Spatial Dependence and the Determinants of Child Births in Swedish Municipalities 1974-2002

Krister Sandberg[1]
International Institute for Applied Systems Analysis
A-2361 Laxenburg, Austria

Thomas Westerberg[2]
Department of Economics
Umeå University
SE-901 87 Umeå, Sweden

2005-03-23

Abstract

The Swedish total fertility rate (TFR) has been shown to fluctuate quite dramatically during the last 40 years, from 2.4 in 1965 reaching down as low as 1.5 in 1999. Although numerous studies in the past have tried to explain this fluctuation, there has been surprisingly little focus on impacts from local or even regional differences. Thus, the purpose of this paper is to analyse, for the first time, whether there exists a significant dispersion in the number of child births normalised by the number of fertile women among Swedish municipalities for the period 1974-2002 and then provide an explanation for occurred differences. Regional differences, as well as the presence of spatial dependence, are first confirmed in an exploratory spatial data analysis. Particularly noticeable is the volatility and how local hotspots emerge and disappear during the study period. Subsequent regression analyses are performed for a selected number of years where we solve for spatial dependence and use economic-, institutional-, sociological-, and geographical characteristics of the municipalities as explanatory variables.

JEL classification: J13, R23, C21

Key words: child births, regional differences, spatial dependence, Sweden

[1] sandberg@iiasa.ac.at
[2] thomas.westerberg@ac.lst.se

Introduction

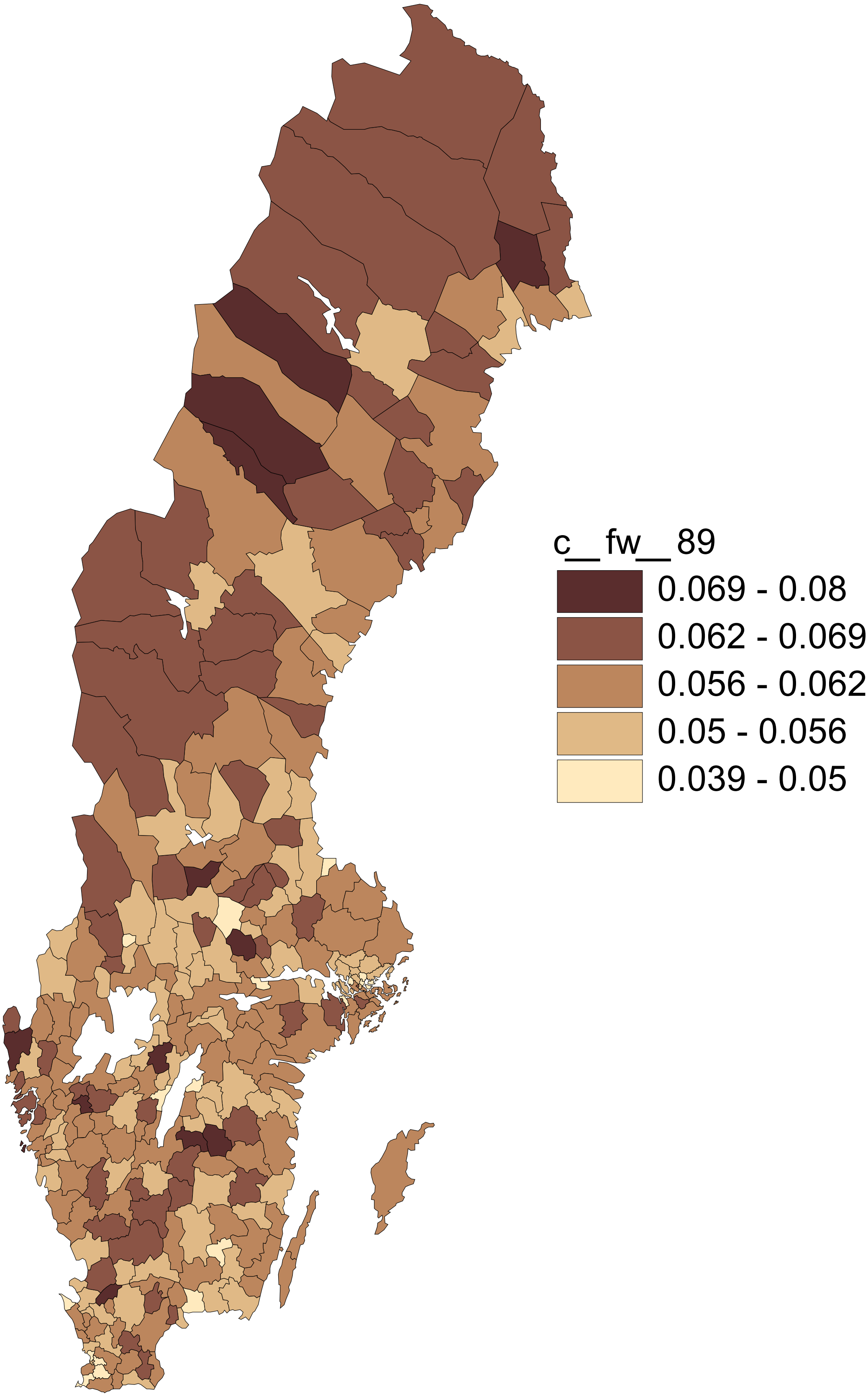
The total fertility rate[3] in Sweden has shown a varying pattern during the 20th century. From 1900 until the mid 1930s the Swedish birth rate decreased from approximately four children per woman to less than two children per woman. The decline was followed by an increase and the 1950s showed a fairly constant birth rate, just above the replacement rate. In the 1960s the fertility rate increased and reached a peak in 1965. Thereafter, the birth rate started to decrease again and in the early 1980s the level was as low as it was in the 1930s. However, the remaining part of the 1980s, the birth rate increased dramatically and was equal to 2.14 in 1991. The birth rate has since then gradually decreased and at the end of the 1990s, the Swedish birth rate was lower than ever. During the first years of the new millennium the negative trend has been broken which means that the fertility rate has gradually started to increase. The discussion concerning fertility swings has in Sweden, as in most other countries so far, been discussed at the national perspective although there might be reason to believe that the fertility rate has differed between Swedish regions over the years.

Thus, the purpose of this paper is to estimate spatial deviations considering municipal (regional) birth rates in Sweden during the period 1974-2002 and discuss economic, demographic and social factors that may have influenced the fertility development at the disaggregated level. The main contributions of the study are (i) a regional analysis of the Swedish fertility rate during the last 25 years, (ii) identify how "regional Hot and Cold Spots", where the annual birth rate has been higher/lower compared to other regions, emerge and disappear during the period and (iii) explain the regional fertility swings from economic, demographical and sociological indicators for selected years.

Common and central aspects of most economic markets are interaction, externalities, spill-overs, copy-catting, etc. Thus, in a similar way as units depend upon each other in a time setting, they may also influence each other over the geographical space; an important complication one has to consider when dealing with cross-sectional data. A further complication that calls for special estimation techniques is that spatial dependence (spatial autocorrelation) does not limit itself to influence in one direction, but is spreading in all directions. Disregarded spatial dependence when actually present will lead to inefficient and in some cases also biased estimates.

[3] The total fertility rate (TFR) is the total number of children a woman would have by the end of her reproductive period (usually defined as between age 15 and 44 or between 15 and 49) given the current age-specific fertility rates throughout her childbearing life.

Children born per fertile woman, 1989



Economic analysis of fertility focus mainly on the number of children in the household see e.g. Ermisch (1991); Tasiran (1993); Hotz et al (1997) or the timing and spacing of children during the woman's life cycle see e.g. Heckman and Willis (1976); Wolpin (1984); Moffit (1984); Cigno and Ermisch (1989); Tasiran (1993); James (1996). Such studies are mostly based on individual data. Models of the relationship between macroeconomic variables and the aggregated fertility rate over time are not as frequent, especially not on a regional level. An exception is found in Hoem (2000a) who analyses whether the decision to have a first child was influenced by the dramatic changes in the labor market between 1986 and 1997. The result shows that first-fertility rates for women rise and fall with employment in different municipalities. A similar result, a strong negative relationship between the local unemployment rate and the probability of a first and second birth, was found in Norwegian data for women aged 20-39 during the period 1989-1995 (Naz, 2000). Some other papers that also study the relationship between fertility swings and macro data are Wilkinson (1973), Gauthier and Hatzius (1997) and Löfström and Westerberg (2002). However, those papers are based on national data. The Wilkinson study uses a non-linear autoregressive estimator and focuses on the period 1870-1965. The results of the analysis are for example a positive relationship between the male wage rate and fertility and a negative correlation between the female wage rate and fertility from 1940 to 1965. The underlying model in the paper by Gauthier and Hatzius uses variables on the *first difference* form and a one-year lag structure between the independent variables and the fertility rate. In their paper, the relationship between family benefits and fertility in different groups of countries is analyzed for the period 1973-1990. The groups of countries are Anglo-Saxon, Southern Europe, Continental Europe and Scandinavia. The variables used are child allowances, maternity leave benefits and duration of maternity leave, and the control variables are female and male wages, unemployment and changes in the level of unemployment. The paper concludes that child allowance was positively related to fertility, although its effect was limited; the effect of benefits appear to be greatest and most consistent in the Scandinavian countries, and maternity leave did not have any significant impact in any of the country groups. Finally, the study by Löfström and Westerberg analyze the fertility swings in Sweden between 1965 and 1998. The results from the study are for example a strong negative correlation between changes in the proportion of female students relative to the total number of women aged 19-30 and the fertility rate as well as a positive (but weak) relation between the total fertility rate and the female labor market participation.

The paper is organized in the following way. In the next chapter, the previous studies on economics and fertility are presented. A presentation of our data and a exploratory analysis is given in chapter three. The econometric model is presented in the fourth chapter while the results are analyzed in chapter five. The findings are summarized in the final chapter.

Economics and births of children

The underlying assumption in the economic analysis of fertility is that households are rational economic units that act optimally in any given situation and that the intentions of the two partners in the household will therefore always be equal. The child is assumed to provide a utility to the household which is compared to that of other goods via the family utility function. The parents' expenditures for their children's upbringing can be divided into two parts: *direct costs* associated with the children's food, clothes, toys etc., *indirect costs* related to the time-consuming effect of children on their parents' time. Population economic focuses mainly on the indirect costs, which are correlated with the income the parents must give up by spending time bringing up their children.

Income, education and labor market conditions

In the early 1960s, Leibenstein (1957), Becker (1960, 1965) and Mincer (1963) stress that the female wage rate is negatively correlated with the households' number of children while the male wage rate is positively related to the family's number of children.[4] The theory assumes that the male partner is the family breadwinner and that the female spouse has the main responsibility for the children. Therefore, the higher female wage rate, the more expensive it will be to stay out of work. Since the wage rate normally depends on the employee's human capital it is also assumed that greater work experience and higher education levels among women results in fewer children. The negative relationship is also well documented empirically, see e.g. Winegarden (1984); Lee and Chuen (1989) and Wang and Famoye (1997).

However, the support for this negative relationship has become weaker in most countries over recent decades. Sweden is one example where the female labor market participation rate and the total fertility rate did simultaneously increase in the 1980s and decrease in the 1990s. One reason, often mentioned as an explanation for the positive relationship in Sweden during this period is the extensive family reforms since the 1970s and when they were withdrawn again during the 1990s, see e.g. Sundström and Stafford (1992); Hoem (1993); Sundström (1994); and Hoem and Hoem (1996). Andersson (1999) for example suggests that the positive relationship depends on the income-replacement character of the parental leave system, which creates strong incentives for women to acquire as high a level of income as possible before having children. Moreover, the norm for young people has long been first education, secondly a secure position in the labor market, and thirdly starts a family. Since this "norm" comprises both men and women, this has also reinforced the idea of sharing responsibility for children. Young women no longer accept the role of sole responsibility for the children and an increasing proportion of young men are reluctant to be the main economic provider.

The discussion implies that increasing labor income for both men and women would result in more children. High labor income in a region may therefore imply good economic conditions which in turn may encourage young people to start a family. The effect on fertility from high labor income is, however, not clear cut since people with high labor income will also face higher opportunity costs, at least in the short run.

[4] Another important contribution to this discussion is Butz and Ward (1979).

Following the assumption that good economic conditions may encourage young people to start a family, we would expect the contrary to be discouraging. Gauthier and Hatzius (1997) state that high unemployment in general may also have a discouraging effect on (especially) women already in permanent jobs, since the risk of not being re-employed on the same terms as before childbirth or not having the same career opportunities will be too high. The discussion is supported by Hoem (2000a) who shows that first-fertility rates for women rise and fall with employment in different Swedish municipalities between 1986 and 1997. A similar result, a strong negative relationship between the local unemployment rate and the probability of a first and second birth, was found in Norwegian data for women aged 20-39 during the period 1989-1995, see Naz, (2000).

Investment in human capital is always a way of improving competitiveness in the labor market, but education is also chosen as an alternative in times of higher unemployment. Irrespective of the reason, education may make women (and men) postpone their first child; see Gustafsson, Kenjoh and Wetzels, 2001. Depending on a woman's age, postponing may mean fewer children and in some cases no children at all. A study by Löfström and Westerberg (2002) support the negative relationship between higher education and the fertility rate. In their study they found a strong negative correlation between changes in the Swedish total fertility rate and changes in the number of female students during the period 1965-1998. However, the expected negative effect of higher education on the fertility rate must be treated with caution. Cigno and Ermish (1989) show in a theoretical model that women with relatively greater human capital will have their first child later in life, but this does not necessarily mean that they will have fewer children. They can have as many children as those women who became mothers at a younger age, only the spacing between children is less.

Geography and Demography

The urbanization trend in Sweden has been quite strong during the last decades. Young people - in fertile ages - move from rural areas to the cities. This pattern will, from a demographically perspective imply better conditions for fertility growth in cities simply because the number of individuals in fertile ages become larger in these regions. However, since the supply of career- and leisure opportunities normally is larger in cities compare with rural areas the birth rate may in fact be lower in city regions and higher in rural areas cause to larger alternative cost for family building in the city regions, see Löfström, (2000). There is neither a clear-cut between a positive net migration to the cities (or a positive net migration in general) and the fertility rate in the short-run. The main reason is that few people move just for having children. The reasons for migration are more likely to get on in the world, having a new job, to study or just fulfill oneself.

Surroundings

Besides the families' incomes, costs and labor market participation there are of course a lot of other factors that affect the households' fertility decision. Some of these factors may be given by living surroundings or services in the villages.

Municipalities that invest in activities for parents and their children might be more attractive for parents to be, than municipalities with a low share of investments in such activities. Further, the birth rate would for that reason also be relatively higher in the places of provincial government (county capitals) compared with other municipalities in the region since the services (of all kind) normally is higher in those cities. On the other hand, these cities – like all larger cities - do normally also show a larger supply of career- and leisure opportunities for potential parents, compare with rural areas and these opportunities may increase the alternative cost for family building, see Löfström, (2000).

Data and Exploratory Spatial Data Analysis

Data in this study are drawn from official statistics (primarily Statistics Sweden) and cover the period 1974-2002 at the municipality level. A problem with the Swedish municipality data, apart from lack of data especially for the earlier years, is the fact that the number of municipalities has changed 6 times, from 277 in 1974 to 289 in 2002 and makes panel data analysis very hard. We have therefore decided to analyze the Swedish fertility data to compare the differences in regional fertility from a set of independent variables. We estimate the regional fertility rate in 1989[5], 1991 and 1993 from a set of independent variables one year earlier. The years 1989, 1991 and 1993 are chosen because they mirror increasing fertility rates as well as decreasing fertility rates.

This paper define municipality fertility rate (MFR) as the ratio between annual number of children born in each municipality respectively and annual number of women in fertile ages in thousands per municipality respectively. MFR is used as the dependent variable in our estimations. The independent variables are first those related to income and labor market participation. A variable of prime interest in our analysis is the average income per municipality (***Avg.inc***). The expected sign is positive even though there is no clear-cut information about the sign. The second variable measures the general economic condition in the municipalities, here given by the change in municipality unemployment rate (***Unemp***). The expected sign is negative.

The second category of variables is used to illustrate the attractiveness of the municipality and the size of the municipality respectively. We use net migration (***Net.mig***) to indicate whether the municipality attracts people or not. The expected sign is positive even though there is no clear-cut information about the sign. To illustrate the relationship between rural areas and the fertility rate a dummy variable (***Rur. area***). The expected sign is positive.

Finally, the third category of variables is used to measure the municipalities' investments in social, educational and leisure activities respectively. The variables are given as the ratio between the social (***Soc***), the educational, (***Educat***) and the leisure activity budget (***Leisure***) and the population in the municipality respectively. The expected sign is positive.

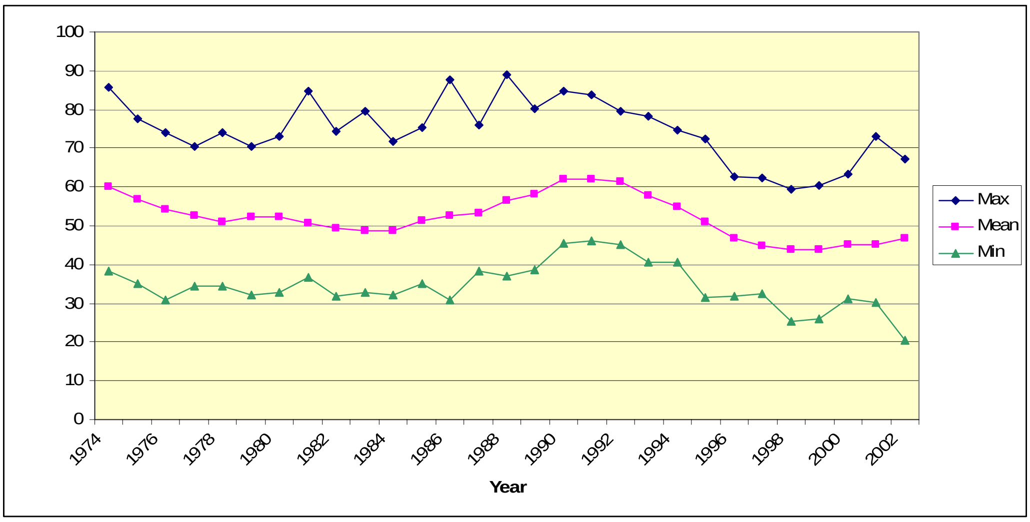
[5] We have so far analyzed the year 1989 in detail, and will include additional years in a final version.

The average, standard deviation, minimum and maximum values for each variable are summarized in Table 1 below.

Table 1. Descriptive Statistics 1989

Variable	*Unit*	*Mean*	*St. Dev.*	*Min*	*Max*
MFR-89	Children/K women	58.07	6.65	38.68	80.25
MFR-88	Children/K women	56.51	7.04	37.00	88.82
Avg.inc	1000 SEK	101.23	12.29	83.00	180.00
Unemp	Change in ratio	-0.27	0.167	-0.65	1.12
Net.mig	Dummy	0.76	0.43	0	1
Rur. area	Dummy	0.19	0.39	0	1
Soc	Ratio	0.28	0.04	0.17	0.39
Educat	Ratio	0.28	0.04	0.12	0.39
Leisure	Ratio	0.07	0.02	0.02	0.11
County Capital	Dummy	0.08	0.28	0	1

Figure 1: Number of children born per 1000 fertile women (16-49) between 1974 and 2002.



In Sweden, the total fertility rate was increasing since the mid-1980s it reached its climax in 1991 and decreased thereafter gradually during the 1990s. The pattern of child births per 1000 fertile women follows the pattern of the TFR for Sweden over the same period. It starts of at a high level in 1974 reaching a low in the mid 1980's. From there it begins to increase again to its peak in the first years of the 1990's. Then it gradually decreases during this decade but with an upward trend when we reach the 21st century. At the same time the municipalities with the highest and lowest birth rates differ quite substantively, but in general follow the average trend.

Exploratory spatial data analysis is a good start in order to test for spatial dependence (spatial autocorrelation). In this way we may confirm or reject the hypothesis that municipalities with similar birth rates are more clustered than would be expected by pure chance. At our disposal are a couple of global tests for spatial autocorrelation, such as the Moran's I (Moran 1948; Cliff and Ord 1973, 1981).

The Moran's I test is defined as[6]:

$$I = \frac{n}{S} \frac{\sum_i \sum_j w_{ij}(x_i - \mu)(x_j - \mu)}{\sum_i (x_i - \mu)^2}$$

$$S = \sum_i \sum_j w_{ij}$$

Where n is the number of observation. x_i and x_j are the observed births rates at locations i and j with a mean μ. S is the sum of all weights, w, and equals n when the weights matrix is row-standardized. This restricts I-values to lie between +/- 1.

It is important to select a proper spatial weighting function when calculating the spatial autocorrelation statistic since the purpose of this function is to assign weights to each pair of observation locations. This means that the value of the autocorrelation will depend on the weights as well as the data for each observation. When constructing the weight function one often has Tobler's first law of geography in mind, Tobler (1970), "everything is related to everything else, but near things are more related than distant things". It therefore becomes a matter of distance or nearness between locations. The simplest of weight functions for area data is a set of binary weights that is assigned the value one if two observations have a common border and zero if they do not. These weights are then summarized in the spatial weights matrix W. Since each observation by convention can't be its own neighbor the diagonal consists of zeros.

Over time it has become more common to instead use the actual geographical distance between objects to define the influence of nearness. To further emphasize this, the use of the inverse or the inversed squared distance is often used. Further, in order to have better properties on the matrix, and thus improve the interpretation of the parameter values in the regression, the matrix is row-standardized so that each row sums to one. The autocorrelation is in this way expressed in relative terms.

For the present study we have constructed and tested a number of different weights matrices. Each municipality has a centroid (two coordinates) to assign it to a geographic place. Based on these centroids a distance grid between the municipalities was created. From this grid we construct a number of row-standardized weights matrices to test for spatial autocorrelation based on the following spatial weights functions, with names in parentheses:

- Inverse distance (Ried)

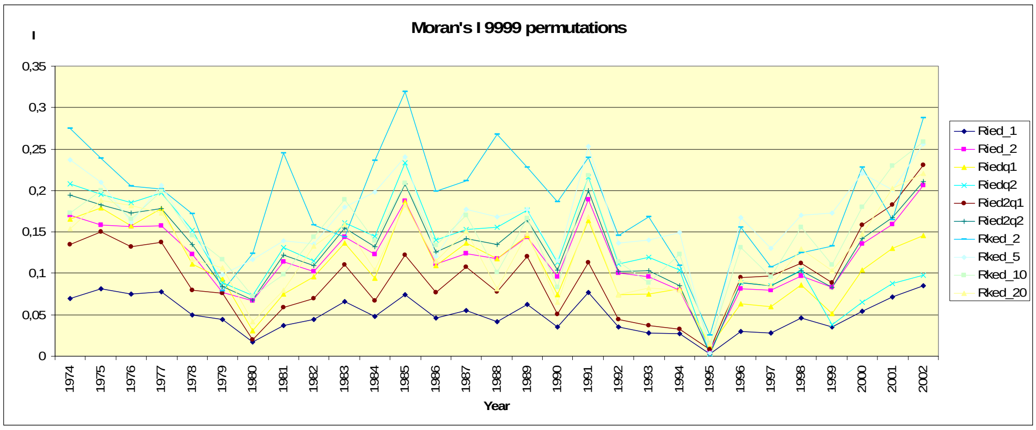
---

[6] The test statistic is compared with its theoretical mean, $I = \frac{-1}{(n-1)}$. $I \to 0$ as $n \to \infty$.

When $H_0 : I = \frac{-1}{(n-1)}$ is rejected and $I > \frac{-1}{(n-1)}$ indicate positive spatial autocorrelation, thus high values and low values are more spatially clustered than would be assumed purely by chance.

- Inverse distance squared (Ried_1)
- Inverse distance, restricted to a distance of the first quartile (Riedq_1)
- Inverse distance, restricted to a distance of the mean (Riedq_2)
- Inverse distance squared, restricted to a distance of the first quartile (Ried2q_1)
- Inverse distance squared, restricted to a distance of the mean (Ried2q_2)
- K-nearest neighbors, (Rked_2, 5, 10, 20)

Figure 2: Moran's I between 1974 and 2002 for the spatial weights matrices



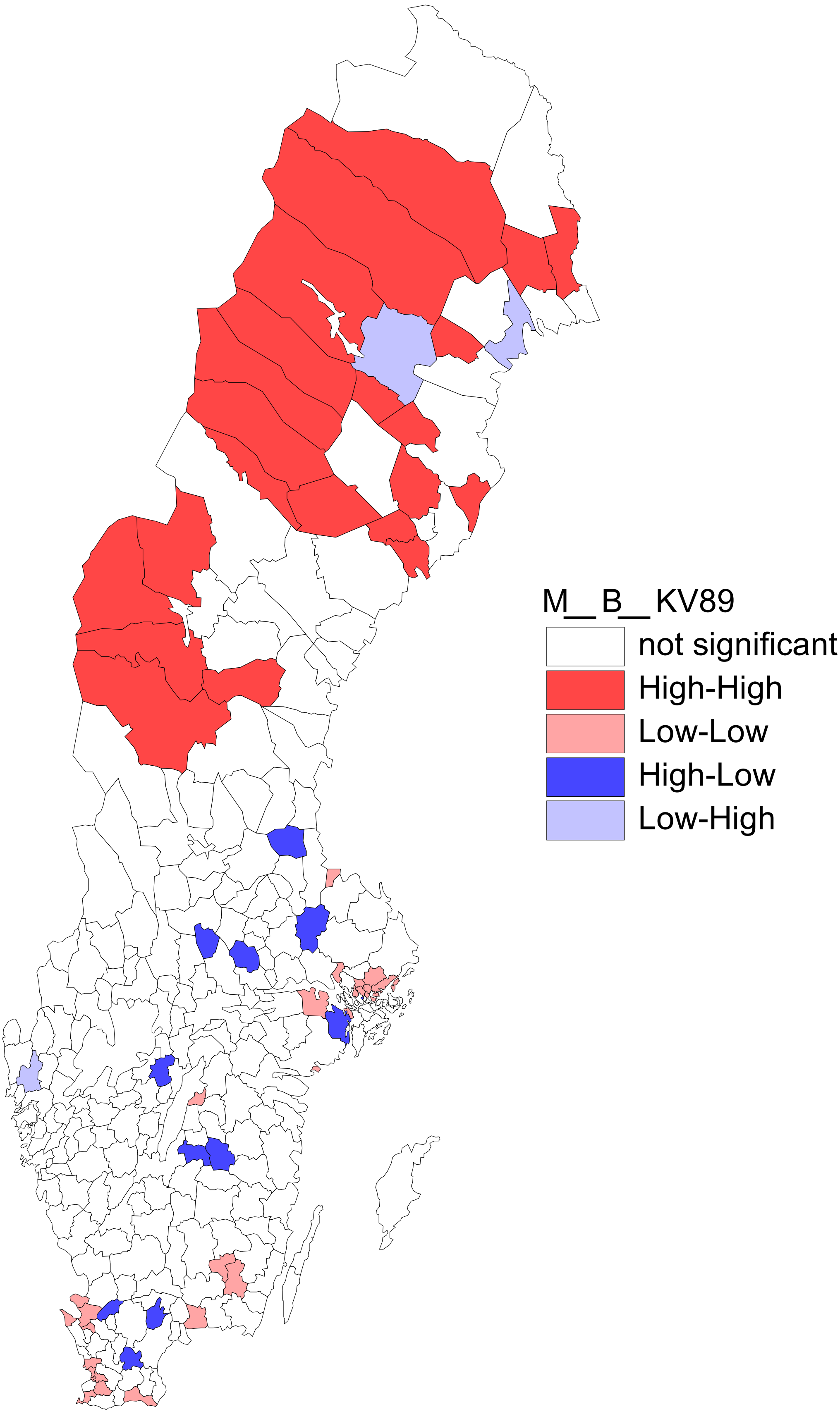
The general picture for the Moran's I is that spatial dependence is present irrespectively of the weights matrix used. From a high value during the 70's it reaches a low in 1980. From this low value it increases to a relatively high value during the most of the 80's and in the first part of the 90's. In the later part of the century and the first years of the 21st century it is once again significant.

The notion of global tests refers to the fact that they consider the overall data pattern and only return a single value which either confirm or rejects the hypothesis of no spatial dependence. No specific information is given about the prevailing pattern. When this is of interest, local tests, such as the local Moran's I test suggested by Anselin (1995) are helpful. The local Moran's I investigates whether the values for each municipality (from the global Moran's I) are significant or not and how influential they are individually for the overall spatial autocorrelation. It is defined as:

$$I_i = \frac{x_i}{\sum_i x_i^2} \sum_j w_{ij} x_j$$

The following three maps show the results of Local Moran's I calculated with the weighs matrix Ried2q_1.

Moran Significance Map, 1989. Ried2q__1



The first map display an interesting pattern with significant clusters of municipalities with large numbers of child births (High-High) in the middle north and far north. Areas with a low number of child births (Low-Low) are apparent in the capital area and in the very south. The clear pattern of High-High municipalities is very similar to the Swedish official H-region of sparsely populated areas[7] and will later be incorporated in the second part of the regression analysis.

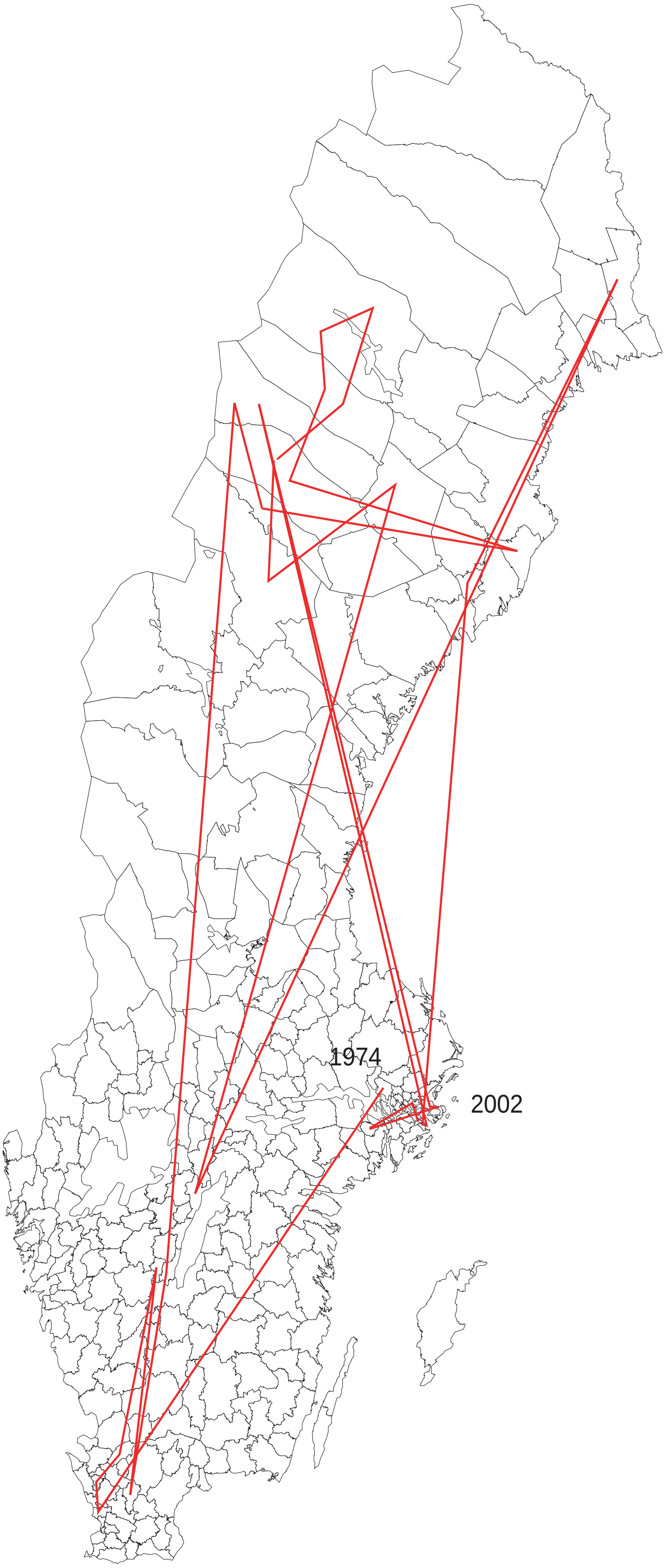
The following two maps show municipalities with the highest/lowest I-value for each of the four quadrants for each year between 1974 and 2002. The lines between each node combine one year with the following year. The first map shows the municipalities in the 1st quadrant (High-High) with each year's highest (and significant) I-value. It begins in 1974 with a Stockholm region municipality. The following years have high values in the southern part of Sweden. In the mid-80s and early 1990s the northern part of Sweden has the highest I-values. The trend of high childbearing municipalities then ends up in the capital region.

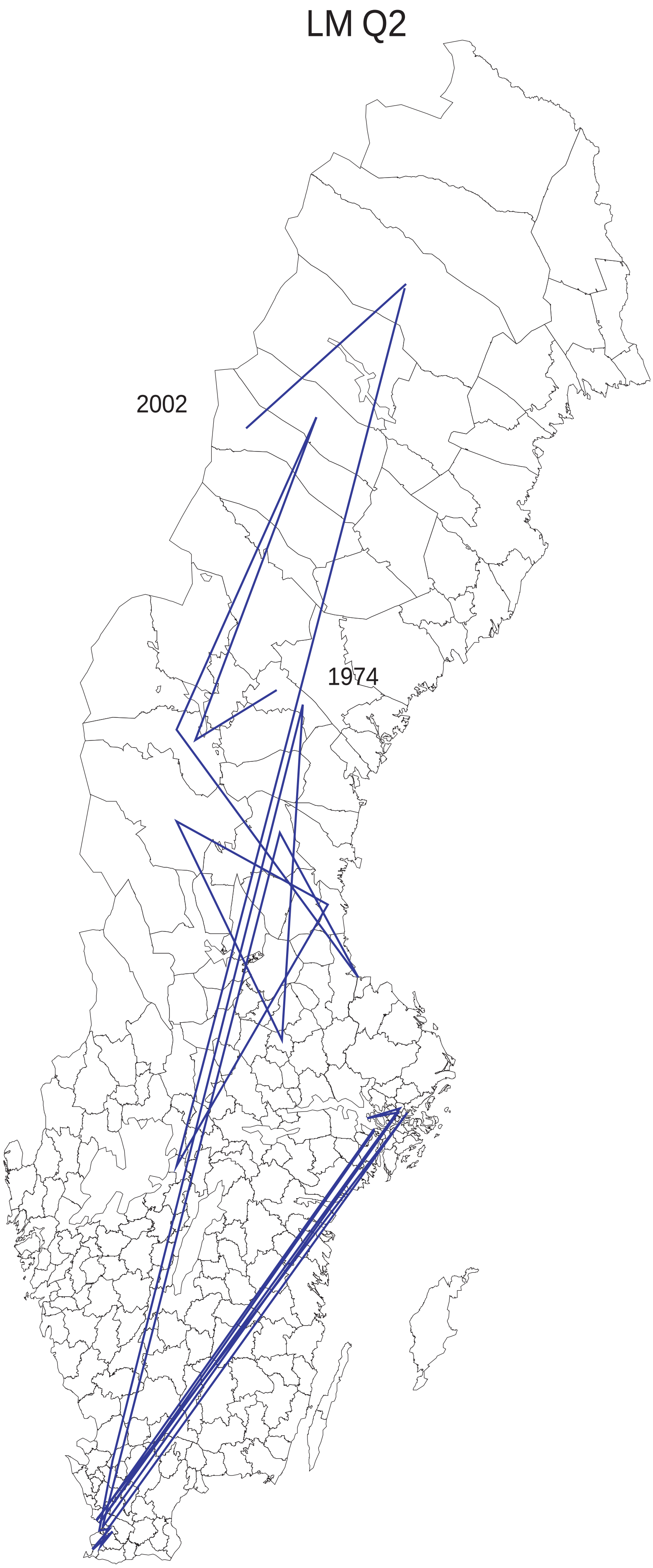
The second map shows the 2nd quadrant (High-Low) and indicates the municipalities with the highest (and significant) I-value for each year. For this map with start in the northern part of Sweden before going to the most south west part in the early 80s. Then a strong tag-of-war starts with the Stockholm area before going back north at the end of the century.

The municipalities with Low-High (3rd quadrant) birth rates are illustrated in the third map. Even though we begin in the north, these municipalities are mostly located in the middle part of Sweden with some temporary deviations. Finally the 4th quadrant (Low-High) show municipalities with low birth rates with neighbours with high birth rates. This seems to be a pattern in the southeast, but mostly in the counties of Västerbotten and Norrbotten in the northern part of Sweden.

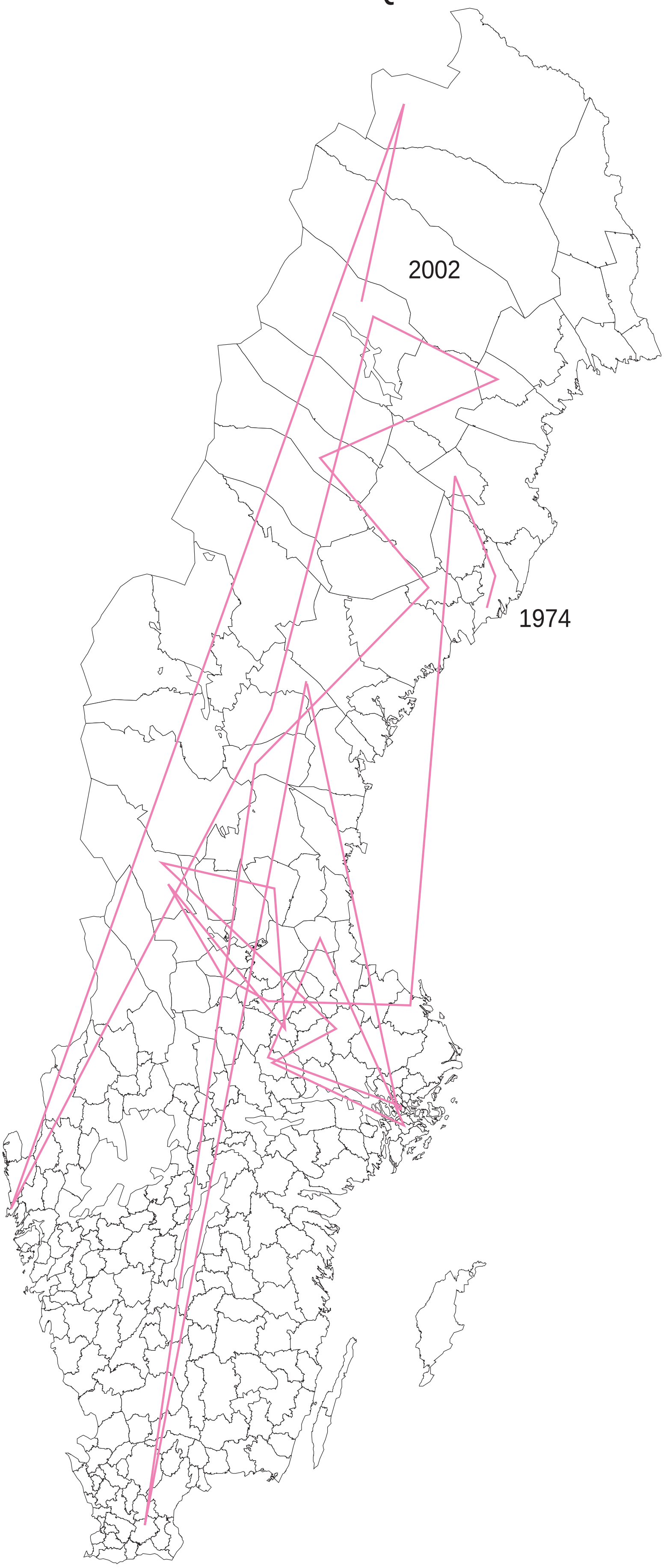
[7] Municipalities with less than 27,000 inhabitants within a radius of 30 km from its centre.

LM Q1

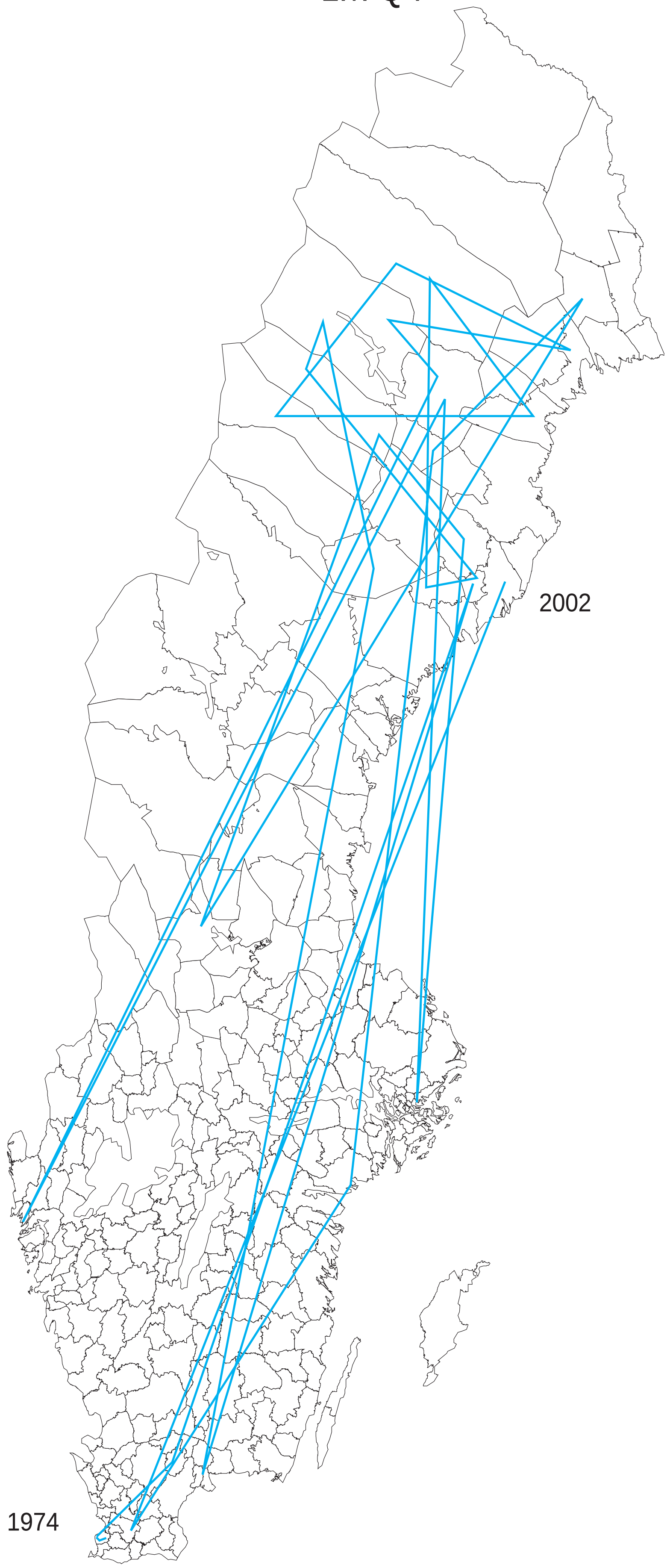



LM Q2
2002
1974

LM Q3



LM Q4



The Econometric Model

A birth begins with the parents' decision to have a child, thereafter it may take some time to conceive, which is followed by the gestation period (see Hoem, 2000a). For this reason, we must distinguish between the time when the parents decide to have a child and the time when the birth is actually recorded. When estimating the total fertility rate at time ***t*** most optimal information is given by a set of covariates (***x***) that describe the variables that may affect the fertility decision some nine months or so before period ***t***. Unfortunately, it is not possible to create such a lag structure from our data since the information we have about the variables is on a yearly basis.

Therefore, as an approximation we use the value of the independent variables from the preceding calendar year (see Gauthier and Hatzius 1997 and Poot and Siegers 2001). This means that we will estimate the fertility rate at time period ***t*** from a set of covariates at time period ***(t-1)***. The discussion can be formalized as:

$$Y_t = \vartheta + K'\mathbf{x_{t-1}} + \Pi_t \qquad (1)$$

where Y is the dependent variable, ϑ is the constant term intercept coefficient, K is a vector of coefficients, **x** a vector of control variables, Π the residual and ***t*** a time indicator.

Gauthius and Hatzius (1997) discuss that changes in unemployment may be more relevant than the level of unemployment for the fertility decision. If the parents expect high unemployment tomorrow and experiences high unemployment today then nothing has changed. On the other hand, if unemployment is expected to rise during the year, then the probability of having a child may be lower due to a higher risk of losing one's job.

Of that reason we use changes in the municipality unemployment rate instead of the level of unemployment. Equation (1) can then be written as:

$$Y_t = \vartheta + K'\mathbf{x_{t-1}} + PAz + \Pi_t \qquad (2)$$

where Az is the change in unemployment between *t* and t-1 and P is a coefficient.

Slow-moving influences may cause serial correlation of the residuals in time series estimation. Therefore, we will also estimate equation (2) with an autoregressive component included. The autoregressive component captures potential sluggish adjustment of fertility to changes in the independent variables (see Gauthius and Hatzius, 1997). Equation (2) can be written as:

$$Y_t = \vartheta + \Omega Y_{t-1} + K'\mathbf{x_{t-1}} + PAz + \Pi_t \qquad (3)$$

If we apply equation (3) on spatial data, dependences between spatially distributed observations may still remain caused by omitted variables or unidentified links between actors. Naturally, the obvious sources of interactions, such as barriers, municipality belonging etc. are controlled for in a normal way. Dummy variables are an obvious choice. However, even after this, there may still be spatial dependence between observations due to omitted variables or tacit relations that is hard to pin down.

To correct for remaining spatial dependencies, Cliff and Ord (1973) introduced the Spatial Autoregressive Error model to control for spatial error dependences while Anselin (1988) suggested the Spatial Autoregressive Lag model (SAR Lag) that may be used to control for spatial autoregressive lags. Slowly, the two models have managed to be recognized in the econometrics literature in such different subjects as hedonic prices, economic growth, political voting, and the spread of deceases.

Spatial error dependence may arise when the number of child births in adjacent municipalities move in the same direction due to correlated unobservable variables. That is, a lack of stochastic independence between observations exists. In this case, the parameter estimates received from the OLS estimation will be inefficient. In order to correct this, the spatial error dependence may be controlled for by the following spatial autoregressive error formulation:

$$Y_t = \vartheta + \Omega Y_{t-1} + K'\mathbf{x_{t-1}} + PAz + \Pi_t \qquad (4)$$

$$\Pi_t = \lambda W \Pi_t + \zeta$$

In (4), λ is the spatial autoregressive parameter, W is a weights matrix used to simplify and mimic the interactions between the municipalities, and ζ is a vector of i.i.d. errors with variance σ^2.

Spatial lag dependence is present when there is correlation between observations in the dependent variable, i.e., the number of child births in one municipality is directly influenced by the number of child births in adjacent municipalities. If spatial lag dependence is not corrected for the presented parameter estimates are both biased and inefficient.

In Anselin (1988), the SAR Lag model was introduced as a solution to the spatial dependence problem. It is defined as an extension of (3) into

$$Y_t = \rho W Y_t + \vartheta + \Omega Y_{t-1} + K'\mathbf{x_{t-1}} + PAz + \Pi_t \qquad (5)$$

Above ρ is the autoregressive coefficient and WY_t is the spatially lagged dependent variable.

In order to know what kind of spatial dependence that is present and how to correct it, the following rule of thumb has been suggested by Florax et al. (2003):

- Run the OLS.
- Use the residuals to test the null hypothesis of no spatial dependence caused by an omitted spatial lag or error term with two Lagrange Multiplier tests (test for error and lag dependence). See Anselin (1988) and Burridge (1980).
- If the null hypothesis is not rejected, use the OLS parameter estimates.
- If both LM test reject the null hypothesis, confirm the test by use of their robust counterparts. The rule is to use the spatial model that receives the highest value.
- When only one of the LM-test rejects the null hypothesis, use that model in the subsequent analysis.

The Empirical Results

As stated in the previous chapter we first run an OLS regression. It results in the parameter estimates in the first column of Table 2 below. The second column shows the probability for the parameter estimates. The adjusted fit is 40% with a number of significant parameters. One of them is the birth rate from the previous year that has a positive effect on this year's birth rate. The average income in the municipality has a negative effect on child births. The parameter for the municipality spending on social welfare is positive and significant at 10%. This is also the case for spending on leisure, although with a negative sign. The dummy variable for municipalities in the rural area has a positive coefficient and thus implies higher birth rates in sparsely populated areas. From the test on spatial dependence we conclude that the model could be improved upon if a spatial lag is included. We have used the Ried2q_1 weights matrix. Since we reject the null hypothesis of normally distributed error terms we are unable to use Maximum Likelihood estimation. Instead we use the Instrument Variable estimation. The instruments are spatially lagged variables, lagged with the same weights matrix. In order to also solve problems of heteroscedasticity we divide the data in two groups according to the rural area dummy. The third and fourth columns show the results from the spatial lag regression. The coefficients all have the same signs but their magnitude have changed. The impact from the average income is reduced but is still significant. The importance of municipality spending on leisure has increased and is now significant at the 5% level. Also the social welfare system has an increased parameter value. The coefficient for the rural municipalities is still positive but is only significant at the 10% level. The spatial lag is quite high and has the expected positive sign. The goodness of fit has increased somewhat although it is hard to compare a spatial model with a non-spatial model.

Table 2. Regression results for the year 1989.

Variable	*OLS*	*OLS (prob)*	*SAR Lag (IV)*	*SAR (prob)*
P			0.56	*0.004*
Constant	44.34	*0.000*	9.68	*0.445*
Y_{t-1}	0.40	*0.000*	0.39	*0.000*
County capital	-0.59	*0.619*	-0.99	*0.378*
Avg.inc	-0.09	*0.002*	-0.07	*0.020*
Leisure	-38.33	*0.071*	-46.78	*0.025*
Educat	-6.10	*0.434*	-9.39	*0.215*
Soc	15.77	*0.081*	20.59	*0.020*
Unemp	1.70	*0.359*	0.08	*0.963*
Net.mig	0.32	*0.667*	0.72	*0.317*
Rur. Area	3.02	*0.001*	1.78	*0.091*
Category_1			22.99	
Category_2			34.55	
R_2	0.42		0.45	
R_2-adj	0.40			
Sig-sq	26.44			
Sq.corr			0.43	
Normality of errors	24.16	*0.000*		
Koenker-Basset with category	2.64	*0.10*		
LM-error (Ried2q_1)	1.06	*0.30*		
LM-lag (Ried2q_1)	5.55	*0.018*		

Heterogeneity (Structural Change in Space)

We divide the data in two groups based on results from the Local Moran's I test and the previously defined H-region that had a significant and positive coefficient in the previous regressions. The first group consists of most municipalities while the municipalities that correspond to a sparsely populated area are collected in the second group. This division means that the regression returns two parameter estimates for each variable. It is then possible to test whether each parameter value is significantly different from its counterpart. The regression results are presented in Table 3 below.

The variable ***County capital*** had to be removed since all those municipalities belong to the first group, and would in that case lead to perfect co linearity. The LM-tests was insignificant for all weights matrices, but only slightly for the Ried2q_1, and since the test results are sensitive to residual properties we still decided to perform a spatial structural change regression. It turns out that the spatial lag parameter is highly significant and has a suspected positive sign. Compared with the previous regressions in Table 2 the overall goodness of fit is slightly higher. This is explained by the small evidence of structural change found in the data. For the tests on stability of individual coefficients we only find a significant difference between the two groups for the ***Unemp***, and a weak difference for the ***Net.mig*** parameter. Interestingly they both had insignificant parameter values in the previous regressions.

Table 3. Structural change regression results for the year 1989

Variable	*OLS_1*	*OLS (prob)*	*Coeff. Stability*	*SAR_1 (IV)*	*SAR (prob)*	*Coeff. Stability (prob)*
P				0.54	*0.005*	
Constant_1	40.34	*0.000*	*0.137*	12.94	*0.543*	*0.241*
Constant_2	68.07	*0.000*		30.60	*0.179*	
Y_{t-1}_1	0.43	*0.000*	*0.262*	0.42	*0.000*	*0.283*
Y_{t-1}_2	0.31	*0.001*		0.30	*0.002*	
Avg.inc _1	-0.08	*0.014*	*0.493*	-0.06	*0.063*	*0.419*
Avg.inc _2	-0.18	*0.226*		-0.19	*0.237*	
Leis _1	-62.18	*0.012*	*0.427*	-65.02	*0.005*	*0.432*
Leis _2	-24.48	*0.545*		-27.00	*0.524*	
Educat _1	-3.18	*0.685*	*0.395*	-5.99	*0.423*	*0.498*
Educat _2	-21.68	*0.286*		-21.26	*0.317*	
Soc _1	16.67	*0.093*	*0.802*	18.95	*0.044*	*0.372*
Soc _2	22.72	*0.304*		41.88	*0.083*	
Unemp _1	-1.12	*0.575*	*0.006*	-1.90	*0.321*	*0.007*
Unemp _2	13.50	*0.007*		12.93	*0.012*	
Net.mig _1	1.13	*0.173*	*0.021*	1.13	*0.015*	*0.056*
Net.mig _2	-3.07	*0.058*		-2.45	*0.150*	
Category_1				22.80		
Category_2				27.86		
R_2	0.46			0.47		
R_2-adj	0.43					
Sig-sq	25.38					
Sq.corr				0.46		
Normality of errors	20.02	*0.000*				
Koenker-Basset with category	0.74	*0.389*				
LM-error (Ried2q_1)	0.68	*0.411*				
LM-lag (Ried2q_1)	3.45	*0.063*				
Chow-test			*0.000*		*0.016*	

Conclusions

At the municipality level, the variables with significant parameter estimates were the average income level and the municipality expenditure on social welfare and leisure. We may also conclude that there exists path dependence in child births from one year to the other. For the year 1989, we may also say that a north-south or sparse-densely populated divide is present with higher birth rates in the sparsely populated municipalities in the northern part of Sweden. This is something that is highlighted in the significant parameter values for the change in unemployment and the net migration.

This divide is important to acknowledge, both when one tries to estimate the influential factors for birth rates, since it means that spatial dependence significantly influence the results and the structural change, unless taken into account, may also lead to a false rejection of variables with otherwise significant parameter estimates. In the long run, since we have shown that regional differences among the Swedish municipalities exists it is also important for the policy makers at the national level to have this in mind since one municipality might benefit from a national policy while others may loose.

This paper has studied an unchartered part of the demographic literature and much research still remains before we can state that we have a satisfying Swedish birth rate pattern. We now have a first rudimentary toolbox that must be expanded to include additional years as well as a more detailed variable selection such as the population distribution in the municipalities, for both men and women. The time-lag aspect should also be developed.

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.

Andersson, G., (1999), *"Trends in Childbearing and Nuptiality in Sweden: A Period Analysis"*, Demography Unit, Stockholm University.

Becker, G, S., (1960), *"An Economic Analysis of Fertility"*. National Bureu of Economic Research, pp 209-231.

Becker, Gary, S., (1965), "A Theory of the Allocation of Time". *The Economic Journal*, vol 75, nr 299, s 493-517.

Burrigde, P. (1980). On the Cliff-Ord test for spatial correlation. *Journal of the Royal Statistical Society B* 42, 107-108.

Butz, William, P., Ward, Michael, P., (1979), "The Emergence of Countercyclical U.S Fertility". *American Economic Review*, vol 69, s 318-328.

Cigno, A., Ermisch, J., (1989), "A Microeconomic Analysis of the Timing of Births", *European Economic Review*, vol. 33, pp. 737-760.

Cigno, Alessandro., Ermisch, John, (1989), "A Microeconomic Analysis of the Timing of Births". *European Economic Review*, vol 33, s 737-760.

Cliff, A. D. and J. K. Ord (1973). *Spatial Autocorrelation.* Pion Limited, London.

Cliff, A. D. and J. K. Ord (1981). *Spatial Processes, Models and Applications.* Pion Limited.

Florax, R. J. G.M., H. Folmer., and S. J. Rey (2003). Specification Searches in Spatial Econometrics: The Relevance of Hendry's Econometric Methodology. *Regional Science and Urban Economic* 33, 557-579.

Gauthier, A. H., Hatzius, J., (1997), "Family Benefits and Fertility: An Econometric Analysis", *"Population Studies"*, vol. 51, pp. 295-306.

Gustafsson, S., Kenjoh, E., Wetzels, C., (2001), "The Role of Education in Postponement of Maternity in Britain, Germany, the Netherlands and Sweden", *Working Paper*, Amsterdam University, Dep. of Economics.

Heckman, James, J., Willis, Robert, J., (1976), "Estimation of a Stochastic Model of Reproduction: An Econometric Approach". *Household Production and Consumption*, Ed., Terleckyj, Nestor, E., Columbia University Press, National Bureau of Economic Research, New York, s 99-138.

Hoem, B., (2000), "Entry into Motherhood in Sweden: the Influence of Economic Factors on the Rise and Fall in Fertility, 1986-1997. "*Demographic Research*", vol. 2, art. 4.

Hoem, Britta, Hoem, Jan, M., (1996), "Sweden´s Family Policies and Roller-Coaster Fertility". *Working Paper*, Stockholm.

Hoem, Britta, Hoem, Jan, M., (1997), "Fertility trends in Sweden up to 1996". *Working Paper*. Stockholm.

Hoem, Jan, M., (1993), "Public Policy as the Fuel of Fertility: Effects of a Policy Reform on the Pace of Childbearing in Sweden in the 1980s". *Acta Sociologica*, vol 36, s 19-31.

James, Harvey, S., Jr., (1996), ”The Impact of Female Employment on the Likelihood and Timing of Second and Higher Order Pregnancies”. *Working Paper*, Department of Economics, University of Hartford, England.

Lee, David, Kuo, Chuen., Gan, Chin, Lee., (1989), ”An Economic Analysis of Fertility, Market Participation and Marriage Behaviour in Recent Japan”. *Applied Economics*, vol 21, s 59-68.

Löfström, Å., (2000), “Att sätta barn till världen – Födelsetal och arbetsmarknad, en förstudie”, *Umeå Economic Studies*.

Löfström, Å., Westerberg, W., (2002), Economic Fluctuations and Fertility Swings in Sweden, 1965-1998, Umeå Economic Studies.

Leibenstein, Harvey., (1962), *”Economic Backwardness and Economic Growth”*. John Wiley & Sons, Inc, USA.

Moran, P. (1948). The interpretation on statistical maps. *Journal of the Royal Statistical Society B* 10, 243-251.

Mincer, Jacob., (1963), ”Market Prices, Opportunity Costs, and Income Effects”. *Measuremen in Economics: Studies in Mathematical Economics and Econometrics in Memory of Yehud Grunfeld*, ed. Carl Christ et al, Stanford, s 67-82.

Moffitt, Robert, (1984), ”Profiles of Fertility, Labour Supply and Wages of Married Women: A Complete Life-Cycle Model”. *Review of Economic Studies*, s 263-278.

Naz, G., (2000), "Determinants of Fertility in Norway", Working Papers in Economics, Department of Economics, Bergen University, Norway.

Sundström, Marianne., Stafford, Frank, P., (1992), ”Female Labour Force Participation, Fertility and Public Policy in Sweden”. *European Journal of Population*, vol 8, s 199-215.

Tasiran, Ali, Cevat, (1993), ”Wage and Income Effects on the Timing and Spacing of Birth in Sweden and the United States”. *Ekonomiska Studier*, nr 35, Handelshögskolan vid Göteborgs Universitet.

Tobler, W.R. (1979). Cellular geography. In: Philosphy and Geography, edited by S. Gale and G.Olsson, 379-386. Dordrecht, The Netherlands. Reidel

Wang, W., Famoye, F., (1997), "Modelling Household Fertility Decisions with a Generalized Poisson Regression, *”Journal of Population Economics”*, vol. 10, pp. 273-283.

Wilkinson, Maurice., (1973), ”An Econometric Analysis of Fertility in Sweden 1870-1965”. *Econometrica*, vol 41, s 633-642.

Winegarden, C., R., (1984), ”Women´s Fertility, Market Work and Marital Status: A Test of the New Household Economics with International Data”. *Economica*, vol 51, s 447-456.

Wolpin, Kenneth, I., (1984), ”An Estimable Dynamic Stichastic Model of Fertility and Child Mortality”. *Journal of Political Economy*, vol 92, s 852-875.