# Alternative Measures of Weight Matrices

Dr. Şebnem Er
(with Prof. Anil K. Bera & Dr. Neslihan Fidan)

University of Cape Town
Statistical Sciences Department
Sebnem.Er@uct.ac.za

June 5, 2014

# Spatial Dependence

Spatial dependence occurs in a situation where values of a variable observed at one geographical location, labelled with $i = 1, ..., n$, depend on the values of its neighbouring data points or nearby locations labelled with $j$, where $j \neq i$.

Why does spatial dependence occur? Two main reasons (LeSage, 1998, pp.4-5):

# Spatial Dependence

Spatial dependence occurs in a situation where values of a variable observed at one geographical location, labelled with $i = 1, ..., n$, depend on the values of its neighbouring data points or nearby locations labelled with $j$, where $j \neq i$.

Why does spatial dependence occur? Two main reasons (LeSage, 1998, pp.4-5):

- measurement error related with the data collection process of geographical units

# Spatial Dependence

Spatial dependence occurs in a situation where values of a variable observed at one geographical location, labelled with $i = 1, ..., n$, depend on the values of its neighbouring data points or nearby locations labelled with $j$, where $j \neq i$.

Why does spatial dependence occur? Two main reasons (LeSage, 1998, pp.4-5):

- measurement error related with the data collection process of geographical units
- the spatial dimension of the research problem may play an important role in analysing the rationale behind the problem

# Exploring Spatial Dependence

The level of spatial dependence can be explored with graphical techniques or some statistical measures.

# Exploring Spatial Dependence

The level of spatial dependence can be explored with graphical techniques or some statistical measures.

- Graphical representation of spatial dependence

# Exploring Spatial Dependence

The level of spatial dependence can be explored with graphical techniques or some statistical measures.

- Graphical representation of spatial dependence
- Moran's I statistic,

# Exploring Spatial Dependence

The level of spatial dependence can be explored with graphical techniques or some statistical measures.

- Graphical representation of spatial dependence
- Moran's I statistic,
- Geary's C Statistic

# Exploring Spatial Dependence

The level of spatial dependence can be explored with graphical techniques or some statistical measures.

- Graphical representation of spatial dependence
- Moran's I statistic,
- Geary's C Statistic

Cliff and Ord (1972, 1975, 1981) have proved that Moran's I statistic is consistently a better measurement than the Geary's C Statistic in many cases.

# Moran's I Statistic

$$I = \frac{\mathbf{z}'\mathbf{W}^s\mathbf{z}}{\mathbf{z}'\mathbf{z}}$$

# Moran's I Statistic

$$I = \frac{\mathbf{z}'\mathbf{W}^s\mathbf{z}}{\mathbf{z}'\mathbf{z}}$$

Here,
$\mathbf{z}'$ is the mean standardized Y variable
$\mathbf{W}^s$ is the row standardized weights matrix that represents the weights of spatial relationships between data points.
$\mathbf{W}^s$ is constructed using the connectivity matrix $\mathbf{W}$ which is in a binary format, where $\sum_{i,j=1}^{n} w_{ij}^s = n$

# Moran's I Statistic

$$I = \frac{\mathbf{z}'\mathbf{W}^s\mathbf{z}}{\mathbf{z}'\mathbf{z}}$$

Here,
$\mathbf{z}'$ is the mean standardized Y variable
$\mathbf{W}^s$ is the row standardized weights matrix that represents the weights of spatial relationships between data points.
$\mathbf{W}^s$ is constructed using the connectivity matrix $\mathbf{W}$ which is in a binary format, where $\sum_{i,j=1}^{n} w_{ij}^s = n$

$$-1 < I < 1$$

# How to Construct Weights Matrices

For geographic data, though not trivial, there are two main choices for W:

# How to Construct Weights Matrices

For geographic data, though not trivial, there are two main choices for W:

- Spatial Weights Based on Distances

# How to Construct Weights Matrices

For geographic data, though not trivial, there are two main choices for W:

- Spatial Weights Based on Distances
- Spatial Weights Based on Boundaries

For financial data, there is not an obvious way of constructing weights matrices.

# Weights Based on Distances

Binary or real valued spatial weights are based on the distances between the centroids of each observation. There are several approaches in defining the observations that are connected based on distances between them (Arbia, 2006, pp.37-38, Fischer et al., 2010, pp.66-68)

# Critical cut-off based connectivity

1. Critical cut-off based connectivity

$$w_{ij} = \begin{cases} 1 & 0 \leq d_{ij} < d^* \\ 0 & otherwise \end{cases}$$

# K-nearest neighbour connectivity

2. K-nearest neighbour connectivity

$$w_{ij} = \begin{cases} 1 & d_{ij} \in min_k \left\{ d_{ik} \right\} \\ 0 & otherwise \end{cases}$$

# Inverse distance connectivity

3. Inverse distance connectivity

$$w_{ij} = d_{ij}^{-1}$$

# Distance decay function connectivity

4. Distance decay function

$$w_{ij} = exp\left(-\alpha d_{ij}\right)$$

5. Combination of the several

# Weights Based on Boundaries

These types of spatial weights define two observations connected if they share a common boundary. Contiguity based measures are more relevant for a polygon type data set or for observations that have a polygon attribute. There are three different ways to represent the connectivity based on the boundaries (Arbia, 2006, pp.37-38, Fischer et al., 2010, pp.66-68, LeSage, 1999, pp.11-13),such as:

1. rook
2. queen
3. bishop

# Rook

1. Rook: Rook based contiguity assigns two observations connected if they share a side in common

# Bishop & Queen

2. Bishop: Bishop based contiguity assigns two observations connected if they share an edge in common
3. Queen: Queen based contiguity assigns two observations connected if they share a side or an edge in common

# For Economic and Financial Data

Several research papers:

- Caporin and Paruolo, (2005). Multivariate ARCH with spatial effects for stock sector and size.
    - Weights matrices according to industrial sector similarity
    - Weights matrices according to same capitalization size
    - An interaction of both - such that neighbours if they have simultaneously the same sector and size
- Fernandez, (2011). Spatial Linkages in International Financial Markets.
    - They consider four financial indicators to quantify the distance between "neighboring" firms: market capitalization (relative to firm size), the market-to-book, EV/EBITDA and debt ratios.

# For Social and Financial Data

- Asgharian, Hess and Liu, (2011-2013). A spatial analysis of international stock market linkages
  - They analyze a number of different linkages: geographical neighborhood, similarity in industrial structure, economic integration (measured by the degree of countries bilateral trade and bilateral foreign direct investment) and monetary integration (measured by the closeness of countries inflation rates, two measures based on interest rate differences and the stability of their bilateral exchange rate).
- Moon and LeSage, (2011). Revisiting the question Does corporate headquarters location matter for stock returns? and Moon and LeSage. (2011). Simultaneous Dependence between Firm-Level Stock Returns
  - Metropolitan Statistical Area
  - Sectoral similarities

# For Economic and Financial Data

- Avilès, Montero, Orlov, (2012). Spatial modeling of stock market comovements.
    - Physical distance
    - proximities in an economic-relations space and to measure distances using FDI linkages between countries
- Arnold, Stahlberg and Wied, (2013). Modeling different kinds of spatial dependence in stock returns.
    - The weights of the firms in the Euro Stoxx 50.
    - Stocks belonging to the same branch (Wb) or same country (Wl).

# Application

- Istanbul Chamber of Industry (ICI) firms' sales and net asset profitability measures for the period 1997-2011.
- $N = 179$ (only those firms that appear at a continuous base from 1997-2011)
- The firms are operating in several cities of Turkey and mainly in the manufacturing sector.

# Defining the Weights Matrices

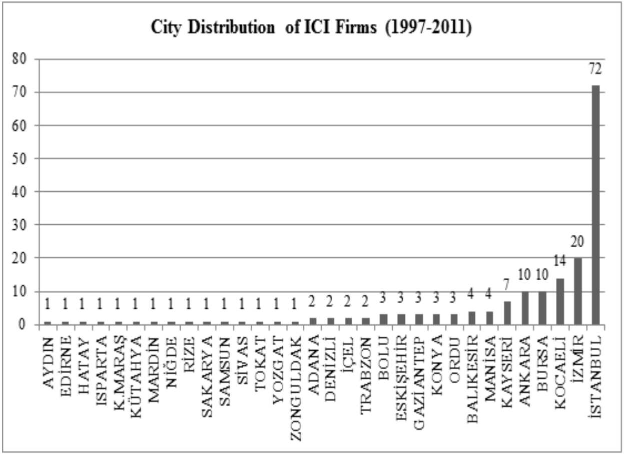
- We used two sets of weights matrices:
  1. Contiguity based weights matrices
  2. "**Financial distance**" based weights matrices

# Contiguity based weights matrices

The contiguity based weights matrices are formed according to the NUTS definitions of the cities that the firms are located. We used three different NUTS level in order to define which firms are connected:

1. NUTS-3 Level Contiguity (81 cities)
2. NUTS-2 Level Contiguity (26 sub-regions)
3. NUTS-1 Level Contiguity (12 regions)

# Distribution of the firms at NUTS-3 Level



# Financial distance based weights matrices

We used three different financial distances:

1. The first one is the sector information of the firms assuming that firms operating in the same sector would be interacting with each other more than the ones that are in different sectors.
2. The second and the third distances are based on firm size information which is measured using firm equity and productivity levels of the firms.

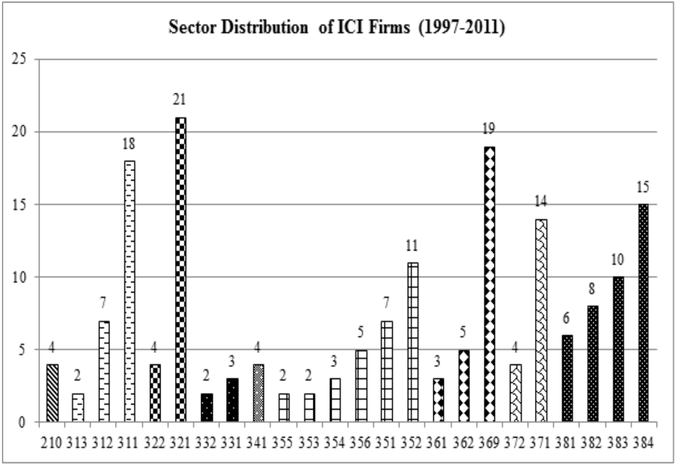
# Sectoral based weights

We have used three different levels of sector divisions:

1. Sub-sector division (29 subsectors)
2. Main-sector division (9 main sectors)
3. A combination of (1) and (2), where,

$$w_{ij} = \begin{cases} 1 & \text{if i and j are in the same subsector} \\ 0.5 & \text{if i and j are in the same main sector} \\ 0 & \textit{otherwise} \end{cases}$$

# Sectoral based weights



# Firm equity as a measure of size

- Firm size based on firm equity levels (10 levels)
- Firm size based on firm equity levels (5 levels)

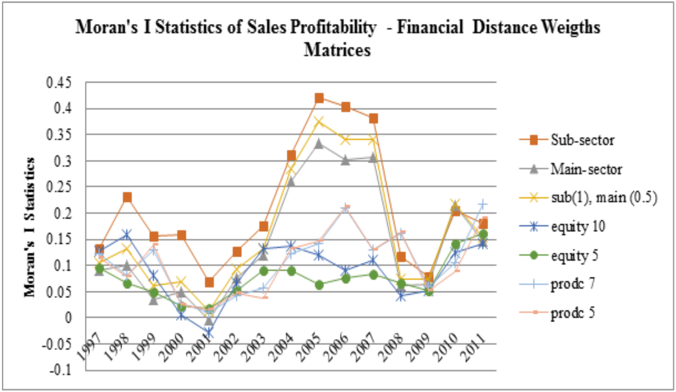
# Firm productivity as a measure of size

- Firm size based on firm productivity levels (7 levels)
- Firm size based on firm productivity levels (5 levels)

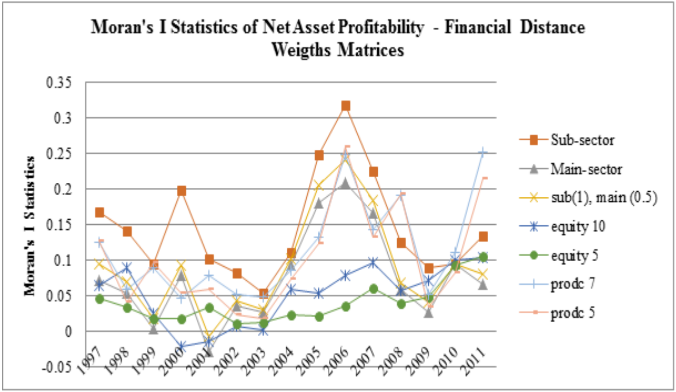
# Results

- No significant Moran's I statistics for geographical contiguity based weights matrices
- Significant dependence is found when "neighborhoods" are defined based on the sectoral, equity or productivity similarities of the firms which were considered as financial distances between firms.
- Moran's I statistics are insignificant for the two economic crisis years in Turkey (2001 and 2008) and sometimes even negative.
- For both sales and net asset productivity measures, the Moran's I statistics that are calculated using the weights matrices based on sectoral similarities outperform all the other Moran's I statistics among financial distances with the highest spatial dependency levels.

# Sales Profitability



# Net Asset Profitability



# Conclusion

- For geographic data, the choice of weights matrices **W** is obvious, since there are distance or boundary based weights matrices that could be used.
- However, we claim that for financial data geographical weights matrices may not be relevant for measuring the spatial dependencies.
- We analysed spatial dependencies of Istanbul Chamber of Industry (ICI) firms' sales and net asset profitability measures for the period 1997-2011 using ten different weights matrices.

# Conclusion

- When we use the geographical location of the firms (the regional contiguity based neighborhoods) to construct **W**, we hardly find any evidence of spatial dependence with insignificant and very small Moran's I statistics.
- Significant spatial dependence is found when "neighborhoods" are defined based on the sector or productivity similarities of the firm (except the economic crisis years in Turkey).
- Among all different financial distance weights matrices, the weights matrices based on sectoral similarities provided the highest level of spatial dependencies.
- This finding supports the fact that with financial data, spatial units might be spatial dependent if they are financially closer rather than being geographically closer.

# Acknowledgements

My special thanks to Prof. Bera and Prof. Arbia and my friend and colleague Dr.Fidan

# Thank you!

Teşekkürler...
Baie dankie...
Enkosi...
Grazie...