, z) = \sum_{j=1}^{k} \sum_{l=1}^{L} \delta_{jl} d_l(z)\beta_i(x) \tag{13}$$

implementing this would give the following grpah

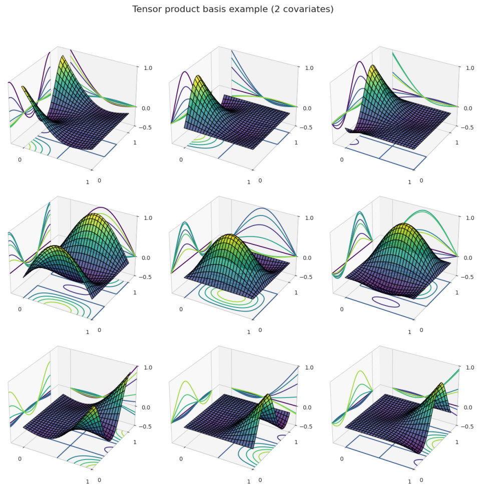

Figure 5: Basis spline functions

We express the tensor–product smoother using the tensor (Kronecker) product design matrix

$$X_{\text{tensor}} = B \otimes D,$$

where B is the spline basis for x and D is the spline basis for z.

The penalized estimation is then obtained by solving the ridge–regularized least-squares criterion

$$\min_{\delta} \|y - X_{\text{tensor}}\delta\|_2^2 + \alpha\,\|\delta\|_2^2.$$

5.3 Cobb-Douglas Implementation

An implementation of this method were it would be of interest is in the production function Cobb-Douglas where there is not much litereature in the effects of spatial component. The Cobb-Douglas is given by the following equation:

$$Y = AK^{\alpha}L^{\beta}$$

Where:

(a) Y is the total output (real GDP) produced.

(b) A is total factor productivity (TFP), capturing the efficiency with which inputs are used.

(c) K is the quantity of physical capital used in production.

(d) L is the quantity of labor employed.

(e) α is the output elasticity of capital, representing the percentage change in output resulting from a 1% change in capital, holding labor constant.

(f) β is the output elasticity of labor, representing the percentage change in output resulting from a 1% change in labor, holding capital constant.

Then we add an error term ϵ where it comes from a normal distribution with mean 0 and constant variance σ^2

$$\log(Y) = \log(A) + \alpha\log(K) + \beta\log(L) + \epsilon; \quad \epsilon \sim N(0, \sigma^2) \tag{14}$$

This functional form assumes constant returns to scale if $\alpha+\beta = 1$, increasing returns to scale if $\alpha+\beta > 1$, and decreasing returns to scale if $\alpha + \beta < 1$.

Turnig this to a stochastic model we convert log both side to get the following:

$$\begin{aligned} Y &= AK^{\alpha}L^{\beta} \\ \log(Y) &= \log\left(AK^{\alpha}L^{\beta}\right) \\ \log(Y) &= \log(A) + \alpha\log(K) + \beta\log(L) + \epsilon \end{aligned}$$

Returning to the original form would give:

$$Y = AK^{\alpha}L^{\beta}e^{\epsilon} \tag{15}$$

6 Prototype Design and Implementation

We will simulate data using the following structure:

$$y \sim \alpha + \sum_{i=1}^{p} x_{it}\beta_i + \rho \sum_{j=1}^{N} w_{ij} y_{jt} + \epsilon \tag{16}$$

With $\epsilon \sim N(0, \sigma^2)$ and multiple forms of W (rook, queen, KNN, etc.) used for robustness testing.

7 Success and Impact

SDM models are highly sensitive to the choice of W, for which no standardized selection method exists. Often, domain expertise is required. The semi-parametric model, however, does not depend on W, making it potentially more robust.

We will evaluate model performance using mean squared error (MSE) of predicted outcomes:

$$\frac{\sum_{n=1}^{N}(y_n - \hat{y}_{nSDM})^2}{N}; \quad \frac{\sum_{n=1}^{N}(y - \hat{y}_{GAM})^2}{N} \tag{17}$$

And by comparing estimated vs. true coefficients:

$$\frac{\sum_{n=1}^{N}(\beta - \hat{\beta}_{SDM})^2}{N}; \quad \frac{\sum_{n=1}^{N}(\beta - \hat{\beta}_{GAM})^2}{N} \tag{18}$$

8 Preliminary resulats

In the following results are for a reguar regression model specified in the following way:

We consider a set of N observations generated from a spatial regression framework. Let y denote the response vector, $X = (X_1, X_2, X_3)$ the explanatory variables, and W a spatial weights matrix (Rook, Queen, or k-nearest neighbors with $k = 6$). The spatial Durbin model (SDM) used in the simulation study is given by

$$y = \rho W y + X\beta + \varepsilon, \tag{19}$$

where ρ is the spatial autoregressive parameter, β are the direct effects

The error term follows

$$\varepsilon \sim \mathcal{N}(0, \sigma^2 I_N).$$

The true parameter values in the data-generating process are

$$\beta = (4, 5, 6, 7)^\top, \qquad \rho = 0.7.$$

Looking over the results we see that the best preforming model is the Queens model as this is the model used to generate the data. The second best preforming model is the penalized tensor products. This shows that if there is no reasonable argument for picking the spatial weights matrix the penalized tensor products are a reasonable starting point for controlling for spatial variability.

Table 1: Simulation results for 1000 runs comparing spatial specifications

	Rook	**Queen**	**KNN (6)**	**Tensor**	**Base**
Intercept	3973.28	0.40	92.49	8252.62	7177.34
X_1	0.74	0.0032	0.30	0.21	1.22
X_2	0.90	0.0032	0.30	0.26	1.57
X_3	1.04	0.0031	0.39	0.27	1.92
ρ	0.43	0.000024	0.0060	–	–