

GEOGRAPHICAL DISPARITIES AND DETERMINANTS OF FERTILITY IN INDIA: AN APPROACH BY SPATIAL ECONOMETRIC

Mohamed HAZEM

QUARG, ESCT, University of Manouba, Tunisia.

*Corresponding Author

Mohamed HAZEM

QUARG, ESCT, University of Manouba,
Tunisia.

Article History

Received: 22.07.2023

Accepted: 02.09.2023

Published: 23.09.2023

Abstract: *The aim of this paper is to analyze the disparities of fertility by the 591 Indian districts, according the population census in 2001. We seek to detect the spatial autocorrelation and heterogeneity of fertility rates relying on methods developed by the Exploratory Spatial Data Analysis and Spatial Econometrics. Factors such as the female literacy rate, the number of secondary schools, the urban or rural character appear largely explain the fertility rate disparity among Indian districts.*

Keywords: *Regional fertility rate, India, Exploratory Spatial Data Analysis, Spatial Econometrics.*

1. INTRODUCTION

During one century, India has seen a multiplication by 4.5 of its population and it is expected that it will increase by 50 percent in the next fifty years. The United Nations estimate that India could exceed 1.5 billion in 2050 (Hazem, 2021a). This is a country that illustrates the paradox of global demographic: the population is increasing rapidly; although the birth control is largely widespread and some states have a fertility rate below the level of replacement generations (Véron, 2006, Hazem, 2021a).

Population growth is explained by the structure of population relatively young by age and by an average fertility rate of 3.1 children per woman. The relative decline in Indian fertility in the past twenty-five years is mainly attributable to the reduction in number of the births of three or more children, indicating that the model of the family with two children begins to spread throughout the country (Spoorenberg, 2001, Hazem, 2021). The Indian population is now significantly urbanizing even if the rural population still above 65 percent. Kerala, one of the most developed state in the south of the country, has made an exception by the precocity and speed of its demographic transition: the crude birth rate declined from 44 per thousand in the 1950's to 17 per thousand in 2002 and the fertility which was 5.6 children per woman in the 1950s, is now below that of France (1.8 against 1.9). (Veron, 2006; Hazem, 2021a; Hazem, 2021b). Our purpose in this work is to analyze the disparities of fertility by 591 Indian districts in 2001; date of last census, and to reveal the relationship between the location of districts and their situations in terms of fertility. In this sense, we seek to detect the forms of spatial autocorrelation and heterogeneity basing on the newly developed methods of Exploratory Spatial Data Analysis as well as on the models of spatial econometrics. We seek to reveal the factors that may explain the differences in fertility rates among districts.

The article is organized as follows: section 2 suggests a study of the geographic disparity in the fertility rate in the Indian districts in 2001 using methods of Exploratory Spatial Data Analysis. In this context, we propose in section 3, a spatial econometric model

to determine the socioeconomic factors that explain the geographical disparities of fertility in India districts.

2. GEOGRAPHICAL DISPARITIES OF FERTILITY RATE

We use data of fertility rate of the last census (2001) for the 591 Indian districts¹. Figure (1) clearly shows the inequality in demographic regime in Indian districts. Districts with a fertility level below the national Indian average², 3.31 children per woman, are more widely observed in the southern part of the country (particularly, in the states of Kerala and Karnataka, the district Chennai in Tamil Nadu state). In these states, couples use contraception for spacing and limiting births and therefore the fertility rate revolves around the replacement level (two children per women) (Hazem, 2021a)). The poors are particularly affected by this practice. When they apply for job or request a credit, they are explained that they would have more chance of accepting their applications if they had a certificate of sterilization (Oyarzun, 2010). The coastal regions often stand out from the inside (ancient Christianization and Islamization, fishermen caste). Trades with the rest of the world, especially during the colonization, were the most intense and the progress of education was much faster (Guilmoto, 2007, Hazem, 2021a; Hazem, 2021b). The highest fertility rates are recorded mainly in the districts located in the northern half of the country (except in Jammu Kashmir state, particularly in the Leh district): districts in the states of Rajasthan, Uttar Pradesh Meghalaya, Madhya Pradesh, Bihar and especially in the Jaisalmer's district (5.79) in the state of Rajasthan, where the major campaigns of birth control, beginning in the late 1960s³, had little influence on demographic behavior in India (Hazem, 2021).

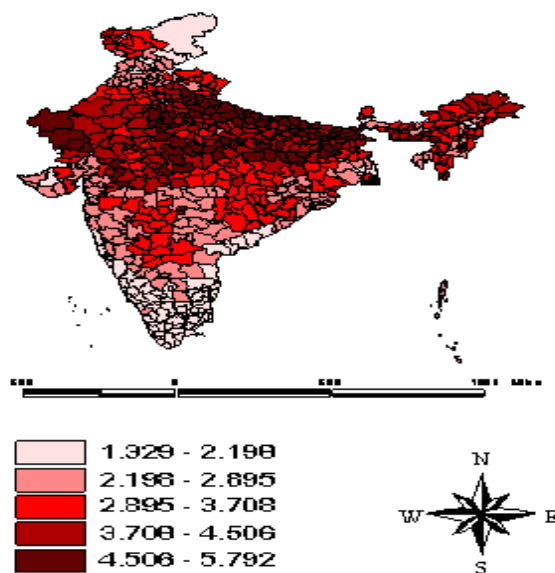
¹ Source: Register General & Census Commissioner, New Delhi, India.

² By taking the arithmetic mean of all sample observations.

³ In an attempt to curb population growth, India has held numerous political birth since the mid-twentieth century, starting with the famous slogan "Hum do Hamare Do" ("We're two, we do have two") in 1960 (Oyarzun, 2010).

The social and cultural diversity that characterize the Indian population explain that the introduction of new policies on birth (spacing and birth control) cannot be done uniformly by the states. For example, Hindus hardly conceive a family without at least two sons, and Muslims are opposed to the family planning. Besides, their beliefs are rejected and insulted to further reduce their minority. The values of fertility are almost homogenous in some states like Uttar Pradesh (it has the largest number of districts, 70 districts in 2001), Bihar, Kerala, and Tamil Nadu. Conversely, states as Jammu and Kashmir, Orissa, Andhra Pradesh, and Maharashtra are particularly heterogeneous. Generally, the distribution of fertility shows that 17 districts have below 50 percent of the national average (3.30), 136 have a level between 50 percent and 75 percent, 159 between 75 percent and 100 percent. Finally, 279 districts have levels above the national rate. Note that the highest fertility rate (1.75) is 4.36 times higher than the lowest one (0.40) (Hazem, 2021a).

Figure 1: The fertility rate in India by district in 2001



The highest fertility rates are recorded mainly in the districts located in the northern half of the country (except in Jammu Kashmir state, particularly in the Leh district); districts in the states of Rajasthan, Uttar Pradesh Meghalaya, Madhya Pradesh, Bihar and especially in the Jaisalmer's district (5.79) in the state of Rajasthan, where the major campaigns of birth control, beginning in the late 1960s⁴, had little influence on demographic behavior in India.

The social and cultural diversity that characterize the Indian population explain that the introduction of new policies on birth (spacing and birth control) can not be done uniformly by the states. For example, Hindus hardly conceive a family without at least two sons, and Muslims are opposed to the family planning. Besides, their beliefs are rejected and insulted to further reduce their

minority⁵. The values of fertility are almost homogenous in some states like Uttar Pradesh (it has the largest number of districts, 70 districts in 2001), Bihar, Kerala, and Tamil Nadu. Conversely, states as Jammu and Kashmir, Orissa, Andhra Pradesh, and Maharashtra are particularly heterogeneous. Generally, the distribution of fertility shows that 17 districts have below 50 percent of the national average (3.30), 136 have a level between 50 percent and 75 percent, 159 between 75 percent and 100 percent. Finally, 279 districts have levels above the national rate. Note that the highest fertility rate (1.75) is 4.36 times higher than the lowest one (0.40) (Hazem, 2021a). Conversely, modernization" explained by the progress of education and urbanization, can directly affect the fertility of the regions. This can create new forms of regional disparities since the access to education; the improvement of infrastructures and the degree of urbanization are related to the economic development of regions.

The map in figure 1 reveals a demographic diversity among the states of North and South of India, but sometimes within states. The values of fertility are almost homogenous in some states like Uttar Pradesh (it has the largest number of districts, 70 districts in 2001), Bihar, Kerala, and Tamil Nadu. Conversely, states as Jammu and Kashmir, Orissa, Andhra Pradesh, and Maharashtra are particularly heterogeneous. Generally, the distribution of fertility shows that 17 districts have below 50 percent of the national average (3.30), 136 have a level between 50 percent and 75 percent, 159 between 75 percent and 100 percent. Finally, 279 districts have levels above the national rate. Note that the highest fertility rate (1.75) is 4.36 times higher than the lowest one (0.40). The map in figure (1) leads to the question of the existence of spatial interdependences between districts (Veron, 2006; Hazem, 2021a).

2.1 AN EXPLORATORY SPATIAL DATA ANALYSIS

Exploratory Spatial Data Analysis brings together a body of techniques that enables the visualization of spatial distributions, identifies atypical localizations, detects the extreme observations and reveals configurations of spatial association (Bailey and Gatrell, 1995; Anselin, 1998; Boumont & al, 2000; Le Gallo, 2002; Ertur and Koch, 2004, Hazem, 2021a). These methods are frequently used for measuring both local and global spatial autocorrelation.

Spatial autocorrelation can be broadly defined as the correspondence between the similarity values of variables and proximity spatial units where these values are observed. In this sense, the evaluation of the global spatial autocorrelation is still the preliminary stage of exploratory spatial data analysis (Anselin and Bera, 1998; Ertur and Thiaw, 2005; Catin and Hazem, Catin and al., 2013, Hazem, 2021a). It is said to be positive if the values of the variable (low or high) form a concentration in space. By contrast, it is called negative if the variable takes different values around each different location. Finally, spatial autocorrelation is said to be null if the observations are randomly distributed in space.

Formally, for two regional units i and j , the presence of spatial autocorrelation is determined by the non-zero values covariance taken by the variable of interest. The structural form is defined as:

⁴ In an attempt to curb population growth, India has held numerous political birth since the mid-twentieth century, starting with the famous slogan "Hum do Hamare Do " ("We're two, we do have two") in 1960 (Oyarzun, 2010).

⁵ See Richard Meury:
http://www.inecine.com/inde_demographie.html

$$\text{cov}(x_i, x_j) = E(x_i x_j) - E(x_i)E(x_j) \neq 0 \quad \forall x_i \neq x_j$$

where x_i and x_j are the values of the study variable recorded in the two regional units i and j . However, it should be noted that this covariance doesn't show a spatial character only if when the distribution of units considered can be interpreted in terms of structure, of interaction or of spatial arrangement of observation (Ertur and Thiaw, 2005; Catin and Hazem 2012, Hazem, 2021a, Hazem, 2021b).

Observations distributed in space are frequently interdependent: what happens in a particular location depends on what happens in others. The "Spatial Weight Matrix" is the commonly used tool in modeling spatial phenomena taking into account the regional proximity. It is a square matrix having many rows and columns as there are regional units. For the two region i and j , the term w_{ij} reveals the degree of interaction and reflects the accessibility of region i to region j . In this sense, the application of Exploratory Spatial Data Analysis techniques requires a priori specification of the spatial weight matrix. In this work, we still seek a spatial weight matrix based on inverse distance as well as on the number (expressed by k) of the nearest districts⁶.

The Moran statistic is the most robust test for the presence of global spatial autocorrelation (Cliff and Ord, 1981; Upton and Fingleton, 1985; Anselin, 1992; Catin and Hazem, 2012; Hazem, 2021). Moran statistic (I_M) is defined by:

$$I_M = \frac{N}{K} \frac{\sum_i \sum_j w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2} \quad (2)$$

where, x_i is the value of the fertility rate in the district i , $\bar{x}$ is the general average, N is the number of districts within the space of India, w_{ij} measures the intensity of spatial interaction between the two districts i and j , and $K = \sum w_{ij}$ is the sum of interaction coefficients.

Table (1) shows that all Moran statistics are significant and the standardized values appear to be very high. This could indicate a problem of spatial levels in our sample: a part of the detected autocorrelation could be a result of administrative subdivision (Le Gallo, 2002; Catin and al, 2013, Hazem, 2021a) or the result of the presence of extreme observations that could affect the value of the global spatial autocorrelation. Through the presentation of two techniques for detecting extreme observation (Figure 2 and 3), we can say that the strong positive spatial autocorrelation is not the result of a non-spatial distribution since no extreme point is recorded. So, there is a real spatial interaction of Indian districts in terms of fertility rate.

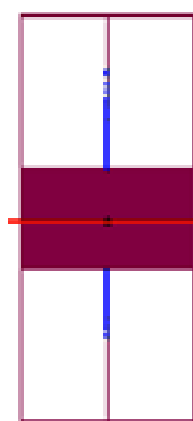
Table 1: Moran statistics: fertility rates in India in 2001

Matrix	I-Moran	Standardized Value	Cr-Prob
k=4	0,8094	23,197	10^{-4}
k=6	0,7791	26,823	10^{-4}
k=8	0,7548	29,644	10^{-4}
Distance	0,4786	53,287	10^{-4}

However, there is evidence that global spatial analysis can sometimes hide local particularities. To determine the significant spatial clusters around each district and to analyze the local instability in the form of atypical locations, it is common to use both the Moran Scatter Plot and LISA statistic (Anselin, 1995). Indeed, using these two techniques provides more information on the state of regional unity as well as on the interactions which can have with the surrounding neighborhoods. Moran scatter plots are established and represented in figure (4) using the matrices used in the previous analysis. In abscissa, the value of the standardized fertility rate (FER) appears and in ordinate the standardized lagged value appears (WFER)⁷. This diagram has four different quadrants. Each quadrant corresponds to a specific spatial affiliation that may exist between a district and its neighbors.

Table (2) shows that 72.5%, 429 districts (respectively 27.41% or 162 districts) are characterized by a positive spatial association (respectively negative). For the positive spatial clusters, 37.23% of the districts are located in a quadrant HH and 35.36% located in quadrant LL. Whereas the negative spatial associations that include outliers, 9.81% are districts with high fertility in an environment with low fertility rates and 17.56% present the opposite case.

Figure 2: Box plot: Detection of outliers



⁶ There are some remote districts which have no borders with other districts. Isolated observations must be eliminated in calculating the global and local autocorrelation since the corresponding lines in the spatial weight matrix will contain null values. For this reason, we cannot use the contiguity matrices.

⁷ The i^{th} element of this variable is defined by the weighted mean of the observations for the neighboring districts of the district i . Also, since W is standardized, the statistics of Moran the slope of the linear regression of WFER with respect to FER. This relationship is frequently represented in Moran's diagram.

Figure 3: Boxmap: Detection of outliers

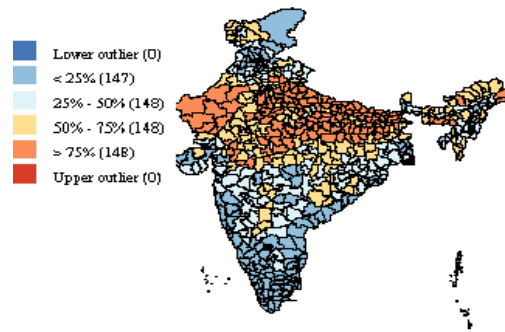
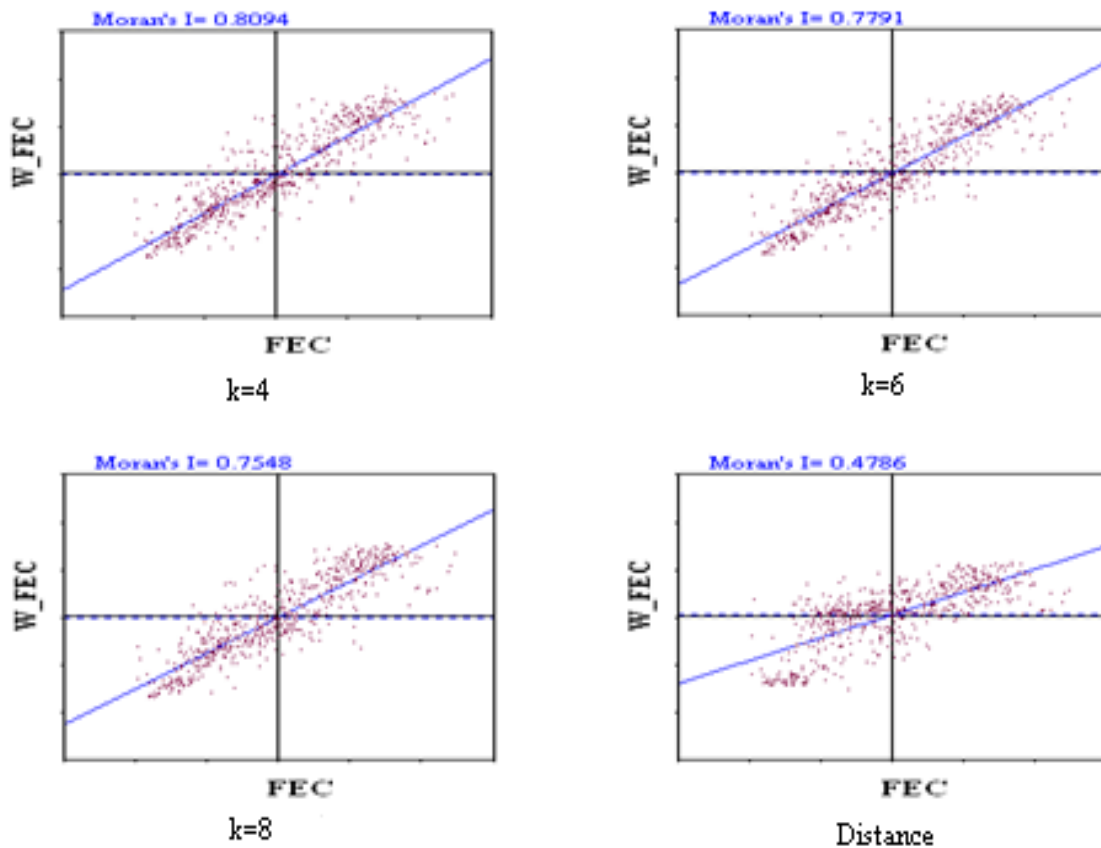


Table 2: Moran scatter plot: spatial associations in Indian districts

Matrix	Quadrant HH	Quadrant HL	Quadrant LH	Quadrant LL
Distances	37,225%	9,813%	17,597%	35,363%

Notes: The quadrant HH includes the districts which are associated with a high value of fertility rate, surrounded by districts which also have high values. Similarly, the quadrant LL contains only the districts that are associated with low fertility rate surrounded by districts that also have low values. Whereas, the quadrant HL contains districts that are associated with high values surrounded by districts associated with low values. Finally, the quadrant LH corresponds to districts associated with low values surrounded by districts s associated with high values.

Figure 4: Moran scatter plot: matrix of (4, 6, 8) nearest neighbor and inverse distances matrix



After determining the distributions of spatial associations using inverse distances matrix, we specify those that are significant by means of LISA statistics. For the district i , this statistic is defined as⁸:

⁸ Where the spatial weight matrix is standardized such that $K = N$, we have the following equation: $I_M = \frac{\sum_i LISA_i}{N}$

$$LISA_i = \frac{(x_i - \bar{x})}{G} \sum_{j \text{ neighbor } (i)} w_{ij} (x_j - \bar{x})^2 \quad \text{with } G = \sum_i (x_i - \bar{x})^2 / N \quad (3)$$

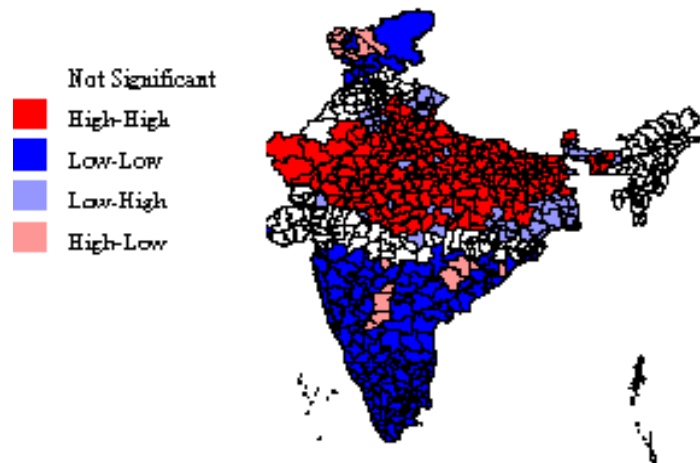
Table (3) shows two patterns of significant local spatial concentration, especially localized in both quadrants HH and LL. Respectively, the percentages are equal to 94.1 % (207 districts from 220) and 97.4% (166 districts from 209). The percentage of significant clusters is equal to 69.2% in quadrant LH and to 60.3% in quadrant HL. This shows the intensity of the positive global spatial autocorrelation found previously using the Moran test.

Table 3: Significant LISA statistics

Matrix	LISA HH	LISA HL	LISA LH	LISA LL
Distances	94,090%	60,344%	69,230%	79,425%

Figure (5) shows the significant clusters of quadrant HH located in districts of the northern part of the country. In addition, significant associations of type LL are located in the districts of the southern part of the country and some northern districts. These results are consistent with the intuitions of Dyson and Moore (1983) who first reported the fertility differentials in India at the social and cultural structure of the country and they pointed out the crucial role of the north-south virtual border separating sanscritized and Islamic north India from north Dravidian India with based specific cultural traditions (Guilomoto, 1997). Similarly, districts that are associated with HL significant LISA statistics are located in the north and south of country. Finally, the most significant LH clusters concern districts that surrounded significant spatial HH associations.

Figure 5: LISA map



2. 2. LINEAR AND QUADRATIC DYNAMIC

The “Trend-surface” is a representation in which the exploratory variables consist of polynomials of geographical coordinates of districts. An expression with three dimensions is generally used, the abscissa (x) and ordinate (y) are respectively represented by the geographical coordinates of the observations and the value of the variable represents the side (z). Linear (x and y) and quadratic (x, y, x², y², xy) forms are the most frequently used. Both models are estimated using the method of the ordinary least square (OLS). Table (4) shows that the coordinates (x and y) are positively significant. This proves that there is a growing trend in the level of fertility from West to East and from South to North.

Table 4: Linear trend

Variable	Coefficient	T-statistic	Critical-Prob
cste	-0,3097	-0,6590	0,5101034
x	0,0200	3,4642	0,0005703
y	0,0856	3,7277	0,0000000

Table (5) indicates that the introduction of quadratic variables (x², y² et xy) clearly improves the model fit (R² goes from 0.27 to 0.43). Excepting the variable y, all variables are statistically significant: the two variables x and y are positively significant while both variables x² and y² are negatively significant. This means that the model reveals a “bowl-like Shaped” for the trend surface, where the level of fertility is high in the middle and decays by approaching to the borders. This result is to be compared with results from Exploratory Spatial Data Analysis previously made.

Table 5: Quadratic trend

Variable	Coefficient	T-statistic	Critical-Prob
<i>cste</i>	-34,3720	-6,1576	0,0000000
<i>x</i>	0,8822	6,6480	0,0000000
<i>Y</i>	0,0448	0,3585	0,7200000
<i>xy</i>	-0,0059	-7,1256	0,0000000
<i>x</i> ²	-0,0078	-9,5594	0,0000000
<i>y</i> ²	0,0048	3,2916	0,0010559

To go further in the analysis, a spatial econometric model can be specified, basing, in first type, on the Moran test and in the second step, on the Lagrange multiplier test. This work is detailed in section 3.

3. THE DETERMINANTS OF FERTILITY IN INDIA

In this section, the goal is to show the determinants of fertility in India districts with socioeconomic nature. Due to the constraints of statistical information at this geographical level, we have identified four possible exploratory variables: the female literacy rate, the number of secondary school, the density of transport infrastructure in bus which describes urban space and the number of cooperatives banks. So, the model is defined by:

$$FER = \alpha S + \beta_1 LITF + \beta_2 SS + \beta_3 BS + \beta_4 COPB + \varepsilon \quad (1) \quad \varepsilon N(0, \sigma^2)$$

where, FER is the fertility rate in 2001, LITF is the female literacy rate, SS is the number of secondary schools, BS is the number of bus divided by the area district, COPB is the number of cooperative banks, S is the sum vector, and α and β are parameters to be estimated.

“Modernization” explained by the progress of education and urbanization, can directly affect the fertility of the regions. This can create new forms of regional disparities since the access to education; the improvement of infrastructures and the degree of urbanization are related to the economic development of regions.

The estimation results of equation (1) adopting the method of ordinary least squares (OLS) are reported in table (6). We note that the parameters β_1 and β_2 are significant with the confidence level of 95%. The parameter that corresponds to the density of transport infrastructure is not statistical significant while the parameter β_4 is significant with an error risk of 10%. However, the estimation of this model rejects the assumptions of normality and heteroskedasticity of error terms: the Jarque-Bera and Brush Pagan statistics are respectively equal to 25.650 ($P=0.0000027$) and 11.964 ($P=0.017166$). According to the table (6), using the inverse distance matrix, we recorded a high value of the Moran's statistics ($I_M = 0.211$) with a critical probability $p=0.0000000$. So there is a strong spatial autocorrelation among the error terms derived from the OLS estimation, which partly explain the value of the Jarque- Bera statistics. Estimations of the model (1) are biased and non-convergent and statistical inference based on OLS is unreliable. Therefore, this estimation method is released and it is necessary to be replaced by maximum likelihood method or generalized moments method.

Table 6: Estimation Results of the classical model for 2001 with spatial weight matrix of the inverse distances

Estimation results		Tests	
$\hat{\alpha}$	6,31 (0,0000000)	JB	23,1327 (0,0000095)
$\hat{\beta}_1$	-0,049 (0,0000000)	PBP	2,845644 (0,5839797)
$\hat{\beta}_2$	-0,002 (0,0000000)	KB	1,848962 (0,7635133)
$\hat{\beta}_3$	-0,07673574 (0,1105304)	WH	19,02485 (0,1640018)
$\hat{\beta}_4$	0,0015 (0,0855877)	I_M	0,216416 (0,0000000)
R^2	0,654824	LM-Lag	296,23559 (0,0000000)
LogL	-369,215	RLM-Lag	112,605958 (0,0000000)
AIC	748,431	LM-Err	244,361779 (0,0000000)
SC	768,425	RLM-Err	102,7321422 (0,0000000)

Notes: we use the following notations: LogL is the value of the log likelihood function. AIC is the Akaike Information Criterion (1974). SC is the Schwarz information criterion (1978). J is the test of normality of residuals of Jarque-Bera. BP, KB, WH, are the tests of heteroskedasticity of error terms respectively Breush-Pagan (1979), Koenker-Basset (1982) and White (1980). I_M is the Moran (Cliff and Ord, 1981). LM-Lag RLM-Lag denotes respectively the Lagrange multiplier test applied on the lagged endogenous variable and its robust version. LM-Err is the Lagrange multiplier applied on the error terms and RLM-Err denotes its robust version. The critical values are in parentheses.

Using the Lagrange multiplier test and its robust version could give a good specification model. In fact, the two most commonly models used to describe the spatial dependence are the model with error spatial autocorrelation and the lag endogenous variable model. A table (6) shows that the spatial lag model is the most appropriate specification to describe data-generating processes. So, the model (1) suffers from misspecification due to the existence of spatial autocorrelation between districts in terms of fertility rates.

The next step is to estimate the new model defined by:

$$\begin{cases} FER = \alpha S + \rho W FER + \beta_1 LITF + \beta_2 SS + \beta_3 BS + \beta_4 COPB + \varepsilon \\ \varepsilon \sim N(0, \sigma^2) \end{cases} \quad (2)$$

where ρ is the spatial autoregressive parameter that measures the intensity of spatial interaction that exists between each district and its neighbors in terms of fertility and W is the spatial weight matrix⁹.

The model (2) can be written as the following¹⁰:

$$FER = (I - \rho W)^{-1}(\alpha S + \beta_1 LITF + \beta_2 SS + \beta_3 BS + \beta_4 COPB) + (I - \rho W)^{-1}\varepsilon \quad (3)$$

or:

$$FER = (I + \rho W + \rho^2 W^2 + \rho^3 W^3 + \rho^4 W^4 + \dots)(\alpha S + \beta_1 LITF + \beta_2 SS + \beta_3 BS + \beta_4 COPB) + (I + \rho W + \rho^2 W^2 + \rho^3 W^3 + \rho^4 W^4 + \dots)\varepsilon \quad (4)$$

This expression indicates that the fertility rate recorded in a district is not only explained by the values of the female literacy rate, the number of secondary schools, the number of buses per km² and the number of cooperative banks recorded in this district, but also by the values associated with all districts (neighboring or not) through the inverse spatial transformation $(I - \rho W)^{-1}$. This is a spatial multiplier effect that affects spatial variables. The effect of spatial diffusion indicates that if a random shock in one district i affects the value of fertility rate in this district, it will have also an impact on fertility rates in other districts (neighboring or not). Obviously, these two effects weaken with distance, with the order of the contiguity or with the number of closets neighborhoods.

The results of the estimation of model (2) by the maximum likelihood method are reported in table (7)¹¹. The coefficient of determination is not appropriate to compare the OLS classical regression model and the spatial model or even spatial models themselves¹². In contrast, when we compare the value of Log-Likelihood (LogL) with that drawn from the OLS estimation, we find an increase (from -369.215 to -289.346) using the matrix of inverse distance (respectively, 247.275 and -269.554 for the case of 10 and 20 nearest neighbors' matrices). Compensating the improved fit for the added variable (the spatially lagged dependent variable), the AIC and SC both decrease relative to OLS, again suggesting an improvement of fit for the spatial lag specification. The estimation of the new model shows a significant positive spatial autocorrelation: the parameter ρ is estimated to 0.632 (respectively 0,628 and 0,615) with a very high critical probability ($p = 0.000000$). These results are confirmed by the significance of the likelihood ratio test (LRT) and justify the magnitude of spatial interactions that exist between the values of fertility rates in neighboring districts. This reflects the idea that the fertility rate decline proceeds typically by contagion, by the spatial diffusion of new behaviors. Note that the problem of heteroskedasticity of error terms has been deleted after this new modeling, which justifies the best specification of spatial model.

Fundamentally, it appears that the four used variables combine themselves to explain the fertility rate in 2001 in the Indian districts.

⁹ The standardization of the spatial weight matrix facilitates the comparison of the magnitude of the parameter when the model is estimated by several spatial weight matrices.

¹⁰ In this case $(I - \rho W)^{-1}$ must be non-singular, ρ is nonzero and the inverse is not an eigenvalue of W .

¹¹ In this estimate, the inverse distance matrix and those of k-neighbors ($k = 10, 20$) are used.

¹² The value listed in the spatial lag output is not a real R^2 , but a so-called pseudo- R^2 , which is not directly comparable with the measure given for OLS results.

Table 7: Estimation Results of the spatial lag model for the year 2001

	$W(dist)$	$W(k=10)$	$W(k=20)$
$\hat{\alpha}$	3,345461 (0,0000000)	2,907001 (0,0000000)	3,09977 (0,0000000)
$\hat{\beta}_1$	-0,03593514 (0,0000000)	-0,0276998 (0,0000000)	-0,03057605 (0,0000000)
$\hat{\beta}_2$	-0,001734427 (0,0000000)	-0,001359905 (0,0000000)	-0,001430823 (0,0000000)
$\hat{\beta}_3$	-0,1195452 (0,0021821)	-0,0658061 (0,0152860)	-0,09526523 0,0095196
$\hat{\beta}_4$	0,002082081 (0,0027493)	0,00155128 (0,0110593)	0,001551949 0,0176993
$\hat{\rho}$	0,6318247 (0,0000000)	0,6285336 (0,0000000)	0,6149714 (0,0000000)
R^2	0,770573	0,821771	0,795095
LogL	-289,346	-247,275	-269,554
AIC	590,693	506,55	551,108
SC	614,687	530,544	575,101
BP	5,088335 (0,2783533)	5,122247 (0,2749826)	4,695524 (0,3199886)
LRT	159,7378 (0,0000000)	243,8803 (0,0000000)	199,3229 (0,0000000)

The female literacy rate and the number of secondary schools tend to reduce the fertility rate, the parameters β_1 and β_2 , are negatively significant. Becker (1981), shows that female literacy is considered a key factor in fertility and child health, the idea is that the woman maximizes her choice of child number (McCray et Royer, 2005). Education improves recourse of care behaviors, such a consulting screening cancer service, prenatal supervision services or sexual transmitted diseases services (OMS, 2008). It is a factor that greatly contributes to the emancipation and empowerment of women about their daily live in general and the child nutrition in particular (Roser, 2003; Hazem, 2021a). Women who have finished at least primary education easily resort to a medical monitoring and access to echography during pregnancy than illiterate mothers. Educated mothers are aware that this contraceptive use has important implications for reproductive and maternal health for which the unwanted pregnancy can cause an unsafe abortion, and lack of contraception and lead to more pregnancies that may have a negative effect on the mother health (OMS, 2008). In fact, in addition to the economic repercussions linked to the level of education (employment and income opportunities), the education of women contributes to changing attitudes towards the sociological and psychological conditions that perpetuate poverty (UNDP, 1997, Hazem, 2021a). It is a factor that strongly contributes to emancipation and women's empowerment (Roser, 2003). In the majority of Indian regions, a high level of

education of women has a direct impact on fertility. In fact, the heterogeneity of Indian states and their individual autonomy translates into a diversity of statistics. In 1956, the scalar curriculum reform was enacted to promote gender equity (Roser, 2003). Literate women are better able to make responsible decisions in matters that affect themselves, their families, and children, including health practices, household income, education and household well-being. their children 15. Thus, the research results demonstrate that a child's ability to develop properly is closely related to the educational level of its mother (UNESCO, 2006; Hazem, 2021a).

Otherwise, large regional disparities exist in India in terms of basic infrastructures. The presence of secondary school can be considered as an indicator of "offer education". If we calculate the distance to school for the entire population, we obtain important values. Thus, Kerala, which is illustrated with a highly populated cities but relatively few, is characterized by an average distance often less than 1km and there are few cities that have not a school (Oliveau and Chasles, 2005). Conversely, the number of primary schools is closely correlated with the size of the population. This follows the policy of the state: make schools where there is population.

There is a very negative significant correlation between the number of buses per km² and the fertility rate of district

(parameter β_3). The density of public transport reflects the urban character of the districts. The level of urbanization in India can be characterized as relatively low, 29% in 2007, despite a very demanding definition of the city in terms of minimum population, more than 5000 people gathered in one city (Dupont, 2008). Three big cities currently have 12 to 17 million people (New Delhi, Calcuta and Bombay), and 35 cities with over one million inhabitants. For major part of the population, urbanization has increased the number and length of daily displacement, including those made using public transport (Pucher and al, 2007). The highest densities of buses number are registered in urban districts where the fertility rate recorded its lowest values, even if there is a high rural migration with high fertility rate of district.

In India, the banking sector is substantially controlled by the government and the cooperative banks. The number of cooperative banks is presented as a criterion of large rural population and is associated with higher fertility rates: the parameter β_4 is positively significant. To finance rural development, cooperatives credit institutions have issued large amounts of credit (Sawai, 2005). Their visibility is very important and their cooperative nature is widely promoted in their communication strategies (Boucher, 2011). According to a study done by the federation of Indian cooperative banks, a total of their loans and deposits are more important than that of the private sector. Currently, the cooperative credit provides 43% of total rural credit. Some studies made in India revealed that microcredit which's provided by the mechanism of self-assistance groups linked to bank credit is associated with a higher level of loan repayment and also helps the small farmers and marginal farmers to receive loans for the development of their activities and consumption needs (Boucher, 2001).

4. CONCLUSION

In this work, we studied the disparities in fertility rates between the Indian districts based on the latest census conducted in 2001. The application of Exploratory Spatial Data Analysis techniques has allowed us to reveal the geography of fertility level in these districts. Despite the cleavage north (high fertility) / south (lower fertility) of the country, the results show that there is a strong global and local spatial autocorrelation; no outlier in fertility rates is detected.

Then, we tried to explain the regional disparity of the fertility depending on education and infrastructure conditions. Estimation of the model by OLS method has shown that there is a strong spatial autocorrelation among the error terms. By using the Lagrange multiplier and its robust version and processing by the estimation method of maximum likelihood, we conclude that the spatial lag model specification is the most appropriate to describe the disparity and spatial interactions of fertility rates.

Fertility in India is lower in the districts where the female literacy rates and the number of the secondary schools are high. Similarly, the number of buses per km² which is considered as a criterion of urbanization is negatively correlated with the level of fertility. Conversely, the number of cooperative banks is very high in districts with high fertility rates. It is the effect of population size which encourages banks to expand their branches in rural districts.

This work deserves to be extended by a dynamic analysis of fertility based on the last two censuses, 1991 and 2001.

Analytically, the determination of steady state of fertility in Indian districts is an idea to explore.

References

1. Anselin L. (1988). Spatial Dependence and Spatial Structural Instability in Applied Regression Analysis, *Journal of Regional Science*, 30, 185-207.
2. Anselin L. (1992). Spatial data analysis with GIS : An introduction to application in the social sciences, Technical Report, 92-10.
3. Anselin L. (1995). Local Indicators of Spatial Association-LISA, *Geographical Analysis*, 27, 93-115.
4. Anselin L. (1998). *Interactive techniques and exploratory spatial data analysis, Techniques, Management and Applications*, Wiley, New York.
5. Anselin L., Bera A. (1998). Spatial dependence in linear regression models with an introduction to spatial econometrics. Chapter 7 (pp. 237-289) in Aman Ullah and David Giles (eds.). *Handbook of Applied Economic Statistics*, NewYork: Marcel Dekker.
6. Arbia G. (1988). *Spatial data configuration in statistical analysis of regional economic and related problems*. Kluwer Academic Publisher, Dordrecht, Boston.
7. Arbia G, Basile R., Salvatore M. (2003). Measuring spatial effects in parametric and non parametric modelling of regional growth and convergence. Paper prepared for the UNU/WIDER Project Meeting on Spatial Inequality in Development Helsinki, 29 May 2003.
8. Bailey T., Gatrell A. (1995). *Interactive Spatial Data Analysis*, Longman, Harlow.
9. Becker G. S. (1981). *A Treatise on the Family*. Cambridge, MA, Harvard University Press.
10. Boucher J. (2011). Tour d'horizon du milieu coopératif en Inde. Coopérative de développement régional. Bas-Saint-Laurent/ Côte-Nord. Coopérative de Solidarité.
11. Cliff A., Ord K. (1981). *Spatial Processes, Models and Applications*. Pion, London.
12. Dall'erba S., Le Gallo J. (2005). Dynamique du processus de convergence régionale en Europe, *Région et Développement*, n°21, 119-139.
13. Dupont V. (2008). L'urbanisation en Inde : révision de quelques mythes et vrais défis, *Rayonnement du CNRS*, n° 47.
14. Dyson T., Moore M. (1983). Kinship Structure, Female Autonomy, and Demographic Behaviour in India, *Population and Development Review*, 9, 1, 1983.

15. Ertur C., Koch W. (2004). Analyse spatiale des disparités régionales dans l'Europe élargie. Pôle d'Economie et de Gestion, Université de Bourgogne.
16. Ertur C., Thiaw K. (2005). Croissance, capital humain et interactions spatiales : une étude économétrique, Pôle d'Economie et de Gestion, Université de Bourgogne.
17. Guilmoto C. (1997). La géographie de la fécondité en Inde (1981-1991). *Espace, Populations. Sociétés*, 145-159.
18. Guilmoto C. (2005). L'Inde et son milliard. Démographie en début du siècle, Document de recherche n°8, Série Santé de la reproduction, fécondité et développement.
19. Haining R. (1990). *Spatial Data Analysis in the Social and Environmental Sciences*, Cambridge University Press, Cambridge.
20. Hazem M. (2021a). Long-Term Spatial Associations and Regional Convergence of Rural
21. Literacy Rate in India, *International Journal of Multidisciplinary Research and Analysis*, 4, 348-361.
22. Hazem M. (2021b). Disparities and Regional Convergence of Literacy Rate in India:
23. A Spatial Econometric Approach, *International Journal of Social Science and Human Research*, 4, 326-336.
24. Le Gallo J. (2002). Econometrie spatiale : L'autocorrélation spatiale dans les modèles régression linéaire, *Economie et prévision*, 155, 139-157.
25. Le Gallo J., Dall'erba S. (2005). Croissance, convergence et interactions régionales : les outils récents de l'analyse spatiale quantitative. *Région et Développement*, n° 21, 5-11.
26. Manier B. (2006). Quand les femmes auront disparu. L'élimination des filles en Inde et en Asie, Portail d'information et de ressources sur Genre et Développement.
27. Mankiw N, Romer D., Weil D. (1992). A contribution to the empirics of economic growth, *Quarterly Journal of Economics*, 107, 407-437.
28. Marie B. (1999). L'Inde et la Chine: problèmes démographiques et fécondité. 8^{ème} Congrès International, Liver development, gene regulation and disease, Palazzo del Popolo Orvieto, Italie, 2-5 juin 1999.
29. McCrary J., Royer H. (2005). The effect of maternal education on fertility and enfant health : evidence school entry policies using exact date of birth. *American Economic Review*, vol. 100, n° 10 158-195.
30. Morin H. (2006). L'Inde en déficit troublant de petites filles. Genre et action, Portail d'information et ressources sur Genre et Développement.
31. Oliveau S., Chasles V. (2005). Sur les chemins de l'école : accessibilité et alphabétisation dans le monde rural sud-indien, *Espace, Population, Sociétés*, 3, 453-464.
32. OMS (2008). Différence entre les sexes et santé de la femme dans la région de la Méditerranée orientale-perspective intersectorielle.
33. Oyarzun E. (2010). L'Inde, deuxième berceau du monde. Journée mondiale de la population : égalité entre les genres cette année, <http://inde.aujourdhuilemonde.com/1%E2%80%99inde-deuxieme-berceau-du-monde>.
34. Pucher J., Korattyswaroopam N., Ittyerah N. (2007). Le transport public urbain en Inde : Tendances, enjeux et innovations, *Revue Union internationale des transports publics*, Bruxelles.
35. Roser C. (2003). L'impact des politiques de scolarisation des filles, Mauritanie, Tunisie, Inde, Bangladesh et Sénégal, Institut National de Planification de l'Education, UNESCO.
36. Sawai S. (2005). Un bref commentaire sur la microfinanace: L'expérience de l'Inde, Union Nationale Coopérative de l'Inde, New Delhi, Inde.
37. Serge C. (1997). Les disparités régionales au Canada : diagnostic, tendances et leçons pour la politique économique, Document de travail, 18.
38. Spoorenberg T. (2010). La transition de la fécondité en Inde entre 1977 et 2004. Analyse des probabilités d'agrandissement, *Population*, 2, vol. 65.
39. Stetzer F. (1982). Specifying Weights in Spatial Forecasting Models: The results of some experiments, *Environment and Planning*, 14, 576-584.
40. UNESCO (2006). Initiative pour l'alphabétisation: savoir pour pouvoir, Perspectives et stratégies, Division de l'éducation de base.
41. Upton G., Fingleton B. (1985). *Spatial Data Analysis by Example*, vol. 1, John Wiley, New York.
42. Véron J. (2006). Stabiliser la population de l'Inde : plus facile à dire qu'à faire, *Population et Sociétés*, Bulletin mensuel d'information de l'institut national d'études démographiques.
43. William M., Anthea B. (2004). Spatial dependance in regional unemployment in Australia, Centre of Full Employment and Equity. The University of Nerwcastle, Australia. Working Paper.