Région et Développement
n° 52-2020

www.regionetdeveloppement.org

Class Size and Educational Achievement in Tunisia: A Spatial Econometric Approach

Faten DARWEZ*
Khaled JEGUIRIM**
Christopher R. BRYANT***

Abstract - This paper addresses a new conceptual framework related to the study of the class size effect on student achievement. It proposes a methodologically different approach from the studies carried out so far. Based on regional observations on schools in Tunisia, the analysis is based on a spatial exploration of analyzed data (ESDA) and estimates the effect of class size through a spatial Durbin model. The results are in line with studies that defend the reduction of class size to yield better academic performance. However, the reduction in class size has significant effects only with 25 pupils, and little beyond.

JEL Classification
I21, C23, C31

Key words
Education
Class size
Spatial autocorrelation
Spatial Durbin model
Tunisia

* Doctor in Economics at the Faculty of Economics and Management of Sousse, University of Sousse, Tunisia ; faten.darwez@yahoo.fr
** Professor in Economics at the Faculty of Economics and Management of Monastir, University of Monastir, Tunisia ; khaled.jeguirim@u-monastir.tn
*** Adjunct Professor, School of Environmental Design and Rural Development, University of Guelph, Canada, PhD in Economics in the London School of Economics and Political Science, University of London, England ; christopher.robin.bryant@gmail.com

1. INTRODUCTION

Successful schooling and access to university is, for many students and their parents, an objective in itself. Achieving a certain level of education is often associated, in the more or less long term, with higher incomes (compared to the average) and a good ranking at the level of the social scale. Several factors can contribute, or otherwise hinder, this success; in particular, socio-economic factors, demographic factors and parental commitment.

Moreover, the factors related to school (or college or high school) are also important. One of these factors, which is the class size, has long attracted the attention of educators as well as policy makers (and politicians in general). The challenge is to know how to offer maximum knowledge to students and thus enable their success with the least cost to society. The debate then takes place between educators who advocate small classes, and donors (the State for public schools, private investors for private schools) who, for budgetary reasons, are rather favorable towards large classes.

Studies attempting to evaluate the effect of class size on school performance are manifold. Most of them have focused on primary schools in the United States (STAR project; Hanushek, 1999; Hoxby, 2000; Krueger and Whitmore, 2001) and other developed countries (Hanushek, 1986). Although the generally held view is that small classes have not only positive effects on student learning and success but also positive economic effects (Krueger, 2003). In Japan, there does not seem to be a consensus on the impact of reducing class size (Ito et al., 2020). Other studies bring into question these results and reopen discussions about the importance of this factor (Dobbie and Fryer, 2013; Denny and Oppedisano, 2013).

This article is part of this debate by proposing a complementary analysis that addresses this issue in a methodologically different way from the studies carried out so far. Based on regional observations about some schools in Tunisia, this academic work is derived from a spatial exploration of analyzed data (ESDA) and estimates the effect of class size through a spatial Durbin model (Anselin, 1988; LeSage and Pace, 2009). We provide theoretical and empirical justifications for the relevance of these tools in the context of the issue raised. As the data are spatially localized, the analysis exploits the potential proximity effects that can influence school performance in neighboring regions. A theoretical framework is advanced where assumptions about the influence of class size and its relationship to other variables are structured and lead to a particular spatial model. The latter is then estimated from comprehensive data relating to small regions (delegations) of Tunisia.

The next section presents the theoretical model and its spatial ramifications. The analyzed data are then introduced in the third section. The fourth section presents an analysis of the data through a spatial exploration. The empirical results of the estimates made are presented in the fifth section. Finally, the conclusion focuses on the scope of these findings in relation to a class size reduction policy and the value of integrating the spatial effect into the processing of educational data in future analyses.

2. THEORETICAL FORMALIZATION AND SPATIAL ASPECTS

This section presents a theoretical model that links class size to school performance by using proximity effects. The formalization of the relationship between school performance and class size is inspired by the work on the production functions of education (Hanushek, 1986; Krueger, 1999). Integrating the effects of proximity draws from the works of the regional economy and in particular those that model spatial interactions (Anselin, 2003).

2.1. The model

Theoretical analysis focuses on schools as analytical entities. The idea behind the model that will be presented is that the closer these establishments are (geographically) to each other, the more their characteristics are as well. The characteristics refer to the size of the classes, and as well to other relatively unobservable or difficult aspects. We can think of the infrastructure quality, the training (experience, pedagogy, etc.) of teachers, amongst other characteristics. These aspects which are difficult to observe can be neutralized by properly modeling their interactions in space and with the observable characteristic (class size) within institutions. It is then a question of attributing differences in school performance between schools to differences in class sizes, taking into account all the possible interactions with the other unobservable variables.

Formally, it is assumed that the school performance (assimilated here to a success rate) observed at the school level and rated r depends on a set of observable factors and others that cannot be omitted from the analysis. Assume that the only observable factor is the size of the classes we record by c. Unobservable factors are grouped in a variable matrix c^*. The relationship between these variables can be expressed as:

$$r = \alpha c + c^* \alpha^* \tag{1}$$

where α is a parameter that expresses the effect of class size variations and α^* is a parameter vector associated with variables that are unobservable.

Since the observations concern schools, it is then possible, once again, that observable or non-observable characteristics of schools are similar for schools close to each other (schools in priority education zones, in rural or urban areas, in crowded conurbations, in wealthy neighborhoods or in working-class neighborhoods, and so on). Noting by $W = \left[w_{ij}\right]$ an adjacency matrix that quantifies how each observation relative to one institution is related to another observation from another institution j, we can express these spatial dependencies between the n observations in the equations from (2) and (3):

$$c = \varphi W c + u \text{ with } u \sim N(0, \sigma_u^2 I_n) \tag{2}$$

$$c^* = \rho W c^* + v \text{ with } v \sim N(0, \sigma_v^2 I_n) \tag{3}$$

where u and v are random terms and I_n is a matrix identity of order n. These terms are therefore zero means and constant variances, respectively equal to σ_u^2 and σ_v^2.

The parameters φ and ρ describe the importance of spatial dependencies. Wc refers here to the characteristics observed, and more particularly to class size, at the level of neighboring establishments[1]. Similarly, Wc^* describes the unobserved characteristics of neighborhoods. For example, if $\varphi = 0$, then we admit that the size of classes in one institution is independent of that in another geographically close institution. Similarly, if $\rho = 0$ then the characteristics not observed in one institution are not related to those of other institutions in the neighborhood.

The hazards u and v express the fact that, for any establishment, the sizes of the classes on the one hand and the other unobservable characteristics on the other hand do not systematically differ from those of the neighboring establishments.

[1] It is more specifically a weighted average of class sizes in neighboring schools.

However, these uncertainties are not necessarily independent and we postulate the following relation:

$$v = \gamma u + \varepsilon \text{ with } \varepsilon \sim N(0, \sigma_\varepsilon^2 I_n) \quad (4)$$

where ε is a random error term of zero mean, of constant variance and whose terms are independent.

Thus, if the parameter γ is non-zero, the random shocks that can influence class size (government policy, Syndical decision applied in some institutions, and so forth) are linked to those affecting the unobserved variables (for example, teachers' pedagogy, student monitoring, parental involvement, as well as other variables). This condition also implies that class size (considered here as the only observed variable) is correlated with unobserved variables if γ is nonzero[2]. Consequently, its effect is possibly influenced by that of unobserved variables. Rivkin et al. (2005) point out that the effect of class size on school performance is not independent of that of teachers. The latter is difficult to assess because of the difficulty in measuring teachers' skills.

The parameter γ thus plays an important role in the analysis. It indirectly links class sizes to other unobservable variables. Its effect is all the more important if we observe, for example, that a policy of class size reduction applies to certain institutions and not to others (pilot high schools in Tunisia for example) or in certain regions and not in others (rural vs. urban for example). If an institution opts for small classes, it probably offers teaching conditions, materials, and so forth, different from establishments with overcrowded classrooms which can be voluntary or not. It is possible that the most qualified teachers prefer institutions with small classes (or vice versa). Jepsen (2015) reports that a class size reduction program in California has resulted in a significant demand for teachers in that state. As a result, several teachers who were in neighboring and less efficient educational structures moved to institutions offering better working conditions and higher performance, which necessitated, for institutions that have seen teachers leave, and the use of new teachers occurs, probably with less experience and probably with less teaching skills.

It would therefore seem that the hypothesis $\gamma = 0$ is the most credible. But back to our model and starting by successively substituting the expressions in (2), (3) and (4) in equation (1). The exact derivation of this result is found in the Appendix and follows the approach proposed by LeSage and Pace (2009). We conclude that:

$$r = \rho W r + \beta c + \theta W c + \varepsilon \quad (5)$$

with $\beta = \alpha + \gamma$ and $\theta = -\rho\alpha - \varphi\gamma$.

The expression in (5) describes a model that Anselin (1988) described as a spatial Durbin model (SDM). It reflects the fact that, for each school, success depends on the size of the classes within that school, but also implicitly on the characteristics of neighboring schools. In order to better understand the interactions involved, a discussion of the assumptions underlying this model is presented in the next section.

2.2 Spatial Interaction Hypotheses

The SDM model presented in (5) is an unconstrained version of other spatial models and in particular the Spatial Auto-Regressive (SAR) model and the Spatial

[2] The correlation between the hazards u and v implies that between c and c^*.

Error Model (SEM) (Anselin, 1988). A brief presentation of these models in the context of our problematic makes it possible to better understand the hypotheses formulated in this work.

In the SAR model, it is postulated that spatial dependencies are manifested only at the level of the endogenous variable (here, school success). Otherwise, academic performance (or achievement) in one institution is directly related to that in another nearby institution.

In this case, it is formally written as follows:

$$r = \rho Wr + \beta c + \varepsilon \tag{6}$$

The SAR model is thus a special case of the SDM model given by the expression in (5) when $\theta = 0$. This hypothesis can be technically tested by testing the equality $\theta = 0$.

This model allows us to consider a hypothesis that we have formulated: school performance in neighboring institutions is closer than those in distant schools. Such a hypothesis is a special case of the so-called first law of geography that all places are related to each other but that nearby places are more related. As stated by Tobler (1970) “everything is related to everything else, but near things are more related than distant things”.

However, the wording is naive in this case. It seems unwise to admit that school performance in different (even very close) schools is linked unless one assumes that there are factors common to both schools. These factors, whether they are observed or not, imply that (when $\theta = 0$), the random error terms included in ε can no longer preserve the independence structure assigned to them up to this point.

The SEM model integrate this alternative by considering that the (spatial) dependence between establishments is modeled through the error term, let:

$$r = \beta c + \varepsilon \text{ with } \varepsilon = \lambda W\varepsilon + \eta \text{ and } \eta \sim N(0, \sigma_\eta^2 I_n) \tag{7}$$

where λ is a parameter that describes spillover effects: school performance in nearby institutions is interrelated and mutually interacting because of the presence of random (unobservable) effects that interact in space themselves.

Such an interpretation is however more restrictive than that envisaged in the case of the SDM model. This is not surprising since the SEM model is, in reality, a special case of the spatial Durbin model. Indeed, the expression in (7) can also be written as:

$$r = \beta c + (I_n - \lambda W)^{-1}\eta \text{ since } \varepsilon = (I_n - \lambda W)^{-1}\eta$$

$$(I_n - \lambda W)r = (I_n - \lambda W)\beta c + \eta$$

$$r = \lambda Wr + \beta c - \lambda\beta Wc + \eta$$

$$r = \rho Wr + \beta c - \rho\beta Wc + \eta \text{ and } \eta \sim N(0, \sigma_\eta^2 I_n) \tag{8}$$

and the SEM model is a special case of the SDM model when the common factor restriction $\theta = -\rho\beta$ is imposed (Anselin, 1988). However, if this restriction is justified, it necessarily imposes $\gamma = 0$ unless we admit that $\rho = \varphi$ that is to say that the explanatory variables observed behave in the same way as the unobserved variables. The SEM model can therefore only admit that $\gamma \neq 0$ by assuming that if the sizes of the classes are identical between neighboring establishments, then they also have the same characteristics at the level of teachers for example. However, there is no justification for the fact that establishments offer the same class sizes as they also offer the same or other services, or that they generally have the same characteristics, especially since they are by their creation unobservable.

As we observed before, the condition $\gamma \neq 0$ indicates that the size of the classes is correlated with the explanatory variables not included in the model, admitting the opposite is not credible enough. As a result, only the SDM model makes it possible to avoid imposing undue restrictions (Elhorst, 2010).

These developments, in particular the omission of a variable explaining school performance that is correlated with class size, lead to the need to include the spatially delayed dependent variable (success) (the term Wr) and the explanatory variable (class size) also spatially delayed (Wc). The effect of the variable c on the success r cannot, therefore, be apprehended simply through the parameter β (see expression (5))[3], since there exists (probably) an indirect effect evaluated as a function of θ.

Quantifying the effect of class sizes on school performance must therefore take into account two effects: a direct effect and an indirect effect. Obviously, it is possible that the direct effect outweighs the indirect one, which may be negligible. The question is not this. The problem is that one cannot evaluate the direct effect by neglecting the presence of a possible indirect effect. The sum of these two effects gives the total effect of a variation in class size on school performance. The presented model can be used to test if only the class size variable influences school performance by testing the restrictions $\rho = 0$ and $\theta = 0$, in the expression in (5). In this case, it is a simple regression where the application of the OLS method is possible. Technically, it is possible to test models from the most restrictive to the most general. However, it seems inappropriate to test whether school performance is related solely to class size, as the possible influence of other factors is evident. We therefore retain the SDM model (the most general) as the model of choice, although statistical test results are conducted to judge whether more restrictive specifications are more appropriate.

The estimation of the model in question requires (at least) the observation of the success rate and the number of pupils per class for different schools. This statistical information surely exists but is often difficult to access. The empirical analysis, which will follow, will be based on an aggregation of data. The idea is that, instead of observing each school, there is a group of schools and the average success rate and the average class size within each group are calculated. The choice of these groups is made here exogenously since it will correspond to clearly defined administrative areas. This makes it possible, in particular, to overcome the endogeneity bias that could arise in this type of analysis (Wöessmann and West, 2006). The data and a better justification for this choice are now presented.

3. THE DATA

The following empirical analysis is based on data related to Tunisia's 264 delegations. These are administrative entities that have a geographical or demographic coherence.

The data collected relate to the pass rates for the Baccalaureate (Bac) test in 2013 and the average number of pupils per class in the schools that provide student training during the seven years before the baccalaureate exam (preparatory

[3] And even less, from the parameter α in the expression (1).

schools and high schools). The data analyzed here relate only to public schools. They are collected from data provided by the Regional Directorates of Education and Training attached to the Ministry of Education.

This choice is obviously imposed by the available data which do not allow (most often) to distinguish, at the level of all the delegations, the number of classes in the high school only and even less the number of classes at the year level of the Baccalaureate. Still, we have chosen to consider that the success of the baccalaureate is the result of a school performance that is acquired over several years prior to this national examination and that, therefore, the size of the Baccalaureate class is not the most important, but it is those that have marked the learning process of the students in the years prior to the Baccalaureate. In other words, the average of class sizes during these years.

The success rate in relation to the baccalaureate examination is considered as a (frustrated) indicator of school performance. The observation of this variable, as well as the average number of pupils per class at the level of each delegation, is an aggregation which may seem, at first glance, detrimental to the analysis. However, we considered that the variations in pass rates and class sizes among schools in each delegation (in case of many) are less important than those between institutions from different delegations. For the success rate, this is a hypothesis that remains to be verified, but does not seem to be too strong. Indeed, the average number of teaching structures (preparatory schools and high schools) in each delegation is of the order of 5 and three-quarters of all delegations have less than 7 institutions. Dispersions in the Baccalaureate pass rates and class sizes cannot be too great between so few closely related institutions belonging to the same delegation. On the other hand, these differences are much more striking between the delegations. In particular, and as shown by the values reported in Table 1, the average class size varies from less than 15 pupils to 30 pupils per class according to the delegations.

Table 1. Statistical Summaries (depending on delegations)

	Min.	1st Qu.	Med.	Avg.	3rd Qu	Max.
Number of schools	1	3	5	5.4	7	17
Number of students per class	14.2	22.7	24.4	24.2	26.4	30.4
Success rate	0	47.3	56.2	56.1	65.3	85.9

Source: Authors.

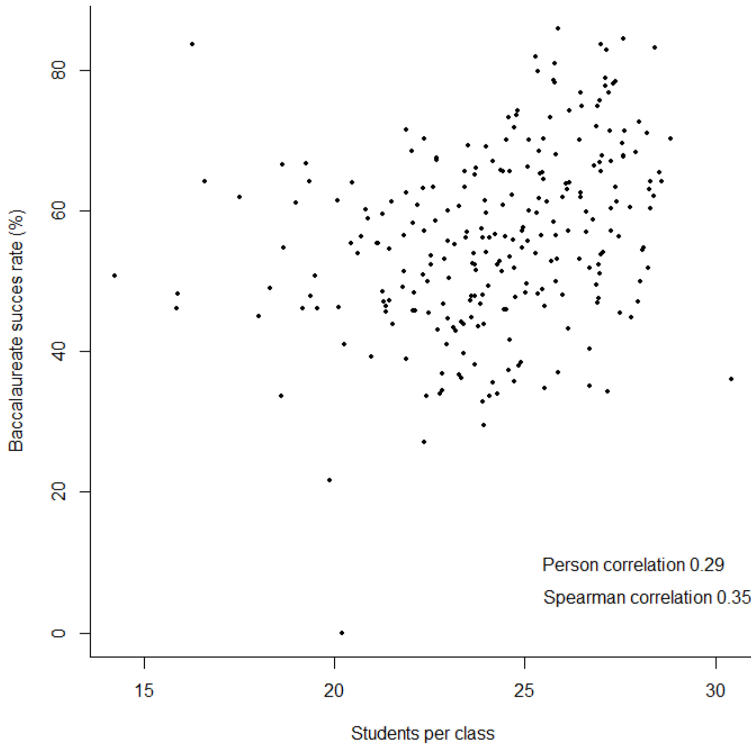
Figure 1 represents, for each delegation, the pass rate for the baccalaureate based on the average number of students per class.[4] The resulting point cloud is far from suggesting the expected effect. At best, the configuration of the cloud is completely random, at worst, it would suggest that the higher the size of a class, the higher the success rate in the Baccalaureate will be. The correlation coefficients (Pearson and Spearman), albeit weak, are positive and significant at the 1% level.

Is this a general configuration (i.e. valid for all the delegations)? Can we imagine that the relationship is not as simple (or maybe it is finally)? Is there any information concealed by the (seemingly random) scattering of points about delegations? A more detailed analysis should provide us some clarification. In particular, given the spatial nature of the observations, a space exploration – precisely – of the

[4] Due to the fact that some observations come down to two different pairs of delegations, the number of observations in this study is 260 and not 264 (which is the number of delegations).

analyzed data (ESDA) is required. The term ESDA is commonly associated with the expression Exploratory Spatial Data Analysis (Anselin, 1988). This is detailed in the following section.

Figure 1. Baccalaureate success rate by average number of students per class



Source: Authors.

4. EXPLORATORY SPATIAL ANALYSIS OF THE DATA

4.1. Preliminary Findings

In general, spatial exploratory analysis of the data allows detecting a possible spatial dependence of the observations. Several tools can be mobilized for this purpose (Anselin et al., 2004). In fact, we will apply some of them in order to illustrate the characteristics that seem to us to be the most interesting in our analysis and also the more specific to the variables treated here.

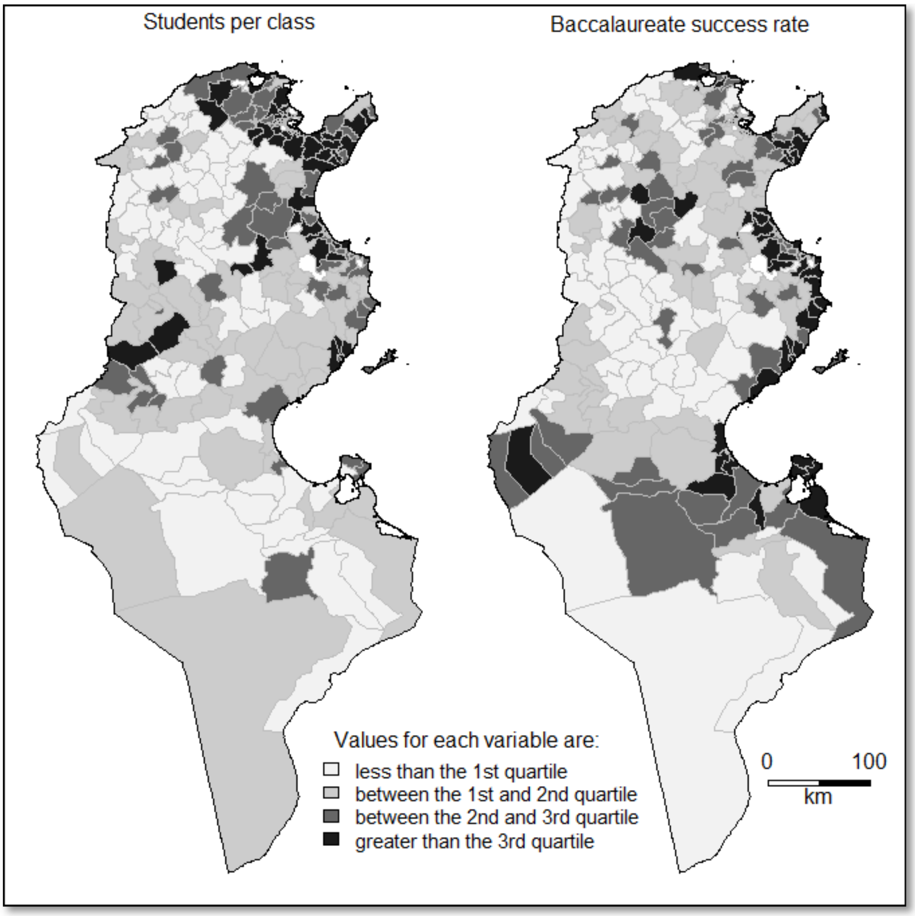
A first step is to chart the data. The maps of Tunisia presented in Figure 2 give a reasonable idea about the two key variables of our analysis, by quartiles distribution, and according to the delegations.

Beyond the usual divisions of the North-South, East-West, coastal-interior regions, which do not seem to be completely respected, two observations are more obvious.

First, the spatial distribution (by delegation) of the two considered variables does not seem to be random. For proof, it is sufficient to visually compare these distributions with random distributions illustrated in Figure A.1 (in the Annex).

Some autocorrelation of geographically close values seems to be emerging. To verify this, a measure of this phenomenon is presented (in the next paragraph) based on a Moran spatial autocorrelation coefficient (Cliff and Ord, 1973).

Figure 2. Baccalaureate success rate and average number of students per class per delegation



Source : Authors.

Second, it can be seen that there are local concentrations in the distribution of the two variables. Some regions (probably) have more pronounced similarities than others. For example, low (or high) class sizes characterize close delegations. It seems that the delegations around Tunis (but also at the level of the Governorate of Sfax) are characterized by high values. A certain number of delegations from the East Coast Governorates (Sfax, Monastir, Mahdia, but also those from the island of Djerba) seem to indicate high rates of growth. It follows that there are probably "clusters" or clusters of delegations locally characterized by similar values. These types of local associations will also be quantified from an LISA indicator (for Local Association of Spatial Autocorrelation, Anselin, 1995).

In order to carry out these investigations, we first have to describe how the spatial proximities between the observations are apprehended. This is done through a contiguity matrix (Anselin, 1988). We retain here a matrix of standardized binary contiguity by line $W = \left[\frac{w_{ij}}{\sum_j w_{ij}}\right]$, where $w_{ij} = 1$ if two delegations have a common border and 0 otherwise. It is thus a square matrix $n \times n$ where n is the number of observations. Other types of matrices can obviously be envisaged. However, our results have proven to be robust to alternative specifications of this matrix.

Note also that because of the standardization of this matrix, the multiplication, by W, of a variable y including n observations, gives a vector of observations Wy whose elements are of the form $Wy_i = \sum_j y_j \left(\frac{w_{ij}}{\sum_j w_{ij}}\right)$. The variable Wy described as an offset spatial variable therefore contains the weighted average of the observations of the neighboring regions at a region i. These are the types of variables that appear in the previous equations.

4.2 Results of the spatial autocorrelation

One of the most used measures to quantify the global autocorrelation of spatial data is Moran's I statistic (Cliff and Ord, 1973) given by:

$$I = \frac{n}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}} \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{9}$$

where y_i indicates the observed value of a variable y for the delegation i, y_j the observation for an (other) neighboring delegation and $\bar{y}$ is the average for all n delegations.

The statistic I given in (9) is interpreted as a « traditional » autocorrelation coefficient, with the difference that the absence of spatial autocorrelation implies that the value of I is equal to 1 its expected average $E(I) = -\frac{1}{n-1}$ (equal here to -0.0039) and not to 0. A value of I greater than this value reflects a positive spatial autocorrelation, while a lower value (negative) indicates a negative spatial autocorrelation.

Table 2. Moran global spatial I autocorrelation

	Moran I	standard deviation	P-value[(a)]
Number of pupils per class	0.5457	13.78	0.000
Bac success rate	0.4448	11.25	0.000

[(a)]*The inference is based on the asymptotically normal distribution of I. The alternative hypothesis is the presence of positive autocorrelation of success in Baccalaureate rather high (compared to others). Source : Authors.*

Table 2 gives the Moran I calculated for the two variables: number of students per class and success rate in the Baccalaureate. Positive and largely significant values indicate the presence of positive spatial autocorrelation in relation to these two variables. The assumption of a random distribution of observations among the delegations is to be rejected. Delegations with a high average number of pupils per class (above average) are surrounded by others with a high number of pupils per class. Similarly for the variable success rate in the Baccalaureate. The delegations with high (or alternatively low) rates are surrounded by delegations with high (or alternatively low) rates.

Moran's statistic, in its totality, does not make it possible to distinguish the types of autocorrelation. It is possible that with positive spatial autocorrelations, the spatial configurations are different. For example, it is possible that only a few delegations are at the origin of this autocorrelation while most other delegations are "normal" (have values close to the average). Similarly, the importance of neighboring groups of delegations with high values cannot be distinguished from neighboring delegations with low values. The LISA indicator is a local spatial autocorrelation statistic based on Moran's *I* (Anselin, 1995). It is a question of distinguishing, for a given variable, a delegation according to whether it is:

- characterized by a high value of this variable and surrounded by delegations with high values. This type of association is commonly noted HH (for Hight-Hight, see Anselin, 1995);
- characterized by a low value and surrounded by delegations with low values: LL;
- characterized by a high value and surrounded by delegations with low values: HL;
- characterized by a low value and surrounded by delegations with high values: LH.

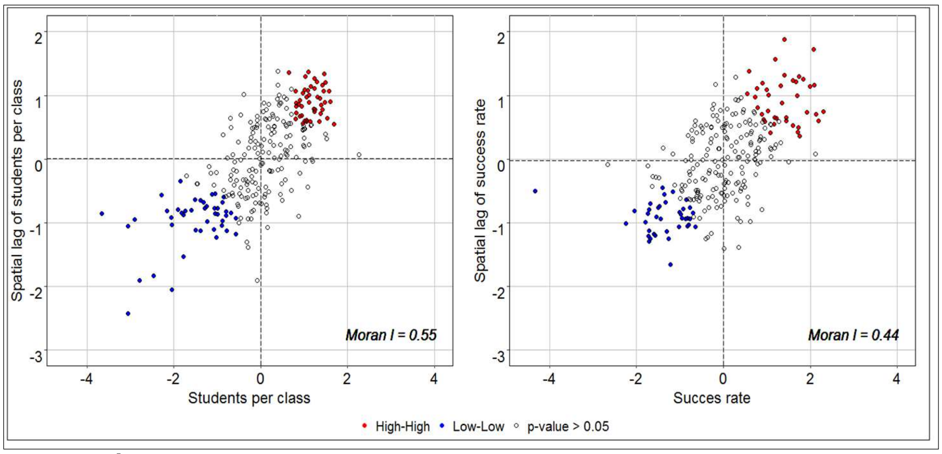
These different types of correlations must be statistically significant; otherwise there is no local autocorrelation. Table 3 shows the local spatial associations between each delegation and its neighbors for the two variables that are the subject of the study. Moran scatterplots (divided into four quadrants according to the types of spatial associations: HH, LL, HL and LH) are also shown in Figure 3 for these variables.

Table 3. Local spatial associations

Types of Associations	Number of pupils per class		Bac success rate	
HH: High-High	47	51.1%	36	53.7%
LL: Low-Low	45	48.9%	31	46.3%
HL: High-Low	0	0%	0	0%
LH: Low-High	0	0%	0	0%
Not significant	168	-	193	-

Source: Authors.

Figure 3. Moran local scatterplots

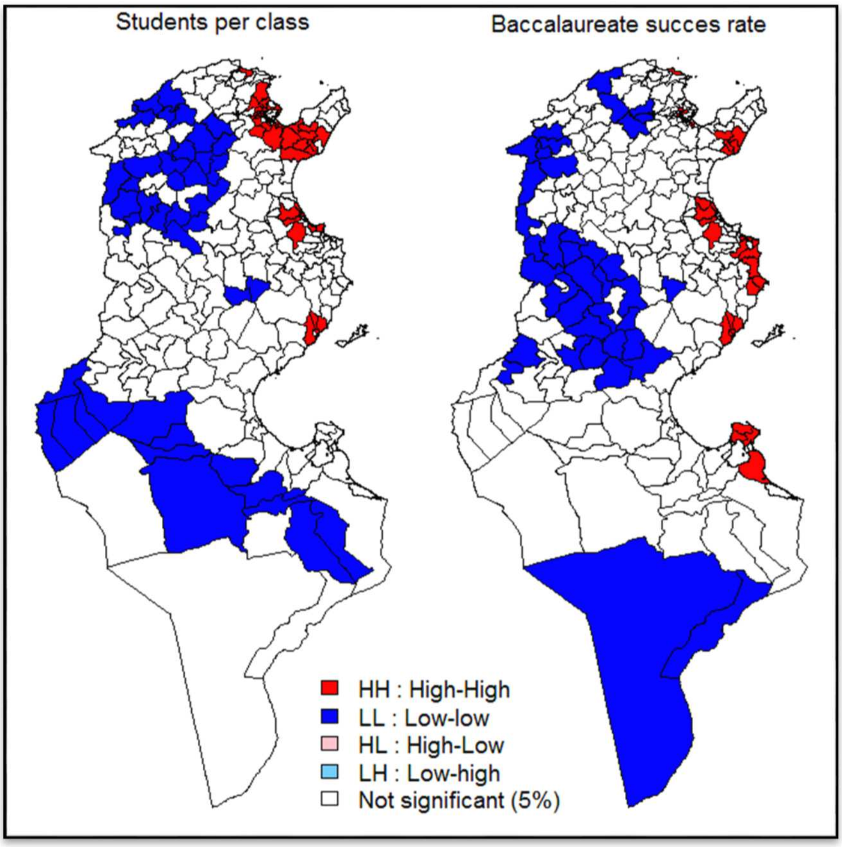


Source: Authors.

All significant correlations are positive for both variables. It can be seen (Figure 3) that, with respect to the variable students per class, significant positive spatial association characterizes all Tunisian delegations: 51.1% in quadrant HH, and 48.9% in quadrant LL. The same applies to the variable success rate: 53.7% in quadrant HH and 46.3% in quadrant LL. In addition, significant spatial correlations are more numerous for the variable number of students per class compared to the variable success rate. However, the importance of HH or LL correlations is almost similar for both variables.

The results presented in Table 3 and Figure 3 do not make it possible to know if the delegations characterized by associations HH (or LL) for the number of pupils by classes are identical or not to those bearing the same characteristics for the variable success in Baccalaureate. To ensure this, a mapping of these associations is presented in the LISA maps shown in Figure 4.

Figure 4. Local spatial autocorrelations: LISA maps



Source: Authors.

The local associations of the low-low type (LL) are mainly located at the level of the delegations of the North-West and also South regions, for the student variable by class, as well as in the Center-West for the Baccalaureate pass rate variable. The high-high associations (HH) are mainly located in some coastal governorates. In particular, the classes are "overloaded" in areas with a high population density: "Greater Tunis" (including the neighboring delegations of the South of Cap-Bon), delegations of the Governorates of Sousse and Sfax.

These delegations (with the exception of those of "Greater Tunis") also have High-High associations in terms of pass rates (by comparing the two maps). Nevertheless, other delegations present the same type of HH association at the level of the success rate, although being characterized by HH associations at the level of the number of pupils per class. The same finding can also be observed for the associations of LL type. On the other hand, no delegation is characterized by a high success rate and a low number of pupils per class (or vice versa). Table A in the Annex details the number of delegations by crossing the types of associations.

In Tunisia, regional disparities are often related to disparities in urbanization as well as in economic performance between geographical areas. Broadly speaking, there is a dual structure in these variables between coastal and inland areas. The effect of different variables explaining disparities between delegations on school performance is certainly a crucial question. However, the objective of this study is not to explain theses effects but rather to consider them in the analysis. We concentrate then our analysis on the correlation between class size and school performance.

The positive correlation between the two variables, which was observed previously, seems to be the consequence of a positive local spatial correlations presence (HH or LL) in well-defined regions. Is such an argument enough to conclude that an overcrowded class has a positive effect on student success? Nothing is less certain. But what is certain is that these spatial autocorrelations must be taken into account in the econometric regressions that will be applied now.

5. SPATIAL REGRESSIONS: SPECIFICATIONS AND RESULTS

5.1. Model specifications, estimates and comparisons

The theoretical model presented above considers the size of classes as the central variable of the analysis, while taking into account the influence of other unobservable variables. It is undeniable that such a formulation is rather restrictive. More explicit factors would certainly explain the differences between delegations in terms of student success rates.

We introduce some of these aspects into the estimated model as control variables. In particular, they take into account the specific economic, social and / or demographic characteristics that differentiate Tunisian delegations and that may also have an impact on academic success. Many of these aspects, such as the lack of access to schools because of economic or even geographical considerations for some students can explain, at least in part, their failure at school. Actually, some of these aspects, such as parents' education level and poverty, have been criticized as the possible causes of abundant schooling, particularly at the level of colleges and high schools in Tunisia (Boughzou, 2016). It is quite possible then that they also have a negative effect on the possibility of a successful bachelor's degree.

Taking all the possible factors into account goes beyond the scope of this analysis, especially since regional data in Tunisia (here at the delegation level) are still very sparse. As a result, we have limited ourselves to the following variables in order to take into account the aspects mentioned:[5]

- The unemployment rate as proxy for the economic situation within a delegation: it has been considered that a delegation with a "bad" economic situation would have a high unemployment rate and a significant drop in schooling among its school population. The expected impact of this variable would therefore be negative.

- The education level of the population, measured by the percentage of individuals with higher education: This variable can tell us how much the parents are

[5] The relative data for these variables are taken from the 2014 General Population and Habitat Census.

involved in their children's education. The higher it is in the population of a delegation, the more such a delegation would be characterized by a high pass rate.

- The percentage of the population of a delegation living in urban areas as proxy for the level of urbanization of a region: It seems that in rural areas, access to secondary education is more difficult than in an urban setting. As a result, it is possible that delegations with (relatively) large populations living in rural areas have lower success rates in the Baccalaureate in comparison with « urban » delegations.

With these included variables, the results of estimates of the SDM model, but also other specifications with spatial effects (SAR and SEM models) or without these effects (MCO), are presented in Table 4. The spatial model estimates are made using the maximum likelihood method and implemented in the R software (R Core Team, 2016) with the spdep package (Bivand and Piras, 2015 ; Bivand et al., 2013). Alternative models estimation (other than SDM) aims to compare our data to more restrictive specifications and thus statistically test the assumptions made previously. Other statistics were also calculated to discriminate between the estimated models.

Table 4. Results of Estimates

	MCO	*SAR*	*SEM*	*SDM*
Constant	73.533***	58.693***	66.348***	65.001***
	(7.434)	(7.175)	(6.967)	(8.707)
Students by Class	−0.553*	−0.550**	−0.360	−0.355
	(0.281)	(0.243)	(0.269)	(0.270)
Higher education	0.417***	0.378***	0.571***	0.769***
	(0.112)	(0.116)	(0.133)	(0.149)
Urban	0.114***	0.092***	0.092***	0.081***
	(0.024)	(0.021)	(0.023)	(0.025)
Unemployment	−0.923***	−0.776***	−0.788***	−0.650***
	(0.126)	(0.099)	(0.108)	(0.126)
Spatially lagged variables				
Students by Class				−0.508*
				(0.302)
Higher education				−0.753***
				(0.221)
Urban				0.015
				(0.039)
Unemployment				−0.237
				(0.183)
ρ (spatially lagged endogenous variable)		0.253 ***		0.367***
		(0.000)		(0.072)
λ (spatial errors)			0.412***	
R^2	0.486	0.515	0.528	0.555
Akaike Information Criterion (AIC)	1907.311	1894.356	1886.816	1879.973
Log-likelihood	−947.655	−940.178	−936.408	−928.986
Wald test		15.23***	29.77***	26.11***
Likelihood ratio test (LR test)		14.955***	22.494***	22.949***
Lagrange multiplier test (LM test)	23.392 ***	4.488**		2.023
Moran's I of residual	0.194 ***	0.056	−0.016	−0.025

Note: The values which are between parentheses below the coefficients are the estimated standard deviations (robust to heteroscedasticity for the MCO model).
** Significant at 10%, ** Significant at 5% and *** Significant at 1%.*
For spatial models, R^2 denotes the Nagelkerke pseudo R^2 (see Nagelkerke, 1991).
Source: Authors.

For all the specifications, overall variables have the expected signs. However, the comparisons of estimated effects of these variables on the success rate cannot be made from one model to another because of the different specifications. In particular, the effect of the class size variable on the success rate cannot be directly observed from the results of the SDM model presented in this table, since the direct and indirect effects must be taken into account (Elhorst, 2010; LeSage and Pace, 2011); this will be discussed later.

It should be noted that the need to take spatial effects into account is – firstly – confirmed by the Moran's I test for residual spatial autocorrelation. The a-spatial specification (MCO) has spatially correlated residuals. The null hypothesis of no spatial autocorrelation at the residues of this model is to be rejected at a 1% risk of error with a Moran I of residues equal to 0.194. The Lagrange multiplier test confirms the presence of spatial autocorrelation in the residuals of the MCO model. On the other hand, Moran'I being insignificant, there is no spatial autocorrelation for the residues resulting from the other specifications. The Wald and likelihood ratio (LR) tests also confirm the need to consider spatial interactions at the data level.

Traditionally, the choice of the appropriate spatial model can go through an approach from a particular to the general type. This approach, recommended by Anselin (1988), consists of applying a battery of tests (especially Lagrange multiplier types) and then choosing the appropriate spatial specification (see Anselin et al., 1996 for details). Elhorst (2010) inspired by LeSage and Pace (2009), recommends rather moving from the general to the particular, that is to say – in this case – from the SDM model to the SAR and SEM models. However, in both cases, these choices are conditioned by the absence of a prior spatial specification, which is not the case here. The SDM model presented in this study is the result of a theoretical approach built around the link between class size and academic success. However, it is always useful to ensure that statistical data support this choice.

The values of R^2, log-likelihood and the AIC selection criterion show that the SDM model is the one that best fits the data. We have also tested the hypothesis $\theta = 0$ (see equation (5)) in order to discriminate between the SDM model and the SAR model. The statistic of the likelihood ratio test performed is equal to 22.38 and follows a law of χ^2 with 4 degrees of freedom.

The hypothesis is rejected with a negligible p-value. The test of the common factor restriction assumption $\theta = -\rho\beta$ is also tested to discriminate between SEM and SDM models. The likelihood ratio test statistic is equal to 14.84 and a p-value equal to 0.005. In both cases, the best specification is the spatial Durbin model.

To complete this first diagnosis, it is interesting to note that the spatial dependence coefficient $\rho = 0.367$ is positive and significant. Success rates at the delegation level are not independent. In fact, this result therefore supports the exploratory analysis (ESDA) carried out previously.

5.2. Estimates from the spatial Durbin model

The effect of a variation in class size on the success rate combines direct and indirect effects. Formally, this effect is deduced from (5) by $\frac{\partial r}{\partial c}$. Equation (5) can also be written:

$$r = (I_n - \rho W)^{-1}\beta c + (I_n - \rho W)^{-1}\theta W c + (I_n - \rho W)^{-1}\varepsilon$$

Consequently, the effect of a variation of c (class size) on r (success rate) is:

$$\partial r / \partial c = (I_n - \rho W)^{-1}(\beta + \theta W)$$

This results in a matrix of $n \times n$ dimension. The direct effect is the mean of n terms on the diagonal $\left(1/n \sum_{i=1}^{n} \frac{\partial r_i}{\partial c_i}\right)$ while the indirect effect is the average of n

terms on the rows (or columns) outside those of the main diagonal $\left(\frac{1}{n}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\frac{\partial r_j}{\partial c_i}\right)$. The sum of these two terms gives the total effect (see LeSage and Pace, 2009). The sum of these two effects leads to the total effect that we estimated and presented in Table 5 for all exogenous variables. Other alternative specifications are also presented. They are applied to data concerning either coastal or interior regions or even delegations where the average class size is less than 24 pupils or greater than this value. The maps in Figure A.2, in the Annex, specify the geographical positions of these delegations.

For all observations, class size would have a negative and significant effect on the success rate. On average, redesigning a student's class size would increase the success rate by more than 1.3 percentage points, all other things being equal. The interpretation of this effect in the context of cross-sectional data must be done with great care, as is always the case for this type of data. It is therefore certainly only an order of magnitude that carries the traditional disadvantages of an average observed on several individuals (here delegations).

To mitigate this effect, we replicated the estimates by referring to delegations belonging to geographical entities that are more "homogenous" in terms of their level of economic and social development. The regional disparities in Tunisia are well known (World Bank, 2014). The so-called "interior" regions (North-West, Central-West and South) are generally less prosperous with levels of development below those of the coastal regions (mainly Greater-Tunis and Center-East). Estimates made for these regions separately show that the effect of class size remains negative but is no longer significant. this is certainly due to the little differentiated values in these regions, and where other explanatory factors appear?

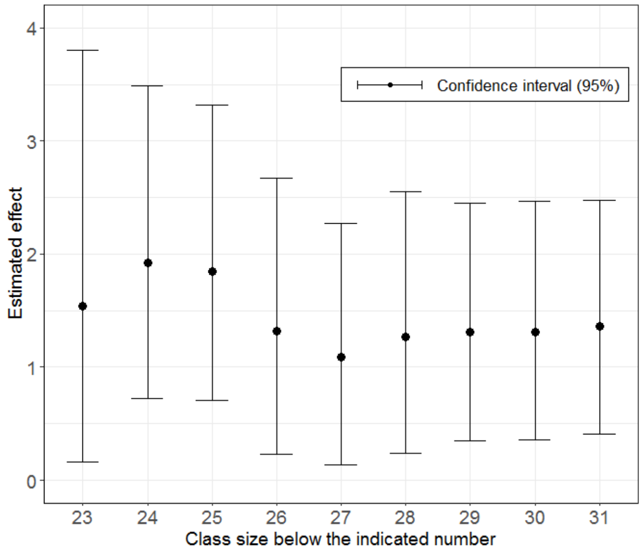
Table 5. Estimates of the total effect of the explanatory variables on the success rate (spatial Durbin model)

	Total effect				
	All regions	Coastal regions	Interior regions	Classes with size inferior to 24	Classes with size superior to 24
Students by classes	-1.363^{***}	-0.572	-1.147	-1.921^{***}	-1.114
	(0.522)	(0.910)	(0.711)	(0.696)	(0.866)
Higher education	0.025	0.073	0.589	1.400	-0.074
	(0.261)	(0.278)	(1.256)	(0.957)	(0.231)
Urban	0.153^{***}	0.093	0.180^{*}	0.040	0.150^{***}
	(0.050)	(0.064)	(0.100)	(0.092)	(0.052)
Unemployment	-1.402^{***}	-1.174^{***}	-1.307^{***}	-0.944^{***}	-1.219^{***}
	(0.218)	(0.360)	(0.361)	(0.318)	(0.207)
Number of obs.	260	137	123	116	144
R^2	0.555	0.466	0.493	0.359	0.599

Note: The values in parentheses are the standard deviations estimated on the basis of 10.000 simulations of the underlying SDM model parameters' estimations.
** Significant at 10%, ** Significant at 5% and *** Significant at 1%.*
R^2 denotes the Nagelkerke pseudo R^2.
Source : Authors.

To exploit the heterogeneity of the data, however, we separated the observations by class size. Two groups of observations are formed: those with a class size below the median (which coincides with 24 pupils per class) and the others. The estimates made revealed that for the first group, the effect of class size on success is even greater than for all observations. It is not significant for delegations where class sizes exceed 24 pupils (see last two columns in Table 5).

Figure 5. The estimated effect of class size reduction on success rate based on the observed number of students per class



Source: Authors.

Figure 5 shows that the effect of reducing the size of a class decreases significantly from 24 (or 25) students. For example, considering only those delegations where class size is less than (or equal to) 24, the effect of reducing the size of a class would be 1.9 percentage points on the pass rate. With classes smaller than 27 students, the effect would be only 1.1 percentage points. This number of 24 (to 25) students seems to be a threshold since when only delegations with classes of less than 23 pupils are considered, the marginal effect is also less important and is only 1.5 percentage points.

Finally, note that beyond class size, the importance of economic considerations (approximated here by the unemployment rate) is crucial. In the regressions carried out, the more the economic situation is unfavorable (high unemployment rate), the lower the success rate. This variable somehow reflects the effect of the level of household income in delegations and their spending on education (see Catin and Hazem, 2012). It also seems that the parents' level of knowledge is neither a stimulant nor a handicap for the baccalaureate success. The effect of the "higher education" variable is insignificant when the direct and indirect spatial interactions from the SDM model are taken into account. Finally, the success rate in the Baccalaureate is more important when one is in the most urbanized areas, which have greater facilities and whose environment is more favorable for studying.

6. CONCLUSION

Reducing the size of a class to improve students' academic performance is an approach that has been experimented with, debated and analyzed for several decades. This policy seems attractive, simple to apply and enjoys the support of teachers. It is, however, quite expensive which raises the problem of its effectiveness and therefore of the interest of its application especially as its effect on the success of the pupils does not seem obvious.

This study aimed to quantify precisely this effect from data relating to Tunisia.

Our results point to a positive and significant effect of class size reduction on student success. The results here are limited to the case of post-primary education (second cycle of basic and secondary education) and are therefore to be considered complementary to other results whose conclusions are not always unanimous as to have a positive and significant effect of class size on student performance.

Beyond the quantitative and statistically significant impact estimated in this study, some remarks emerge from our approach.

First, there is a threshold effect inherent in the effect of class size. It seems that this policy is not effective, or at least very weak, when the classes are "too overloaded". The impact of reducing the size of a class is relative to the number of students in a class. From this point of view, a more in-depth analysis could be carried out in order to estimate more appropriately the expected effect of such an educational policy in relation to the size of classes analyzed.

Secondly, the tools of statistics and spatial econometrics implemented during the analysis and the estimation of the results are essential for a better study of this policy. The spatial interdependence of school characteristics is an important factor when the study is conducted on several institutions belonging to the same country. In this work, we provide theoretical and empirical justifications that go in this direction. It seems to us that a future analysis of this educational policy, in a given school system, cannot be dissociated from the tools of statistics and spatial econometrics.

Thirdly, the potential benefits of reducing the size of a class are also relative to the costs of such a policy (see Yeh, 2010). Reducing the size of a class often requires more teachers, more premises, more materials, and so on. In particular, recruiting more teachers with certain qualification requirements requires expenditure on both backward training and forward remuneration. The question arises whether it would be better to have overcrowded classes with highly qualified teachers or rather small classes but with low-qualified teachers. Wöessmann and West (2006) look at the first alternative by comparing educational systems in different countries and conclude that teacher qualifications, rather than class size, are the most influential in student performance.

Finally, it should be noted that the factors, other than those measuring the size of the classes, were only implicitly analyzed in this study. However, integrating the effect of certain factors, including teacher qualifications, would make our results more robust. Indeed, and on the basis of complementary data more aggregated than those presented in this study, it was found that large size classes are most often associated with teachers with higher qualifications than average (see Figure A.3 in the Annex). This would probably explain the insignificant effect of class size on educational outcomes for this type of overcrowded class observed in our empirical estimates, due to the presence of fairly qualified teachers who compensate for class size effect.

This particularity of the educational system in Tunisia deserves to be more deeply analyzed. Spatial disparities in teacher qualifications at the school level are another factor to be explicitly taken into account in future analyses and spatial statistics tools should once again be solicited.

REFERENCES

Anselin L. 1988, *Spatial Econometrics: Methods and Models*, Kluwer Academic, Dordrecht.

Anselin L. 1995, Local Indicators of Spatial Association – LISA, *Geographical Analysis*, 27, 2, 93–115.

Anselin L. 2003, Spatial Externalities, Spatial Multipliers, and Spatial Econometrics, *International Regional Science Review*, 26, 2, 153–166.

Anselin L., Bera K., Florax R., Yoon M.J., 1996, Simple diagnostic tests for spatial dependence, *Regional Science and Urban Economics*, 26, 1, 77–104.

Anselin L., Florax R., Rey S., 2004, *Advances in Spatial Econometrics: Methodology, Tools and Applications*, Springer, Berlin.

Bivand R., Hauke J., Kossowski T., 2013, Computing the Jacobian in gaussian spatial autoregressive models: An illustrated comparison of available methods, *Geographical Analysis*, 45, 2, 150–179.

Bivand R., Piras G., 2015, Comparing implementations of estimation methods for spatial econometrics, *Journal of Statistical Software*, 63, 18, 1–36.

Boughzou K., 2016, L'abandon scolaire en Tunisie : Etat des lieux, caractéristiques et perspectives, *L'Education en débats : analyse comparée*, 7, 47–25.

Catin M., Hazem M., 2012, Les disparités de taux d'alphabétisation selon les genres dans les délégations tunisiennes : une approche par l'économétrie spatiale, *Région et Développement*, 35, 177-193.

Cliff A.D., Ord J.K., 1973, *Spatial autocorrelation*, Pion, London.

Dee T., West M., 2011, The Non-Cognitive Returns to Class Size, *Educational Evaluation and Policy Analysis*, 33, 1, 23-46.

Denny K., Oppedisano V., 2013, The surprising effect of larger class sizes: Evidence using two identification strategies, *Labour Economics*, 23, 57–65.

Dobbie W., Fryer J., Roland G. 2013, Getting beneath the Veil of Effective Schools: Evidence from New York City, *American Economic Journal: Applied Economics*, 5, 4, 28–60.

Elhorst J.P., 2010, Applied spatial econometrics: Raising the bar, *Spatial Economic Analysis*, 5, 1, 9–28.

Hanushek E.A., 1986, The Economics of Schooling: Production and Efficiency in Public Schools, *Journal of Economic Literature*, 24, 3, 1141–1177.

Hanushek E.A., 1999, Some Finding from an Independent Investigation of the Tennessee STAR Experiment and from Other Investigations of Class Size Effects, *Educational Evaluation and Policy Analysis*, 21, 2, 143–163.

Hoxby C.M., 2000, The effects of class size on student achievement: New evidence from population variation, *The Quarterly Journal of Economics*, 115, 4, 1239–1285.

Ito H., Nakamuro M., Yamaguchi S., 2020, Effects of class-size reduction on cognitive and non-cognitive skills, *Japan and the World Economy*, 53.

Jepsen C., 2015, Class size: Does it matter for student achievement?, IZA World of Labor, p.190.

Krueger A.B., 1999, Experimental Estimates of Education Production Functions, *The Quarterly Journal of Economics*, 114, 2, 497.

Krueger A.B., 2003, Economic considerations and class size, *The Economic Journal*, 113, 485, 34-63.

Krueger A.B., Whitmore D.M., 2001, The Effect of Attending a Small Class in the Early Grades on College-Test and Middle School Test Results: Evidence from Project STAR, *The Economic Journal*, 111, 468, 1–28.

LeSage J., Pace R.K., 2009, *Introduction to Spatial Econometrics*, Chapman & Hall / CRC Press.

LeSage J., Pace R.K., 2011, Pitfalls in Higher Order Model Extensions of Basic Spatial Regression Methodology, *The Review of Regional Studies*, 41, 1, 13–26.

Nagelkerke N.J.D., 1991, A note on a general definition of the coefficient of determination, *Biometrika*, 78, 3, 691–692.

R Core Team, 2016, *R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing*, Vienna, Austria.

Rivkin S.G., Hanusheket E.A., Kain J.F., 2005, Teachers, schools, and academic achievement, *Econometrica*, 73, 2, 417–458.

Tobler W.R., 1970, A Computer Movie Simulating Urban Growth in the Detroit Region, *Economic Geography*, 46, Supplement: Proceedings. International Geographical Union. Commission on Quantitative Methods, 234-240.

Wöessmann L., West M., 2006, Class-size effects in school systems around the world: Evidence from between-grade variations in TIMSS, *European Economic Review*, 50, 3, 695–736.

World Bank, 2014, *The Unfinished Revolution: Bringing Opportunity, Good jobs and Greater Wealth to all Tunisians*, Washington, DC: World Bank Group.

Yeh S.S. 2010,The cost effectiveness of 22 approaches for raising student achievement, *Journal of Education Finance*, 36, 1, 38–75.

ANNEX

Appendix

Derive the expression in (5):

$r = \alpha c + \rho W c^{*} \alpha^{*} + v$ using (3)

$= \alpha c + \rho W c^{*} \alpha^{*} + \gamma u + \varepsilon$ using (4)

$= \alpha c + \rho W r - \rho \alpha W c + \gamma u + \varepsilon$

$= \rho W r + \alpha c + (-\rho\alpha - \varphi\gamma) W c + \varphi\gamma W c + \gamma u + \varepsilon$

$= \rho W r + \alpha c + (-\rho\alpha - \varphi\gamma) W c + \varphi\gamma W c + \gamma c - \gamma c + \gamma u + \varepsilon$

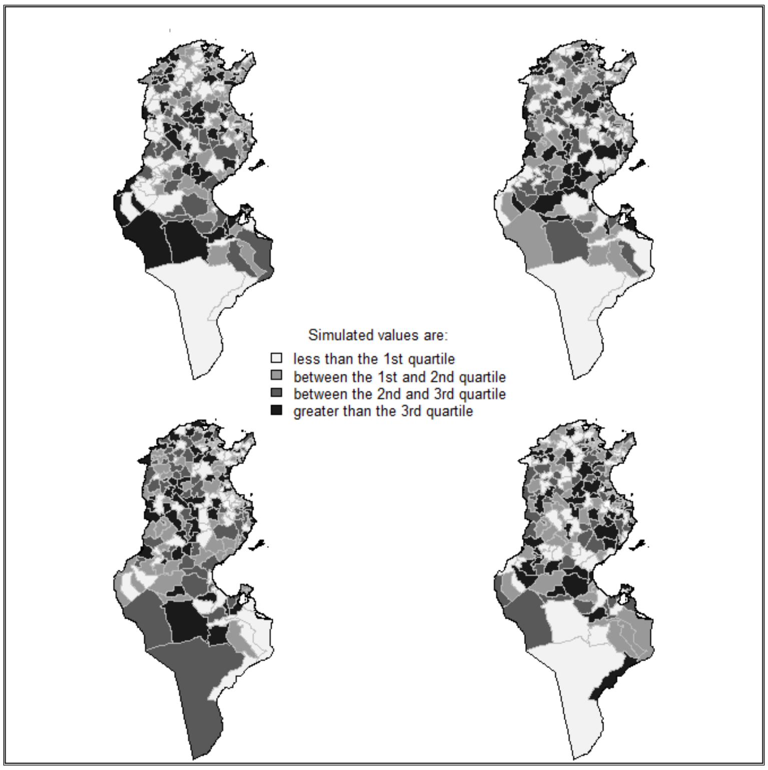
$= \rho W r + (\alpha + \gamma) c + (-\rho\alpha - \varphi\gamma) W c + \gamma(\varphi W c + u) - \gamma c + \varepsilon$

$= \rho W r + (\alpha + \gamma) c + (-\rho\alpha - \varphi\gamma) W c + \varepsilon$ using (2)

$= \rho W r + \beta c + \theta W c + \varepsilon$

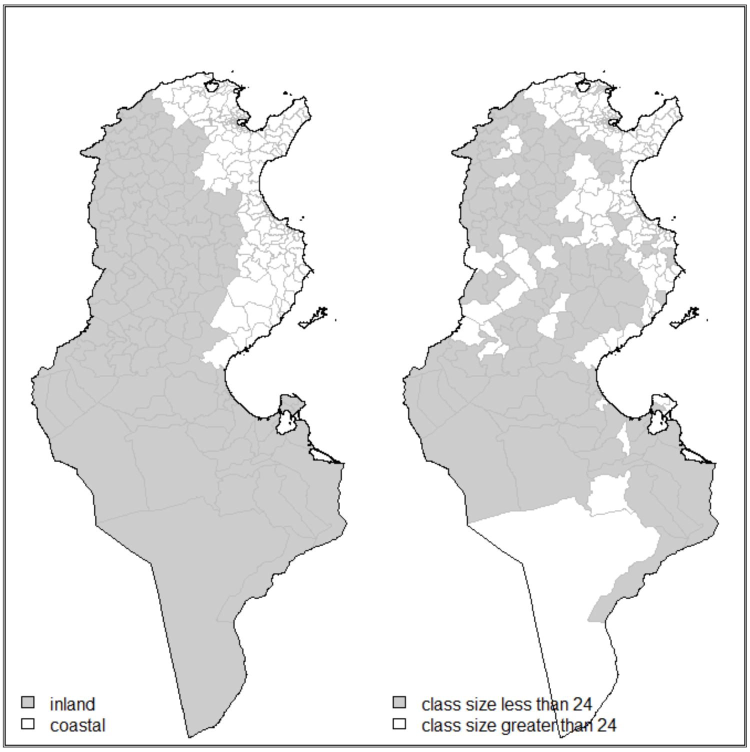
with $\beta = \alpha + \gamma$ and $\theta = -\rho\alpha - \varphi\gamma$.

Figure A.1. Illustration of four cases of random distributions from the generation of simulated values from a normal distribution



Source: Authors.

Figure A.2. Delegations considered in the study as "inland" or "coastal" (left) and delegations with an average number of students per class higher than 24 or lower or equal to 24 (right)



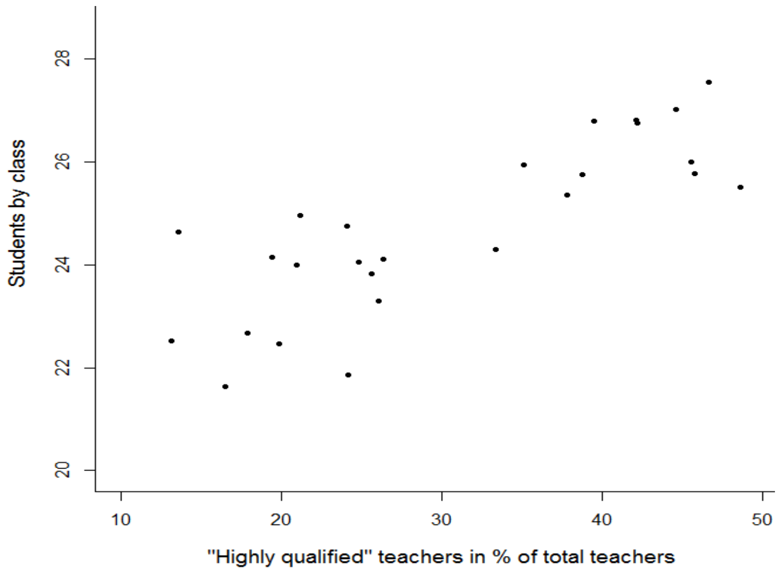
Source: Authors.

Table A. Crossing local associations

		Success rate		
		HH	LL	Other
	HH	18	0	29
Students per class	LL	0	4	41
	Other	18	27	123
(number of delegations by type of association).				

Source: Authors.

Figure A.3. Size of classes related to teachers' qualifications



Source: Authors.
This is the percentage of practicing teachers (2013-2014) in the general preparatory and secondary cycles having a Senior Professor title, Professor Emeritus, Senior Adjunct Pro-fessor, Senior Professor Emeritus, Associate Teacher or Senior Associate Teacher in relati-on to the total number of teachers in the preparatory and secondary cycle. The data are published on the website of the Ministry of National Education (www.education.gov.tn) and related to the 24 Governorates of Tunisia (the Tunis Governorate is subdivided into Tunis 1 and Tunis 2 and Sfax into Sfax 1 and Sfax 2), hence the number of observations being equal to 26. This percentage is 32% for the whole of Tunisia. The pupil class ratios at the level of the governorates come from the same source.

Taille des classes et réussite scolaire en Tunisie : une approche économétrique spatiale

Résumé - Cette étude s'inscrit dans le débat autour de l'influence de la taille des classes sur les performances scolaires. L'article propose une approche méthodologiquement différente des études réalisées jusqu'ici. À partir d'observations régionales portant sur des établissements scolaires en Tunisie, l'analyse s'appuie sur une exploration spatiale des données analysées (ESDA) et estime l'effet de la taille des classes à travers un modèle de Durbin spatial. Les résultats vont dans le sens des études qui défendent la réduction de la taille des classes pour de meilleures performances scolaires. Cependant, la réduction de la taille des classes n'a d'effets sensibles qu'avec 25 élèves, et peu au-delà.

Mots-Clés
Education
Taille des classes
Autocorrélation spatiale
Modèle de Durbin spatial
Tunisie