The matching process among regional Tunisians labor markets

Abstract

This paper extends the application of spatial econometrics models to analyze interactions and heterogeneity among governorates in the Tunisian labor market based on the matching function. Due to the limited information provided by the contiguity matrix, we suggest a spatial weight matrix based on multidimensional scaling to tackle heterogeneity and spatial agglomeration; The socio-economic weight matrix summarizes interactions among the 24 governorates given the characteristics of each one (socio demographic environment, employment and economic environment, and quality of life) and the flow weight matrix that takes into account the flows of migrations among regions, using the technique proposed in Hondroyiannis (2009). To test the proposed matrices, we have used spatial econometrics techniques on panel data (Spatial Durbin model, spatial error model, and spatial autoregressive model) to take into account spatial effects (proximity interaction and spatial heterogeneity). Our results show that jobseekers in neighboring governorates cause a difficulty in the matching process in the local market.

Keywords: Matching function, Weight matrix, Heterogeneity, Spatial econometrics, Multidimensional scaling

1. Introduction

The most worrying dysfunction of labor markets at the international level is the consequence of high unemployment rate. In Tunisia, since 1986, the date of the adoption of the structural adjustment program, the country has recorded some important macroeconomic performances. The average growth rate was about 5% over the past 30 years, these performances contrast with an unemployment rate that persists at high levels of about 15%; in a corridor of plus or minus one point. Due to the demographic transition which reinforce the active age groups. Also, another characterization of unemployment is the heterogeneity for example (governorates on the coastline are less affected by unemployment). In addition, the rate of unemployed graduates in 2016 accounts for roughly one-third (31.9%) of all unemployed persons, of whom 61.7% are females. Despite, the previously depicted issues cause a decline in the Tunisian economic activities. This recession fueled the stagnation of Tunisian economy following the revolution of January 14, 2011, it has largely affected the employment market where the unemployment rate recorded its highest value reaching 18.9% in 2011. Conversely, this sudden increase exponentially rose the pressure ratio achieving 8 points at the end of 2012, in addition the unsatisfied demand has increased sharply. Analogously, this marked increase can be explained in large part by the economic growth difficulties and the decline in such activities since the revolution, even the elected government has failed to adopt a perform employment policy following the launch of "active job search system" program which prompted many applicants to enroll in the employment offices to benefit from the financial advantages.

At the same time, unemployment rates are differentiated in space, in Tunisia the gap between interiors and littoral regions on the scale of unemployment rates is wide. However, it is clear that the current knowledge of the problem does not allow the implementation of adequate measures to curb its development. This finding calls more than ever for research on the functioning of the labor market, with a view to better understanding the mechanisms of the origin of unemployment and to improve the effectiveness of economic policy in this field. Furthermore, the labor market has a spatial dimension that must be taken into account. Indeed, workers and firms are scattered in space, which goes beyond the hypothesis of a centralized confrontation of offers and demand of work of the standard theory. As a result, adjustments are slow due to the very imperfect geographical mobility of this factor of production. To bare on the problem of heterogeneity. Following, the scientific progression in geographic information systems highly motivates us to introduce this dimension to our study. Therefore, the techniques of spatial econometrics adopted, in order to more accurately measure the extent of the effects of interdependence in space. However, the spatial amount of knowledge distributed more directly. In the first part, our proposition of 3 spatial models to test the most appropriate one: spatial autoregressive (SAR), spatial error model (SEM) ,and spatial Durbin model (SDM), to the matching function, we also include control variables as determinants of matching efficiency and heterogeneity, where “quality_life” denotes the quality of life indicator, “demographic_env” presents the socio demographic environment indicator for each governorate, and “economic_env” the indicator of employment and economic environment for each governorate.

This work is organized as follows. The first section reviews the literature on the matching function. The second section provides details of the datasets and some comments on the policy employed in Tunisia to reduce the unemployment rate and regional disparities an explanation of the estimation methodology for the spatial effects model and outline the alternative measures of distance that we aim to test. In the next section, we present results and discussions; particularly, we shed the light to the different measure of distance to model the spatial dependency, and the last section holds the most relevant conclusions and future research.

2. ***Literature review***

The issue of matching process between jobseekers and vacancies has generated a large and growing literature.

We notice several attempts to tackle the unemployment problem, in this paper our attention is shifted to those who tries to analysis whether the spillover effect, spatial externalities influence or both, in the matching process.

Several researches valorize the spatial dimension, by assuming that regions or even countries which should not be considered as isolated geographical units. In our case, the potential links between the three components of the matching function should be taken into account, which are the number of hired jobseekers, the number of jobseekers and the number of vacancies in a governorate with those of its neighbors. **Upton and Fingelton (1985)** define the spatial dependence by the existence of a systematic spatial variation between the values distributed on a map. Spatial dependence may be due to externalities effects whose frequency or intensity depends on the distance between cities and the mobility of the labor factor. Indeed, these different dynamics imply a particular organization of the efficiency of matching through space.

The first study in this context is conducted by ***Burda and Profit (1996)*** the aim of their study is to test the geographic instability in the labor matching process, they consider the search intensity endogenous and the outflows from unemployment is essentially due to local and neighboring labor market conditions. The main finding from their work is that external unemployment affects the local labor matching process, where shorter distances causes a positive externality whereas longer distances causes a negative externality. An extended version of the latter analysis is supposed in ***Burgess and Profit (1998)*** they study the effect of unemployment and vacancies inflows on labor matching process by suggesting the methodology of travel-to-work area (TTWA). They explored the impact of neighboring regions on local labor markets. In particular, the results figure out that high unemployment in nearby regions implies the increase of the number of filled vacancies in local areas, but fall the outflow from unemployment. Thereby the high number of vacancies in surrounding regions increases the local outflow from unemployment and local outflow of filled vacancies.

A common finding is provided by ***Ilmakunnas and Pesola (2003)*** and ***Kosfeld (2007)*** who find also that spatial dependence stimulates (dependent) cyclical volatility. They also show that unemployment in neighboring regions generate a negative effect on local labor market, while vacancies expend a positive effect. Besides ***Kosfeld (2007)*** has shown that these effects are imbalanced across space.

Dmitrijeva and Mihails Hazans (2007) conduct a stock flow model estimation, his finding that vacancy inflows exert a negative externality for local matches in Latvia, another relevant result in Slovenia shown that unemployment inflows in neighboring areas increases local job creation.

To incorporate spatial spillover through borders ***Hynnien (2005)*** uses exogenous variables lagged in space, he points out that the congestion effect occurs between jobseekers in local labor markets and the spatial spillover intensify.

The issue of return to scale is studied in ***Lopez Tamayo et al. (2000)*** at different level of data aggregation, by taking into account the spatial interactions between the different areas by using the inverse distance. the SPDM is implied to study the labor market matching process, the results show that at lower data aggregation the hypothesis of constant return to scale is rejected.

The spatial dependency problem in labor market is analyzed in ***Lottman (2013)*** using a spatial autoregressive model to regional labor market in Germany. He concludes that spatial dependency captured by geographic distance, and for the matching function the dynamic modeling by including a spatially lagged independent variable is more convenient.

The matching process in labor market in Germany with fixed effect and random effect was been applied in ***Stops (2011)*** he proposes an "occupational topology" and tested the assumption of non-separated occupational labor markets using a modified version of SDM by including spatial lags for regressors. the results depict the existence of dependency among similar occupational groups in the matching process.

3. Data and variables

3.1. Data

The data was collected from 109 employment agencies distributed on the 23 governorates of Tunisia. their role is essentially to provide information on employment and professional qualifications for companies and jobseekers. In order to estimate the aggregated Tunisian matching function and to deduce the nature of the returns to scale, we use annual data covering the period 1995-2014. The data is usable as an indicator of the matching flow and the number of annual placements made. It represents the total number of offers during this period and transmitted by ANETI at the level of each governorate.

The job application and job offers provide the U and V stock data for each governorate. Inventories U include persons who are seeking an employment and registered in the employment office files for each pool. As in the hiring series (H), these figures may be overestimated because there is nothing that can be done to prevent an employee from registering as a jobseeker, as it is difficult to estimate intensity with which each individual conducts his or her job search. Also, it should be noted that these figures can be overestimated with the disconnection between the different regional employment offices, since an unemployed person can be registered in various offices simultaneity. The series of vacancies represent the number of offers received by each office belonging to one of the 24 labor pools.

[insert table 1]

3.2. Variables

The introduction of control variables [1]as determinants of matching process and heterogeneity. In other words, theses variables characterize each region, where "quality_life" denotes the quality of life indicator, "demographic_env" presents the socio demographic environment

[1] The modeling procedures of these variables is reviewed in appendix A.

indicator for each governorate, and “economic_env” the indicator of employment and economic environment for each governorate.

3.3. Descriptive statistics

We give an overview of the most relevant variables used through a descriptive statistics table presented below:

[insert table 2]

This table summarizes variables used for 24 studied governates from 1994 to 2014.
However, the time variable **“year”** indicates the period of our study which lasts 20 years, and the space variable **“ID_1”** refers to the 24 governates of Tunisia. In fact, the variable **“U”** labeled unemployed workers,” V**”** vacant jobs and **“H”** the hired workers. however, by the first glance we can say that the mean of unemployed workers represents the quadrupled of vacant jobs which express the high pressure on the Tunisian labor market. In addition, for U and V the **SD** seems to high compared to the mean. Likewise, the gap between the maximum and minimum looks significant mainly for U and V.
As a matter of fact, these gaps maybe express the differences of U and V distribution among regions or their distribution over time or both.

3.4.Regional analysis

For regional analysis, we adopt a regional analysis by examining the relationship between **U, V, H** and **"regions",** to deeply describe our data. The figure below overviews the most relevant governorates.

[insert figure 1]

The maps above summarize the mean and SD for each governorate during the period 1994 to 2014.

The highest unemployed workers mean from 1994 to 2014 is recorded in **“Tunis”** which known by the highest density of population. while it comes in the second range in the number of vacant jobs after **“Nabeul”** which reached a mean of 12588,90 vacant jobs during 21 years. Although the touristic activities have provided these vacant jobs, the majority of these activities are cyclical. Thereby, those cyclical activities don’t contribute employment. Besides, **“Sfax”** comprise the maximum hiring mean about 9978,10 from 1994 to 2014. due to manufacturing and trading of oil and the development of the economic infrastructure: the creation of a modern port and Industrial zone. Make of “**Sfax”** known as the economic capital of Tunisia.

Whereas coastal zones are characterized by the highest averages of vacant jobs and hired workers, for the regions of northeast, east and south east the gap between unemployment and vacancies are significant compared to coastal zones however, these facts can be expressed by the agglomeration of the economic activities in the coastal zones. whereas, the regions of south are more affected by this phenomenon.

The analysis of the table above makes significant the inequality among regions. however, industrial areas and the development of infrastructure which attracts investors make the coastline regions less touched by unemployment. while the regions of northeast, east and south east suffers from the extreme unemployment and the less vacant jobs.

3.5. Moran test

The table above display Moran's statistic from 1994 to 2014 in order to detect spatial autocorrelation.

[insert table 5]

The governorates of Tunisia were negatively autocorrelated. Besides, the spatial interaction between regions were stronger for vacancies than unemployment. At the first glance, we can note that the unemployment stock became non-significant since 2011 , this marked decline by 182% can be explained in large part by the economic growth difficulties and the decline in such activities since the revolution of 14 January 2011, and this recession persists in 2012 where the elected government has failed to adopt a perform employment policy following the launch of "active job search system" program which prompted many applicants to enroll in the employment offices to benefit from the financial advantages.

During the period of study (1994 to 2014) the spatial interaction decreased by 236% for unemployment stock whereas it increased by 186% for vacancy stock.

3.6. Spatial models

It is broadly perceived that information gathered from geologically close elements are not independent, but rather spatially associated, which implies that perceptions of nearer units have a tendency to be more comparative than further ones **Tobler (1970)**.
The question of agents' location in space suppose a major problem in economic theory. we focus on the most relevant works presented in geographical economic theory starting by

Ricardo (1823) who defines the theory of localization, **weber (1909)** when he studies the ideal location of firms.

The work of **Paelinck and Klaseaen (1979)** is considered the cornerstone of spatial econometrics. However, they introduce spatial factors in the economic reasoning and defining the first spatial econometrics model.

Besides, analysis of international trade theory considers the basis of the new geographical economy defined by Paul Krugman. He was awarded the Nobel prize in economic sciences in 2008 on his pioneering work **Paul Krugman (1991)** by adding spatial factors to answer the question of free trade and globalization influences in economics.

In fact, those researches highly motivate researchers to incorporate spatial factors in several domains as well as industrial economy, agriculture, education and employment market.

However, spatial models are essentially used to handle spatial spillover among neighboring regions. This effect depends of the distance among regions. In other words, it intensifies with neighboring regions and reduces if the distance increases.

3.6.1. Spatial autoregressive model approach

The SAR model poses the question of dependency between the dependent variable in region i and y of neighbors $j \neq i$ and on characteristics of each region**. Anselin et al. (2006, p.6)** consider "the spatial lag model is typically considered as the formal specification for the equilibrium outcome of a spatial or social interaction process ".

$$\log H_{it} = \alpha + \rho \sum_{j=1}^{n} w_{ij} \log H_{jt} + \beta_1 \log(U_{i,t}) + \beta_2 \log(V_{i,t}) + \beta_3 \text{quality–life} + \beta_4 demographic + \beta_5 economic + \mu_i \gamma_t + \varepsilon_{it} \quad (3.3)$$

3.6.2. Spatial Durbin model approach

The spatial Durbin model (SDM) extends the spatial autoregressive model (SAR) by taking into account spatially lagged independent variables.

$$\log H_{it} = \alpha + \rho \sum_{j=1}^{n} w_{ij} \log H_{jt} + \beta_1 \log(U_{i,t}) + \beta_2 \log(V_{i,t}) + \beta_3 \text{quality– life} + \beta_4 demographic$$
$$+ \beta_5 economic + \ \beta_6 \text{ w1xlog}(U_{i,t}) + \beta_7 \text{ w1xlog}(V_{i,t}) + \beta_8 \text{w1xquality– life}$$
$$+ \beta_9 w1xdemographic + \beta_{10} w1xeconomic + \mu_i \gamma_t + \varepsilon_{it} \qquad (3.5)$$

Where, α the constant term, β the spatial interaction effects term, $\mu_i \gamma_t$spatial specific effect term which controls all specific space variable invariant in time, $w1x\beta$ denotes the spatially lagged independent variables, ρ the parameter of spatial autoregression, $w_{ij} y_{jt}$ spatial lag variable that grant interactions between spatial entities and ε_{it} iid term.

3.6.3. Spatial error model approach

The spatial error model suggests the problem of dependency on characteristics and disturbance term is correlated across space. In other words, a random shock in a region i affects not only the dependent variable of this region but also the dependent variable in neighboring regions.

$$\log H_{it} = \alpha + \beta_1 \log(U_{i,t}) + \beta_2 \log(V_{i,t}) + \beta_3 \text{quality– life} + \beta_4 demographic + \beta_5 economic$$
$$+ \mu_i \gamma_t + \lambda \sum_{j=1}^{n} m_{ij} v_{it} + \varepsilon_{it} \quad (3.4)$$

Where, λ symbolizes the spatial autocorrelation component, and ν_{it} represents the spatially autocorrelated error term.

3.7. Migration flow matrix

Spatial statistics differentiate to traditional by introducing a new dimension that tackle the issue of bias and perform results. Interactions between spatial entities are plotted in a spatial weight matrix. the most common one is the binary weight matrix it contains information of every regions combination regardless of whether they are considered neighbors or not. Several kinds of spatial neighboring are used to construct the spatial weight matrix [2] such as queen contiguity, rook contiguity and k-nearest neighbor.

From another point of view, we notice the concentration of the economic activities in such regions. besides, to the agglomeration of population in certain governorates which highly motivate us to suggest a spatial weight matrix that take into account the flows of migrations among regions. It provides us both handling the spatial agglomeration in such regions and control the outflow of employment migration which differentiates among the governorates as specified by **Hondroyiannis (2009)**. For example, the governorate of Tunis accounted for 1 073 640 in 2015 in an area of 288 Km². Contrariwise, Tataouine amounted 149 870 in 2015 in an area of 38 889 Km².

$$w_{ij} = \left(\frac{1}{d_{ij}}\right)\left(\frac{x_{ij} + x_{ji}}{x_{i.} + x_{j.}}\right)\left[1 - \frac{\left|\frac{x_{ij}}{x_{i.}} - \frac{x_{ji}}{x_{j.}}\right|}{\frac{x_{ij}}{x_{i\cdot}} + \frac{x_{ji}}{x_{j.}}}\right] \quad (3.10)$$

Where, d_{ij} denotes the distance between the governorates i and j. x_i and x_j symbolizes the outflows of the migration of employees from i to j, $x_{i.}$and $x_{j.}$ designates the sum of outflow from the governorates i and j respectively. Alternatively, the first term expresses the inverse distance which is considered as an interpolation method. The second term valorizes the connectivity between i and j by measuring the outflows between i and j in relation to that with the other governorates to compute the importance of flow between i and j. Finally, the third term gauge the relative importance between i and j.

[insert table 3]

The matrix above [3]summarizes the flow of working migrants among Tunisians governorates, where the row presents the exposure. In other words, its incorporate the effect of the inflow from other regions and the column depicts the pressure which measure the effect of the outflow of region i to others regions.

in other words, it quantifies the degree of connection between the governorates via the outflows and inflows of migrants between them. On one hand, littoral governorates are the most affected by the inflows from interiors regions, where the highest inflows recorded in Tunis which accounted for 1

[2] Modeling details in appendix B
[3] The constructed weight matrix is too sizable 24X24 we present a sample of 5X5 in this section the full weight matrix is loaded in the appendix C

073 640 in 2015 in an area of 288 Km². On the other hand, north-western and south regions recorded the highest outflows of migrants.

We can explain the wave of incoming flows to the coastal regions and the high population density in these regions by the concentration of economic activities in these governorates (industry, tourism), which forces job applicants to migrate to these regions.

4. Model estimation and discussions

In this analysis, we will start by building a model based on 3 weights matrices. Using the mixed approach presented by **Elhorst (2010)** based on 3 models SDM, SAR, SEM[4]. The dependent variable "ln_H" which denotes the log of hired jobseekers in each governorate from 1994 to 2014 whereas, the independent variables are "ln_U" and "ln_V" refers to the log of jobseekers and vacant jobs respectively at each governorate from 1994 to 2014. Moreover, each indicator has a facet of regional development in the governorate. Where, "quality_life", "demographic_env", "economic_env" denotes quality of life, Socio-demographic environment, Economic environment and employment correspondingly.

4.1 Model specification

[insert figure 1]

The choice of the model. Can be done through 3 approaches. First, the bottom-up approach starting from a non-spatial model **Gallo (2002)** in addition to the Lagrange multiplier test as specified by **Anselin (1996)**. Second, the top down approach starts from the spatial Durbin model and conducting the likelihood ratio tests to choose the most appropriate model as specified in **LeSage and Pace (2009)**.

Finally, **Elhorst (2010)**, define the mixed approach which begins with "bottom-up" in the case of spatial interactions $\rho \neq 0$ *and* $\lambda \neq 0$, instead of choosing directly SAR or SEM, we study the spatial model of Durbin tests (Lagrange multiplier and likelihood ratio) to confirm the relevance of the model, this allows us to integrate exogenous interactions in the model. In case of uncertainty we retain SDM.

[insert table 4]

[4] SDM refers to spatial Durbin model, SAR: spatial autoregressive, SEM: spatial error model

To determine the form of spatial autocorrelation (endogenous, exogenous or unobserved), the approach is pragmatic. **Elhorst (2010)** approach would lead to retaining the SDM model, the set of spatial models is nevertheless estimated, for 2 neighborhood matrices: contiguity, and flow of migrants' workers. Which has the strongest explanatory character (AIC the weakest) and whose interpretation economy is the most intuitive.

Concerning the choice of the model, we can retain the following points in Table 4.

The statistical criteria would lead to an SDM model. All the spatial autocorrelation tests carried out from the residuals of the MCO model are rejected. Also, Lagrange ratio tests to SDM are rejected which confirm our choice of SDM model for the 3 weights matrix defined.

Comparing the SDM for the 2 proposed matrices. The flow of migrants' matrix seems the most appropriate one, by comparing the log likelihood where the latter lowest value recorded for the flow matrix.

The analysis in Table 4 aims to examine the robustness of the results according to the spatial specifications, to highlight the possible sensitivity of the effects according to the structure of spatial interactions and specify the underlying spatial processes in terms of direct, indirect and external effects.

The spatial autocorrelation parameter $\hat{\rho}$ is significant and positive for the 2 matrices. Tunisians governorates are not isolated from each other. In particular, the assumption of no dependency between their hiring process over the period 1994-2014 cannot be rejected. The log of hired jobseekers of a governorate increases by $\hat{\rho}$ % (with $-2.4890 < \hat{\rho} < 1$) when the hiring capacity of all its neighboring governorates is on average 1%.

We see from the results that all explanatory variables exhibit a significant effect on the variable we wished to explain except "demographic_env" and the spatially lagged variables "w1x_ln_U", "w1x_Quality_life". In other words, the coefficients β_1 is negative whereas β_2 is positive which implies that 1% increase of β_1 will reduce the hiring capacity by 0.09% and 1% increase of β_2 will rise the hiring capacity by 0.99%. The coefficient β_2 is significant despite the proximity matrix specified. The differences in elasticity of the number of matches to the number of vacancies and job seekers respectively can explains the inability of occupational integration policies to absorb large numbers of jobseekers as well as the effects of the stagnation of the Tunisian economy in 2011 following political instability. In addition, the fairly low and negative value of β_1 due to the existence of an informal labor markets operating in Tunisia especially by temporary workers, and the phenomenon of smuggling which have aggravated the situation of the Tunisian employment market. Also, the number of jobseekers registered in the ANETI employment agencies is not really representative of the total population in search of employment it covers only 37.5% of the total unemployed. Whereas, the share of higher educated reached 62.5%.all these factors may justify the low value of β_1.

Besides, spatially lagged dependent variable "w1x_ln_V", "w1x_economic_env", and "w1x_demographic_env" are statistically significant. In other words, for region i the flow of migrants' workers from governorates $i \neq j$ affects negatively the numbers of vacancies in region i. Which can be explained by the wave of migrants' workers from interior governorates to littoral regions which negatively affects the hiring in these governorates. While, a 1% increase in "w1x_economic_env", or "w1x_demographic_env" in regions $j \neq i$ will increase the hiring in region i by 6.3% or 3.3% respectively. In other words, the results display the magnitude of the demographic environment and economic environment to attracts jobseekers which rose its hiring capacity.

6.3. *Decomposition of spatial effects*

The table [5]below summarizes the decomposition of spatial effects into direct and indirect effects. Indeed, this decomposition allows us to determine the extent of the effect of externalities, measured by the indirect effect, in the hiring capacity.

[insert table 6]

The direct and indirect effects of "ln_V" are significant and positive. For the direct effect, an increase in the number of vacancies in a region i will improve its hiring capacity while stimulating its exposure to the flow of migrants. Moreover, the indirect effect implies that the hiring capacity (ln_H) of other regions ($j \neq i$) increases, thus relaxing the flow of migrants' pressure while the feedback effect [6]of the rising the hiring capacity (ln_H) of other regions ($j \neq i$) implies the saturation of labor market in region i, which is under a severe exposure because of the wave of incomings flows. This further confirm, that there is a positive spillover effect (indirect effect) for the number of vacancies (ln_V). For example, an unemployed resident in a governorate has two options: either he agrees to remain unemployed in his governorate of residence, or decides to move to the neighboring governorates to maximize his chances of finding a job. this decision is strongly linked to the transportation constraints; the distance,

[5] Note: p-value: * p<0,1 , ** p<0,05, *** p<0,01

[6] The feedback effect represents the difference between direct effect and the coefficient beta. It will be positive if Direct > Beta, otherwise negative

accessibility, transportation and other relevant costs. Indeed, such constraints are the cause of unemployment in the governorates which already has a low integration capacity.

In addition, the direct effect of “ln_U” is negative and non-significant. An increase in the number of jobseekers in a region i will not have an impact to its hiring capacity, whereas the indirect effect is also negative but significant which implies that an increase in the number of jobseekers in a region i weaken the hiring capacity (ln_H) of other regions ($j \neq i$) , thus relaxing the flow of migrants’ pressure. In fact, the feedback effect of stimulating the hiring capacity (ln_H) of other regions ($j \neq i$) intensify the number of jobseekers in i which strengthen the flow of migrants’ pressure to i. In other words, an increase in the number of jobseekers didn’t stimulate the hiring capacity, contrariwise its decline the pressure on labor market which explain the difficulty faced by jobseekers to migrate due to transportation costs and other relevant costs.

Besides, the total effect of quality of life indicator “quality_life” , “demographic_env” ,and “economic_env” are significant. While, the direct and indirect effects are both non significants. In fact, an increase in the indicator “quality_life” , “demographic_env” ,and “economic_env” will not have an impact to the hiring capacity neither in a region i nor other regions ($j \neq i$).

The direct and indirect effects of spatially lagged dependent variables “w1x_ln_V” is significant and negative, the direct effect can be explained by the increase in the number of job vacancies in regions $j \neq i$ contribute negatively the capacity of hiring of these regions which can due to their exposition to the flow of migrants’ workers from region i. In addition, the indirect effect clarifies that a fall in the hiring capacity of governorates $j \neq i$ will decrease the hiring capacity in i due to the incoming flow from other regions. while, the feedback effect is positive, the decline in hiring capacity of region i, will implies an increase in “ln_V” in regions $j \neq i$ which stimulate the exposure of i to the inflow of migrants. In other words, job creation will stimulate negatively the hiring capacity due to the wave of incoming inflow from neighbors which exceed the capacity of the governorate to absorb this additional demand.

In addition, the direct effect of “w1x_demographic_env”, and “w1x_economic_env” are both non-significant and positive. Whereas, the indirect effect is positive and significant for both which clarifies that rising the hiring capacity in regions $j \neq i$ will implies the rise of the hiring capacity in a region i. Afterwards the feedback effect implies that increase of “ln_H” in region i cause the fall of hiring capacity in regions $j \neq i$.

variables	Description
H	The number of hired jobseekers for each governorate from 1994 to 2014
U	The number of jobseekers for each governorate from 1994 to 2014 registered in employment agency.
V	The number of vacancies for each governorate for each governorate registered in employment agency.
Quality_life	Quality of life indicator for each governorate
Deomgrphic_env	Socio demographic environment indicator for each governorate
Economic_env	Indicator of employment and economic environment for each governorate

Table 1: Data dictionary

	N	Minimum	Maximum	Mean	Std. Deviation
Year	504	1994	2014	2004,00	6,061
ID_1	504	1	24	12,50	6,929
U	497	2138	68348	19135,71	12203,949
V	497	189	20205	4814,52	4177,106
H	497	148	16797	3869,54	3520,845
Valid N (listwise)	497				

Table 2:Data description

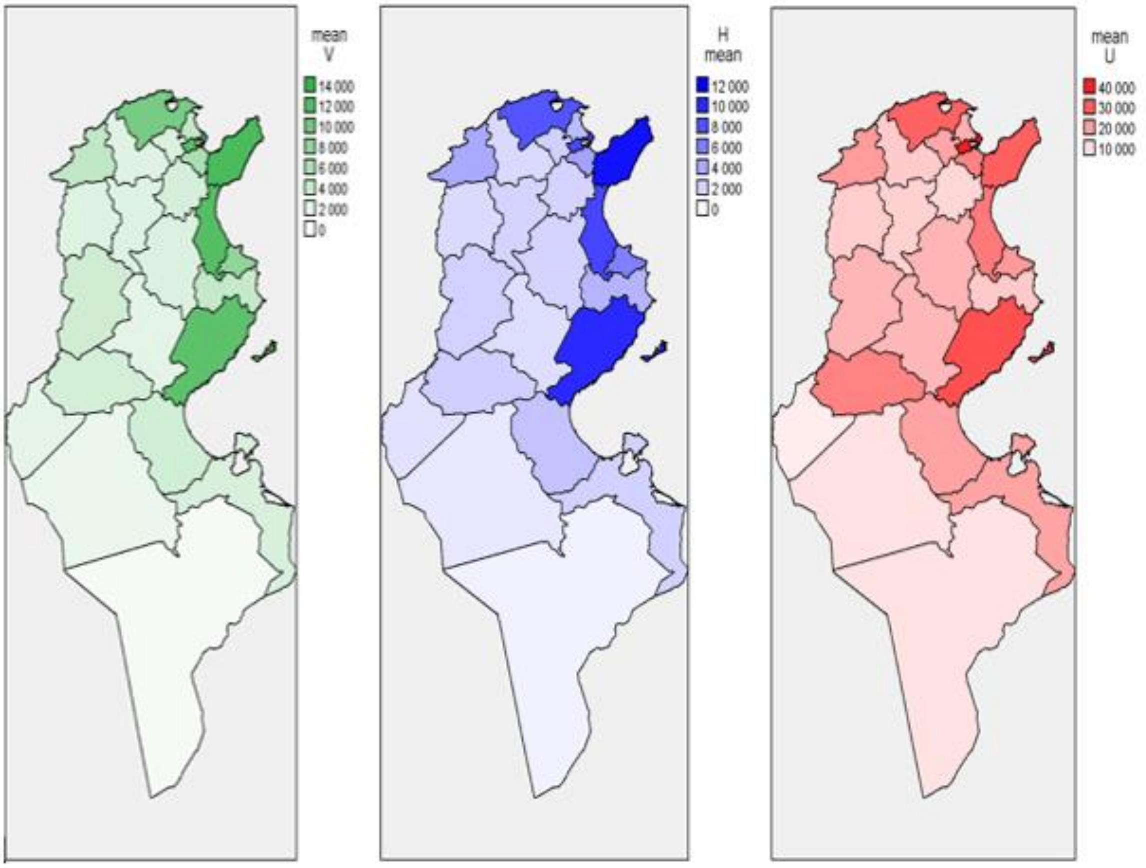

Figure 1: Regional analysis maps

	Ariana	Beja	Ben Arous	Bizerte	Gabes
Ariana	0	1,020	0,112	0,537	2,980
Beja	1,020	0	1,039	0,884	2,999
Ben Arous	0,112	1,039	0	0,642	2,872
Bizerte	0,537	0,884	0,642	0	3,416
Gabes	2,980	2,999	2,872	3,416	0

Table 3: Migration flow matrix between Tunisians governorates

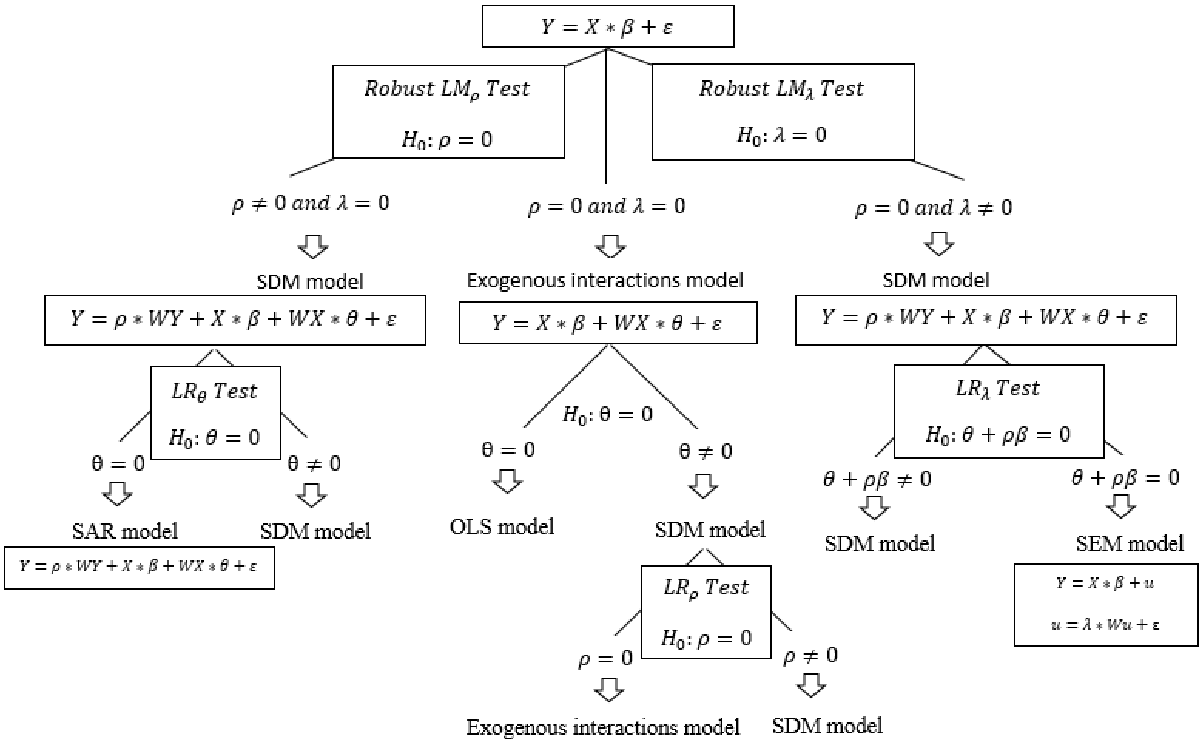

Figure 1: Model specification mixed approach Elhorst (2010)

	Flow matrix			Contiguity matrix		
LN_H	**SAR*** **	**SEM** **	**SDM** **	**SAR*** **	**SEM*** **	**SDM** **
LN_U ($\boldsymbol{\beta_1}$)	-0.2371904	-0.105522	-0.0916585	-0.244523	-0.1161051	0.0320168
	(0.000)	(0.000)	(0.001)	(0.000)	(0.013)	(0.245)
LN_V ($\boldsymbol{\beta_2}$)	1.109441	1.028469	0.9958402	1.075575	1.039787	0.9087016
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
QUALITY_LIFE($\boldsymbol{\beta_3}$)	0.0122953	0.0160286	0.0370961	0.0338158	0.0154593	0.0319742
	(0.393)	(0.300)	(0.031)	(0.022)	(0.275)	(0.030)
DEMOGRAPHIC_ENV($\boldsymbol{\beta_4}$)	0.0325054	0.025149	0.0685676	-0.043127	0.0218808	-0.0279888
	(0.143)	(0.254)	(0.065)	(0.063)	(0.338)	(0.279)
ECONOMIC_ENV($\boldsymbol{\beta_5}$)	-0.0392515	-0.0093537	0.1537182	0.1221078	0.0802647	0.0759136
	(0.405)	(0.842)	(0.023)	(0.016)	(0.075)	(0.152)
W1X_LN_U($\boldsymbol{\beta_6}$)			-0.0009563			-0.2605367
			(0.982)			(0.000)
W1X_LN_V($\boldsymbol{\beta_7}$)			-0.8930244			-0.6434627
			(0.000)			(0.000)
W1X_QUALITY_LIFE($\boldsymbol{\beta_8}$)			0.2133054			-0.174762
			(0.542)			(0.000)
W1X_DEMOGRAPHIC_ENV($\boldsymbol{\beta_9}$)			3.339474			0.1041192
			(0.002)			(0.053)
W1X_ECONOMIC_ENV($\boldsymbol{\beta_{10}}$)			6.316938			-0.4110245
			(0.002)			(0.000)
$\boldsymbol{\alpha}$	-0.8350028	0.9927586	9.771198	0.3796951	0.6441778	2.593665
	(0.021)	(0.018)	(0.000)	(0.171)	(0.023)	(0.000)
/RHO	0.2541967		0.7515584	0.1529024		0.6383371
	(0.000)		(0.000)	(0.000)		(0.000)
/LAMBDA		0.9626171			0.5069911	
		(0.000)			(0.000)	
/SIGMA	0.2689039	0.2939908	0.2816877	0.2752622	0.2578198	0.2648287
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
LOG LIKELIHOOD FUNCTION	-51.9376	-144.6407	-87.9962	-63.6882	-48.7589	-74.9228
R2 (R-SQUARED)	0.8986	0.1811	0.7903	0.8989	0.8889	0.7215
LR TEST SAR VS. OLS (RHO=0)	(0.000)	(0.000)	(0.0000)	(0.0000)	(0.000)	(0.0000)
GLOBAL MORAN MI	(0.000)	(0.0000)	(0.0000)	(0.0000)	(0.000)	(0.0000)
LM ERROR (BURRIDGE)	(0.000)	(0.0000)	(0.0000)	(0.0746)	(0.0746)	(0.0000)
LM ERROR (ROBUST)	(0.000)	(0.0000)	(0.0000)	(0.000)	(0.000)	(0.000)
LM LAG (ANSELIN)	(0.000)	(0.6348)	(0.0000)	(0.0000)	(0.000)	(0.0000)
LM LAG (ROBUST)	(0.000)	(0.0000)	(0.0006)	(0.0000)	(0.000)	(0.0000)

Table 4: model estimation

year/var	U	V
1994	-0,168***	-0,111***
1995	-0,153***	-0,121***
1996	-0,171***	-0,154***
1997	-0,174***	-0,156***
1998	-0,138***	-0,146***
1999	-0,148***	-0,147***
2000	-0,139***	-0,173***
2001	-0,144***	-0,131***
2002	-0,167***	-0,163***
2003	-0,169***	-0,16***
2004	-0,135***	-0,153***
2005	-0,15***	-0,167***
2006	-0,151***	-0,169***
2007	-0,152***	-0,169***
2008	-0,124***	-0,171***
2009	-0,113***	-0,195***
2010	-0,12***	-0,197***
2011	-0,066	-0,165***
2012	-0,043	-0,16***
2013	-0,069	-0,199***
2014	-0,071	-0,206***

Table 5: Moran test

Variable	*Beta*	*Total*	*Direct*	*Indirect*	*Feedback*	*Mean*
Ln_H						
Ln_U	-0.0917	-0.0831***	-0.0228	-0.0603**	0.0689	9.2433
Ln_V	0.9958	0.9027***	0.2474***	0.6553***	-0.7484	7.7665
Quality_life	0.0371	0.0336*	0.0092	0.0244	-0.0279	-0.9543
Demographic_env	0.0686	0.0622*	0.0170	0.0451	-0.0516	-0.8144
Economic_env	0.1537	0.1393**	0.0382	0.1011	-0.1155	-0.7408
W1x_ln_U	-0.0010	-0.0009	-0.0002	-0.0006	0.0008	9.2070
W1x_ln_V	-0.8930	-0.8095***	-0.2219***	-0.5876***	0.6711	7.6769
W1x_quality_life	0.2133	0.1934	0.0530	0.1404	-0.1603	-0.9333
W1x_demographic_env	3.3395	3.0271***	0.8297	2.1974**	-2.5098	-0.8753
W1x_economic_env	6.3169	5.7260***	1.5694	4.1566**	-4.7475	-0.7021

Table 6:decompositttion of spatial efffect

Appendix B

5. Weight matrix construction

In order, to incorporate the spatial effect of the models as presented in the previous section, we shed the light to various attempts to handle this effect by suggesting a weight matrix such as: rook contiguity, queen contiguity, and nearest neighbors that resumes the distance between governorates. Thus, the contiguity matrix has shown its limit. In fact, regions on borders are characterized by a few number of neighbors, furthermore it didn't represent the characteristics of each spatial entities, and neglects the weight of each neighbor.

Thereby, to overcome these problems we suggest two weight matrices: the economic weight matrix and migration flow matrix using the multidimensional scaling techniques.

2.1 Multidimensional scaling

The aim of the pioneering work of **Richardson (1938)** in psychometrics is to understand people judgement of similarity among set of objects and to establish dimensionality from data. Further **Torgerson (1952)** was extended Richardson's finding to introduce the classical multidimensional scaling which tackle the issue of mapping data, by determining the coordinates of each item using euclidean distance, using dissimilarity among each pair of objects as input and the coordinate matrix as output which minimizes the loss function presented below.

$$D = \left(\frac{\sum_{i,j} \left(b_{ij} - x_i x_j \right)^2}{\sum_{ij} b_{ij}^2} \right)^{\frac{1}{2}} \qquad (3.8)$$

Borg and Gronen (1997) have relaxed the assumption of Euclidean distance where the MDS has becoming a proximities distance which are derived from confusion data and correlation, another extension by **Krustal (1964,a,b)** which includes the least square to the loss function his technique is to "allow a direct fitting of Euclidean distances to possibly transformed dissimilarity " this technique facilitate the estimation of unknown constant by which we obtain a comparative distances. Krustal's contribution was conduct **De Leeuw(1972)** to one of the most robust algorithm for MDS. He defines a raw stress function adding a weight matrix w_{ij} in the numerator to handle the problem of missing data. He obtains a general optimization approach based on majorization labeled scaling by majorizing a convex function. Many multidimensional scaling algorithms have been proposed we figure out the proxscal algorithm which perform MDS of proximities.

2.1.1 Proxscal Algorithm

The main objective of using Proxscal algorithm is to locate a least squares presentation of items in a low dimensionality. The majorization specification allows a monotone convergence[7].
The following expression presents the loss function which could be reduced to get the coordinate matrix.

$$\sigma^2 = \frac{1}{m}\sum_{k=1}^{m}\sum_{i<j}^{n} w_{ijk}\left[\tilde{d}_{ijk} - d_{ij}(X_k)\right]^2 \quad (3.9)$$

Where n and m are the numbers of items and sources respectively, $d_{ij}(X_k)$ the Euclidean distance between the different items and $\tilde{d}_{ijk}$ a nondecreasing values of transformed proximities. Conversely, σ^2 mentions the averaged mean squared error between $d_{ij}(X_k)$ and $\tilde{d}_{ijk}$.
The algorithm is composed by 4 important steps

- Initialize X_k and evaluate σ^2
- Update X_k
- Update $\tilde{d}_{ijk}$
- Evaluate σ^2if the stop criteria are verified, stop; else return to step 2

References

Dmitrijeva, J., & Hazans, M. (2007). A Stock–Flow Matching Approach to Evaluation of Public Training Programme in a High Unemployment Environment. *Labour*, *21*(3), 503-540.

[7] More mathematical details in **Heiser (1993)**.