6. Spatial Heterogeneity
6.1 Types and Forms of Spatial Heterogeneity

Spatial heterogeneity:
Structural instability or nonstationarity of economic relationships over space

Heterogeneity can be related to spatial structure or the spatial process generating data.

Two types of heterogeneity

/ \ 
Structural instability  Heteroscedasticity

Structural instability pertains to varying structural parameters (= regression coefficients) over space. Heteroscedasticity stands for different error variances of the spatial units.

Forms and modelling of spatial heterogeneity

• Systematic variation of individual coefficients (→ spatial expansion method, geographically weighted regression)
• Spatial regimes (→ switching regression)
• Random coefficient variation (→ random coefficient model)
• Heteroscedasticity (→ models with heteroscedastic errors)
Testing for heterogeneity ignoring spatial dependence
   (Tests known from Econometrics I)

Tests for heteroscedasticity                  Tests for structural instability
   - White test                                  - Chow test
   - Breusch-Pagan test

Testing for heterogeneity in presence of spatial dependence

   /                                                /
   - Spatial Breusch-Pagan test                    - Spatial Chow test
      (Spatially adjusted Breusch-       (Spatially adjusted Chow test)
       Pagan test)
6.2 The Spatial Expansion Method

Systematical variation of regression coefficients over space can be modelled by a function of a locational variables. The **spatial expansion method** (Cassetti, 1972; Cassetti and Jones, 1987) considers spatially varying regression coefficients as a function of such expansion variables. The locational variables are usually **longitude and latitude coordinates** \((x=z_1\text{ and } y=z_2)\) that are used as continuous variables with decimal points.

**Example:**

Longitude and latitude coordinates of Berlin (Airport Tempelhof)

Longitude: 13°40’ \(\rightarrow\) \(x = z_1 = 13° + (40/60)° = 13.667°\)
Latitude: 52°47’ \(\rightarrow\) \(y = z_2 = 52° + (47/60)° = 52.783°\)

In order to avoid confusing the symbol \(x\) and \(y\) with the fundamental variables of a regression model, we denote the longitude coordinate in the following with \(z_1\) and the latitude coordinate with \(z_2\).
The regression coefficients in $\bm{\beta}_i$ can vary in an economic relationship across all spatial units 1, 2, ..., n:

$$ y_i = \bm{X}_i \cdot \bm{\beta}_i + \varepsilon_i , \quad i=1,2,\ldots,n. $$

Equation (6.1) is called the initial model. For the sake of clarity, we will illustrate the spatial expansion method for a simple regression. When we assume a constant intercept, $\beta_1$, and spatially varying slopes, $\beta_{2i}$, the initial model reads

$$ y_i = \beta_1 + \beta_{2i} \cdot x_i + \varepsilon_i . $$

The regression coefficients $\beta_{2i}$ are viewed as functions of the longitude and latitude coordinates $z_1$ and $z_2$:

$$ \beta_{2i} = \gamma_0 + \gamma_1 \cdot z_1 + \gamma_2 \cdot z_2 $$

We substitute the spatially expanded parameter $\beta_{2i}$ in (6.2) to obtain

$$ y_i = \beta_1 + (\gamma_0 + \gamma_1 \cdot z_1 + \gamma_2 \cdot z_2) \cdot x_i + \varepsilon_i $$

and after rearranging the terms on the right side
(6.4b) \[ y_i = \beta_1 + \gamma_0 \cdot x_i + \gamma_1 \cdot z_1 \cdot x_i + \gamma_2 \cdot z_2 \cdot x_i + \varepsilon_i. \]

Equations (6.4a) and (6.4b) render the terminal model of spatial expansion. On particular the version (6.4b) of the terminal model shows the partitioning of the explanatory variable \( x \) into a common and spatial effects.

**Common and spatial effects**

Regression coefficient \( \gamma_0 \): Strength of the common effect of \( x \) on \( y \)

Regression coefficient \( \gamma_1 \) and \( \gamma_2 \): Spatially varying effects of \( x \) on \( y \)

Variables \( z_1 \cdot x \) and \( z_2 \cdot x \): Spatial interactions (interactions between location and \( x \))

Regression coefficient \( \gamma_1 \): Strength of interaction between longitude and \( x \) on \( y \)

Regression coefficient \( \gamma_2 \): Strength of interaction between latitude and \( x \) on \( y \)
6.3 Geographically weighted regression (GWR)

Geographical weighted regression (GWR) is another method that allows regression coefficients systematically vary over space. With this approach, separate regression models are estimated for each areal unit. All observations are included in each regression but with different weights. The farer away a spatial unit from the region for which the regression is fitted, the lower its data is weighted.

The GWR method accounts for spatial nonstationarity in the economic relationship. Response parameters of an economic model may be location-specific. From a theoretic point of the relationships are intrinsically different across space.

The GWR equation for the \( i \)th region is given by

\[
y_i = \beta_1(u_i,v_i) + \beta_2(u_i,v_i) \cdot x_{i2} + \ldots + \beta_k(u_i,v_i) \cdot x_{ik} + \varepsilon_i.
\]

The location-specific regression coefficients \( \beta_t(u_i,v_i) \) are functions of longitude and latitude coordinates \( u_i \) and \( v_i \). More precisely, \( \beta_t(u_i,v_i) \) is a realisation of the continuous function \( \beta_t(u,v) \) at point \( i \).
The local parameters $\beta_j(u_i, v_i)$ are estimated by a **weighted least-squares procedure**. The weights $w_{ij}$, $j=1,2,...,n$, at each location $(u_i, v_i)$ are defined by a function of the distance $d_{ij}$ between the centre of the region $i$ and those of the other regions.

We have to estimate $n$ unknown vectors of local regression coefficients:

$$
\begin{bmatrix}
\beta_1(u_1, v_1) \\
\beta_2(u_1, v_1) \\
\vdots \\
\beta_k(u_1, v_1)
\end{bmatrix}, \quad
\begin{bmatrix}
\beta_1(u_2, v_2) \\
\beta_2(u_2, v_2) \\
\vdots \\
\beta_k(u_2, v_2)
\end{bmatrix}, \ldots, \quad
\begin{bmatrix}
\beta_1(u_n, v_n) \\
\beta_2(u_n, v_n) \\
\vdots \\
\beta_k(u_n, v_n)
\end{bmatrix},
$$

The GWR estimates of the unknown local parameter vectors $\hat{\beta}(i)$ are given by

$$
\hat{\beta}(i) = [X'W(i)X]^{-1} \cdot X'W(i)y, \quad i=1,2,...,n.
$$

$W(i)$ is the nxn spatial weight matrix which has the form

$$
W(i) =
\begin{bmatrix}
w_{i1} & 0 & \cdots & 0 \\
0 & w_{i2} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & w_{in}
\end{bmatrix}
$$
• Weighting schemes

Several weighting functions ("kernels") are available for calculating the weighted least-squares estimators (6.7). The kernels generally define weights such that observations of regions near a point in space, for instance the centre of a region, have more influence on the estimator than observations of areas further away.

The most common kernels are
- the Gaussian function and
- the bi-square function.

Fixed Gaussian weighting function:

\[(6.9) \quad w_{ij} = \exp[-(d_{ij}/b)^2/2]\]

Fixed bi-square weighting function:

\[(6.10) \quad w_{ij} = \begin{cases} [1 - (d_{ij}/b)^2]^2 & \text{if } d_{ij} < h \\ 0 & \text{otherwise} \end{cases}\]

b is the so-called bandwidth and controls the degree of distance-decay:
- b large → smoothing effect high (relative higher weight for more remote regions)
- b small → smoothing effect low (relative lower weight for more remote regions)
Figure: Fixed weighting function

- **Regression point**
- **Data point**

\( w_{ij} \) is the weight of data point \( j \) at regression point \( i \)

\( d_{ij} \) is the distance between regression point \( i \) and data point \( j \)
Problems with fixed bandwidth:
- Undersmoothing in areas with only small observations
- Oversmoothing in areas with a high density of data points

**Adaptive kernels** (spatially varying kernels)

Kernels (bandwidth) smaller in regions with high density of data points and larger in regions with low density of data points

Figure: Adaptive kernel

X regression point
● data point
Spatially adaptive kernels (weighting functions)

Adaptive **bi-square weighting function**

\[
(6.11) \quad w_{ij} = \begin{cases} 
[l - (d_{ij} / b)^2]^2 & \text{if } j \text{ is one of the } s \text{th nearest neighbour of } i \\
0 & \text{otherwise}
\end{cases}
\]

Adaptive **weighting function with ranked distances**

\[
(6.12) \quad w_{ij} = \exp(-R_{ij} / b)
\]

\(R_{ij}\): Rank of the \(j\)th point from \(i\) according to the distance \(d_{ij}\)
6.4 Spatial regimes

Spatial regimes are defined by groups of regions for which different responses of a set of explanatory variables on the dependent variable are assumed. The simplest case is a **two-regime model**:

\[
\begin{bmatrix}
  y_1 \\
  y_2 
\end{bmatrix} = \begin{bmatrix}
  X_1 & 0 \\
  0 & X_2 
\end{bmatrix} \begin{bmatrix}
  \beta_1 \\
  \beta_2 
\end{bmatrix} + \begin{bmatrix}
  \varepsilon_1 \\
  \varepsilon_2 
\end{bmatrix}
\]

(6.13)

The error variance is usually assumed to be different in the two-regime model (**groupwise heteroscedasticity**).

Income convergence or production elasticities can, for instance, be different for West and East Germany. In this case, the convergence equation and the production function should be modelled by an West-East regimes model which can be estimated by the method of maximum likelihood.

A **multi-regimes model** could consist of German macroregions such as the German states. Not only the regression coefficients can be different in **switching regressions** but also the explanatory variables.

In spatially switching regressions often spatial dependence is accounted for by a spatially autoregressive error process.