# Hedonic Prices, Economic Growth, and Spatial Dependence

by

Krister Sandberg

## Acknowledgements

It is now time to show how grateful I am to those that in one way or another have helped me get to this point were I am today.

First of all I would like to give a big hand to my supervisor Lars Westin, who has been there for many years - a man with firm support, thousand ideas, and a great sense of humor. Come to think of if it, I have to thank his wife Kristin too. Why? Well, when finalizing my MSc paper I was invited to their home for some final adjustments. She then confronted me with a simple and, I suppose, harmless question - so are you applying for graduate school? Not knowing exactly what it was at the time, or what to answer, I panicked and said yes! Thanks.

Jörgen Johansson, my co-author, deserves some recognition too. It was he that pushed me into the research direction in conjunction with the first paper. He was also a great help during the first year of the graduate studies. Robert Sörensson deserves some recognition for sharing a room with me for the last couple of years (actually some more). It has been mostly fun. Patrik Asplund used to quote Creedence Clearwater Revival "Someday Never Comes". For once you were wrong! Other colleagues that made life easy during this time were Thomas Westerberg, Mattias Ankarhem, Ulf Jonsson, Anna Norin, and Niklas Nordman. I would also like to thank the rest of the people at CERUM and at the Department of Economics, past and present. I have not forgotten my non-work related friends either, thank you all.

During my visits, first to IIASA in Vienna and later to UC Berkeley I had the fortune to get to know some nice people. Maarit, Rud, Monica, James, Ann Kristin, Catherine, Johannes, Keishiro, Jonas, Thomas, and all other unmentioned but not forgotten really know how to combine work with alternative activities. Life would truly be boring without you all. Thanks also to the staff at both places.

To conclude, I would also like to take this opportunity to thank my parents Kjell and Eva, siblings Kristina and Kristian, and grandparents Erik, Nanny, Nils, and Astrid. I could not have done it without you.

Dixi et salvavi animam meam

Umeå in April 2004
Krister Sandberg

**Abstract**

This thesis consists of three papers on econometric modeling of spatial dependence. The awareness of interactions between actors is fundamental for understanding property markets as well as the growth of regions. In both cases, neighbors and neighboring markets may stimulate or hamper growth of values. From a modeling point of view, these interdependencies calls for spatial econometric models. In the thesis we introduce such methods in the analysis of regional property markets as well as in a comparative regional growth analysis.

[1] In the first paper, we estimate hedonic prices in the market for co-operative flats in the city of Umeå, Sweden, during 1998 and 1999. Structural, neighborhood, and accessibility characteristics are used as attributes in the hedonic price function. Important attributes were the rent, floor space, age, and population density. Two attractive nodes, although with different characteristics, were found. Thus there are signs supporting the view that Umeå has developed into a multi-nodal structure for property values. SAR-GM estimation was used due to signs of spatial error dependence.

[2] In the second paper, hedonic prices for single-family homes in two Swedish counties are estimated for two years. Parameter estimates are compared and changes in space and time analyzed. Spatial lag dependence is found to influence the results. Hence, four independent variables are lagged with a spatial weights matrix. Additional spatial error dependence is treated by SAR-GM estimation. Structural, neighborhood, and accessibility characteristics are used as attributes. The regional price pattern and its changes over time, is illustrated and identified with GIS maps. Proximity to the two county capitals, as well as the other municipality centers, influence property prices positively. This is also noticable over time, where values have risen for homes located near major population centers and those which have water provided by the municipality. Values are in addition largely a function of the quality of each home.

[3] The third paper examines the provincial pattern of growth in China during the period 1985–2000, testing the hypothesis that provinces with similar growth rates are more spatially clustered than would be expected by chance. The provincial economic growth is explained by the distribution of industrial enterprises, foreign direct investment, infrastructure, and governmental preferential policies. The neoclassical hypothesis of convergence is also tested. Indications of unconditional convergence does occur during the periods 1985–2000 and 1985–1990. In addition, conditional convergence is found during the sub-period 1990–1995. Evidence of spatial dependence between adjacent provinces has also been established, and in the econometric part, solved by a spatial lag, or alternatively a spatial error term, in the growth equation.

***Keywords* :** Hedonic prices, Spatial econometrics, Co-operative flats, Spatial dependence, SAR-GM, Single-family homes, Heterogeneity, China, Convergence, Provincial economic growth.
***Classification*[*JEL*] :** D46, D61, R11, R12, R20, R21, O18

The thesis consists of a summary and the following papers.

[I] Spatial Autoregressive Generalized Moment Estimation of Hedonic Prices for Co-operative Flats (Co-authored with Jörgen Johansson)

[2] On Space-Time Changes of Hedonic Prices for Single-Family Homes

[3] Growth of GRP in Chinese Provinces: A test for Spatial Spillovers

# Hedonic Prices, Economic Growth, and Spatial Dependence

## Background and Summary

Krister Sandberg

## 1 Introduction

Actors in most markets depend on collaboration with, competition from, and the decisions made by other actors. In the housing market, it is obvious that appearance and quality of individual homes is important for value creation and price determination. Market prices of homes in the neighborhood, prices that indicate accessibility and status of the area may, in addition, have an impact on individual property prices. Property prices are also closely connected with the overall economic development in the region of location.

On the regional level, wealth and income growth depends on the dynamics within each region, e.g. how the market for homes operates. Additionally, regional growth may be dependent on the development in other regions. As in the property market, this dependence is a function of the type and strength of the interaction between markets in the different regions. Hence, in both the property market, and at the regional level, one may assume that proximity to other markets contributes to the explanation of values and growth.

When economic behavior is modeled, spatial dependence of this type calls for spatial econometric methods. In this thesis, the economic behavior of spatially located actors are analyzed with data from two regional property markets in Sweden and from data on regional economic growth in China.

The first two papers in the thesis deals with hedonic prices for two important categories of housing in the Swedish real estate market. The first paper examines the market for co-operative flats in the city of Umeå. In the second paper, hedonic prices for single-family homes in the two Swedish counties of Västerbotten and Västernorrland are estimated for the years 1994 and 1999. Space-time changes are illustrated and analyzed. The third paper is concerned with provincial (regional) economic growth in China during the fifteen-year period between 1985 and 2000.

The rest of this summary is devoted to an overview of the issues and previous literature of relevance for the three papers. Section 2 provides an introduction to spatial econometrics followed by a section on hedonic prices related to the treatment of co-operative flats and single-family homes. Regional economic growth theory and the regional situation in China is treated in Section 4. Summaries of the included papers are given in the concluding section.

## 2 Spatial Data and Spatial Econometrics

What is so special about spatial data analysis? In the property market especially, but also more generally in discussions on urban and regional growth, one saying often heard is "Location, Location, and Location is everything". Add to this Toblers first law of geography – everything depends on everything else, but closer things more so – and we begin to grasp what spatial analysis is all about. A central aspect of most economic markets is spatial interaction, externalities, spill-overs, copy-catting, etc. An implication is that in the same way as economic units are examined in time series for time dependence, they may be examined for dependence in geographical space. However, this comparison does not only contain similarities. A major difference between spatial and time dependencies is how the dependence operates. In a spatial context the dependence may, and usually does, operate in both directions. This makes it much harder to deal with. Initially, dependence due to measurable relations such as distance, barriers, and congestions are easily controlled for. However, there may still be signs of dependence due to omitted variables or tacit relations. One common simplification is first to assume equally strong dependence in both directions. A spatial weights matrix is thereafter constructed to proxy for these multiple dependencies between observations that are to be included in the estimation. There are various ways to construct this matrix, but the most common, so far, is a binary approach based on unit contiguity or a matrix based on some distance decay function.

Thoughts on the character of these spatial relationships in a statistical sense dates back at least 60 years. The Swedish statistician Bertil Matérn (1947) acknowledged the existence of spatial variation, or as he called it, typological variation to approximate overall forest growth. That is, trees at different locations are associated with different growth rates, or more generally, he thought that there might be an inherent systematic dependence between observations that could not be explained by traditional variables.

Exploratory 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 objects of similar values are more clustered than by pure chance. At our disposal are a couple of global tests for spatial autocorrelation, such as Moran's I and Geary's C (Moran 1948; Geary 1954; Cliff and Ord 1973, 1981). 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. No specific information is given about the prevailing pattern. When this is of interest, local tests are used instead, such as those suggested by Getis and Ord (1992), Ord and Getis (1995, 2001), and Anselin (1995).

The next step is usually to further examine and solve for spatial dependence in a regression analysis. Two kinds of spatial dependence are commonly assumed to potentially contaminate the analysis. The first arises when, for instance, prices of adjacent observations move together due to common or correlated unobservable variables, i.e. lack of stochastic independence between observations. This dependence leads to inefficient estimates if left unsolved. The problem is discussed at length in Cliff and Ord (1972, 1973). A partition of the error term into two parts, together with a spatial weights matrix, solves this spatial dependence problem. The model is known as the Spatial Error Model.

The second and more serious problem of spatial dependence is present when spatial correlation in the dependent variable between observations exists. An example of this is how the growth rate in one region is influenced by the growth rates in nearby regions and vice versa. Such dependence leads to both biased and inefficient estimates. For a further discussion of this problem see Anselin (1988). This problem may be solved for in various ways. The most common is to include the dependent variable of the other observations on the right hand side of the equation lagged by a spatial weights matrix. This model is known as the Spatial Lag Model.

An alternative to the Spatial Lag Model is the Spatial Cross-Regressive Model suggested by Florax and Folmer (1992), where instead spatially lagged independent variables of the adjacent observations are included in order to avoid otherwise induced heteroscedasticity when the dependence is of a more local variety. As usual, additional problems during estimation, such as heterogeneity and heteroscedasticity may occur. These problems are solved by standard econometric methods.

The classical estimation routine towards a proper model specification under the potential influence of spatial dependence is, for instance, given in Florax et al. (2003). The initial model is estimated by means of OLS. The residuals are then used to test the hypothesis of no spatial dependence caused by an omitted spatial lag or by spatially autoregressive errors by use of two Lagrange Multiplier tests (the LM-lag test and the LM-error test), e.g., Anselin (1988) and Burridge (1980). When the hypothesis cannot be rejected (no spatial dependence is at hand) meaning that the results from the OLS may be used. However, in the event that the hypothesis is, by both tests rejected, a new model should be estimated. The proper model is indicated by the most significant LM test. In case that only the LM-lag test is significant, the next step would be to estimate a Spatial Lag Model, or, consequently, a Spatial Error Model if the opposite is indicated.

# 3 The Swedish Housing Market and its Hedonic Prices

## 3.1 Bid rent and Hedonic Prices

A comprehensive exposition of modern urban economic theory is provided by e.g. Fujita (1989). Fujita observes that land is a commodity like any other, apart from that it is completely immobile. A parcel of land may differ in size but most importantly it is associated with a unique location in geographical space. These two characteristics of land imply strong non-convexities in consumers' preferences, and concave household indifference curves for distance and parcel size. In urban economic theory, this is avoided by the following two assumptions.

Each household is assumed to choose one and only one location. Thus, the household consumption space may be defined separately as consumption of land at that location and consumption of a composite good.

The households of each type are assumed to be so large that they may be represented in terms of a density function throughout the monocentric city. Not only does this assumption solve the problem above, it also implies that the bid rent function approach is applicable for the determination of household equilibrium location.

The bid rent $\gamma(d, u)$ is defined as the maximum rent per unit of land that the household is able to pay for residing at distance $d$ from the central business district (CBD) in a monocentric city given a fixed utility level $u$. In other words, the bid rent transforms consumption space indifference curves into corresponding indifference curves in urban space (bid rent curves) with the dimensions of location and land rent.

The bid rent function approach, first introduced by von Thünen (1826) in his agricultural land use model, plays a significant role in the first two papers of the thesis. The application for urban land use was later developed by Alonso (1964) and Solow (1973).

In order to connect the theory of land use with a theory of attribute valuation of individual homes, we use the theory of hedonic prices. Thus, heterogeneous goods may be seen as bundles of valued attributes that match the household utility function as discussed by e.g., Haas (1922), Lancaster (1966), and Rosen (1974). In other words, it is assumed that the buyer implicitly reveals his or her preferences for attributes through the price paid. Thus, given the assumption that the highest bidder purchases each home, the market prices yield the outer envelope of each attribute valuation by all households in the market. This optimal choice is characterized by equality between the slope of the bid rent and the hedonic price, with respect to each attribute.

### 3.2 Supply and Demand for Housing

The interaction between supply and demand for housing in the regional property market obviously is important for the determination of the hedonic prices related to the overall attraction of a property. But transaction prices also provide detailed information on the attractiveness of local areas, individually designed homes and their standards. If the supply of property in all those aspects don't match the utility driven demand for housing by consumers or the profit driven demand for nonresidential space by firms, the value of the regional property and the attractiveness of the city-region is put under pressure by competition from other cities. In DiPasquale and Wheaton (1996) the dynamics of supply and demand for property is to some extent modeled at an aggregated level. However, since the actual match making takes place at the level of individual properties and since this is the level where many of the hedonic prices still are determined, it may be appropriate to briefly sketch the process here.

The seller is assumed to release a home on the market when a potential buyer is thought to exist. If not, the seller must improve the object before advertising it, or withdraw from the market and wait for a future increase in demand. The buyer, given his preferences related to the characteristics that fulfill his utility, searches for an attractive home restricted by his budget constraint. When an object is found to be within both the choice set and the budget constraint, negotiations may begin. One may observe here that the choice set as such may be more or less dynamic as the buyer modifies his set in the process of object evaluation.

Thus, hedonic prices are market determined through buyer/seller negotiations. In these negotiations the seller tries to maximize the price of the object subject to a time constraint given, e.g. by the time when his new residence is available or changes in tax legislations. The buyer naturally tries to reduce the price given his experience on possible bids from competing buyers. If an agreement is not met, the buyer must either modify his choice set, wait for another object, or completely withdraw from the process. Although interesting, the dynamics of the negotiations or auctions between buyers and sellers and thus which objects in the property market that actually are transacted is not considered in the following.

Given that we currently only have transaction data at hand, changes in aggregate regional supply of objects with given attributes are neither possible to control for in a dynamic analysis. Nor is it possible to fine-tune our models with respect to all neighborhood attributes since we lack information about the stock of nontransacted objects and their attributes.

The major housing categories in Sweden are co-operative flats, single-family homes, and flats with right of tenancy. Since market prices may only be observed for co-operative flats and single-family homes, this thesis is focused on these markets. Brief descriptions are given below.

### 3.3 The Role of Co-operative Flats in the Swedish Housing Market

Contrary to the international housing market, co-operative flats are fairly common in Sweden. According to Statistics Sweden (2003a; 2003b) about 15% of the Swedish population resides in co-operative flats, and about 70,000 flats were transacted in the year 1999. Of those, 2,000 flats, or 2.8%, were sold in the county of Västerbotten and about half of those were sold in the municipality of Umeå. The average price for a flat was 170,500 SEK in Umeå, 110,000 in the county of Västerbotten, and 284,500 in Sweden, (Statistics Sweden, 2001).

In the first paper in this thesis our sample consists of 194 flats, or 18% of those sold in the municipality of Umeå. Since our sample is limited geographically to flats sold in the actual city of Umeå, the sample percentage is somewhat higher. The average price in our sample is 145,000 SEK. This low average transaction price may come as a surprise, but may be explained by data selection. Our sample consists of observations from many areas in order to get as complete a geographic picture as possible of the market for co-operative flats in Umeå. Most observations are located at some distance from the city center, while most transacted flats in the high price segment are actually found in the immediate vicinity of the city center.

Since most co-operative flats are located in a similar type of multifamily housing as flats with right of tenancy, it is at this point appropriate to distinguish between these two categories in the Swedish housing market. The most obvious difference is the kind of ownership they represent. The co-operative flat is coupled with a membership and a share in a housing co-operative; the legal owner of the building. This provides the owner of the flat an indirect ownership of the building, while a right of tenancy does not represent any ownership at all. This share in the housing co-operative has an economic value and may, together with the flat, be sold on the market.

The owners of co-operative flats also have more rights and responsibilities than renters. They have, for instance, wider possibilities to change the quality of the flat or alter its appearance. Ownership also makes it possible to influence the maintenance of the building, and thus the rent per month. The real estate tax is paid by the co-operative, thus indirect by the members. For tax purposes a co-operative flat is considered to be an asset, so members must additionally pay wealth tax for it. On the other hand, he or she may also make tax deduction for the interest on housing mortgages. In a flat rented by tenure, all these responsibilities and benefits belong to the housing company owners.

Different strategies may be chosen to finance the co-operative when it is started. Dependent of the chosen strategy, risk and capital costs are distributed differently between the co-operative and the individual owners. One alternative is for the co-operative to assume a large portion of the loans and sell the flats to its future members with a high rent per month charge. Another alternative is to let the members themselves raise the money and thereby directly finance the co-operative loans. These flats are then sold with a relatively high price while the rent per month is kept at a lower level.

## 3.4 Single-Family Homes as the Major Housing Category

The second paper deals with sales of Single-family homes. This is the most important category of housing in Sweden. About 57% of the Swedish population resides in single family homes, (Statistics Sweden, 2003a). In the year 1994 the total stock of single-family homes in Sweden were 1,937,000. Five years later the stock was 1,957,000, a modest increase of 1%. The three city regions, Stockholm, Göteborg, and Malmö, contributed half of that increase, (Statistics Sweden, 2004). The average price rose by 29% in Västerbotten, 9% in Västernorrland, and 49% in Sweden between the years 1990 and 2002. During the study period in the second paper, 1994–1999, the increase was 15% in Västerbotten, a mere 5% in Västernorrland, and 27% in Sweden, (Statistics Sweden (2003b). However, in the market for homes, the regional price indicators are very crude since the price of individual homes may vary in a broad interval around the average, a topic that will be analyzed further in the paper below.

## 4 Regional Economic Growth in China

Knowledge about which factors that determines national and regional GDP growth is important in order to determine how regions grow, how the regional economy works, and to understand how regions interact with each other. One hypothesis often tested in studies of regional economic growth is the convergence hypothesis. That is, the lower the initial level of per capita income a region has, the higher should its growth be. This is predicted by the neoclassical growth theory, e.g., Ramsey (1928), Kuznets (1955), Solow (1956), Swan(1956), Cass (1965), and Koopmans (1965). The prediction of convergence is derived from the assumption of diminishing returns to capital. When the capital/labor ratio within a region is below its long run value, that region tends to have a higher rate of return than other regions. Hence, if all regions were essentially the same, with the exception of their initial capital/labor ratios, the convergence would be unconditional so that poorer regions would eventually "catch up" with the wealthier regions. However, if the regions are different in various aspects the convergence would instead be conditional and the regions would strive towards their own steady state growth.

The hypothesis of convergence has, however, been criticized in many studies of nations in favor of the endogenous growth theory; where the growth is instead driven by the accumulation of human and physical capital. Standard references are Romer (1986) and Lucas (1988).

The convergence and growth discussion provides us with a general background to the analysis of the Chinese development in paper three of the thesis. In many aspects, China is far from a homogeneous country, a fact that only makes it more interesting to study. The Chinese history and a rough characteristic of its provinces may serve as an introduction to the third paper.

China was forced to cede Hong Kong to British rule and to open up their ports to foreign control after the Opium war loss in the 1840's. China kept on waging war against its enemies but lost more and more of its independence to foreign countries. In 1911, revolution broke out and the Qing dynasty collapsed. Different warlords ruled the country until 1927 when China was united followed by a brief period of relative peace. Manchuria was seized by Japan in 1931, followed by the war in 1937 that eventually transformed into the well-known civil war, which ended in 1949 with communist victory.

The Japanese had during their occupation invested in heavy industry mainly for export to Japan so by the end of the war, the majority of China's industrial capacity was in Manchuria. Apart from that, China was mainly a country inhabited by peasants. With Manchuria as a base, and with help from the Soviet Union, China started its industrial transformation. The new leaders established a strong government (but contrary to the Soviet model acknowledged that regions differ in various ways, and thus allowed different regional initiatives) and progressed towards a socialistic industrial complex. Planners, however, neglected labor-intensive sectors suitable for the large population, and instead poured resources into capital-intensive factories to produce metals, machinery, and chemicals. During this period most of the GDP growth may be attributed to capital input. The "big pushes" that additionally drew resources from the labor-intensive industry to the capital intensive seems only to have made things worse.

In 1978 China began to reform their command economy and have since then gradually created the framework of a socialistic market economy. The major forces behind this change was partly the fear of not being able to support its growing population (memories of the great famine in the early 1960's were still vivid), and partly the threat from their neighbors' new economic success. Positive effects were visible on economic growth almost immediately. Loop-holes and other possibilities made it possible for individuals with initiative to take advantage of the new situation and make handsome profits. When inflation and corruption grew, people began to protest, which culminated in the massacre at Tiananmen Square in 1989. This led to a slow down of the reform process. International commentators, as well as the Chinese community, had strong doubts whether the reforms would continue or not. Deng Xiaoping gave a clear sign in 1992 during his "Southern Tour" to the most successful provinces and made sure that the reform process would continue. Not only did the reforms continue, they also changed direction. This time more effort was inserted to level the playing field for the market actors. Other reforms included fiscal reforms and a reconstruction of the state owned enterprise sector. A recent and important step was the World Trade Organization (WTO) membership in December 2001.


Figure 1: The Chinese Provinces.

The three metropolises, Beijing, Shanghai, and Tianjin, are highly industrialized and also the provinces with the highest Gross Regional Product (GRP) per capita. The coastal provinces in the southeast, previously a poor and isolated region with a rugged coastline and an inhospitable climate, have since the reforms started experienced a rapid growth in GRP per capita and are now among the richest provinces in the country. These provinces have a special status due to the preferential policies levied on them by the government and are generally considered to be the engines of growth in the Chinese economy.

In the northeast, we find the old industrialized center, Manchuria. This area had, during the pre-reform era, the highest GRP per capita level. Even though they have not experienced such a rapid growth as the coastal provinces in the southeast, the GRP per capita is still among the highest in China.

The central provinces, between the rivers Yellow and Yangtze, have a high population density and are well suited for agriculture. The southwestern provinces are also, from a climatic perspective, suited for agriculture but have limited access due to the mountains. These provinces have, in general, had a low annual GRP per capita growth since the start of the reforms.

In the northwestern part of China we find quite isolated provinces like Tibet, Xinjiang, and Qinghai, characterized by high elevation and a low transport infrastructure capacity.

In the third paper of the thesis, the above mentioned aspects are considered when the Chinese provincial growth and its pattern are explored and explained. As is the case in the first two papers, problems of spatial dependence play a major role.

## 5 Summaries of the Papers

[1] In the paper "Spatial Autoregressive Generalized Moment Estimation of Hedonic Prices for Co-operative Flats" the important characteristics and their magnitude for the price determination of Co-operative flats in the city of Umeå during the years 1998 and 1999 are estimated. Due to spatial dependence, the model in the final estimation is adjusted by a spatially lagged error term.

The results show that central business district and university accessibility influences the prices in a positive direction. Other variables that are important for the price are age, rent per month, and square meter area of the flats. Included are also the examination of neighborhood importance through the population density, the share of single-family homes, and the rate of turnover in the co-operatives.

[2] The second paper "On Space-Time Changes of Hedonic Prices for Single-Family Homes" studies the prices and important characteristics in the real estate market for single family homes in the counties of Västerbotten and Västernorrland for two years, 1994 and 1999. Included are all sales of single-family homes for both years. Apart from the usual object specific characteristics such as age and floor size, a major effort was made to deal with the spatial aspects of this market. This is not only taken care of by a gravity approach to capture the importance of accessibility and size of population nodes simultaneously, but also by the introduction of a spatial weighs matrix for four of the characteristics based on distance between the observations. To investigate and control for heterogeneity, these variables were also divided into four groups based on municipality location. Group 1 and 2 consists of homes in the municipalities of Umeå and Sundsvall. The third group consists of homes in the other coastal municipalities, and group 4 consists of homes in the inland municipalities. The final model includes all these aspects mentioned above with the addition of a spatial error component. The paper concludes with an analysis of the temporal adjustment between the two years. It is found that prices have increased most for homes with high accessibility to population, water provided by the municipality, and high housing quality. The predicted prices and their spatial distribution are also illustrated in GIS maps.

[3] The paper "Growth of GRP in Chines Provinces: A test for Spatial Spillovers" concludes the thesis. Two aspects of provincial

economic growth in China during the period 1985–2000 are investigated. Part one consists of a search for global spatial autocorrelation as well as local spatial association (hot/cold spots) in an exploratory data analysis. Clusters of provinces with high growth, especially in the coastal region in the southeast, and provinces with low growth in the center and western parts of China were found.

This is in the second part followed by a regression analysis. The hypothesis of unconditional convergence could not be rejected for the periods 1985–2000 and 1985–1990, while positive spatial dependence was found between the provinces for the periods 1990–1995 and 1985–2000.

In the conditional convergence regressions, preferential policy, enterprise structure, transport capacity, and foreign direct investments, were all important variables used in order to explain provincial economic growth. Spatial dependence, however, was not that significant. Thus, the provinces are in fact relatively independent of each other. One possible explanation for this could be that the aggregation at province level is too high, with a sample of only 30 observations so that differences in growth and spillover-effects actually takes place within each province, hinted at by e.g. Yao and Liu (1998) and Oi (1999). A similar investigation at the more detailed county level might have left us with different results. On the other hand, major problems of non existing data restricted us in this case. Even if data would be available there would be questions about data quality.

## References


[1] Alonso, W. (1964). *Location and Land Use*. Harvard University Press, Cambridge.

[2] Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Kluwer Academic, Dordrecht.

[3] Anselin, L. (1995). Local Indicators of Spatial Association–LISA. *Geographical Analysis* 27 (2), 93–115.

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

[5] Cass, D. (1965). Optimum Growth in an Aggregative Model of Capital Accumulation. *Review of Economic Studies* 32, 233–240.

[6] Cliff, A., and K. Ord (1972). Testing for Spatial Autocorrelation Among Regression Residuals. *Geographical Analysis* 4, 267–284.

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

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

[9] DiPasquale, D., and W. C. Wheaton (1996). *Urban Economics and Real Estate Markets*. Prentice Hall, Englewood Cliffs, NJ.

[10] Florax, R. J. G. M., and H. Folmer (1992). Specification and Estimation of of Spatial Linear Regression Models: Monte Carlo Evaluation of Pre-test Estimators. *Regional Science and Urban Economics* 22, 405–432.

[11] Florax, R. J. G. M., H. Folmer., and S. J. Rey (2003). Specification Searches in Spatial Econometrics: The Relevance of Hendry's Econometric Methodlogy. *Regional Science and Urban Economics* 33, 557–579.

[12] Fujita, M. (1989). *Urban Economic Theory. Land Use and City Size*. Cambridge University Press, Cambridge.

[13] Geary, R. (1954). The contiguity ratio and statistical mapping. *The Incorporated Statistician* 5, 115–145.

[14] Getis, A., and Ord, J. K. (1992). The Analysis of Spatial Association by Use of Distance Statistics. *Geographical Analysis*, 24, 189–206.

[15] Haas, G. C. (1922). *A Statistical Analysis of Farm Sales in Blue Earth County, Minnesota, as a Basis for Farm Land Appraisal.* Master Thesis, The University of Minnesota.

[16] Koopmans, T. C. (1965). *On the Concept of Optimal Economic Growth. The Econometric Approach to Development Planning.* North Holland, Amsterdam.

[17] Kuznets, S. (1955). Economic Growth and Income Inequality. *American Economic Review* 45 (1), 1–28.

[18] Lancaster, K. (1966). A new approach to consumer theory. *Journal of Political Economy* 74, 132–157.

[19] Lucas, R. E. (1988) On the Mechanics of Economic Development. *Journal of Monetary Economics* 22, 3–42.

[20] Matérn, B. (1947). Metoder att uppskatta noggranheten vid linje- och provytetaxering. In: Näslund, M. (ed) *Meddelanden från Statens Skogsforskningsinstitut* 36, 1–117.

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

[22] Oi, J. C. (1999). *Rural China Takes Off. Institutional Foundations of Economic Reforms*. University of California Press, Berkeley and Los Angeles.

[23] Ord, J. K., and A. Getis (1995). Local Spatial Autocorrelation Statistics: Distributional Issues and an Application. *Geographical Analysis* 27 (4), 286–306.

[24] Ord, J. K., and A. Getis (2001). Testing for Local Spatial Autocorrelation in the Presence of Global Autocorrelation. *Journal of Regional Science* 41 (3), 411–432.

[25] Ramsey, F. (1928). A Mathematical Theory of Saving. *Economic Journal* 38, 543–559.

[26] Romer, P. M. (1986). Increasing Returns and Long-Run Growth. *Journal of Political Economy* 94, 1002-1037.

[27] Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy*, 34–55.

[28] Solow, R. M. (1956). A Contribution to the Theory of Economic Growth. *Quarterly Journal of Economics* 70, 65–94.

[29] Solow, R. M. (1973). On Equilibrium Models of Urban Locations. In: Parkin, J. M. (ed) *Essays in Modern Economics* 2–16. Longman, London.

[30] Statistics Sweden (2001). Pressmeddelande från SCB. 2001-05-14. Nr 2001:127. (In Swedish)

[31] Statistics Sweden (2003a). Levnadsnivåindikatorer i ULF efter region, kön, ålder och indikator. Tvåårsgenomsnitt 1988-1989-2000-2001. URL: www.scb.se.

[32] Statistics Sweden (2003b). Sveriges Offentliga Statistic. Statistiska Meddelanden, BO 41 SM 0301.

[33] Statistics Sweden (2004). Sweden's Statistical Databases. The dwelling stock, projections by the metropolitan areas and type of building. 1990-2002. URL: www.scb.se

[34] Swan, T. V. (1956). Economic Growth and Capital Accumulation. *Economic Record* 32, 334–361.

[35] von Thünen, J. H. (1826). *Der Isolierte Staat in Beziehung auf Landtschaft und Nationallkonomie*. Hamburg. (English translation by C. M. Wartenberg (1966), *von Thünen's Isolated State*. Pergamon Press, Oxford.

[36] Yao, S., and J. Liu (1998). Economic Reforms and Regional Segmentation in Rural China. *Regional Studies* 32 (8), 735–746.

# Spatial Autoregressive Generalized Moments Estimation of Hedonic Prices for Co-operative Flats

Krister Sandberg and Jörgen Johansson*
Department of Economics and CERUM
Umeå University
SE-901 87 Umeå
Sweden
krister.sandberg@econ.umu.se

**Abstract**

In this paper we estimate hedonic prices in the market for co-operative flats in the city of Umeå, Sweden, during 1998 and 1999. Structural, neighborhood, and accessibility characteristics are used as attributes in the hedonic price function. Important attributes were the rent, floor space, age, and population density. Two attractive nodes, although with different characteristics, were found. Thus there are signs supporting the view that Umeå has developed into a multi-nodal structure for property values. Problems of spatial dependence made Ordinary Least Squares estimation inappropriate. Instead Spatial Autoregressive Generalized Moments estimation was used.

***Keywords* :** SAR-GM, Hedonic prices, Co-operative flats, Spatial dependence, Heterogeneity
***Classification*[*JEL*] :** D46, D61, R20, R21

*The authors wishes to thank HSB and Riksbyggen in Umeå for providing the data. We would also like to thank Associated Professor Lars Westin, chairman Professor Arthur Getis and discussant Professor Leo J. G. Van Wissen at the 41st Congress of the European Regional Science Association in Zagreb, where the paper won the Epainos Prize 2001, for valuable comments. This paper was funded by the EU project Urban Design.

# 1 Introduction

The value of attributes of public and private real estate in a city develops through processes that involves actors on both the supply and the demand sides. If supply does not match the utility driven demand for housing by consumers or the profit driven demand for nonresidential space by firms, then the value of the property and the attractiveness of the city is placed under pressure by competition from other cities or other parts of the city. In the current movement of the economy towards a knowledge based society the competition for movable labor through supply of “attractiveness” has been even more emphasized. Clearly, the actors on the supply side are interested in how attributes associated with their real estate are valued on the demand side. Hedonic price theory is based on the assumption that attractiveness of a city, areas within the city, and individual residential and nonresidential units may be measured through the appraisal of the attributes associated with the supply.

With this as a general background, the purpose of this paper is to determine and quantify the appraisal of housing attributes on the market for co-operative flats on the basis of hedonic price theory under the influence of spatial dependence using a Spatial Autoregressive Generalized Moments (SAR-GM) estimator. This is the first study to use these methods in the appraisal of housing attributes of co-operative flats.

The theory of hedonic prices assumes that a good, in this case a co-operative flat, should not be seen as a good, but rather as a bundle of characteristics that match the household’s utility function, Lancaster (1966). It is assumed that the buyer implicitly reveals his preferences through the price paid for these attributes. Since each flat is purchased by the highest bid it is also assumed that prices represent the outer envelope of the valuation of the attributes market value.

Knowledge about characteristics and how they are valued on the Swedish real estate market is still quite fragmentized. This is particularly true for the co-operative flat market. Studies based on multifamily houses and co-operative flats are, compared with studies on single-family homes, very few. Among these are, Eriksson (1997) who studied how owners of multifamily houses value their attributes, Andersson (1997) who conducted a study of the Singapore market for condominium housing, and Werner (2000) who analyzed the co-operative flat architecture and its influence on prices. Also, the Swedish Association of Municipal Housing Companies (SABO) presented one of the first studies on the subject in a discussing format, SABO (1997). In 1999 the Swedish government initiated a study where the importance of distance to the CBD (Central Business District) on the rent per month for multifamily houses was investigated for eight municipalities, SOU (2000:33). Simultaneously, various consultants offer estimates of attribute values for segments of the market, although generally without a transparent methodological criteria.

In our case the analysis is applied on the city of Umeå, situated at the Baltic coast in the northern part of Sweden. Since the nearest city is located almost 100 kilometers away, border effects ought not to influence the valuation of housing in any part of the city. The city of Umeå has experienced a rapid population growth in the last 50 years and is now a medium sized city with a rank size among the twenty largest in Sweden with 104,000 inhabitants in the municipality and 137,000 in the city region, SOU (2000:87).

The largest and most important accumulation of work places in the city region is located in an area some distance away from the CBD. This area consists of two universities, a science park, and a regional hospital. The impact from this cluster of work places on the prices for co-operative flats has not previously been investigated.

In the next section, a brief introduction to the characteristics of co-operative flats is provided. The third section is devoted to the theory of hedonic prices and spatial dependence while the forth section describes the data for co-operative flats in Umeå. The empirical examination is outlined in the fifth section followed by conclusions in the final section of the paper.

## 2 Characteristics of Co-operative Flats

Co-operative flats are not very common internationally, but is a category of housing that is quite important in Sweden. In the late 1990's, about 15 percent of the Swedish population resided in co-operative flats, Statistics Sweden (2003). In this section, we provide a short presentation of some special features connected with these flats.

First of all, it is important to distinguish a co-operative flat from a rented flat with right of tenancy, since many of the attributes are quite similar. The most fundamental difference is the kind of ownership they represent. The co-operative flat is bought and sold on a competitive market. Through the purchase, the buyer also receives a membership and a share in the housing co-operative. This means that the buyer has an indirect ownership of the flat.

This increases the possibilities to alter the flat appearance, adjust the overall level of maintenance in the co-operative, and in the long run the determination of rent per month levels together with the other members in the co-operative. The real estate tax is paid by the co-operative (thus indirect by the members). Since a co-operative flat is considered an asset the owner must additionally pay wealth tax, but may in return deduct interest payments on his or her taxes. In a flat rented by tenure, all the later responsibilities belong to the owners of the housing company. The rent is also surrounded by regulations.

The financial strategy chosen when the co-operative is first established has implications for the future rent per month level. One alternative is that the co-operative assumes a large portion of the loans. The flats are then sold with an assigned high rent per month level. Another alternative is to let the members themselves borrow the money and thereby directly finance the flats. This means that the flats are sold at a relatively high price while the rent per month is kept at a low level. These different strategies distribute risk and capital costs differently between the co-operative and the buyers. Thus, the rent per month level works as a signal to the potential buyer about co-operative's financial situation. A low degree of debt ought to be mirrored by a low rent level.

## 3 The theory of hedonic prices and spatial dependence

Hedonic prices are, according to Lancaster (1966), defined as implicit prices of attributes and are revealed by observed prices on differentiated goods and the specific amounts of characteristics associated with them.

The concept of implicit or hedonic prices was first formalized by

Rosen (1974). A good, e.g. a co-operative flat, may according to the theory be described by $m$ characteristics. Each flat is then represented by the vector $\boldsymbol{z} = (z_1, \ldots, z_m)$ where element $z_i$ reflects the amount of the $i$th characteristic attached to the flat. A market clearing price function that equates supply and demand, driven by buyer and seller utility maximization, on the basis of their location and quantity decisions for this vector of characteristics, is defined as the hedonic price function $p(\boldsymbol{z}) = p(z_1, \ldots, z_m)$. Thus, the hedonic price function is market determined through buyer/seller transaction negotiations were the seller tries to maximize the price of the flat subject to a time constraint given by the time when an alternative residence is available.

The preferences of the household may be represented by the utility function:

$$U = u(\boldsymbol{z}, c; \boldsymbol{\alpha}) \tag{1}$$

Above, $\boldsymbol{z}$ is consumption of the co-operative flat, $c$ is consumption of a composite good, and $\boldsymbol{\alpha}$ is a vector of parameters that characterize the household preferences. The price that a household is willing to pay for the flat is derived from (1) as a function of the embedded characteristics, a given household income ($M$), and the achieved utility level. Together this comprises the household's bid rent function:

$$\gamma(\boldsymbol{z}, M, U; \boldsymbol{\alpha}) \tag{2}$$

and implicitly:

$$U = u(\boldsymbol{z}, M - \gamma; \boldsymbol{\alpha}) \tag{3}$$

The derivative of the bid rent function with respect to $z_i$, $\frac{\partial \gamma}{\partial z_i}$, gives the rate at which the consumer is willing to change it's expenditure on a co-operative flat when characteristic $i$ increases, other levels remaining constant.

The consumer chooses a co-operative flat with characteristic $\boldsymbol{z}$, and its consumption of the composite good $c$ by solving:

$$\max_{\substack{\boldsymbol{z}, y \\ \text{s.t. } M \geq p(\boldsymbol{z}) + y}} \quad u(\boldsymbol{z}, c; \boldsymbol{\alpha}) \tag{4}$$

The Lagrangian to (4) with the Lagrangian parameter $\vartheta$ is:

$$L = u(\boldsymbol{z}, c; \boldsymbol{\alpha}) + \vartheta \left[ p(\boldsymbol{z}) + c - M \right] \tag{5}$$

The first order conditions are:

$$\begin{aligned}\frac{\partial L}{\partial z_i} &= \frac{\partial u}{\partial z_i} + \vartheta \frac{\partial p}{\partial z_i} = 0 \ \forall i \\ \frac{\partial L}{\partial c} &= \frac{\partial u}{\partial c} + \vartheta = 0 \\ \frac{\partial L}{\partial \vartheta} &= p(\boldsymbol{z}) + c - M = 0\end{aligned}$$

Rearranging gives:

$$\frac{u_i}{u_c} = p_i \ \forall i \tag{6}$$

where $u_i = \frac{\partial u}{\partial z_i}$, $u_c = \frac{\partial u}{\partial c}$ , and $p_i = \frac{\partial p}{\partial z_i}$ the hedonic price of characteristic $i$.

A combination of the first order condition (6) and the implicit differentiation of (3) yields that the consumer's optimal choice of a flat is characterized by equality between the slope of the bid rent and the hedonic price with respect to each characteristic:

$$\frac{\partial \gamma}{\partial z_i} = \frac{\partial c}{\partial z_i} = p_i \ \forall i \tag{7}$$

Under the assumption of optimizing behavior, equation (7) indicates that the hedonic price for a characteristic provides local information about the consumer's preferences or willingness to pay for the attribute in the vicinity of the observed choice.

It is common in hedonic price studies to divide the vector $\boldsymbol{z}$ into structural ($\boldsymbol{s}$), neighborhood ($\boldsymbol{n}$), and accessibility ($\boldsymbol{a}$) attributes with $\boldsymbol{\omega}$, $\boldsymbol{\eta}$, and $\boldsymbol{\psi}$ as the corresponding parameter vectors. The hedonic price function of a general regression model is then formulated as:

$$p(\boldsymbol{z}) = f(\boldsymbol{s}, \boldsymbol{n}, \boldsymbol{a}; \boldsymbol{\omega}, \boldsymbol{\eta}, \boldsymbol{\psi}) + \boldsymbol{\varepsilon} \tag{8}$$

or expressed in linear vector form:

$$\boldsymbol{p} = \boldsymbol{Z}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{9}$$

where, $\boldsymbol{p}$ is a ($n$ by 1) vector of observations on the dependent variable, $\boldsymbol{Z}$ is a ($n$ by $k$) matrix of observations on the exogenous variables with $\boldsymbol{\beta}$ as the associated ($k$ by 1) vector of regression coefficients. $\boldsymbol{\varepsilon}$ is a ($n$ by 1) vector of random error terms.

In most hedonic price studies, as in the present one, the regression analysis is based on cross-sectional data that are prone to problems of spatial dependence. Spatial regression analysis must therefore be used to avoid potentially biased estimates and faulty inference due to violation of the basic assumptions in the classic regression analysis, Anselin (1988) and Anselin and Bera (1998). Two different cases of spatial dependence are considered in this paper.

The first case, Spatial Lag Dependence, relates to theoretically driven interactions between agents in geographical space, cf. Anselin (1988) and Can (1992). If neglected the results would be biased and inefficient. The spatial autoregressive lag model is formally expressed as:

$$\boldsymbol{p} = \rho \boldsymbol{W}\boldsymbol{p} + \boldsymbol{Z}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{10}$$

where $\boldsymbol{W}$ is a ($n$ by $n$) spatial weights matrix, with elements $w_{rs}$ that correspond to observation pair $r$ and $s$. $\boldsymbol{W}\boldsymbol{p}$ is the spatially lagged dependent variable that captures the average values of neighboring observations and $\rho$ is the spatial autoregressive coefficient.

In the second case, Spatial Error Dependence, the error terms show correlation with the error terms of observations located nearby, i.e. lack of stochastic independence between observations. This dependence is a nuisance dependence due to spillover effects caused by problems such as spatial scale, cf. Cliff and Ord (1972, 1973). Unsolved spatial dependence yields, in this case, inefficient parameter estimates. The spatial error dependence is incorporated in (9) via an autoregressive error term:

$$\begin{aligned} \boldsymbol{p} &= \boldsymbol{Z}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \\ \boldsymbol{\varepsilon} &= \lambda \boldsymbol{W}\boldsymbol{\varepsilon} + \boldsymbol{\xi} \end{aligned} \tag{11}$$

where $\boldsymbol{W}\boldsymbol{\varepsilon}$ is the spatial lagged error term, $\lambda$ is the autoregressive coefficient and $\boldsymbol{\xi}$ is an ($n$ by 1) vector of well-behaved error terms, $\boldsymbol{\xi} \backsim N(0, \sigma^2 \boldsymbol{I})$. The autoregressive coefficient is in both cases unknown and must therefore be estimated jointly with the regression coefficients.

# 4 Attributes of the observed co-operative flats in Umeå the years 1998–99

The data set consists of 194 observations on realized sales of co-operative flats in eleven co-operatives in the city of Umeå from late 1998 and 1999. The locations of the co-operatives are given in the city map in Figure 1. The two major housing associations in Umeå, HSB and Riksbyggen, have provided the data on individual flats and their approximate location.


Figure 1: The Umeå City Map 1999

Access to information about the individual flats is restricted due to confidentiality and thus restricts the choice of independent variables. The structural variables available for each flat are:

- Rent per month
- Floor area
- Number of rooms
- Age

The neighborhood attributes for each area are:

- Population density
- Single-family house density
- Rate of turnover in each co-operative

Three accessibility measures are used for each flat. They measure the accessibility to the:

- CBD
- University area
- Nearest of two major shopping centres

To enhance the understanding of the data, the city of Umeå, and the problem at hand, some descriptive statistics are presented in Table 1. The dependent variable in hedonic price models is usually the transaction price, or some transformation of it. However, the flats in our sample are more heterogeneous than single-family homes in the aspects of price, rent, and floor area. Hence, against convention, but in order to get a more scale neutral dependent variable, that facilitates estimation as well as comparisons, we instead use the normalized measure price per square meter, $p/\text{m}^2$, a frequently used measure at the Swedish real estate market.

Table 1: Descriptive Statistics for Co-operative Flats in Umeå 1998/99.

| Variable | Unit | Mean | St. dev | Min | Max |
|---|---|---|---|---|---|
| *PRICE* ($p$) | SEK | 145,000 | 151,889.9 | 1 | 1,195,000 |
| *AREA* | $\text{m}^2$ | 79.5 | 20.9 | 17.5 | 128 |
| *RENT* | SEK/month | 3,800 | 1,302.2 | 650 | 6,800 |
| *AGE* | year | 22.1 | 11.3 | 6 | 44 |
| *POPDENS* | # Inhab/1,000 $\text{m}^2$ | 4.8 | 3.9 | 0.12 | 9.4 |
| *SHDENS* | % Single f. homes | 60 | 40 | 0 | 100 |
| *TO* | # Sold/stock | 19.7 | 8.9 | 8 | 35 |
| *CBD* | m | 4,350 | 1,147.8 | 600 | 5,700 |
| *UNIV* | m | 3,300 | 2,126.1 | 1,100 | 8,200 |
| *SHOP* | m | 3,550 | 2,500.7 | 1,200 | 9,300 |

The first part, PRICE, is the realized transaction price with a range from 1 to 1,195,000 SEK[1]. The second part of the dependent variable is the variable AREA, which measures the floor area of each flat. The smallest flat in the sample has a floor area of 17.5 $m^2$ while the largest is 128 $m^2$. In the estimations, AREA will also be used as an independent structural variable, a part of a new variable rent per square meter, RENT/$m^2$.

The rent is thereby also normalized in order to improve the comparability between the flats. The variable RENT, rent per month paid to the co-operative, ranges from 650 up to 6,800 SEK per month. This gives the owner access to the private good (the actual flat) as well as the club good (the common property in the co-operative). The age variable AGE is defined in years. The most recent flat is only 6 years old while the oldest have been on the market for 44 years.

The neighborhood attribute population density, POPDENS, has a mean of 4.8 inhabitants per 1,000 $m^2$ for the area in which the co-operative is located. The maximum is nine inhabitants per 1,000 $m^2$, indicating that the multifamily houses in Umeå is dominated by four to five floor buildings. The density of single-family homes in the area nearby the co-operative, SHDENS, is the second neighborhood attribute. The variable has it's minimum at zero, since one area in the sample is without single-family houses. Both variables are computed from data provided by the municipality of Umeå and dates back to 1996.

The third neighborhood attribute is the turnover in the co-operatives for the year 1999. The average turnover, TO, in the sample is 19.7%. The youngest co-operative has the highest turnover, 35%. These flats are relatively cheap to buy, yet have high rent levels. As such, they are most similar to flats with the attributes of right of tenancy.

Accessibility in Umeå is quite good in all directions, and a fair approximation of accessibility is the traveled distance by car from each observation to the major nodes of attraction. Distances are calculated by use of the city map "Stadskartan", Lantmäteriet (1997). The routes used are those that are most likely chosen for travel between the co-operative and each node. The distance to CBD, the

[1] The co-operative flats that are sold for 1 SEK (6 observations) belong to a fairly new co-operative (built in 1993) that have comparatively high rents per month. The other flats sold in this particular co-operative are also sold at a "low" price; the maximum price was 30,000 SEK. Thus, the range is very small compared to the ranges in other co-operatives so inclusion of the six observations are not believed to be a problem.

variable *CBD*, from the flats are at minimum 600 meters and at most 5,700 meters. The university area is located at the east of the city and the maximum travel distance given in the variable, *UNIV,* to the campus area is 8,200 meters. The city of Umeå has two external shopping centres. The maximum distance to any of them, given in the variable *SHOP* is 9,300 meters.

Before we continue to the econometric portion of the paper, it is appropriate to discuss the expected sign of the parameters of the independent variables.

*POPDENS* is included to control for whether households find it attractive to live in a densely populated neighborhood. The city centre is one example of these areas, but there are also other parts of the city that have an equal or higher density, so the variable should not solely be considered as demand for "city life". This demand is taken care of through the CBD variable explained below. A very high density may also be negative for the attractiveness of the area. It is for this reason difficult to have an a priori hypothesis regarding the sign of the coefficient.

In a similar way *SHDENS* measure the attractiveness of areas with a high degree of single-family homes. This measure may be more strongly related with the price and thus have a positive sign. The three variables *POPDENS*, *CBD*, and *SHDENS* are all indirectly related to density. But contrary to the first variable, the two latter variables are more directly related to the attributes in the flat environment.

Finally, the turnover variable *TO* is defined as the percentage of the flats that has been sold, in each co-operative, during the year 1999. The interpretation of the variable is also twofold. It may be seen as a proxy for the liquidity of the asset. A higher liquidity would accordingly yield a higher price per square meter. The other way to interpret the variable is from the household stability point of view. A stable co-operative with long time owners may put more interest in maintenance and thereby increase the value of flats. The expectation is that low turnover/high stability yields a higher price/$m^2$.

The second group of characteristics is the accessibility measures. Since, due to confidentiality, the exact position of the observations within each block is unknown, all observations within a co-operative are given the same accessibility. We have chosen to include three accessibility measures, defined as the squared inverse distances to the CBD ($INCBD^2$), to the university area ($INUNIV^2$), and to the nearest of two external shopping centres[2] (*INSHOP*$^2$). If these nodes are attractive, then the price per square meter increases for locations in the vicinity of the nodes. We expect a positive sign for all three coefficients.

In Figure 2, the $p/\mathrm{m}^2$ is plotted against *RENT*$/\mathrm{m}^2$ where a dot represents a single observation. The correlation between $p/\mathrm{m}^2$ and *RENT*$/\mathrm{m}^2$ seems to be negative. This is in accordance with the presumtion that there exists a trade-off between price and rent in the co-operative flat market. That is, flats with a high rent are sold at a low price. Additionally, the correlation seems to exhibit convexity. To capture this, the inverse of *RENT*$/\mathrm{m}^2$, that is ($\mathrm{m}^2$/*RENT*), is used as an independent variable.

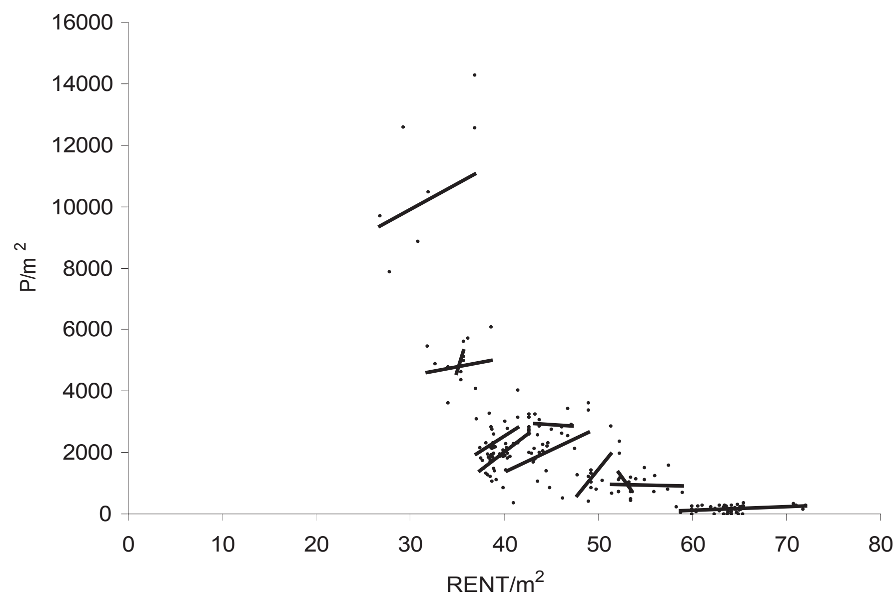


Figure 2: $P/\mathrm{m}^2$ and *RENT*$/\mathrm{m}^2$ for Co-operative Flats in Umeå 1998–1999.

However, this is not the complete picture. The straight lines in Figure 2 illustrate the correlation between *RENT*$/\mathrm{m}^2$ and $p/\mathrm{m}^2$ for the flats within each co-operative. These correlations instead suggests that the trade-off works in the opposite direction, within each co-operative. For most of the co-operatives, higher *RENT*$/\mathrm{m}^2$ generates higher $p/\mathrm{m}^2$.

[2]Hence, we generally assume that each household use the nearest of two shopping areas.

To some extent a fixed cost impact imply that smaller flats within a co-operative both have a higher price and a higher rent per square meter than larger ones, since all flats is required to reach a minimum quality such as kitchen and plumbing. Single room flats may also attract other consumer groups; i.e., there are different sub-markets for co-operative flats. To control for this a structural dummy variable for single room flats *1–ROOM* was introduced. The sign is expected to be positive.

The rent per month that the household pays to the co-operative is moreover in itself determined by factors directly related to information currently not available, but important for the price, such as the quality of the flat, the maintenance status, as well as the pressure from the market on specific segments. Thus, the rent is also a proxy for such structural characteristics. Other factors include the nature of geographic expansion of the co-operatives towards the urban fringe, where the more recent flats have higher quality and thus higher rents. Despite that *1–ROOM* is introduced to take care of scale impacts on small flats, and that accessibility measures control for the location, it is hard to make a sign prediction for ($m^2$/*RENT*) since we lack crucial flat information.

In addition to this, the sign of the age coefficient is also difficult to predict. It would, of course, be natural to assume that younger flats are more expensive than older ones; however that is not necessarily true. Since most of the older co-operatives in general have a lower degree of debt, this makes them more attractive. The other aspect of age is the correlation with the city expansion that Umeå has experienced during the last 50 years. It has followed a spatial diffusion pattern with the older co-operatives close to the CBD, and more recent ones at the outskirts. This would also be in favor for a positive sign on the age coefficient.

## 5 The empirical examination

In this section, the sample is first examined for spatial autocorrelation by the Moran's I test statistic, Moran (1948) and the G-statistic, Getis and Ord (1992). Next, an OLS regression is executed, in order to evaluate whether the indicated spatial autocorrelation has been internalized by the regression parameters or not. If the latter is the case, the OLS regression still provides information on how to proceed in order to make an appropriate final estimation.

Moran's I and the G-statistic depend partly on the chosen spatial weights matrix (that is supposed to mimic the unexplained spatial relationships between the object of investigation). Therefore, five weights matrices are tested in this paper. All matrices used are based on distances between observations within ranges from 0 to 1,000, 1,500, 2,000, 2,500, and 3,000 meter, respectively, and labeled d_1000 up to d_3000. The elements of row $r$ in each weights matrix are set to one for all observations within the specified distance from observation $r$ and zero otherwise, including the diagonal element which is zero by convention. The matrix is finally row standardized by dividing each element with its row sum.

To determine which method to use for the spatial autocorrelation tests the dependent variable must first be assessed as to whether or not it is normally distributed. For this, an asymptotic Wald test, distributed as $\chi^2$ with 2 degrees of freedom, Anselin (1995), is used. The dependent variable $p/\mathrm{m}^2$ gives a W value of 1,190 with a zero probability to reject the null hypothesis of a normal distribution. This means that a permutation approach must be used.

The results from the Moran's I tests for the five weights matrices are presented in Table 2. The dependent variable, $p/\mathrm{m}^2$, indicates significant positive spatial autocorrelation at the 0.1 percent level for all five spatial weights matrices. That is, adjacent located observations tend to have a more similar price per square meter than would be expected purely by chance[3]. The Moran's I value is 0.82 for the d_1000, which is considered to be a high degree of spatial autocorrelation.

For the other weights matrices the Moran's I value is also positive and significant, but decreases with the increased bandwidth from $r$, which is quite reasonable. The question then arises, which weights matrix should then be utilized in the next step, the OLS regression? The high level of spatial autocorrelation for all weights matrices (except for d_3000) makes it, at this stage, difficult to select one weights matrix before the others.

[3]The test statistic is compared with its theoretical mean, $\mathrm{I}=-1/(\mathrm{n}-1)$. So, $\mathrm{I} \rightarrow 0$ as $\mathrm{n} \rightarrow \infty$. The null hypothesis $\mathrm{H}_0 : \mathrm{I} = -1/(\mathrm{n}-1)$ is tested against the alternative hypothesis $\mathrm{H}_a : \mathrm{I} \neq -1/(\mathrm{n}-1)$. If $\mathrm{H}_0$ is rejected and $\mathrm{I} > -1/(\mathrm{n}-1)$, this indicates a positive spatial autocorrelation. That is, high values and low values are more spatially clustered than would be assumed purely by chance. For the other event, if $\mathrm{H}_0$ is again rejected but $\mathrm{I} < -1/(\mathrm{n}-1)$, it indicates negative spatial autocorrelation. Hence observations with high and low values are systematically mixed together.

Table 2: Moran's I test for Spatial Autocorrelation in the Co-operative Flat Data (empirical pseudo significance based on 999 random permutations)

| Variable | Weights matrix | I | Mean | St. dev. | Prob. |
|---|---|---|---|---|---|
| $p/\mathrm{m}^2$ | d_1000 | 0.82 | -0.005 | 0.023 | 0.001 |
| $p/\mathrm{m}^2$ | d_1500 | 0.72 | -0.005 | 0.020 | 0.001 |
| $p/\mathrm{m}^2$ | d_2000 | 0.52 | -0.005 | 0.017 | 0.001 |
| $p/\mathrm{m}^2$ | d_2500 | 0.50 | -0.005 | 0.015 | 0.001 |
| $p/\mathrm{m}^2$ | d_3000 | 0.19 | -0.005 | 0.011 | 0.001 |

The second test for spatial dependence is the G-statistic[4] presented in Table 3. This test does not, at first glance, give such clear cut results as the Moran's I test did. All, except the second weights matrix are insignificant. The reason for the first weights matrix not to be significant is most likely due to the location of the co-operatives. Since the distance between co-operatives often is larger than 1,000 meters, many observations only have surrounding neighbors in their own co-operative. This implies that the difference between observations within this first band width is insignificant. However, the weights matrix with the distance bound 0–1,500 meters indicates positive spatial dependence. This implies that we may expect serious spatial dependence, especially at this distance.

Table 3: G-statistic for Spatial Association

| Variable | Weights matrix | G | Mean | St. dev. | z-value | Prob. |
|---|---|---|---|---|---|---|
| $p/\mathrm{m}^2$ | d_1000 | 0.14 | 0.14 | 0.01 | 0.35 | 0.73 |
| $p/\mathrm{m}^2$ | d_1500 | 0.21 | 0.29 | 0.03 | -2.54 | 0.01 |
| $p/\mathrm{m}^2$ | d_2000 | 0.34 | 0.38 | 0.04 | -0.94 | 0.35 |
| $p/\mathrm{m}^2$ | d_2500 | 0.35 | 0.39 | 0.04 | -1.02 | 0.31 |
| $p/\mathrm{m}^2$ | d_3000 | 0.34 | 0.55 | 0.04 | -0.13 | 0.90 |

Heterogeneity in data, i.e., similar attributes are valued differently at different locations, is another issue that might cause problems. The material was divided into two groups based on Figure 2. Flats in the co-operative closest to the CBD are included in one group while the second group contains all other co-operatives. The variable ($\mathrm{m}^2/\textsc{rent}$) was then tested for heterogeneity with a spatial Chow test, Anselin (1990) that rejected the null hypothesis of a common coefficient for the two groups. Hence, they will from now

[4]The theoretical expected value for Geary's C is 1. A value of C less than 1 indicates positive spatial autocorrelation, and a value above 1 indicates negative spatial autocorrelation.

on be treated as two separate variables, labeled $(\text{m}^2/\textit{RENT})_{:CBD}$ and $(\text{m}^2/\textit{RENT})$.

We now proceed to the estimations and begin with the OLS regression:

$$\boldsymbol{p} = \boldsymbol{Z\beta} + \boldsymbol{\varepsilon} \tag{12}$$

A reason for performing the OLS estimation despite all previous indications of its unsuitability, is first to test if the indications indeed are true. Even if this is the case, the OLS still provides guidance towards a proper model specification. The parameter values are presented in Table 5 in completion.

Based on the OLS residuals two Lagrange Multiplier test are then used to identify remaining spatial dependence and how to solve it. The test results are presented in Table 4 for all five spatial weights matrices.

Table 4: Diagnostics for Spatial Dependence in the Co-operative flat data. Probability measures in parentheses.

| Weights matrix | LM, error | LM, lag |
|---|---|---|
| d_1000 | 13.67 (0.00) | 1.53 (0.22) |
| d_1500 | 20.32 (0.00) | 0.12 (0.73) |
| d_2000 | 21.47 (0.00) | 1.98 (0.16) |
| d_2500 | 1.28 (0.26) | 21.24 (0.00) |
| d_3000 | 0.00 (0.97) | 7.37 (0.01) |

The LM error dependence test for d_1000, d_1500, and d_2000 is highly significant, while the LM lag test is insignificant. Together, these tests indicate problems of spatial error dependence. This would be solved by inclusion of a spatial error correction. The tests for weights matrices d_2500 and d_3000 on the other hand indicate that spatial lag dependence prevails. However, since the tests with the more restrictive band width weights matrices indicated spatial error dependence, these test results are disregarded. This also means that the weights matrices d_2500 and d_3000 are ruled out as inappropriate in the subsequent analysis.

To determine the appropriate estimator for the subsequent regression analysis, a non-normal distribution of the OLS residuals were tested with the Jarque-Bera test statistic; $\chi^2$ distributed with 2 degrees of freedom. With a critical value of 98.6 and a zero probability, the test clearly rejected the null hypothesis of normally distributed residual terms[5]. This rule out ML-estimation. To solve both the problem of spatial autocorrelation and the non-normal distribution of the error terms, we instead suggest a Spatial Autoregressive Generalized Moments estimation, Kelejian and Prucha (1999) together with the three weights matrices presented earlier.

The estimated model has the following form:

$$\begin{aligned} \boldsymbol{p} &= \boldsymbol{Z\beta} + \boldsymbol{\varepsilon} \\ \boldsymbol{\varepsilon} &= \lambda \boldsymbol{W\varepsilon} + \boldsymbol{\xi} \end{aligned} \tag{13}$$

For the regression with the weights matrix d_1000 Sig-sq. is 315,723 (561.9), or 41,000 lower than in the OLS case[6]. In other words, the estimates are improved. Table 5 shows that the estimations with matrices d_1500 and d_2000 improve this further.

Initially, we observe negative signs for the parameters ($m^2$/RENT) :CBD and ($m^2$/RENT), and as predicted by the heterogeity test earlier, a steeper slope for the co-operative closer to the CBD. Thus, the price per square meter actually increases when the rent increases ceteris paribus, and more so for small co-operative flats. The sign may be explained by a re-examination of the fitted lines in Figure 2, where higher rents in fact induce higher prices in most co-operatives. This is probably closely connected with lack of individual flat quality information.

The dummy variable 1–ROOM coefficient is positive and significant indicating that small flats are more expensive to purchase per square meter than larger ones. The age coefficient is highly significant with a positive sign.

The two accessibility measures $INCBD^2$ and $INUNIV^2$ are both significant and positive, indicating that Umeå has a structure with two attractive nodes, although with different characteristics.

Two of the neighborhood coefficients are significant. A high percentage of single-family homes forces the price upwards while the population density works in the opposite direction.

[5]Log-log and semi-log specifications was also tested but yielded similar results.

[6]The normal measures of fit are not applicable for the spatial models. Instead pseudo measures must be used. Unfortunately they can not accurately be compared with the measures of fit from the OLS regressions. The $R^2$ is, in the spatial cases the ratio of the predicted values over the variance of the observed values for the dependent variable. Sq.corr is the squared correlation between the predicted and observed values and Sig-sq. is an estimate for the residual variance.

Table 5: The SAR-GM Regression Results compared with the OLS estimation.***, **, and * indicate a significant value at the 1, 5, or 10% level. § indicate that no inference has been made.

| Variable | d_1000 | d_1500 | d_2000 | *OLS* |
|---|---|---|---|---|
| *CONSTANT* | 619.3 | 661.9 | 78.5 | -206.5 |
| *1–ROOM* | 718.9*** | 701.6*** | 709.2*** | 884.3*** |
| ($m^2$/*RENT*):*CBD* | -254,170*** | -259,163*** | -262,644*** | -182,982*** |
| ($m^2$/*RENT*) | -104,238** | -110,203** | -105,359** | -36,250.1 |
| *AGE* | 138.1*** | 141.6*** | 162.4*** | 115.3*** |
| $INCBD^2$ | $4.5{\cdot}10^9$*** | $4.4{\cdot}10^9$*** | $4.3{\cdot}10^9$*** | $4.2{\cdot}10^9$*** |
| $INUNIV^2$ | $1.7{\cdot}10^9$** | $1.8{\cdot}10^9$** | $1.4{\cdot}10^9$ | $1.4{\cdot}10^9$*** |
| $INSHOP^2$ | $7.7{\cdot}10^8$ | $7.9{\cdot}10^8$ | $1.1{\cdot}10^9$ | $7.3{\cdot}10^8$* |
| *POPDENS* | -173.0** | -176.2** | -200.9** | -172.2*** |
| *SHDENS* | 1,061.0* | 1,089.1* | 898.0 | 976.9*** |
| *TO* | -21.8 | -22.6 | -7.3 | -22.3* |
| $\lambda$ | 0.46(§) | 0.50(§) | 0.61(§) | |
| $R^2$ | 0.96 | 0.96 | 0.96 | 0.93 |
| $R^2$ (Buse) | 0.85 | 0.85 | 0.88 | |
| Sq. Corr | 0.93 | 0.93 | 0.93 | |
| Sig-Sq | 315,723 | 312,620 | 310,262 | 356,809 |
| Sig | 561.9 | 559.1 | 557.0 | 597.3 |
| Iterations | 8 | 8 | 11 | |

Finally, we may observe that the autoregressive parameter $\lambda$ is positive, thus the error terms in locations nearby tend to coincide more than purely by chance. This parameter is considered a nuisance parameter because its sole purpose is to increases the precision of the other regression parameters. Apart from that it is of minor interest. Thus no inference is made on this parameter in this study.

To illustrate the predicted values for $p/m^2$ across the city of Umeå a smoothing map, based on the parameter values from the second spatial regression model, is produced and presented in Figure 3. The reader may clearly detect the concentration of high values around the CBD and, to a lesser degree, around the university area. Low prices are found among the co-operative flats at the outskirts of the city.

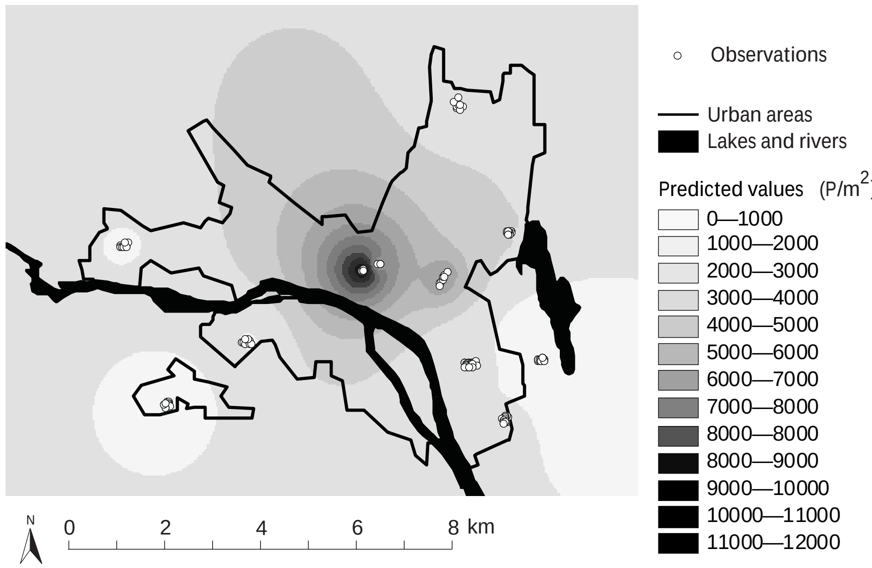


Figure 3: The prediction on $p/\mathrm{m}^2$ for Co-operative flats in Umeå.

# 6 Conclusions

The purpose of this paper was to use hedonic price theory in order to estimate prices on co-operative flat attributes in the city of Umeå. The econometric analysis showed that

- OLS estimation was not applicable due to spatial error dependence. The spatial dependence was controlled for by three spatial weights matrices based on observations with neighbors within 1,000, 1,500, and 2,000 meters and a SAR-GM estimator in the regression analysis. The parameters changes in magnitude as well as significance compared to OLS.

- The coefficient of the attribute that describes rent and area of the flat ($\mathrm{m}^2/\textit{RENT}$) is negative and significant in both the city center and the other co-operatives. Hence, after accessibility etc. has been controlled for, rent and price of the flat are positively correlated. Our interpretation for this correlation is that this reflects unobserved flat quality differences.

- It is also clear that single room flats are significantly more expensive per square meter in comparison with larger flats. This may indicate different markets for small and large co-operative flats in the city.

- The hypotheses of positive parameter estimates for centre accessibility on real estate prices are confirmed. In the case of Umeå, access to the CBD as well as to the university area influence the price in a positive way. Thus, the city of Umeå may be described as a city with a dual nodal structure, although the centres have different characteristics.

The research done in this paper may be used to guide policy makers in their quest for an attractive city. However, since we have not estimated attribute demand functions these results must be used with caution. Even so, a reasonable action based on the signals received from present consumers, in order to increase the attractiveness of the city, would be to develop co-operative flats in the areas between the CBD and the university area, more closely to the former. In other areas, a mix with a high degree of single-family homes might improve the attractiveness.

The fact that we, in this paper, did not have data for the quality attributes for the individual flats makes it difficult to comment on internal attractiveness; and we can therefore not argue for a specific type of housing. Further studies are needed in this field to gain knowledge about consumer preferences regarding flat quality. Another important aspect not investigated in this paper is the consumer's willingness to pay for waterfront location and proximity to park areas. It would also be interesting to make a similar study for single-family homes as well to compare the strength of the preferences for the two different types of housing in a single setting.

## References


[1] Andersson, D. E. (1997). *Hedonic Prices and Center Accessibility: Conceptual Foundations and an Empirical Hedonic Study of the Market for Condominium Housing in Singapore.* Ph. D. Dissertation, Royal Institute of Technology, Stockholm.

[2] Anselin, L. (1988). *Spatial Econometrics: Methods and Models.* Kluwer Academic, Dordrecht.

[3] Anselin, L. (1990). Spatial dependence and spatial structural instability in applied regression analysis. *Journal of Regional Science* 30, 185–207.

[4] Anselin, L. (1995). *SpaceStat, A Software Program for the Analysis of Spatial Data, Version 1.80.* Regional Research Institute, West Virginia University, Morgantown, WV.

[5] Anselin, L., and A. Bera (1998). Spatial dependence in linear regression models with an introduction to spatial econometrics. In: Ullah, A., and D. E. A. Giles (eds). *Handbook of applied economic statistics.*

[6] Can, Y. (1992). Specification and estimation of hedonic housing price models. *Regional Science* 22, 453–474.

[7] Cliff, A., and K. Ord (1972). Testing for Spatial Autocorrelation Among Regression Residuals. *Geographical Analysis* 4, 267–284.

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

[9] Eriksson, B. J. (1997). *Bostäder och Hedoniska Priser. Beställning av flerfamiljshus 1967–1987.* Ph. D. Dissertation, Royal Institute of Technology, Stockholm.

[10] Getis, A., and J. K. Ord (1992). The Analysis of Spatial Association by Use of Distance Statistics. *Geographical Analysis* 24, 189–206.

[11] Kelejian, H. H., and I. R. Prucha (1999). A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model. *International Economic Review* 40 (2), 509–533.

[12] Lancaster, K. (1966). A new approach to consumer theory. *Journal of Political Economy* 74, 132–157.

[13] Lantmäteriet (1997). *Stadskartan*, Kartförlaget, Gävle.

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

[15] Rosen, S. (1974) Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy* 82, 34–55.

[16] SABO. (1997) *Vad värderar vi i vårt boende. En sammanfattning av två undersökningar*. SABO, Stockholm.

[17] SOU. 2000:33 *Bruksvärde, förhandling och hyra - en utvärdering*. Finansdepartementet, Stockholm.

[18] SOU. 2000:87 *Regionalpolitiska utredningens slutbetänkande.* Näringsdepartementet, Stockholm.

[19] Statistics Sweden (2003). Levnadsnivåindikatorer i ULF efter region, kön, ålder och indikator. Tvåårsgenomsnitt 1988-1989-2000-2001. URL: www.scb.se.

[20] Werner, I.B. (2000) *Spelar kvalitet någon roll för priset? - En studie av bostadsrättsköp i StorStockholm.* Ph.D. Dissertation, The Royal Institute of Technology, Stockholm.

# On Space-Time Changes of Hedonic Prices for Single-Family Homes

Krister Sandberg*
Department of Economics and CERUM
Umeå University
SE-901 87 Umeå
Sweden
krister.sandberg@econ.umu.se

**Abstract**

In this paper hedonic prices for single-family homes in two Swedish counties are estimated for two years. Parameter estimates are compared and changes in space and time analyzed. Spatial lag dependence is found to influence the results. Hence, four independent variables are lagged with a spatial weights matrix. Additional spatial error dependence is treated by Spatial Autoregressive Generalized Moment estimation. Structural and neighborhood characteristics together with accessibility measures are used as attributes. The regional price pattern and its changes over time, is illustrated and identified with GIS maps. Proximity to the two county capitals, as well as the other municipality centers, influence property prices positively. This is also noticable over time, were values rise for homes located near major population centers, with water provided by the municipality. Additionally, home value is largely a function of the material condition of the home.

***Keywords* :** Hedonic prices, Single-family homes, Spatial dependence, Heterogeneity
***Classification*[*JEL*] :** D46, D61, R20, R21

*The author wishes to thank The Swedish National Land Survey and Peder Axensten for their contributions providing and preparing data. Thanks as well to Associated Professors Lars Westin and Raymond J.G.M. Florax, and two anonymous referees for their valuable comments on previous versions of the paper. This paper was funded by the EU project Urban Design.

# 1 Introduction

The aim of this paper is to estimate hedonic prices for single-family homes in northern Sweden. Since data for two years within a five-year period are available, we are able to perform a comparative analysis. We may thus for the first time illustrate the price landscape for single-family homes in the region and initiate a discussion on the space-time evolution of the market. Such a discussion may lead to interesting conclusions for actors who want to improve the attractiveness of the region.

Since the seminal work by Haas (1922), Lancaster (1966), and Rosen (1974), the values of attributes associated with heterogeneous goods have been analyzed by hedonic price theory where it is assumed that the buyer implicitly reveals his or her preferences for the attributes through the price paid. A good may then be seen as a bundle of characteristics matching the household utility function. When the bidder with the highest bid purchase each home, the market price gives the outer envelope of the valuation of each attribute by all households in the market.

The empirical literature on hedonic prices for single-family homes is nowadays quite numerous but to a large extent based on American data, e.g. Blomquist et al. (1998) and Sivitanidou (1996). Exceptions that use Swedish and British data are e.g. Wigren (1987), Englund et al. (1998), as well as Cheshire and Sheppard (1995).

Another common feature of many other studies is their focus on a single specific characteristic and its influence on prices. Blomquist (1998), Benson et al. (1998), Shultz and King (2001), Beron et al. (2001), and Bond et al. (2002) analyze the influence of a property's "view" on its value. Des Rosiers et al. (2002) examine the importance of landscaping homes and the surrounding property. Clark and Allison (1999) analyzed the impact of risk perception. The price effect of increased accessibility through a new bridge and internet connections was studied by Smersh and Smith (2000) and Thompson and Hills (1999) respectively. Bogart and Cromwell (2000) analyzed the impact of a re-distribution of schools on home values. Several, albeit qualitative, studies of the Swedish market, among them Lindgren and Rosberg (1992), and Andersson (1998) found that proximity to the Central Business District, to major commercial service outlets, and to the waterfront all positively influence prices.

Recently, the attention in the hedonic price literature towards spatial dependence (spatial autocorrelation) has increased. Can and Megbolugbe (1997), Pace and Gilley (1997), Basu and Thibodeau (1998), Brasington (1999), as well as Tse (2002) are examples in this direction.

The purpose of this paper is to determine, which attributes and to what extent they influence the price of homes in estimations corrected for spatial dependence. The paper is organized as follows. The next section deals with the theory of hedonic prices and spatial econometrics. This is followed by a section on data description. The empirical examination is outlined in the fourth section, while our conclusions may be found in the final section.

## 2 Hedonic price theory and spatial econometrics

Under hedonic price theory, the value of a good may be decomposed into the value of its component attributes. These hedonic prices are revealed through observed prices on differentiated goods and by quantifying the “amount” of each attribute that the good possesses. This permits a value to be assigned to each of the attributes. After enumerating the attributes that differentiate one good from a substitute, and quantifying the contribution that each attribute of the good makes to the good’s overall value, we can derive an implicit price for each of the attributes. The hedonic price approach is best employed when care is taken to select only a mix of attributes which are commonly found in the “average” home in the market. A nice introduction is found in Lancaster (1966).

The concept of implicit or hedonic prices was first formalized in Rosen (1974). The good considered, e.g. a home, may be described by $m$ characteristics. Each home is then represented by a vector $\boldsymbol{z} = (z_1, \dots, z_m)$ where element $z_i$ measures the amount of the $i$th characteristic attached to each home. A market clearing price function that equates supply and demand, driven by buyer and seller utility maximization, for this vector of characteristics, is defined as the hedonic price function $p(\boldsymbol{z}) = p(z_1, \dots, z_m)$. Thus, the hedonic price function is market determined through buyer/seller transaction negotiations were the seller tries to maximize the price of the flat subject to a time constraint given by the time when an alternative residence is available.

The household utility function may be written:

$$U = u(\boldsymbol{z}, c; \boldsymbol{\alpha}) \tag{1}$$

where $\boldsymbol{z}$ is consumption of the single-family home, $c$ is consumption of a composite good, and $\boldsymbol{\alpha}$ is a vector of parameters that characterize the household preferences. The price that a household is willing to pay for the home is derived from (1) as a function of the embedded characteristics, a given household income ($M$), and the achieved utility level. That is, the household bid rent function:

$$\gamma(\boldsymbol{z}, M, U; \boldsymbol{\alpha}) \tag{2}$$

and implicitly:

$$U = u(\boldsymbol{z}, M - \gamma; \boldsymbol{\alpha}) \tag{3}$$

The derivative of the bid rent function, $\frac{\partial \gamma}{\partial z_i}$, gives the rate at which the household would be willing to change its expenditure on a home when the amount of characteristic $i$ increases, while keeping other characteristics constant.

The household chooses a single-family home with characteristics $\boldsymbol{z}$, and its consumption of the composite good $c$ by solving:

$$\max_{\substack{\boldsymbol{z}, c \\ \text{s.t. } M \geq p(\boldsymbol{z}) + c}} \quad u(\boldsymbol{z}, c; \boldsymbol{\alpha}) \tag{4}$$

The equilibrium market price, $p(\boldsymbol{z})$, reflects the market value of a home with a set of attributes given i.e. amortization payment, available interest schemes, and expected costs for repair and improvements during the entire period that the household intends to keep the home.

Through the first order conditions we obtain:

$$\frac{u_i}{u_c} = p_i \ \forall i \tag{5}$$

where $u_i = \frac{\partial u}{\partial z_i}$, $u_c = \frac{\partial u}{\partial c}$ , and $p_i = \frac{\partial p}{\partial z_i}$ is the hedonic price of characteristic $i$.

A combination of the first order condition (5) and the implicit differentiation of (3) reveals the household's optimal choice for a home. It is found where equality occurs between the slope of the bid rent function and the hedonic price with respect to each characteristic. Thus, the household selects such that its indifference curve is tangent to the price gradient.

As mentioned earlier, the vector $\boldsymbol{z}$ consists of a set of characteristics which are subjectively determined by the household. To facilitate the discussion of the attributes, this vector of characteristics is in the hedonic price literature usually divided into three broad groups, structural ($\boldsymbol{s}$), neighborhood ($\boldsymbol{n}$), and accessibility ($\boldsymbol{a}$) attributes with $\boldsymbol{\omega}$, $\boldsymbol{\eta}$, and $\boldsymbol{\psi}$ as the corresponding parameter vectors. Hence, the traditional hedonic price function of a general regression model may be formulated as:

$$p(\boldsymbol{z}) = f(\boldsymbol{s}, \boldsymbol{n}, \boldsymbol{a}; \boldsymbol{\omega}, \boldsymbol{\eta}, \boldsymbol{\psi}) + \boldsymbol{\varepsilon} \tag{6}$$

or expressed in linear vector form:

$$\boldsymbol{p} = \boldsymbol{Z\beta} + \boldsymbol{\varepsilon} \tag{7}$$

where $\boldsymbol{p}$ is a ($n$ by 1) vector of observations on the dependent variable, $\boldsymbol{Z}$ is a ($n$ by $k$) matrix of observations on the exogenous variables with $\boldsymbol{\beta}$ as the associated ($k$ by 1) regression coefficient vector. $\boldsymbol{\varepsilon}$ is a ($n$ by 1) vector of random error terms.

Cross-sectional data is usually prone to problems of spatial dependence. When this in fact is the case, specialized methods of spatial regression analysis must be applied to avoid potentially biased results and faulty inference due to violation of the basic assumptions in classic regression analysis, cf. Anselin (1988) or Anselin and Bera (1998).

Two kinds of spatial dependence specifications are commonly used in the literature. Spatial Error Dependence, arises when the prices of neighboring homes move together due to common or correlated unobservable variables i.e. lack of stochastic independence between observations, cf. Cliff and Ord (1972, 1973). If unsolved, this problem will violate the standard error assumptions under normality of the linear regression model, with inefficient estimates as a result. To control for this the spatial error dependence is incorporated via an autoregressive error term:

$$\begin{aligned} \boldsymbol{p} &= \boldsymbol{Z\beta} + \boldsymbol{\varepsilon} \\ \boldsymbol{\varepsilon} &= \lambda \boldsymbol{W\varepsilon} + \boldsymbol{\xi} \end{aligned} \tag{8}$$

where $\boldsymbol{W}$ is a ($n$ by $n$) spatial weights matrix (that is supposed to mimic the spatial relationships between the objects of investigation) with elements $w_{rs}$ that corresponds to observation pair $r$ and $s$. $\boldsymbol{W\varepsilon}$ is the spatial lag for the error term, $\lambda$ is the autoregressive coefficient, and $\boldsymbol{\xi}$ is a ($n$ by 1) vector of well-behaved error terms $\boldsymbol{\xi} \backsim N(0, \sigma^2\boldsymbol{I})$.

Spatial Lag Dependence, is present if spatial correlation exist between observations of the dependent variable. This means that the transaction price for one home is influenced by the transaction prices for nearby homes and vice versa, cf. Anselin (1988) and Can (1992). If ignored, the OLS estimates will be both biased and inefficient. Inclusion of a spatial lag solves the spatial dependence problem:

$$\boldsymbol{p} = \rho\boldsymbol{W}\boldsymbol{p} + \boldsymbol{Z}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{9}$$

where $\rho$ is an autoregressive coefficient and $\boldsymbol{Wp}$ is the spatially lagged dependent variable.

However, Anselin (2003) observed that this solution to the spatial lag dependence problem is not without flaws unless you believe in global spatial dependence. If the dependence is more or less local, then the specification in equation (9) must be adjusted, due to otherwise undesired induced heteroscedasticity. Since we in our case have reasons to believe, with large distances between the objects, that the influence between the single-family homes is cut-off at some distance we instead use a spatial cross-regressive model suggested by Florax and Folmer (1992), to consider strictly local spillovers:

$$\boldsymbol{p} = \boldsymbol{W}\boldsymbol{Z}\boldsymbol{\rho} + \boldsymbol{Z}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{10}$$

where $\boldsymbol{\rho}$ now is a ($k$-1 by 1) vector. In case spatial dependence among the error terms remains after this correction a spatial error correction is added. Hence, the extended hedonic price model with correction for spatial lag dependence, is either (10) or a combination of (8) and (10):

$$\begin{aligned} \boldsymbol{p} &= \boldsymbol{W}\boldsymbol{Z}\boldsymbol{\rho} + \boldsymbol{Z}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \\ \boldsymbol{\varepsilon} &= \lambda\boldsymbol{W}\boldsymbol{\varepsilon} + \boldsymbol{\xi} \end{aligned} \tag{11}$$

Another issue to consider before we continue is heterogeneity, or structural regimes, in the data. Different regions may place different values on a given attribute. Failure to consider these regional differences and instead examine only the apparent average value for a given attribute (as revealed through hedonic prices) risks excluding attributes which may have significant value for a particular region. This matter is discussed further in the estimation part.

## 3 Single-family home attributes in 1994 and 1999

Our data covers the market for single-family homes in the counties of Västernorrland and Västerbotten, both located in the northern part of Sweden. Data are available for two years, 1994 and 1999, and consists of 2,778 realized transactions in 1994 and 4,538 transactions in 1999. The spatial distributions of sold homes are presented in Figure 1 where each home is indicated by a dot.

While the number of transactions is larger for the second year, the spatial distribution of sales is rather similar and it is readily apparent that the distribution reflects the underlying supply pattern of homes in the two counties. Most transactions occur along the coast. Concentrations exists especially near the regional population centers of Umeå (the cluster west of the small island) and Sundsvall (the south east cluster). Further inspection of the maps reveals that the transaction pattern also follows the inland roads. Compared to a "normal" year during the 1990's, the first year represents a low number of transactions while in the second year, the number of sales is higher than usual. The explanation for this may be found in the different economic situations in Sweden. It is also likely that some of the 1999 transactions took place due to upcoming changes in the taxation regulations. In either case, the effect of supply changes on the overall results have not been invesigated in this paper.

The total population in the region was 519,000 in 1994, dropping to 510,000 inhabitants in 1999. This population is distributed over 22 municipalities, where the densest municipalities are found along the coast. During the period from 1994 to 1999 all municipalities in the region, except the largest, Umeå, faced a population decrease. This pattern follows the overall movement of people towards the larger cities during the period. This leads us to expect to find a general increase of home prices along the coast, while the reverse may be expected in the inland areas. However, the material condition of each home, the accessibility to various services etc. will obviously also have a strong influence on the price and we may not conclude a priori that each part of the coast experienced an increase in home value by the end of the period.

Before we continue it is appropriate to present the descriptive statistics for our data and discuss the expected signs of the variables used in the estimation. Information on the characteristics of each home is obtained from the annual Swedish property taxation data.

Table 1: The Descriptive Statistics for sold Single–family homes in Västerbotten and Västernorrland the years 1994 and 1999. All prices are expressed in values of the year 2000. A bar (-) = implies that no information is registered.

| Variable | Unit | Mean_94 (st.dev) | Range_94 | Mean_99 (st.dev) | Range_99 | Sign |
|---|---|---|---|---|---|---|
| *ln* PRICE | ln K SEK | 5.97(0.79) | 0.74–7.87 | 6.17(0.69) | 2.21–8.02 | |
| *ln* FLOOR SPACE | ln $m^2$ | 4.80(0.35) | 3.22–6.07 | 4.85(0.28) | 3.64–6.55 | (+) |
| *ln* LOT AREA | ln $m^2$ | 7.06(0.86) | 4.36–11.28 | 6.99(0.78) | 4.36–10.72 | (+) |
| *ln* AGE | year | 3.47(0.68) | 0.00–5.28 | 3.59(0.55) | 0.69–5.40 | (-) |
| *Ordinary house* | dummy | 0.86 | 0–1 | 0.86 | 0–1 | (+) |
| *If added floor space* | dummy | 0.07 | 0–1 | 0.08 | 0–1 | (+) |
| *Noise* | dummy | 0.01 | 0–1 | 0.005 | 0–1 | (-) |
| *No electricity* | dummy | 0.0003 | 0–1 | – | – | (-) |
| *Construction error* | dummy | 0.001 | 0–1 | 0.002 | 0–1 | (-) |
| *Moisture* | dummy | – | – | 0.003 | 0–1 | (-) |
| *Difficult lot* | dummy | 0.004 | 0–1 | 0.0002 | 0–1 | (-) |
| *Renovation object* | dummy | – | – | 0.0002 | 0–1 | (-) |
| *Radon* | dummy | – | – | 0.006 | 0–1 | (-) |
| *No maintenance* | dummy | 0.002 | 0–1 | 0.002 | 0–1 | (-) |
| *Less than 50,000* SEK | dummy | – | – | 0.006 | 0–1 | (-) |
| *Other annotations* | dummy | 0.007 | 0–1 | 0.007 | 0–1 | (?) |
| *Own water* | dummy | 0.12 | 0–1 | 0.10 | 0–1 | (-) |
| *No water* | dummy | 0.01 | 0–1 | 0.003 | 0–1 | (-) |
| *Quality points 16–20* | dummy | 0.12 | 0–1 | 0.07 | 0–1 | (+) |
| *Quality points 21–25* | dummy | 0.26 | 0–1 | 0.27 | 0–1 | (+) |
| *Quality points 26–35* | dummy | 0.49 | 0–1 | 0.58 | 0–1 | (+) |
| *Quality points 36–45* | dummy | 0.10 | 0–1 | 0.06 | 0–1 | (+) |
| *Quality points 46-50(52)* | dummy | 0.002 | 0–1 | 0.003 | 0–1 | (+) |
| NET MIGRATION | persons | 207.1(500.2) | -176–1321 | -150.75(155.2) | -347–100 | (+) |
| *ln* (UNEMPL/POP) | quotient | -3.0(0.15) | -3.92–2.76 | -3.4(0.22) | -3.07–3.07 | (-) |
| 2 YEARS INTEREST RATE | % | 10.78(1.13) | 8.50–12.00 | 5.68(0.63) | 4.50–6.45 | (-) |
| MUNICIPALITY TAX | % | 20.17(0.48) | 19.05–20.85 | 22.41(0.46) | 21.29–23.15 | (?) |
| AVERAGE INCOME 20+ | K SEK | 156.38(7.84) | 136.0–165.7 | 183.90(9.19) | 158.68–194.14 | (+) |
| *Beach* | dummy | 0.01 | 0–1 | 0.01 | 0–1 | (+) |
| *Near beach* | dummy | 0.03 | 0–1 | 0.02 | 0–1 | (+) |
| *Built–up area* | dummy | 0.77 | 0–1 | 0.81 | 0–1 | (+) |
| ACCESS | pop./m | 21.49(27.51) | 1.68–454.27 | 23.79(34.53) | 1.56–893.26 | (+) |
| *Within 5 km of* E12 | dummy | 0.17 | 0–1 | 0.21 | 0–1 | (+) |
| *Within 5 km of* E4 | dummy | 0.55 | 0–1 | 0.61 | 0–1 | (+) |

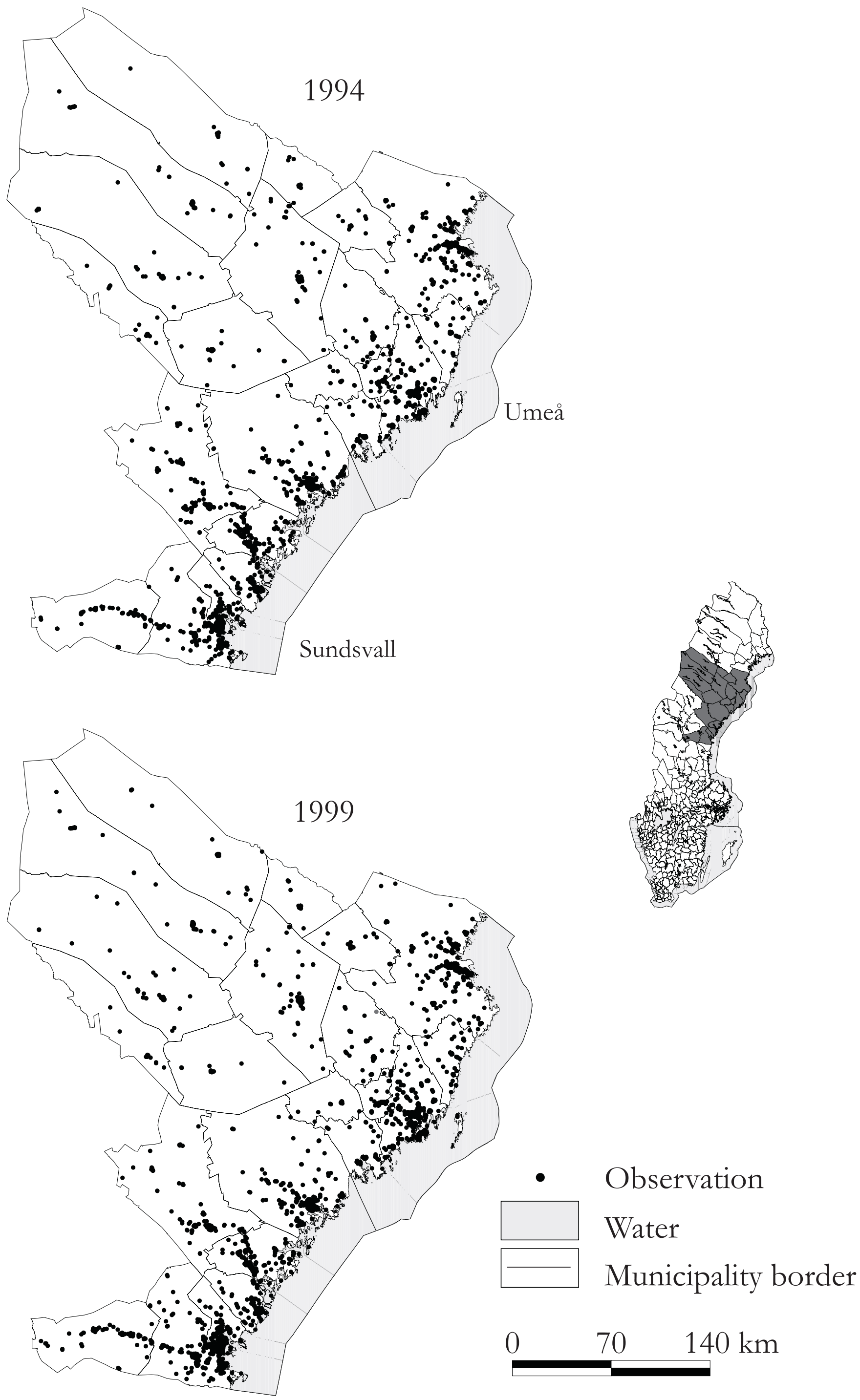


Figure 1: Transactions of Single-Family Homes in Västerbotten and Västernorrland during the years 1994 and 1999.

In the first row, the *dependent variable*, the natural logarithm of the price, *ln PRICE*, is given. The average price has, after correction for inflation, increased from 391,000 SEK in 1994 to 478,000 SEK in five years time (an average annual increase of around 4.5 percent). This is in line with the overall rise in household incomes during the period.

The *independent variables* are, as mentioned earlier, divided into three groups based on their structural and neighborhood characteristics as well as their proximity to major population centers.

The *structural variables* consist of characteristics specific to each home. Obvious attributes to be included as continuous variables are floor space, the area of the surrounding property, and the age of the home. The two former are expected to have a positive impact on the price while the latter is expected to influence prices negatively.

The remaining structural variables are treated as dummies. Semi-detached homes or homes linked to other homes by a garage represents the base case while ordinary homes are represented by a dummy. The coefficient is expected to be positive sign. If a home has been extended during its lifetime, the value may increase and a positive sign is expected. Specific annotations about the lot or the home itself are indicated for some of the observations in the annual real estate taxation. Those are treated as dummies and are assumed to have negative/positive signs for bad/good attributes.

Another set of dummies refer to water supply. The default is municipality supplied water. The coefficients for other forms of water supply ought to have negative signs because of the extra time that must be spent on maintenance etc. and for potential problems with water quality. The interior quality of each home is graded by a system of quality points in the property taxation. The points are here divided into five groups each given a dummy. All coefficients connected with these dummies are expected to have positive signs since they are measured against the lowest group that includes values 0–15. The mean values for all dummy variables indicate the share of observations with these attributes.

The *neighborhood attributes*, consider the neighborhood for each home, in this case the overall situation in the municipality where the home is located. Net migration is a proxy for the demand on the real estate market in each municipality. It does not include fertility nor mortality which means that net migration may capture whether the municipalities are attractive to migrants. A positive net migration should increase the demand and a positive sign on this coefficient is therefore expected. To investigate the importance of imbalances at the labor market, the natural logarithm of the unemployment quotient (unemployed/population) is included. It is expected that a positive quotient, revealing a large share of unemployed in the population, will influence prices downward. A high interest rate at the date of purchase makes borrowing expensive and increases the risk in connection with a purchase. The coefficient sign is therefore expected to be negative.

The parameter sign of the municipality tax level is somewhat harder to predict. The average municipality tax was 20.2% in 1994 and had increased to 22.4 by 1999. A high tax may be an indication of improved public service but it could also be an indication of financial strains. Thus, the sign is ambiguous.

The next variable measures the average income level for people 20 years and older. It is included to reflect the economic situation among the households in the municipality. The average income increased by almost 30,000 SEK in real terms between 1994 and 1999. At the same time, the income spread between municipalities also increased during the period. The sign with respect to home prices is expected to be positive.

For homes located at or near a beach, a positive sign is expected due to the presence of the waterfront. In 1994, 77% of the transactions were transactions in built-up areas. By 1999, this share had grown to 81%. The sign for this dummy coefficient is expected to be positive.

The third and last group contains the *accessibility attributes*. A combination of distance and population is used, defined as:

$$ACCESS_i = \sum_{1}^{22} \frac{POP_k}{D_{ik}} \; ; \quad \forall i \tag{12}$$

Hence, ACCESS is the sum over the 22 municipalities of the quotients between population and the distance, "as the crow flies", between the observation $i$ and each municipality CBD $k$. This implies that a home is valued differently depending on the number of people that has access to the home, or put another way, a proxy for the number of people one may reach from the home. The assumption here is that a large population in the vicinity enhances the value of a home. If this is correct, a positive sign should be expected.

The last two accessibility attributes are dummy variables, indicating whether a home lies within a 5 km range of the European roads E4 and E12. The assumption is that close proximity to major roadways should raise the value of the home.

## 4 The empirical examination

The regression results are presented in Table 2 for the year 1994 and in Table 3 for the year 1999. Each table consists of six regressions enabling the reader to follow the impact of different specifications. The first regression in each table is an OLS regression without any spatial considerations, included for completeness. For both years the regression diagnostics indicates a non-normal distribution of residuals, heteroscedasticity, and spatial dependence in the data.

Three tests on heteroscedasticity are reported. The first of these tests, a Koenker-Basset test (Koenker and Basset, 1982) where the null hypothesis of homoscedasticity, for both years, is rejected. This leads us to the next heteroscedasticity test, which is the same Koenker-Basset test, based on the same OLS residuals, now with a specified possible solution in the form of a category variable. When the null hypothesis of homoscedasticity cannot be rejected indicates that the homoscedasticity would prevail if this category variable is used. The third test is a Wald test that checks if the proposed treatment actually solves the problem in the following regressions. This is the case when the null hypothesis of remaining heteroscedasticity is rejected. The rest of the estimation part is organized as follows. We start by addressing the heteroscedasticity problem, continue with problems of heterogeneity, and in columns five and six deal with these two problems together with the spatial dependence problem simultaneously.

### 4.1 The regression analysis

As discussed above, the diagnostics based on the OLS residuals suggests that the problem of heteroscedasticity is likely to be additive. A category variable is therefore created that consists of four groups based on the geographical location of each home. The first and second groups consists of observations located in the municipalities of Umeå and Sundsvall respectively. Homes located in the other coastal municipalities are assigned to the third group. All homes in inland municipalities are consequently assigned to the forth group. The Feasible Generalized Least Squares regression method is used to estimate the hedonic prices. The results from this regression is found in the second column in each table. Initially, we may observe that the overall model fit[1] is 66% in 1994 and 62% in 1999, which to a large extent continues for the remaining regressions. Moreover, all four category variables are positive and significant for both years which indicate different variances for the locations. This is, as expected, confirmed though the significant value of the Wald test rejecting the null hypothesis of remaining heteroscedasticity.

In the next step of the regression analysis, presented in the third columns, we test for data heterogeneity with a spatial Chow-Wald test, suggested by Anselin (1990). The variables were tested for structural instability and for stability of the individual coefficients with a null hypothesis of a joint common coefficient for all homes. The test results suggested heterogeneity among a couple of variables for each year. The age and access variables, as well as the variables for home quality in 1994 and floor space in 1999 are divided into four structural shift variables based on the location of homes in the same sets as the category variables described earlier. Hence, the previous variables are exchanged for four new variables. The value of homes in the first group (Umeå), is least reduced by aging. Value of homes in the other coastal municipalities are more sensitive to age than in Umeå, but it is in the inland municipalities that we find the greatest impact from aging on the value of a home. As expected, the four access variables are all significant for both years indicating proximity to major population centers increases the home's value. For the year 1994 the quality variables were divided in four groups and as seen the parameter values vary quite a bit. As was the case with the age variable less emphasis on quality is put on properties in Umeå and Sundsvall. Perhaps is this a sign of relative shortage of supply in these locations. For the year 1999 the floor space was

[1]The normal measures of fit are not applicable for the spatial models. Instead pseudo measures must be used. Unfortunately they cannot accurately be compared with the measures of fit from the OLS regressions. The $R^2$ is, in the spatial cases, the ratio of the predicted values over the variance of the observed values for the dependent variable. Sq.corr is the squared correlation between the predicted and observed values.

also heterogeneous. The interpretation is the same as above; home buyers demand higher quality (or greater floor space) when they purchase a home in the hinterland as compared with the locations near the coast or major population centers. Notice also the difference between all these parameter values and those in the previous regression.

As seen in the diagnostics of the regressions in column two there is a clear indication of spatial dependence from the Lagrange Multiplier tests. Normally, the rule of thumb is to choose the remedy based on the test with the most significant value, as suggested by Florax et al. (2003). That would in our case imply that we should include a lag in 1994 and an error component in 1999. However, we would loose some of the comparability between the years with this solution. The decision was therefore to only partially follow this rule and first solve the spatial lag problem for both years. We will then continue to test for remaining spatial error dependence. As mentioned earlier in Section two, the spatial lag on the dependent variable leads to unavoidable heteroscedasticity. Instead we use the spatial cross-regressive model by Florax and Folmer (1992), and thus consider only local spillovers. The variables (*ln FLOORSPACE, ln LOTAREA, ln AGE*, and *ACCESS*) are lagged with a spatial weights matrix, in the regression tables indicated with prefix $\rho$, as the weighted average of neighboring observations.

A number of spatial weights matrices used to transform the variables in question was tested in the process. The one finally chosen was a spatial weights matrix that consists of the row standardized inverse distances between all observations within a 48 kilometer radius. This is the minimum allowable distance between observations in the 1994 data set, given that each observation has at least one neighbor. For comparability between the years, the same distance cut-off was used for the 1999 data set.

To investigate the effect and magnitude of the neighborhood spillovers, first without considering heterogeneity, the FGLS regression in column two is expanded in column four. The results are that the floor space impact is slightly smaller than before. On the other hand we have instead caught the importance of the weighted average floor space of homes transacted within the 48 km range. Homes with large floor sizes in the vicinity are for both years considered attractive. The magnitude of the lot area effect is about the same as before. Large lot areas in the surroundings influence the value negatively, especially in 1999. This may be explained by the fact that larger lots are often found in sparse locations such as in the inland municipalities. The age effect is in both years negative and significant. However, the effect of the age of surrounding homes while negative in 1994, is positive in 1999. The access coefficient that previously had a positive value is now negative for the year 1994, while the lagged access coefficient has a stronger positive sign. The variable operates as a density variable and at the same time as an accessibility variable. A high value means that observation $i$ is surrounded by observations in close proximity to major population centers. It reflects the proximity of the surrounding areas, and not just the proximity of the home itself, to major population centers. Thus, we may conclude that the accessibility of the surrounding area to major population centers is more important than the accessibility of the home itself to major population centers.

Table 2: The 1994 Regression Results for LN PRICE. ***, **, and * indicate a significant value at the 1, 5, or 10% level. § indicate that no inference has been made.

| Variable | OLS | FGLS | FGLS | FGLS | FGLS | SAR–GM |
|---|---|---|---|---|---|---|
| $\lambda$ | | | | | | 0.51(§) |
| *Constant* | -2.82*** | -3.16*** | -0.99 | -2.40*** | -0.78 | -1.62 |
| *ln* FLOOR SPACE | 0.58*** | 0.57*** | 0.55*** | 0.54*** | 0.53*** | 0.54*** |
| $\rho$ *ln* FLOOR SPACE | | | | 0.29*** | 0.25*** | 0.41*** |
| *ln* LOT AREA | 0.10*** | 0.10*** | 0.10*** | 0.12*** | 0.12*** | 0.11*** |
| $\rho$ *ln* LOT AREA | | | | -0.08** | -0.08* | -0.11** |
| *ln* AGE | -0.30*** | -0.27*** | | -0.25*** | | |
| *ln* AGE_1 | | | -0.17*** | | -0.16*** | -0.17*** |
| *ln* AGE_2 | | | -0.23*** | | -0.22*** | -0.23*** |
| *ln* AGE_3 | | | -0.34*** | | -0.31*** | -0.31*** |
| *ln* AGE_4 | | | -0.43*** | | -0.39*** | -0.39*** |
| $\rho$ *ln* AGE | | | | -0.13*** | -0.09** | -0.07 |
| *d_Ordinary house* | -0.003 | 0.002 | -0.02 | 0.007 | -0.02 | -0.01 |
| *d_If added floor space* | 0.10*** | 0.09*** | 0.09*** | 0.08*** | 0.08*** | 0.08*** |
| *d_Noise* | 0.09 | 0.08 | 0.06 | 0.10 | 0.08 | 0.05 |
| *d_No electricity* | -0.04 | -0.12 | -0.19 | -0.13 | -0.19 | -0.20 |
| *d_Construction error* | -0.31 | -0.27 | -0.25 | -0.20 | -0.19 | -0.19 |
| *d_Difficult lot* | -0.05 | -0.05 | -0.04 | -0.05 | -0.04 | -0.05 |
| *d_No maintenance* | -0.82*** | -0.83*** | -0.69*** | -0.80*** | -0.68*** | -0.61*** |
| *d_Other annotations* | -0.04 | -0.08 | -0.09 | -0.11 | -0.12 | -0.14* |
| *d_Own water* | -0.11*** | -0.10*** | -0.08 | -0.11*** | -0.09*** | -0.09*** |
| *d_No water* | -0.12 | -0.13 | -0.12 | -0.16** | -0.14* | -0.13* |
| *d_Quality points (16–20)* | 0.35*** | 0.28*** | | 0.29*** | | |
| *d_Quality points (16–20)_1* | | | 0.18*** | | -0.09 | -0.15 |
| *d_Quality points (16–20)_2* | | | 0.32*** | | 0.35*** | 0.38*** |
| *d_Quality points (16–20)_3* | | | 0.33*** | | 0.32*** | 0.31*** |
| *d_Quality points (16–20)_4* | | | 0.50*** | | 0.45*** | 0.49*** |
| *d_Quality points (21–25)* | 0.61*** | 0.51*** | | 0.50*** | | |
| *d_Quality points (21–25)_1* | | | -0.05 | | 0.03 | 0.007 |
| *d_Quality points (21–25)_2* | | | 0.44*** | | 0.46*** | 0.48*** |
| *d_Quality points (21–25)_3* | | | 0.62*** | | 0.59*** | 0.58*** |
| *d_Quality points (21–25)_4* | | | 0.77*** | | 0.72*** | 0.77*** |
| *d_Quality points (26–35)* | 0.81*** | 0.69*** | | 0.67*** | | |
| *d_Quality points (26–35)_1* | | | 0.14*** | | 0.21* | 0.17 |

*continued on next page*

Table 2: *continued*

| Variable | OLS | FGLS | FGLS | FGLS | FGLS | SAR–GM |
|---|---|---|---|---|---|---|
| *d_Quality points (26–35)_2* | | | 0.57*** | | 0.59*** | 0.60*** |
| *d_Quality points (26–35)_3* | | | 0.82*** | | 0.78*** | 0.78*** |
| *d_Quality points (26–35)_4* | | | 1.00*** | | 0.94*** | 0.97*** |
| *d_Quality points (36–45)* | 0.90*** | 0.77*** | | 0.77*** | | |
| *d_Quality points (36–45)_1* | | | 0.22* | | 0.32** | 0.25* |
| *d_Quality points (36–45)_2* | | | 0.65*** | | 0.66*** | 0.68*** |
| *d_Quality points (36–45)_3* | | | 0.97*** | | 0.93*** | 0.92*** |
| *d_Quality points (36–45)_4* | | | 1.06*** | | 1.00*** | 0.99*** |
| *d_Quality points (46–50)* | 0.85*** | 0.70*** | 0.45*** | 0.59*** | 0.40*** | 0.36** |
| NET MIGRATION | 0.0002*** | 0.0002*** | 0.0003*** | 0.0001*** | 0.0002 | 0.0002* |
| LN UNEMP/POP | -0.05 | 0.03 | -0.13* | 0.01 | -0.13 | -0.09 |
| INTEREST RATE, 2 YEARS | 0.008 | 0.009 | 0.009 | 0.007 | 0.008 | 0.01* |
| MUNICIPALITY TAX | 0.07*** | 0.08*** | 0.05** | 0.07*** | 0.04** | 0.06 |
| AVERAGE INCOME | 0.02*** | 0.03*** | 0.01*** | 0.02*** | 0.01*** | 0.01*** |
| *d_Beach* | 0.31*** | 0.31*** | 0.28*** | 0.31*** | 0.27*** | 0.29*** |
| *d_Near beach* | 0.10** | 0.08* | 0.07* | 0.09** | 0.08** | 0.08** |
| *d_Built–up area* | 0.17*** | 0.20*** | 0.17*** | 0.23*** | 0.20** | 0.19*** |
| ACCESS | 0.003*** | 0.003*** | | -0.002*** | | |
| ACCESS_*1* | | | 0.001*** | | -0.002*** | -0.002*** |
| ACCESS_*2* | | | 0.004*** | | 0.0007 | 0.0002 |
| ACCESS_*3* | | | 0.004*** | | 0.003* | 0.0002 |
| ACCESS_*4* | | | 0.012*** | | 0.004** | 0.0003 |
| $\rho$ ACCESS | | | | 0.01*** | 0.01*** | 0.01*** |
| *d_Within 5 km of* E*12* | 0.17*** | 0.18*** | 0.25*** | 0.08*** | 0.14*** | 0.15*** |
| *d_Within 5 km of* E*4* | 0.19*** | 0.18*** | 0.19*** | 0.10*** | 0.11*** | 0.11*** |
| CATERGORY_*1* | | 0.10*** | 0.08*** | 0.09*** | 0.07*** | 0.07(§) |
| CATERGORY_*2* | | 0.08*** | 0.07*** | 0.08*** | 0.06*** | 0.06(§) |
| CATERGORY_*3* | | 0.19*** | 0.18*** | 0.17*** | 0.17*** | 0.16(§) |
| CATERGORY_*4* | | 0.30*** | 0.28*** | 0.29*** | 0.28*** | 0.26(§) |
| *Normality* | 13312*** | | | | | |
| *Heterosc. K-B* | 225.9*** | | | | | |
| *Heterosc. K-B with Cat.4* | 38.5*** | | | | | |
| *Heterosc.Wald* | | 242.7*** | 330.5*** | 260.1*** | 332.6*** | |
| *Spatial dep. LMerror* | 488.6*** | 435.5*** | 267.4*** | 263.9*** | 191.2*** | |
| *Spatial dep. LMerror Robust* | 108.3*** | | | | | |
| *Spatial dep. LMlag* | 497.8*** | 504.1*** | | | | |
| *Spatial dep. LMlag Robust* | 117.5*** | | | | | |
| $R^2$ | 0.70 | 0.66 | 0.72 | 0.68 | 0.74 | 0.74 |
| $R^2$*-adj.* | 0.70 | | | | | |
| *LIK* | -1575.5 | | | | | |
| *AIC* | 3211.0 | | | | | |
| *SC* | 3388.9 | | | | | |
| *Sq.corr* | | 0.70 | 0.72 | 0.72 | 0.74 | 0.74 |
| *Observations/ Iterations* | 2778 | 2778 | 2778 | 2778 | 2778 | 2778 / 6 |

Table 3: The 1999 Regression Results for LN PRICE. ***, **, and * indicate a significant value at the 1, 5, or 10% level. § indicate that no inference has been made.

| Variable | OLS | FGLS | FGLS | FGLS | FGLS | SAR–GM |
|---|---|---|---|---|---|---|
| $\lambda$ | | | | | | 0.81(§) |
| *Constant* | -4.26*** | -4.52*** | -1.99*** | -3.36*** | -0.21 | -3.66** |
| *ln* FLOOR SPACE | 0.55*** | 0.54*** | | 0.48*** | | |
| *ln* FLOOR SPACE_*1* | | | 0.46*** | | 0.43*** | 0.42*** |
| *ln* FLOOR SPACE_*2* | | | 0.49*** | | 0.49*** | 0.48*** |
| *ln*FLOOR SPACE_*3* | | | 0.53*** | | 0.50*** | 0.53*** |
| *ln* FLOOR SPACE_*4* | | | 0.56*** | | 0.53*** | 0.60*** |
| $\rho$ *ln* FLOOR SPACE | | | | 0.95*** | 0.66*** | 0.49*** |
| *ln* LOT AREA | 0.06*** | 0.06*** | 0.07*** | 0.11*** | 0.11*** | 0.11*** |
| $\rho$ *ln* LOT AREA | | | | -0.52*** | -0.40*** | -0.40*** |
| *ln* AGE | -0.33*** | -0.31*** | | -0.32*** | | |
| *ln* AGE_*1* | | | -0.21*** | | -0.26*** | -0.20*** |
| *ln* AGE_*2* | | | -0.31*** | | -0.38*** | -0.30*** |
| *ln* AGE_*3* | | | -0.41*** | | -0.42*** | -0.37*** |
| *ln* AGE_*4* | | | -0.46*** | | -0.42*** | -0.43*** |
| $\rho$ *ln* AGE | | | | 0.11** | -0.04 | -0.12 |
| *d_Ordinary house* | 0.08*** | 0.09*** | 0.07*** | 0.14*** | 0.11*** | 0.08*** |
| *d_If added floor space* | 0.07*** | 0.07*** | 0.07*** | 0.07*** | 0.07*** | 0.06*** |
| *d_Noise* | -0.22** | -0.18** | -0.07 | -0.15** | -0.13* | -0.07 |
| *d_Construction error* | 0.007 | -0.003 | -0.03 | -0.03 | -0.05 | -0.03 |
| *d_Moisture* | -0.14 | -0.14 | -0.16* | -0.14 | -0.13 | -0.15* |
| *d_Difficult lot* | -0.37 | -0.35 | -0.35 | -0.38 | -0.39 | -0.35 |
| *d_Renovation object* | -0.40 | -0.38 | -0.38 | -0.47 | -0.47 | -0.38 |
| *d_Radon* | -0.01 | 0.002 | -0.04 | -0.03 | -0.06 | -0.08 |
| *d_No maintenance* | -0.01 | -0.08 | -0.14 | -0.12 | -0.14 | -0.15 |
| *d_Less than 50,000* SEK | 0.30*** | 0.24*** | 0.22*** | 0.22*** | 0.21*** | 0.19*** |
| *d_Other annotations* | 0.13* | 0.10 | 0.07 | 0.13** | 0.09 | 0.07 |
| *d_Own water* | -0.13*** | -0.13*** | -0.13*** | -0.14*** | -0.14*** | -0.17*** |
| *d_No water* | -0.06 | -0.04 | -0.06 | -0.08 | -0.10 | -0.11*** |
| *d_Quality points (16–20)* | 0.18** | 0.17** | 0.17** | 0.21** | 0.22*** | 0.20*** |
| *d_Quality points (21–25)* | 0.36*** | 0.34*** | 0.32*** | 0.39*** | 0.38*** | 0.37*** |
| *d_Quality points (26–35)* | 0.53*** | 0.50*** | 0.49*** | 0.54*** | 0.53*** | 0.52*** |
| *d_Quality points (36–45)* | 0.68*** | 0.64*** | 0.64*** | 0.70*** | 0.69*** | 0.66*** |
| *d_Quality points (46–52)* | 0.85*** | 0.85*** | 0.89*** | 0.93*** | 1.00*** | 0.97*** |
| NET MIGRATION | 0.0008*** | 0.0009*** | 0.0005*** | 0.0008*** | 0.0006*** | 0.0009*** |
| LN UNEMP/POP | 0.07* | 0.08* | 0.06 | 0.13*** | 0.12** | 0.11 |
| INTEREST RATE, 2 YEARS | -0.02** | -0.03*** | -0.03*** | -0.03*** | -0.03*** | -0.03*** |
| MUNICIPALITY TAX | 0.13*** | 0.14*** | 0.09*** | 0.07*** | 0.04* | 0.16*** |
| AVERAGE INCOME | 0.03*** | 0.03*** | 0.02*** | 0.02*** | 0.02*** | 0.03*** |
| *d_Beach* | 0.15*** | 0.14*** | 0.15*** | 0.13*** | 0.15*** | 0.16*** |
| *d_Near beach* | 0.02 | 0.007 | 0.03 | 0.01 | 0.02 | 0.02 |
| *d_Built–up area* | 0.20*** | 0.17*** | 0.12*** | 0.17*** | 0.14*** | 0.15*** |
| ACCESS | 0.002*** | 0.002*** | | 0.0005 | | |
| ACCESS_*1* | | | 0.0005*** | | -0.002*** | -0.001** |
| ACCESS_*2* | | | 0.004*** | | 0.001** | 0.002*** |
| ACCESS_*3* | | | 0.006*** | | 0.003*** | 0.002*** |
| ACCESS_*4* | | | 0.07*** | | 0.009*** | 0.006*** |
| $\rho$ ACCESS | | | | 0.0008 | 0.004*** | 0.004*** |
| *d_Within 5 km of* E*12* | 0.26*** | 0.25*** | 0.28*** | 0.12*** | 0.18*** | 0.12*** |
| *d_Within 5 km of* E*4* | 0.20*** | 0.20*** | 0.22*** | 0.13*** | 0.15*** | 0.10*** |
| CATERGORY_*1* | | 0.12*** | 0.10*** | 0.11*** | 0.10*** | 0.09(§) |
| CATERGORY_*2* | | 0.10*** | 0.09*** | 0.09*** | 0.09*** | 0.08(§) |
| CATERGORY_*3* | | 0.20*** | 0.19*** | 0.19*** | 0.17*** | 0.15(§) |
| CATERGORY_*4* | | 0.23*** | 0.21*** | 0.22*** | 0.21*** | 0.16(§) |
| *Normality* | 11683*** | | | | | |
| *Heterosc. K-B* | 385.0*** | | | | | |
| *Heterosc. K-B with Cat.4* | 38.10*** | | | | | |
| *Heterosc.Wald* | | 248.1*** | 276.6*** | 253.9*** | 243.5*** | |
| *Spatial dep. LMerror* | 2760.4*** | 1720.5*** | 1033.9*** | 1278.9*** | 784.0*** | |
| *Spatial dep. LMerror Robust* | 538.1*** | | | | | |
| *Spatial dep. LMlag* | 2502.2*** | 1538.4*** | | | | |
| *Spatial dep. LMlag Robust* | 279.9*** | | | | | |
| $R^2$ | 0.64 | 0.62 | 0.66 | 0.65 | 0.68 | 0.69 |

*continued on next page*

Table 3: *continued*

| Variable | OLS | FGLS | FGLS | FGLS | FGLS | SAR–GM |
|---|---|---|---|---|---|---|
| $R^2$-*adj.* | 0.64 | | | | | |
| *LIK* | -2410.8 | | | | | |
| *AIC* | 4887.6 | | | | | |
| *SC* | 5099.5 | | | | | |
| *Sq.corr* | | 0.64 | 0.67 | 0.67 | 0.69 | 0.68 |
| *Observations/ Iterations* | 4538 | 4538 | 4538 | 4538 | 4538 | 4538 / 9 |

In order to investigate the effects of heterogeneity and neighborhood spillovers simultaneously the features of the models in column three and four are combined in a new model in column five. The influence of neighborhood spillovers from the age of homes becomes insignificant in 1999. The lagged access variable coefficient are for both years positive. The LM-tests for both years still indicates that spatial error dependence remains. The problem seems worse for the second year.

The final regression shown in column six includes all features from previous regressions, but considers as well the spatial dependence in the error terms. For both years, we use a Spatial Autoregressive Generalized Moments (SAR-GM) estimator, cf. Kelejian and Prucha (1999). The motivation for this choice, instead of the more common Maximum Likelihood estimator, is the fact that SAR-GM accepts non-normal distributed errors (indicated by the OLS estimations) and the inclusion of our additive heteroscedastic variable. The spatial weights matrix used for the error terms is the same as before, a row standardized matrix of the inverse distances between all observations with a distance cut-off at 48 km.

The best fit is produced for the 1994 regression. The autoregressive coefficient $\lambda$ is 0.51 in 1994 and quite high, 0.81 in 1999[2]. The floor space and lot area have, as expected, a positive impact on home values in both years. The importance of floor space is also higher in the more rural municipalities in 1999. The lagged floor space and lot area coefficients are positive and negative respectively for both years. The age of the home has a negative impacts on price. For both years the heterogeneity of this variable is evident. The value of a home increases if the home has been extended during its lifetime. The dummy variable indicating lack of maintenance have a negative and significant parameter value for 1994 but not for 1999.

[2]This parameter is considered as a nuisance parameter with the sole purpose to increase the precision of the other regression parameters. Apart from that it is of little interest and no inference is therefore made on this parameter.

The quality point dummy coefficients all have the expected positive sign, with higher estimates for higher points except for the top class that has a slightly reduced value in 1994. As expected a positive net migration improves the values in both years. The interest rate level is highly significant in 1999 and has the expected negative sign while the parameters for average income are both significant with a positive sign. The municipality tax parameter is positive and significant for 1999.

A waterfront location is as indicated valued positively, at least if the home is located close to the beach. The value of a home is further improved when located in a built-up area. When it comes to the accessibility measures, the dummy variable parameters for the two European roads are significant and positive for both years. However, contrary to expectation, the access coefficient is significant negative for the Umeå observations for both years. The other access coefficient are only significant for the year 1999. The lagged access variable coefficient is significant and positive for both years, and continues to capture some of the surrounding accessibility variations within each municipality not accounted for by the individual accessibility variable.

### 4.2 Space-time comparisons

The next step is to compare the estimated hedonic prices for the two years. Comparisons are only made for variables present for both years and if at least one of the years presents a significant coefficient value (the insignificant value is in that case given the value 0). All comparisons are based on the parameter estimates from the sixth regression model from earlier. The predicted prices for the single-family homes are finally presented graphically.

Initially, we may observe that the importance of floor space has been reduced slightly over time. The negative impact of lagged lot area has instead become larger, a sign of divergence. We may also conclude that the importance of the negative age effect is stronger for all four groups for the latter year. All quality point variables have experienced a decrease in importance except for the top class. Consumers have increased their preferences for small homes with small lots while age and quality are less important. In total, there seems to be an increased demand for homes in densely populated areas

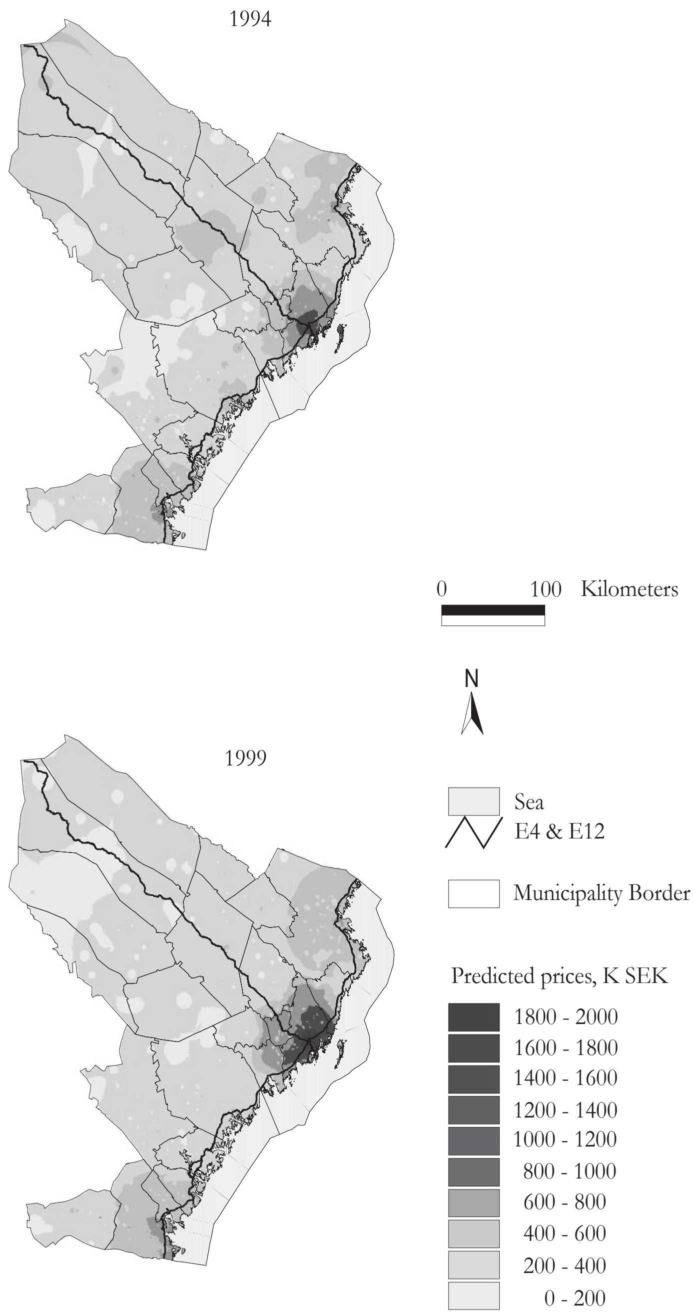


Figure 2: Predicted prices for Single-Family Homes in 1994 and 1999.

with less emphasis on quality. Other parameter estimates that have changed over time are the municipality tax, i.e. the level of public service, and the average income. They are in 1999 more important for the determination of home prices than in 1994. Less important over time are also the attributes that indicate whether homes are located near water or in areas classified as built-up areas. The lagged access parameter value is also reduced in 1999 compared with five years earlier. The importance of proximity to major population centers captured by the access variable has on the other hand increased. The importance of proximity to major roads is still positive but its contribution to the value of a home is reduced.

The regression results are presented in a couple of graphical illustrations. The first two maps in Figure 2 show the spatial pattern of the predicted transaction prices for *ln PRICE* smoothed across the area of investigation, one for each year. The reader may clearly detect the concentration of high prices around the two regional centers and in some parts of the coast line. Lower prices are found in the inland area.

An alternative way of displaying how this real estate market evolved during this period is displayed on a third map, Figure 3. Here, the difference between the predicted prices for the two years is illustrated, in terms of standard deviations from the mean. The darker gray areas in the top map have experienced a significant increase in prices during the period. This is particularly noticeable in the Umeå region, an indication of regional expansion. However, this regional expansion is rather concentrated from a geographic perspective since the infrastructure around Umeå has not kept pace with the region's growth. The regional expansion should instead be viewed in term of increased values and they have in fact increased significantly within the commuting distance. Some minor increases are also found both north and south of Umeå in the inland areas to the southwest of Umeå. The significant increase in this inland region is mostly a sign of a recovery from very low prices during the down swing in the business cycle in the early 1990's.

The opposite applies to the western part of the lower map that show negative price differences. The previous boom in the skiing resort area has now subsided and as a result we see lower prices in this area as well as many other parts of the inland.

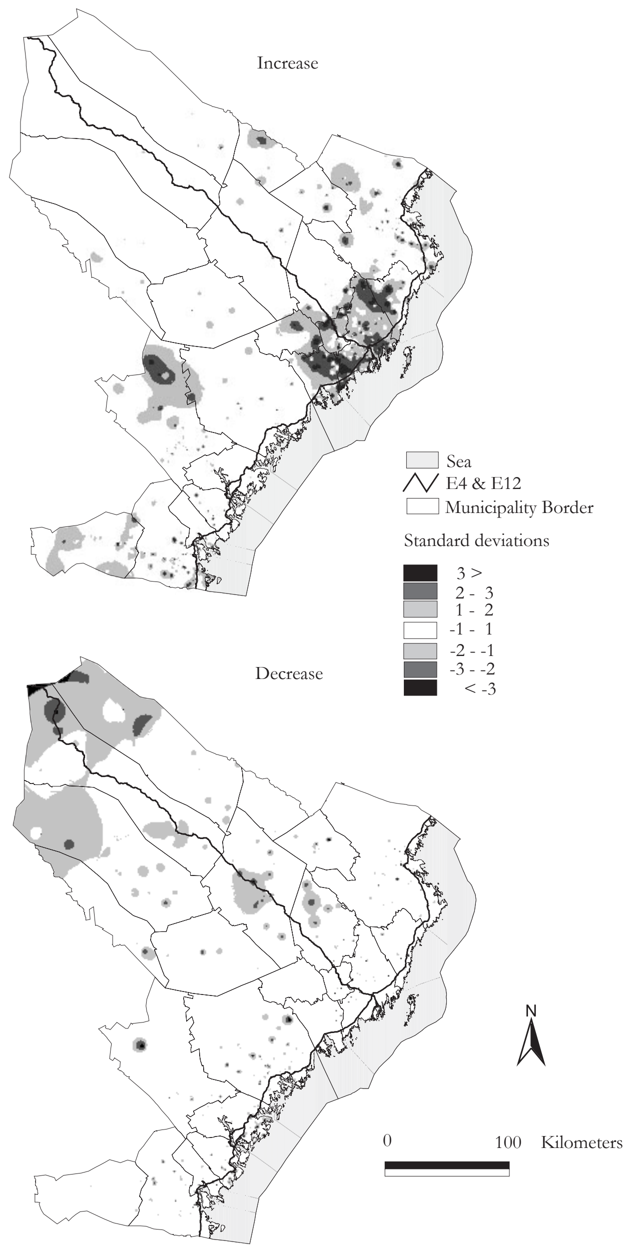


Figure 3: Changes in predicted prices between 1994 and 1999.

## 5 Conclusions

This paper has dealt with hedonic prices, or, the valuation of attributes, connected with purchases of single-family homes in northern Sweden for two years using hedonic price theory and spatial econometrics. The data covered a large part of the real estate market with objects in both densely and sparsely populated areas. The models was despite these heterogeneities able to explain about 70% of the variances for both years. A large part of this paper has been occupied with the treatment of spatial heterogeneity and spatial dependence in the data. When the model specification is expanded to include these effects the parameter estimates changes in sign, level, and significance. The parameter changes between the two years was then compared and illustrated in a series of maps.

The quality of homes at the time of purchase has become less important to the buyer. This is in accordance with the increased negative influence we have observed for the age variable. Instead greater weight is put on location variables, the population migration to the region, and regional household incomes, as important factors determining home values. The observed changes in the price landscape between the two years indicate a concentration of high valued homes in the leading municipal areas of Umeå and Sundsvall. We may also observe a significant increase of home values in the municipalities surrounding Umeå itself, a sign of increased inter-dependence among the municipalities in this area and may conclude that this region has had an outward growth within commuting distance, also referred to as “Urban Sprawl”, during the five years. The other important region in this part of the country is the Sundsvall region. However, Sundsvall does not show any strong signs of expansion nor retraction during the study period. A decrease in prices is found in large parts of the inland west.

It is difficult to explain the reasons behind these price changes using only the data that is available to us. One possibility is of course changes caused by changes in the preferences of homebuyers. Another plausible explanation is increased/decreased demand for some attributes that has not been met by a corresponding supply increase/decrease. Hence, as we don’t have a complete data set for the entire real estate population, for each year in the study, we are unable to more thoroughly explore the cause.

Further research would take these matters into consideration and additionally include transaction data for additional years in order to improve the dynamic analysis of the market for homes. Since the data exhibits spatial dependence, future research must take this under consideration as well. Improvement may also include additional information on municipality expenses as well as governmental grants.

## References


[1] Andersson, R. (1998). *Attraktiva Städer - en samhällsekonomisk analys*. Byggforskningsrådet Svensk Byggtjänst, Stockholm.

[2] Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Kluwer Academic, Dordrecht.

[3] Anselin, L. (1990). Spatial dependence and spatial structural instability in applied regression analysis. *Journal of Regional Science* 30, 185–207.

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

[5] Anselin, L., and A. Bera (1998). Spatial dependence in linear regression models with an introduction to spatial econometrics. In: Ullah, A., and D. E. A. Giles (eds). *Handbook of applied economic statistics.*

[6] Basu, S., and T. G. Thibodeau (1998). Analysis of Spatial Autocorrelation in House Prices. *Journal of Real Estate Finance and Economics* 17 (1).

[7] Benson, E. D., J. L. Hansen, A. L. Schwatz Jr., and G. T. Smersh (1998). Pricing Residential Amenities: The Value of a View. *Journal of Real Estate Finance and Economics* 16 (1).

[8] Beron, K., J. Murdoch, and M. Thayer (2001). The Benefits of Visibility Improvement: New Evidence from the Los Angeles Metropolitan Area. *Journal of Real Estate Finance and Economics* 22 (2/3).

[9] Blomquist, G. C. (1998). Valuing Urban Lakeview Amenities Using Implicit and Contingent Markets. *Urban Studies* 25, 333–340.

[10] Blomquist, G. C., M. C. Berger, and J. P. Hoehn (1998). New Estimates of Quality of Life in Urban Areas. *The American Economic Review* 78 (1), 89–107.

[11] Bogart, W. T., and B. A. Cromwell (2000). How Much Is a Neighborhood School Worth? *Journal of Urban Economics* 47, 280-305.

[12] Bond, M. T., V. L. Seiler., and M. J. Seiler (2002). Residential Real Estate Prices: A Room with a View. *Journal of Real Estate Research* 23 (1/2), 129–137.

[13] Brasington, D. M. (1999). Which Measures of School Quality Does the Housing Market Value? *Journal of Real Estate Research* 18 (3), 395–413.

[14] Can, Y. (1992). Specification and estimation of hedonic housing price models. *Regional Science* 22, 453–474.

[15] Can, A., and I. Megbolugbe (1997). Spatial Dependence and House Price Index Construction. *Journal of Real Estate Finance and Economics* 14 (2).

[16] Cheshire, P., and S. Sheppard (1995). On the Prise of Land and the value of Amenities. *Economica* 62, 247–267.

[17] Clark, D. E., and T. Allison (1999). Spent Nuclear Fuel and Residential Property Values: The Influence of Proximity, Visual Cues and Public Information. *Papers in Regional Science* 78, 403–421.

[18] Cliff, A., and K. Ord (1972). Testing for Spatial Autocorrelation Among Regression Residuals. *Geographical Analysis* 4, 267–284.

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

[20] Des Rosiers, F., M. Thériault, Y. Kestens, and P. Villeneuve (2002). Landscaping and House Values: An Empirical Investigation. *Journal of Real Estate Research* 23 (1/2), 139–161.

[21] Englund, P., J. M. Quigley, and C. L. Redfearn (1998). Improved Price Indexes for Real Estate: Measuring the Cource of Swedish Housing Prices. *Journal of Urban Economics* 44, 171–196.

[22] Florax R. J. G. M., and H. Folmer (1992). Specification and Estimation of of Spatial Linear Regression Models: Monte Carlo Evaluation of Pre-test Estimators. *Regional Science and Urban Economics* 22, 405–432.

[23] Florax R. J. G. M., H. Folmer., and S. J. Rey (2003). Specification Searches in Spatial Econometrics: The Relevance of Hendry's Econometric Methodlogy. *Regional Science and Urban Economics* 33, 557–579.

[24] Haas, G. C. (1922). *A Statistical Analysis of Farm Sales in Blue Earth County, Minnesota, as a Basis for Farm Land Appraisal.* Master Thesis, The University of Minnesota.

[25] Kelejian, H. H., and I. R. Prucha (1999). A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model. *International Economic Review* 40 (2), 509–533.

[26] Koenker, R. and G. Basset (1982). Robust tests for heteroskedasticity based on regression quartiles. *Econometrica* 50, 43-61

[27] Lancaster, K. (1966). A new approach to consumer theory. *Journal of Political Economy* 74, 132–157.

[28] Lindgren, P., and G. Rosberg (1992). *Vem bor till vilket pris?* Boverket, Karlskona.

[29] Pace, R. K., and O. W. Gilley (1997). Using the Spatial Configuration of the Data to Improve Estimation. *Journal of Real Estate Finance and Economics* 14 (3).

[30] Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy* 34–55.

[31] Shultz, S. D., and D. A. King (2001). The Use of Census Data for Hedonic Price Estimation of Open Space Amenities and Land Use. *Journal of Real Estate Finance and Economics* 22 (2/3).

[32] Sivitanidou, R. (1996). Do Office-Commercial Firms Value Access to Service Employment Centers? A Hedonic Value Analysis within Polycentric Los Angeles. *Journal of Urban Economics* 40, 125–149.

[33] Smersh, G. T., and M. T. Smith (2000). Accessibility Changes and Urban House Price Appreciation: A Constrained Optimization Approach to Determining Distance Effects. *Journal of Housing Economics* 9, 187–196.

[34] Thompson, B., and M. Hills (1999). Wired Up for Extra Value. *Journal of Real Estate Research* 17 (1/2), 245–255.

[35] Tse, R. Y. C. (2002). Estimating Neighborhood Effects in House Prices: Towards a New Hedonic Model Approach. *Urban Studies* 39 (7), 1165–1180.

[36] Wigren, R. (1987). House Prices in Sweden: The Significance of Attributes. *Scandinavian Housing and Planning Research* 4, 243–261.

# Growth of GRP in Chinese Provinces: A Test for Spatial Spillovers*

Krister Sandberg†
Department of Economics and CERUM
Umeå University
SE-901 87 Umeå
Sweden
krister.sandberg@econ.umu.se

**Abstract**

This paper examines the provincial pattern of growth in China during the period 1985–2000, testing the hypothesis that provinces with similar growth rates are more spatially clustered than would be expected by chance. The provincial economic growth is explained by the distribution of industrial enterprises, foreign direct investment, infrastructure, and governmental preferential policies. The neoclassical hypothesis of convergence is also tested. Indications of unconditional convergence does occur during the periods 1985–2000 and 1985–1990. In addition, conditional convergence is found during the sub-period 1990–1995. Evidence of spatial dependence between adjacent provinces has also been established, and in the econometric part, solved by a spatial lag, or alternatively a spatial error term, in the growth equation.

***Keywords* :** GRP-growth, Chinese provinces, Spatial dependence
***Classification*[*JEL*] :** O18, R11, R12

---

*An earlier version of this paper was published in IIASA IR-02-080. http://www.iiasa.ac.at/Publications/Documents/IR-02-080.pdf

†The author wishes to thank IIASA and FORMAS for financial support. Moreover Associate Professor Lars Westin and Assistant Professor Johan Lundberg from CERUM, Professor Peter Rogerson, and the participants at the 16th European Advanced Studies Institute in Regional Science, June 23–27, 2003 in Ascona, Switzerland have all given valuable comments on earlier versions of the paper.

## 1 Introduction

The purpose of this paper is to explain differences in economic growth among Chinese provinces where spatial dependence is tested and controlled for. Thus, our task is to determine the factors behind the spatial growth pattern, to test for conditional and unconditional convergence, as well as consider the fact that provinces may be dependent on each other in positive and negative ways in order to avoid biased and inefficient estimates.

During the last decades, China has experienced an exceptionally high economic growth. Expressed in year 2000 prices, the Gross Domestic Product (GDP) per capita has risen from 855 Yuan in 1985 to 7,078 Yuan in 2000. This increase in wealth is, however, unequally distributed. In the year 2000, the Shanghai province had the highest level of Gross Regional Product (GRP) per capita (27,187 Yuan), compared with the poorest province, Guizhou (2,818 Yuan). In Figure 1 below, the income pattern is presented with GRP per capita levels for the year 2000.

The income differences have not always been this large. In fact, they were actually reduced somewhat when the reforms started in 1978, as the now successful provinces began their rapid growth from a lower level. According to Démurger (2001), income disparities did not start to increase until the second part of the 1980's.

When the communists came to power in 1949, one of their particular objectives was to provide equal wealth to the whole population (disparities between urban and rural areas were, however, accepted). This was accomplished through a strong central policy, redistribution of incomes and resources from wealthy to poor provinces, and large-scale investments in the poorer provinces. In 1978, this system was abandoned in favor of reforms such as decentralization of the agricultural production, decentralization of the fiscal system, diversification of the ownership structure, and especially the introduction of the Open Door Policy. The Open Door Policy started on a small scale in the early 1980's when areas within the provinces of Guangdong and Fujian were given the status of Special Economic Zones in order to attract foreign investments. In the mid 1980's, this expanded to other areas opened for increased international trade and foreign investments. New economic zones were created throughout the country in the early 1990's.

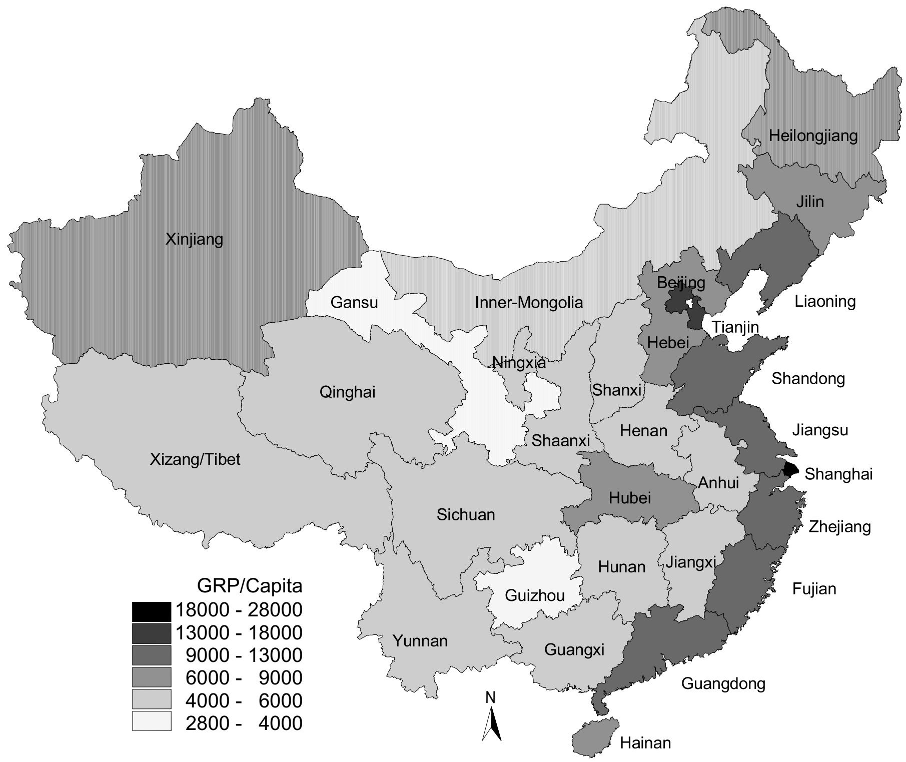


Figure 1: GRP per capita in the Chinese provinces for the year 2000. Source: China Statistical Yearbook 2001.

Today, the three metropolises, Beijing, Shanghai, and Tianjin, are industrialized and have the highest GRP per capita. The coastal provinces in the southeast have experienced a rapid growth in GRP per capita since the reforms started in 1978, and are now among the richest provinces in the country. These provinces have a special status in relation to the other provinces due to the preferential policies levied upon them by the government, and are generally considered the new engines of growth in the Chinese economy. In the northeast we find the three provinces, Heilongiang, Jilin, and Liaoning, collectively called Manchuria. This area used to be China's industrial center with the highest GRP per capita in China. Even though Manchuria has not experienced as rapid of a growth as the southern coastal provinces the GRP per capita is still among the highest in the country. The central provinces, between the rivers Yellow and Yangtze, have a high population density and are well-suited for agriculture. The southwestern provinces are also, from a climatic perspective, suited for agriculture but are hard to access due to the mountainous terrain. These provinces have had, in general, a low annual GRP per capita growth since the start of the reforms. The northwestern part of China consists of the the provinces Tibet, Xinjiang, and Qinghai. These provinces are characterized by high elevation and a low degree of transport infrastructure.

Hence, this short introduction have shown that the regional growth pattern in China may, at least to some extent, be explained by factors related to policy and resource endowments. Additionally, we are also interested in potential impacts of growth spillovers between provinces. The next two sections addresses the theory of economic growth and spatial dependence in connection with previous studies on provincial growth in China, as well as a presentation of supporting data. The forth section consists of an exploratory data analysis in search of spatial dependence. Estimations of the provincial economic growth equations are explored in the fifth section. The final section of this paper concludes with a presentation of our findings.

## 2 Theory of Economic Growth and Spatial Dependence - The China Case

A large part of the empirical literature on regional growth, e.g. Barro and Sala-i-Martin (1992, 1995), Persson (1997), and Sala-i-Martin (1996) are concerned with the convergence hypothesis, as predicted by the neoclassical growth theory, given by Solow (1956), Swan (1956), and Koopmans (1965). That is, provinces with an initial low growth will eventually catch up with the richer ones since their capital/labor ratio is below it's long run value and thus has higher rates of return, therefore growing faster. Given that all provinces are intrinsically the same, apart from their initial capital/labor ratios, convergence would be unconditional. If we allowed, however, the provinces to be different in various aspects, the convergence would instead be conditional. Each province would instead converge toward its own steady state level of growth.

The hypothesis of convergence, however, has been rejected in many studies of nations in favor of endogenous growth theory, Romer

(1986) and Lucas (1988), where the long-term growth rate of output per worker is determined by variables within the model, such as accumulation of human and physical capital.

The analysis of regional economic growth is a recent contribution to the economic growth literature, e.g. Nijkamp and Poot (1998), Bal and Nijkamp (1998), Rey and Montouri (1999), Vayá et al. (2000), Wheeler (2001), and Carrington (2002). Since countries and especially regions, interact with each other in various ways potential estimation problems caused by spatial dependence may occur. These problems are apparent in China, especially with its division of growth between the western and eastern part of the country. Therefore, the presence of two types of spatial dependence (Spatial Lag dependence and Spatial Error Dependence) are tested and controlled for in this paper. Spatial Lag Dependence is present if spatial correlation in the dependent variable exists between observations. This means that the rate of growth in one province influences, and is influenced by, growth rates in nearby provinces, cf. Anselin (1988) and Can (1992). If this problem is ignored, the OLS estimates will be biased and inefficient and hence lead to incorrect inference. The solution is to add a spatial lag to the growth equation:

$$\boldsymbol{g} = \rho \boldsymbol{W} \boldsymbol{g} + \boldsymbol{Z} \boldsymbol{\beta} + \boldsymbol{\varepsilon} \qquad (1)$$

where, $\boldsymbol{g}$ is a ($n$ by 1) vector of observations on the dependent variable, $\boldsymbol{Z}$ is a ($n$ by $k$) matrix of observations on the exogenous variables with $\boldsymbol{\beta}$ as the associated ($k$ by 1) vector of regression coefficients, $\boldsymbol{\varepsilon}$ is a ($n$ by 1) vector of random error terms, $\rho$ is the autoregressive coefficient, $\boldsymbol{W}$ is a ($n$ by $n$) spatial weights matrix, with elements $w_{ij}$ corresponding to observation pair $i$ and $j$. Finally, $\boldsymbol{Wg}$ is the spatially lagged dependent variable, a weighted average of other regions.

Spatial Error Dependence is present when the error terms show correlation with the error terms of adjacent observations, i.e., lack of stochastic independence between observations, e.g. Cliff and Ord (1972, 1973). The standard error assumptions under normality of the linear regression model are violated and as a result inefficient estimates are produced. The solution is to incorporate the spatial dependence in the growth equation via an autoregressive error term:

$$\begin{aligned} \boldsymbol{g} &= \boldsymbol{Z\beta} + \boldsymbol{\varepsilon} \\ \boldsymbol{\varepsilon} &= \lambda \boldsymbol{W\varepsilon} + \boldsymbol{\xi} \end{aligned} \qquad (2)$$

where, $\boldsymbol{W\varepsilon}$ is a spatial lag for the error term, $\lambda$ is the autoregressive coefficient, and $\boldsymbol{\xi}$ is a ($n$ by 1) vector of well behaved error terms $\boldsymbol{\xi} \backsim N(0, \sigma^2 \boldsymbol{I})$. We will return to these two models in Section 5 below.

A review of the literature may serve as an introduction in the search for determinants behind the provincial economic growth. The literature examining the Chinese economy and its spatial income disparities is vast. Many contributions consider the question of convergence, both conditional and unconditional, e.g. Chen and Fleisher (1996), Tian (1999), and Yao and Zhang (2001). Among the explanations behind provincial growth we may find factors related to physical and human capital, institutions, and spatial spillovers.

*The physical and human capital.* The infrastructure as studied by Yao and Zhang (2001) is usually measured as the sum of the length of railway, highway, and waterway per area unit converted into equivalent highways, based on the transport work of each mode. Foreign Direct Investments (FDI) have thus far been important in explaining income disparities in the Chinese economy, Graham and Wada (2001). Zhang and Kristensen (2001) argue that FDI should, in principle, enlarge the disparities, but are unable to find evidence to support their argument. The importance of human capital is also acknowledged in the literature of the Chinese economy. Human capital is often measured as enrollment in higher education divided by the working population or the total population, e.g. Chen and Feng (2000).

*The Chinese institutions.* Geographical differences, accessibility, and governmental policy is often accounted for by dummy variables for the coastal provinces in order to explain growth divergence between coastal and non coastal provinces. An alternative solution is presented by Démurger et al. (2002). Démurger uses a preferential policy index based on the different degrees of openness among the provinces. Additionally, Démurger argues that the topography, measured as the average elevation and slope of the province, is an important factor behind growth.

State Owned Enterprises (SOEs) are generally considered less competitive than other forms of ownership. A large share of these enterprises have had a negative effect on income growth, as shown by e.g. Chen and Feng (2000). This may, however, be explained by the kind of industries they generally are involved in, such as strategically important production and defense related industries. Instead, Démurger (2001) uses the share of collectively owned enterprises of total industrial production to control for the internal reform process. Oi (1999), on the other hand, explores the role of local authorities in the economic transition from 1978 to the mid 1990's and concludes that the most important factor for local growth is the property rights to means of production.

*Spillover effects.* The only study we have found that consider spatial dependence in China, Ying (2000), is limited to an exploratory data analysis of the existence of dispersion or spillover effects from the core to the periphery provinces. The author not only found evidence of economic spillovers from the Guangdong province to nearby provinces, but also a pattern of polarization. It is also concluded that preferential policies play a major role in the direction of this process.

Compared to the studies above, our paper contributes to the literature on economic growth in China not only by identifying the presence of previously overlooked problems of spatial dependence, but also by solving for this problem by inclusion of a spatial lag, alternatively a spatial error term in the growth equation as needed.

## 3 The Chinese Provincial Data

Four time periods are defined in this study; 1985–2000, 1985–1990, 1990–1995, and 1995–2000, with data from China's 30 provinces. All economic variables are measured in year 2000 prices. Data was collected from various China Statistical Yearbooks, (National Bureau of Statistics of China, 1986–2001) and Hsueh et al. (1993). Additional data were gathered from the LUC project database at IIASA[1]. The descriptive statistics for the selected variables are presented in Table 1.

The dependent variable, GRPC, is the average annual per capita growth rate over each specific time period. For the whole period 1985–2000, the average annual per capita growth rate was 6.64%,

[1] Modeling Land-Use and Land-Cover Changes. http://www.iiasa.ac.at/Research/LUC/

Table 1: Descriptive Statistics and Expected Signs

| Variable | Unit | Mean | St.dev | Min | Max | Sign |
|---|---|---|---|---|---|---|
| GRPC_8500 | | 6.64 | 1.69 | 3.41 | 11.30 | |
| GRPC_8590 | % | 2.12 | 2.26 | -5.82 | 6.04 | |
| GRPC_9095 | | 10.20 | 4.02 | 3.73 | 19.42 | |
| GRPC_9500 | | 7.58 | 1.79 | 3.54 | 12.05 | |
| EDUP_1985 | Graduates/capita | 0.0009 | 0.0006 | 0.0004 | 0.003 | (+) |
| EDUP_1990 | | 0.0014 | 0.0010 | 0.0007 | 0.005 | |
| EDUP_1995 | | 0.0016 | 0.0010 | 0.0007 | 0.005 | |
| SOE_TE_1985 | SOE/Total | 0.27 | 0.12 | 0.08 | 0.67 | (-) |
| SOE_TE_1990 | | 0.24 | 0.12 | 0.08 | 0.64 | |
| SOE_TE_1995 | | 0.26 | 0.12 | 0.12 | 0.63 | |
| TPAREA_1985 | $km/km^2$ | 0.24 | 0.15 | 0.014 | 0.63 | (+) |
| TPAREA_1990 | | 0.26 | 0.17 | 0.016 | 0.71 | |
| TPAREA_1995 | | 0.29 | 0.20 | 0.019 | 0.81 | |
| DINVC_8500 | | 1760.51 | 1576.94 | 485.40 | 7928.41 | |
| DINVC_8590 | 10' rmb/capita | 1057.80 | 836.76 | 322.25 | 3929.98 | (+) |
| DINVC_9095 | | 1649.41 | 1462.16 | 403.50 | 7021.10 | |
| DINVC_9500 | | 2574.30 | 2479.02 | 730.45 | 12834.17 | |
| FDIC_8500 | | 22.20 | 33.62 | 0.001 | 124.60 | |
| FDIC_8590 | USD/capita | 4.62 | 8.82 | 0.003 | 32.94 | (+) |
| FDIC_9095 | | 24.02 | 36.58 | 0 | 132.99 | |
| FDIC_9500 | | 37.96 | 57.72 | 0 | 215.93 | |
| PREF_8500 | | 1.35 | 0.81 | 0.56 | 3 | |
| PREF_8590 | Index | 0.78 | 1.10 | 0 | 3 | (+) |
| PREF_9095 | | 1.47 | 0.79 | 0.67 | 3 | |
| PREF_9500 | | 1.80 | 0.66 | 1 | 3 | |

within a range between 3.41 and 11.30. The largest spread of growth for a sub-period, 3.73 to 19.42%, was found in the first five years of the 1990's.

The proxy variable for human capital, *EDUP*, is measured as the number of graduates from Institutions of Higher Education and Specialized Secondary Schools, divided by the total population in each province for the years 1985, 1990, and 1995 respectively. This share increased from an average of 0.0009 in 1985, to 0.0016 ten years later. The coefficient sign is expected to be positive.

The transport capacity in each province is captured by the variable *TPAREA* for the same three years. It is measured as the total length of railways in operation, navigable inland waterways, and highways in kilometers/$km^2$. As expected, the capacity has increased over the years, as has the spread between the best and the worst province. The sign is expected to be positive.

The capital accumulation in each province is captured by two variables. *DINV* is measured as the annual domestic investment in 10,000 rmb/capita averaged over the actual time period. The amount of Foreign Direct Investment, *FDIC*, is measured in USD/capita averaged over the actual time period. Both variables show an increase over time and both are expected to result in positive signs.

The next two variables in Table 1 are included to characterize the institutional structure of the provinces. SOE_TE, is the number of state owned enterprises divided by the total number of enterprises. The average share is about 25%, but in some provinces, more than 60% of the companies are state owned. Since these enterprises are generally considered to be less profitable, the expected sign of the coefficient is negative. The preferential policy, PREF, levied by the government upon each province, is constructed as an index, based on the degree of openness. Following Démurger et al. (2002), the index is constructed in 4 groups with different weights, as shown in Table 2. These weights are then averaged over specific time periods. The coefficient sign is expected to be positive.

Table 2: Preferential Policy Index

| Variable | Weight |
|---|---|
| *No open zone* | 0 |
| *Coastal Open Cities*<br>*Coastal Open Economic Zones*<br>*Open Coastal Belt*<br>*Major Cities along the Yantze river*<br>*Bonded Areas*<br>*Capital Cities of inland provinces and autonomous regions* | 1 |
| *Economic and Technological Development Zones*<br>*Border Economic Cooperation Zones* | 2 |
| *Special Economic Zones*<br>*Shanghai Pudong New Area* | 3 |

In order to measure the impact of population density, or, agglomeration effects in the three metropolis provinces, Beijing, Shanghai, and Tianjin, the dummy variable D_CITY, was introduced. It is assigned the value one for these provinces. A positive coefficient sign is expected. The southeast provinces, Guangdong and Fujian, have historical as well as geographical advantages compared to the other provinces. Guangdong is a neighbor to Hong Kong, and many of the Taiwanese have close ties with the Fujian population. The role of intensive external relations is thus tested with the dummy variable, D_EXTERNAL, with an expected positive sign.

## 4 Spatial Exploratory Data Analysis

Before we proceed and estimate the growth equations, let us first test the hypothesis that provinces with similar growth rates are more spatially clustered than would normally be expected. One test often used to indicate the possibility of global spatial autocorrelation is the Moran's I test. A similar, but less known test, is the Geary's C test. To complement and validate these results, the Local Moran's I test is utilized.

### 4.1 Global Spatial Autocorrelation

The Moran's I test is defined as:

$$I = \frac{\mathrm{n}}{\mathrm{S}} \frac{\sum_i \sum_j w_{ij}(x_i - \mu)(x_j - \mu)}{\sum_i (x_i - \mu)^2} \tag{3}$$

where n is the number of observations and $x_i$ and $x_j$ are the observed growth rates in locations $i$ and $j$ (with mean $\mu$). S is a scaling constant given by the sum of all weights:

$$\mathrm{S} = \sum_i \sum_j w_{ij} \tag{4}$$

When row standardized weights are used, which is preferable, Anselin (1995a), S equals n since the weights of each row adds to one. The test statistic is compared with its theoretical mean, $\mathrm{I} = -1/(\mathrm{n}-1)$. So, $\mathrm{I} \to 0$ as $\mathrm{n} \to \infty$. The null hypothesis $\mathrm{H}_0 : \mathrm{I} = -1/(\mathrm{n}-1)$ is tested against the alternative hypothesis $\mathrm{H}_a : \mathrm{I} \neq -1/(\mathrm{n}-1)$. When $\mathrm{H}_0$ is rejected and $\mathrm{I} > -1/(\mathrm{n}-1)$, indicates positive spatial autocorrelation. That is, high values and low values are more spatially clustered than would be assumed purely by chance. For the other event, if $\mathrm{H}_0$ is again rejected but $\mathrm{I} < -1/(\mathrm{n}-1)$, it indicate negative spatial autocorrelation. Hence observations with high and low values are systematically mixed together.

The second global measure is the Geary's C test, defined as:

$$C = \frac{\mathrm{n\text{-}1}}{2\mathrm{S}} \frac{\sum_i \sum_j w_{ij}(x_i - x_j)^2}{\sum_i (x_i - \mu)^2} \tag{5}$$

The theoretical expected value for Geary's C is 1. A value of C less than 1 indicates positive spatial autocorrelation, and a value above 1 indicates negative spatial autocorrelation.

Obviously, these tests are quite crude. One apparent drawback is the a priori choice of the spatial weights matrices. However, when repeated with different weights matrices, this becomes a test of the robustness of the weights matrix, its performance, and the kind of relationships that may be hidden in the data.

Four different weights matrices are tested in this paper to mimic the economic integration between the provinces; the $1^{st}$ and $2^{nd}$ order contiguity, denoted *QUEEN_1*, and *QUEEN_2* (where neighbors are defined as those that share a common border) respectively; and two inverse distance matrices using distance and squared distance (arc great circle distance between the province capitals), denoted *DIST_1* and *DIST_2*. All matrices are row standardized. The results from the two global tests for the four weights matrices are presented in Table 3.

Table 3: Moran's I test and Geary's C test for Spatial Autocorrelation between the Chinese provinces. (* = using 999 permutations since the normal distribution was in this case rejected by the Wald test and prevented the use of the normal approach.)

| Variable | I | Mean | St.Dev | Prob | C | Mean | St.Dev | Prob |
|---|---|---|---|---|---|---|---|---|
| QUEEN_1_*8500* | 0.22 | -0.03 | 0.12 | 0.03 | 0.70 | 1.00 | 0.13 | 0.02 |
| QUEEN_2_*8500* | 0.14 | -0.03 | 0.08 | 0.02 | 0.81 | 1.00 | 0.10 | 0.05 |
| DIST_1_*8500* | 0.04 | -0.03 | 0.04 | 0.04 | 0.90 | 1.00 | 0.05 | 0.03 |
| DIST_2_*8500* | 0.09 | -0.03 | 0.10 | 0.21 | 0.84 | 1.00 | 0.10 | 0.10 |
| QUEEN_1_*8590** | -0.11 | -0.03 | 0.11 | 0.24 | 1.02 | 1.00 | 0.16 | 0.39 |
| QUEEN_2_*8590** | 0.04 | -0.03 | 0.07 | 0.16 | 0.78 | 1.00 | 0.10 | 0.03 |
| DIST_1_*8590** | -0.08 | -0.03 | 0.04 | 0.08 | 0.93 | 1.00 | 0.05 | 0.14 |
| DIST_2_*8590** | -0.17 | -0.03 | 0.10 | 0.05 | 1.04 | 1.00 | 0.10 | 0.33 |
| QUEEN_1_*9095* | 0.35 | -0.03 | 0.12 | 0.00 | 0.59 | 1.00 | 0.13 | 0.00 |
| QUEEN_2_*9095* | 0.25 | -0.03 | 0.08 | 0.00 | 0.69 | 1.00 | 0.10 | 0.00 |
| DIST_1_*9095* | 0.10 | -0.03 | 0.04 | 0.00 | 0.86 | 1.00 | 0.05 | 0.00 |
| DIST_2_*9095* | 0.23 | -0.03 | 0.10 | 0.01 | 0.73 | 1.00 | 0.10 | 0.01 |
| QUEEN_1_*9500* | -0.00 | -0.03 | 0.12 | 0.80 | 0.89 | 1.00 | 0.13 | 0.41 |
| QUEEN_2_*9500* | -0.01 | -0.03 | 0.08 | 0.76 | 0.92 | 1.00 | 0.10 | 0.42 |
| DIST_1_*9500* | -0.01 | -0.03 | 0.04 | 0.58 | 0.93 | 1.00 | 0.05 | 0.12 |
| DIST_2_*9500* | 0.00 | -0.03 | 0.10 | 0.72 | 0.92 | 1.00 | 0.10 | 0.43 |

The Moran's I values are positive and significant for the two periods 1990–1995 and 1985–2000, to indicate that provinces with similar growth rates are more clustered than may be assumed purely by chance. For the period 1985–2000 the significant I value is limited to the three first weights matrices. This might be an effect of the relatively steep decline of influence as the distance increases. It suggests that the provinces except for direct neighbors are relatively isolated from each other. The results from the Geary's C test (the right hand side of Table 3) confirms the previous results with some exception for the first sub period with the second and forth weights matrices. The main conclusion from this exploratory examination is that there exists clusters of provinces with similar growth rates irrespective of the weights matrix used.

## 4.2 Local Spatial Autocorrelation

With the help of Moran Scatterplots that show the individual I values for each province from the global Moran's I test, and Local Moran's I tests[2], Anselin (1995b), the investigation continues in search of local spatial autocorrelation, or hot/cold spots.

The Local Moran's I test investigates whether the values for each province (from the global Moran's I) are significant or not.

$$I_i = \frac{x_i}{\sum_i x_i^2} \sum_j w_{ij} x_j \tag{6}$$

The results are, in order to improve readability, presented in a series of maps (Figure 2–Figure 8).

The results for the long period, 1985–2000, are shown in Figure 2 and Figure 3. The Moran Scatterplots (individual I values of the global Moran's I test without significance considerations) for the four weights matrices all show the same area of Low-Low values (provinces with low growth values surrounded by provinces with low growth values) from Tibet across China to the northeast, with some minor deviations. A half circle of High-High values (provinces with high growth surrounded by provinces with high growth) is visible in the southeast.

On the other hand, the Local Moran's I in Figure 3 reveals that this pattern is not as strong as first expected. Only two provinces, Fujian and Zhejiang have significant positive values, while the province Qinghai has a significant negative value for the *QUEEN_1* matrix. Local Moran's I, for the *DIST_1*, shows significant High-High values for Fujian, Zhejiang, and Jiangsu, and Low-High for the Shanghai province. For *DIST_2*, the Local Moran's I is even more limited and reports only two High-High provinces, Fujian and Zhejiang.

To see whether this pattern is stable over the whole 15-year period, the material was once again divided into the three five-year sub-periods.

For the first sub-period, 1985–1990, (Figure 4 and Figure 5) the clear pattern from the previous figures has disappeared. The Moran Scatterplot with *QUEEN_1* only reports some clusters of high values in Manchuria and in the southeast. The Local Moran's I shows some

[2]Tests with the *New* $G_i$*, Ord and Getis (1995) were also performed and was found to be in line with the other test results presented below and are therefore not presented here in order to save space. They are available upon request.

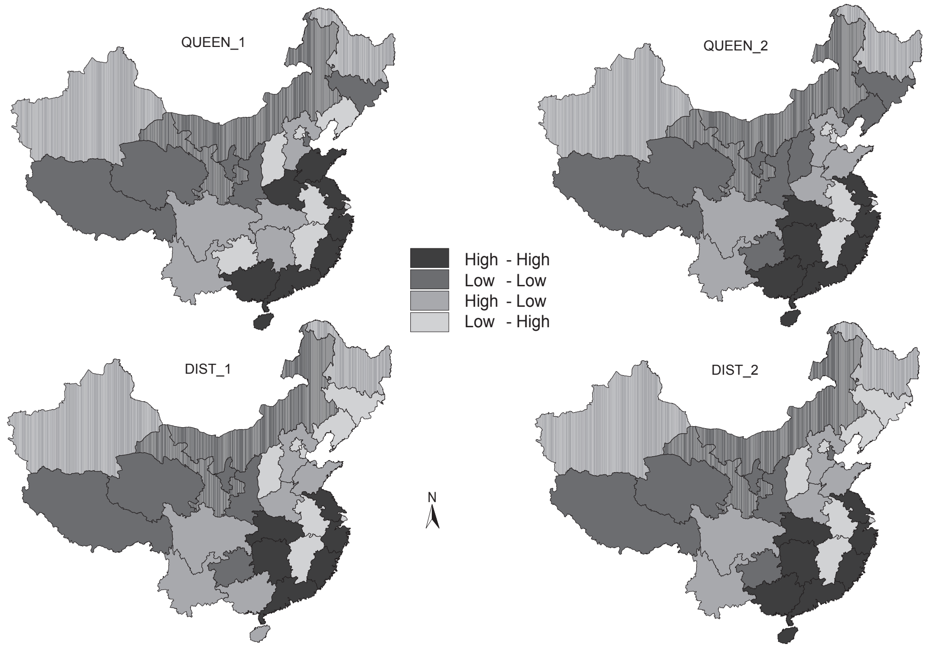


Figure 2: Moran Scatterplot 1985-2000.

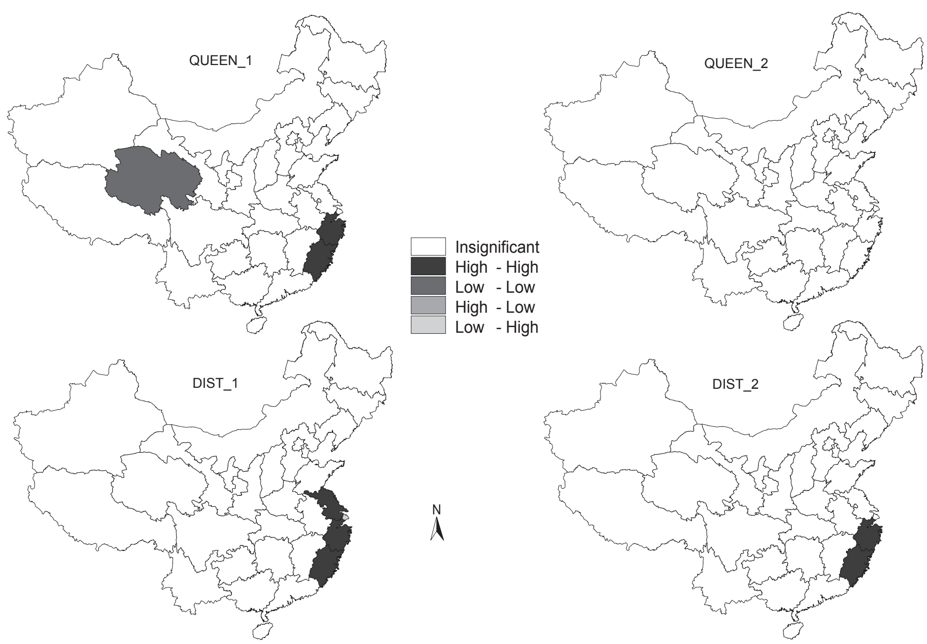


Figure 3: Local Moran's I 1985-2000.

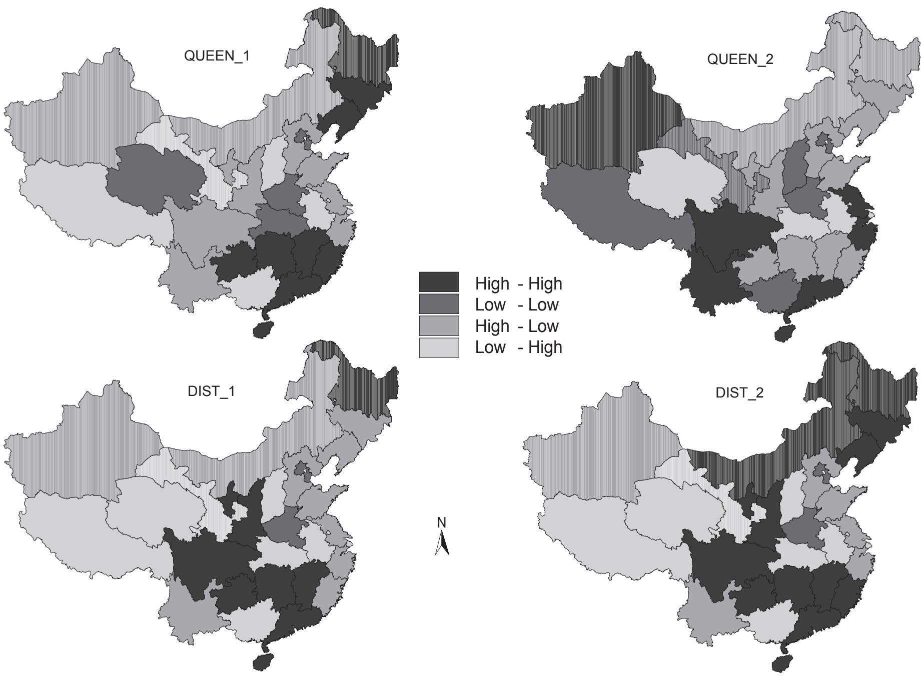


Figure 4: Moran Scatterplot 1985-1990.

significant values. There is the coastal cluster of high values, and then two provinces with high values on either side of the low growth provinces in the interior. There are also two low growth provinces close to the coast. For DIST_1, we see a High-High belt from north to south in the interior of China, and also a belt of High-Low values in Manchuria and down the coast of southeast China. The Local Moran's I, however, shows significant values only for Tibet, Xinjiang, and Shanghai. DIST_2 yields a similar pattern as DIST_1 in the Moran Scatterplot, but the High-High belt is now spread all the way up to the northeast Manchuria. The Local Moran's I is the same as the previous one with the exception of Shanghai. The fact that so few provinces are significant in the Local Moran's I test explains why we were unable to find support of global spatial autocorrelation in the Moran's I test earlier.

In the next five-year period, 1990–1995, the global tests, Moran's I and Geary's C, were significant for all weights matrices. Since the Moran Scatterplot patterns are similar to those in Figure 2, they are left out to save space. For the Local Moran's I test (Figure 6),

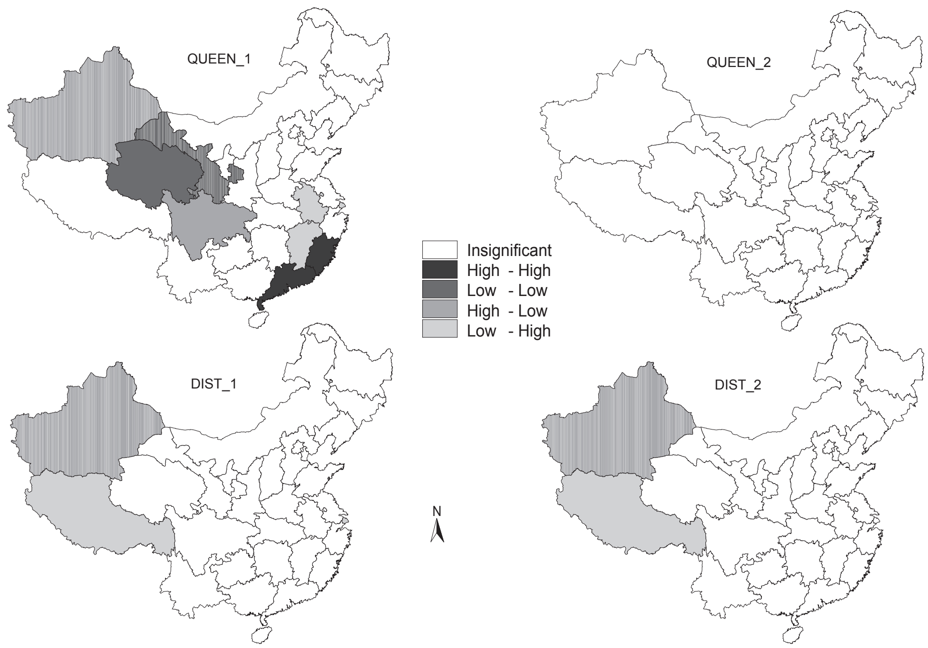


Figure 5: Local Moran's I 1985-1990.

only three hot spot provinces, and two with low growth, are significant with *QUEEN_1*. For *QUEEN_2*, four Low-Low provinces and three High-High provinces are reported. This pattern is repeated for the distance weights matrices. The three interior provinces, Qinghai, Gansu, and Ningxia, are negative and significant; while the three coastal provinces, Jiangsu, Zhejiang, and Fujian, remain significant and positive. Hence, we may conclude that China, during this period, had one hot spot area and one area in the interior with significantly lower growth.

The final period, 1995–2000, (Figure 7 and Figure 8) is of special interest because the pattern has changed dramatically. The Moran Scatterplots for *QUEEN_1* and *QUEEN_2* reports a Low-Low belt in the south and for the Shanghai province, and a High-High area around the capital city of Beijing. The *DIST_1* yields similar results. The Low-Low area is extended from the south up to the interior, and the High-High area is spread from the Beijing area down along the coast. The Local Moran's I presents significant negative values for Hainan, Guangxi, Xinjiang, and Shanxi, while Tibet is significant and positive. Hence, the economies in the former productive coastal

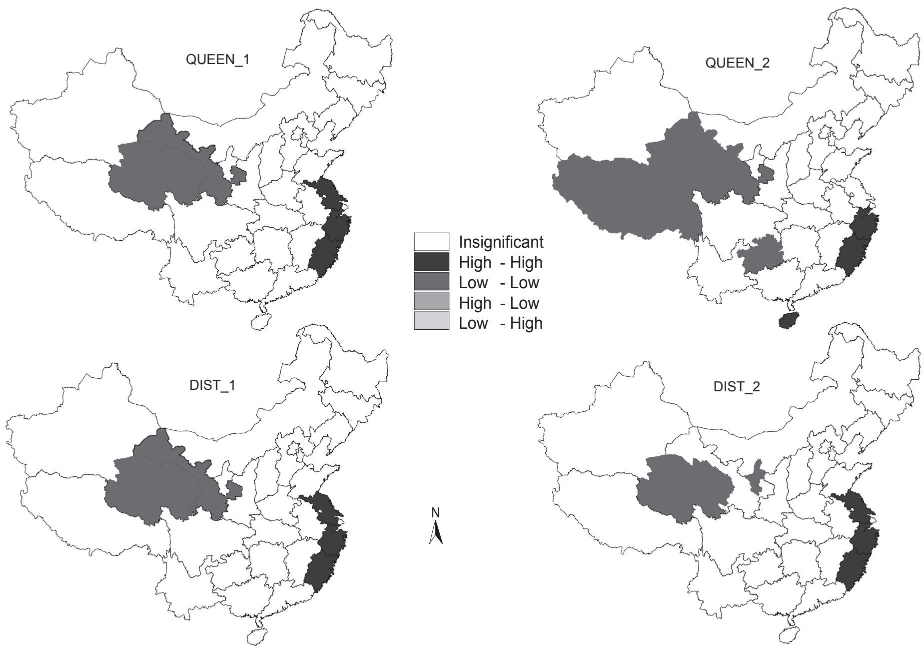


Figure 6: Local Moran's I 1990-1995.

area, such as the provinces of Fujian and Guangdong, are no longer that successful in relative terms. This may be a first sign of convergence between the Chines provinces.

To conclude the findings so far, the global tests indicate spatial autocorrelation for the periods 1985–2000 and 1990–1995. However, local tests indicate that the clusters of similar values are not as strong as first expected. Nevertheless, some hot spots was found for all time periods, especially in the southeast region. The provinces with low growth values are mostly found in the interior. A change of this pattern was hinted at in the last five-year period. With this knowledge in mind, the next step will be to estimate growth equations.

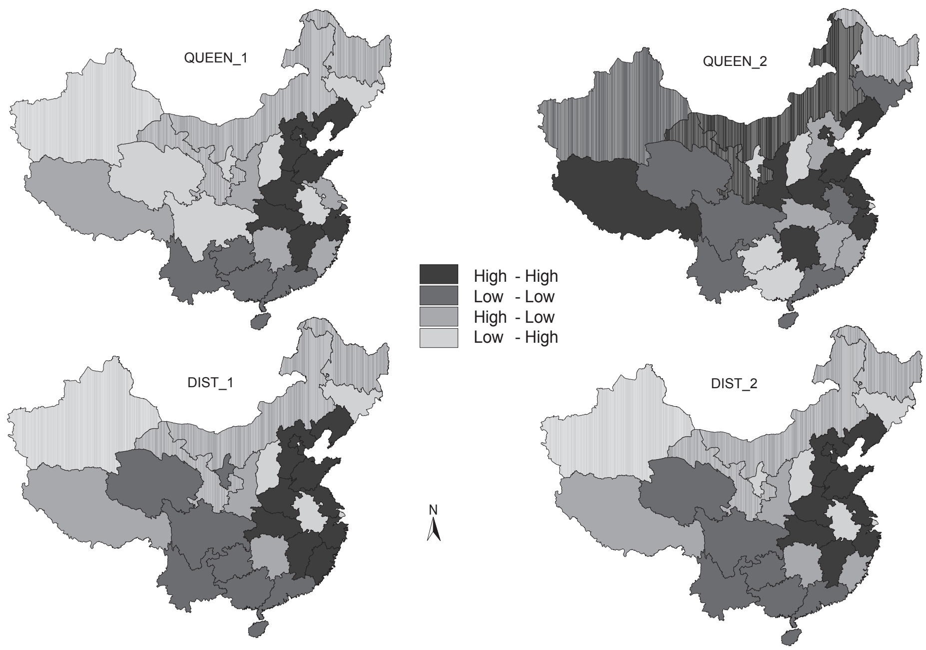


Figure 7: Moran Scatterplot 1995-2000.

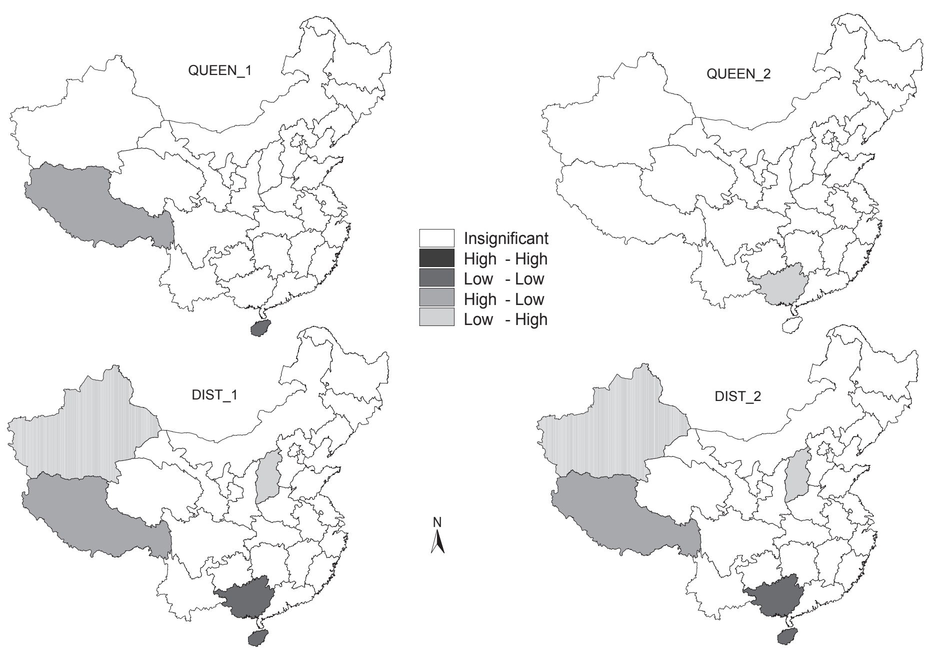


Figure 8: Local Moran's I 1995-2000.

## 5 Estimation of the Provincial Economic Growth Equation

The estimation of provincial economic growth is made in two regressions. The first regression makes it possible to test the hypothesis of unconditional convergence. The conditional convergence hypothesis is tested in the second regression. In both cases spatial dependence adjustments are included when needed, as indicated by the Lagrange Multiplier tests based on the OLS residuals. The Lagrange Multiplier test results are not reported here but are available upon request.

The first growth equation is expressed as:

$$g_{tT} = \beta_0 + \beta_1 GRPC_t + \varepsilon_t \tag{7}$$

where $t$ and $T$ indicate the initial and final year for the period in question.

The independent variable in (7) is the initial level of GRP per capita, *GRPC*, at time $t$ for each period. The results are presented in Table 4. For the entire 15-year period, 1985–2000, the introductory OLS results are given in column one. As expected from the exploratory analysis, and according to the Lagrange Multiplier tests, the results from the OLS regression is subject to spatial error dependence. The test also indicates that the first order contiguity matrix, *QUEEN_1*, is the most appropriate matrix to capture the spatial dependence. The estimation is made with the Spatial Autoregressive Generalized Moments (SAR-GM) estimator, as it accepts non-normal distributed errors. The result is presented in the second column of Table 4. The coefficient for *GRPC* is negative and significant, thus indicating unconditional convergence between the Chinese provinces. The autoregressive coefficient, $\lambda$, is positive[3] with a value of 0.42. The $R^2$ is, as expected, quite low, 13%[4].

The parameter for the *GRPC* is negative and significant at the five percent level for the first five-year period, 1985–1990. Hence,

[3]The estimate for $\lambda$ has no inference since it is treated as a nuisance variable.

[4]The normal measures of fit are not applicable for the spatial models. Instead pseudo measures must be used. Unfortunately they cannot accurately be compared with the measures of fit from the OLS regressions. The $R^2$ is, in the spatial cases, the ratio of the predicted values over the variance of the observed values for the dependent variable. Sq.corr is the squared correlation between the predicted and observed values and Sig-sq. is an estimate for the residual variance.

Table 4: Regression results for unconditional GRP/capita growth. *** and ** indicate significant values at 1% and 5% percent level. § indicate that no inference has been made.

| Variable | 8500 OLS | 8500 SAR–GM | 8590 OLS | 9095 OLS | 9095 SAR–ML | 9500 OLS |
|---|---|---|---|---|---|---|
| WEIGHTS | | QUEEN_1 | | | DIST_2 | |
| $\lambda$ | | 0.42(§) | | | | |
| $\rho$ | | | | | 0.56*** | |
| *Constant* | 7.05*** | 7.42** | 3.43*** | 9.95*** | 4.85** | 7.61*** |
| *GRPC_t* | -0.0001 | -0.0003*** | -0.0004** | $7 \cdot 10^5$ | -0.0001 | $-3 \cdot 10^7$ |
| $R^2$ | 0.03 | 0.13 | 0.18 | 0.00 | 0.06 | 0.00 |
| $R^2$*-adj.* | -0.00 | | 0.15 | -0.03 | | -0.03 |
| *Sq.corr* | | 0.03 | | | 0.25 | |
| SIG-SQ. | 2.85 | 2.12 | 4.33 | 16.73 | 12.71 | 3.31 |
| *Suggestion* | QUEEN_1 | | | DIST_2 | | |

growth convergence across provinces is implied. There is no indication of spatial dependence, so no spatial adjustment is needed. The $R^2$ is 18%. For the second five-year period, 1990–1995, the Lagrange Multiplier tests indicate problems of spatial error dependence. This is solved with a spatial error term, and the more narrow inverse distance weights matrix, *DIST_2*. The results from this regression model estimated with Maximum Likelihood is shown in the fifth column. As in the previous regression, the *GRPC* parameter value is not significant and thus no sign of unconditional convergence appears. On the other hand, the spatial error variable is positive and highly significant and suggests that spillover effects exist between adjacent provinces. Thus, a high GRP per capita growth in the neighboring provinces correlate positively with the growth rate in province $i$. The last period, 1995–2000, differs from the previous periods in so far as that the constant is the only significant parameter. There is no indication of spatial dependence during the last period.

Hence, the unconditional convergence hypothesis is rejected for the two latter sub-periods. The unconditional convergence that was captured for the 15-year period is somewhat misleading and should merely be seen as a reflection of the convergence found in the first five-year period.

Next, we continue the analysis and consider other explanatory variables. In this way, we may test the hypothesis of conditional convergence. That is, whether regions converge towards its own steady state growth rate level. The second regression equation is expressed as:

$$g_{tT} = \beta_0 + \beta_1 EDUP_t + \beta_2 TPAREA_t + \beta_3 PREF_{tT} + \beta_4 DINV_{tT} + \beta_5 FDIC_{tT} + \beta_6 D\_CITY_t + \beta_7 SOE\_TE_{tT} + \beta_8 GRPC_t + \varepsilon_t \quad (8)$$

Table 5: Regression results for conditional GRP/capita growth. ***, **, and * indicate significant values at 1, 5, and 10% percent level.

| Variable | 8500 | 8590 | 8590 | 8590 | 9095 | 9500 |
|---|---|---|---|---|---|---|
| | OLS | OLS | SAR–IV | SAR–IV | OLS | OLS |
| WEIGHTS | | | QUEEN_1 | DIST_2 | | |
| $\rho$ | | | -0.82** | -0.64** | | |
| *Constant* | 5.44*** | 7.16*** | 7.04*** | 8.23*** | 6.90* | 9.34*** |
| EDUP_$t$ | -665.00 | 663.00 | 2284.38 | 860.00 | -1213.71 | -1008.15 |
| TPAREA_$t$ | 4.40 | -7.39 | -5.43 | -7.06 | 12.00** | 0.04 |
| FDIC_$t$-$T$ | -0.06** | 0.18* | 0.22*** | 0.20** | -0.10** | -0.02 |
| DINVC_$t$-$T$ | 0.0009 | -0.001 | -0.003 | -0.001 | 0.003* | -0.0003 |
| SOE_TE_$t$ | -3.08 | -13.14*** | -9.16** | -12.70*** | -5.90 | 0.17 |
| PREF_$t$-$T$ | 2.10*** | 0.58 | 1.05*** | 0.52 | 5.51*** | -1.08 |
| D_CITY | 2.43 | -1.96 | -3.12 | -2.13 | 3.44 | 5.73 |
| D_EXTERNAL | 2.08* | 0.92 | 0.76 | 1.02 | 2.21 | 2.73 |
| GRPC_$t$ | -0.0006 | -0.0008 | 0.0001 | -0.0002 | -0.002* | 0.0005 |
| $R^2$ | 0.84 | 0.69 | 0.76 | 0.75 | 0.86 | 0.19 |
| $R^2$-*adj.* | 0.77 | 0.55 | | | 0.80 | -0.18 |
| *Sq.corr* | | | 0.82 | 0.79 | | |
| SIG-SQ. | 0.65 | 2.33 | 1.44 | 1.67 | 3.26 | 3.77 |
| *Suggestion* | | QUEEN_1<br>DIST_2 | | | | |

As before, the analysis is done in a model for the 15-year period as well as for the three sub periods, 1985–1990, 1990–1995, and 1995–2000. Regression results are given in Table 5. All weights matrices were tested for possible inclusion in the final model.

The results for the period 1985–2000 may be found in the first column. The preferential policy variable is positive and significant at the 1% level. Foreign direct investments have a negative impact on the provincial growth rate. On the other hand, strong external relationships operates in the opposite direction with a positive and significant value. China have three provinces that are categorized as cities, Beijing, Tianjin, and Shanghai. The city dummy variable is not significant and we may therefore not conclude that there exists a positive growth effect from agglomeration alone. Neither is there any sign of conditional convergence or problems of spatial dependence. The fit is radically improved to 84%.

The second column show the OLS results from the period 1985–1990. The share of State Owned Enterprises of the total number of enterprises has, as expected, a negative impact on growth. The average Foreign Direct Investment per capita during the period is positive and significant at a 10% level. The $R^2$ is 69%. There are indications of spatial lag dependence for two of our weights matrices, QUEEN_1 and DIST_2. The spatial regressions that follows, shown in columns three and four, are both estimated with a 2SLS approach, due to the non-normal error distribution, with spatially lagged variables as instruments. The spatial lag is in both cases significant but has an unexpected negative sign with a lower parameter value in the second regression. This means that the provinces do not benefit from closeness to each other. Quite the opposite, they are competitors. The preferential policy variable is positive and significant only for the QUEEN_1 regression. The SOE variable has the expected sign and a larger parameter value for the DIST_2 regression. The magnitude is slightly less than in the previous OLS regression. The foreign direct investments parameter is robust, with a positive and significant value in both regressions. The improvement between the regressions is shown through smaller standard deviations (Sig-sq.) compared with the OLS results.

In the first five-year period of the 1990's, no indication of spatial dependence is evident. The preferential policy is more important than before for the GRP per capita growth rate. The transport capacity and domestic investment per capita are also positive. Foreign direct investments are, however, negative and significant. For the first time we have evidence of conditional convergence due to the fact that the initial level of GRPC coefficient is negative and significant. The regression fit is, again, quite high, 86%.

The latter part of the 1990's is of interest, not because the richness of results, but the lack of results. There are no indications of spatial dependence, and the only significant parameter is the constant. The reason for this unexpected result is hinted at in the maps shown earlier in Figure 7 and Figure 8. The former successful provinces do not grow faster than the rest of the Chinese provinces, and an explanation for provincial economic growth must be found elsewhere. In her book, Oi (1999) provides some interesting thoughts about increased competition from 1995 onwards, increased need of investment capital, and a more efficient company structure where the many collectively owned companies might need to be exchanged for privately owned firms to maintain and increase growth.

## 6 Conclusions

In this paper, the provincial economic growth in China during the period 1985–2000 have been investigated. This was accomplished in two parts. The first part consists of an exploratory data analysis in search of spatial autocorrelation and hot spots. Global spatial autocorrelation was indicated for two time periods. Some clusters of provinces with a high growth, especially in the coastal region in the southeast, as well as low growth clusters in the center and western parts of China were found.

The second part of the paper is comprised of a regression analysis aiming to find explanatory variables for growth and to check for convergence. Positive spatial lag dependence was found for the periods 1990–1995 and 1985–2000. When conditional convergence was tested, several important variables were found that explain provincial economic growth, such as preferential policy, enterprise structure, and external relations. The effect of foreign direct investments changed sign over time and seems to have had the most important effect in the late 1980's. Evidence of conditional convergence was only found for the sub period 1990–1995. The influence of spatial dependence does not seem to be high. One possible explanation for this low degree of spatial dependence could be that the aggregation to province level is to high, with a sample of only 30 observations. Another possible explanation is that the provinces are, in fact, more independent of each other than we have suspected and that the major forces behind growth are found within each province between the urban and rural areas. A similar investigation of this present paper, at the more disaggregated county level (not available yet though) may shed new light on growth relationships within and between the Chineses provinces.

## References

[1] Anselin, L. (1988). *Spatial Econometrics: Methods and Models.* Kluwer Academic, Dordrecht.

[2] Anselin, L. (1995). *SpaceStat, A Software Program for the Analysis of Spatial Data, Version 1.80.* Regional Research Institute, West Virginia University, Morgantown, WV.

[3] Anselin, L. (1995b). Local Indicators of Spatial Association–LISA. *Geographical Analysis* 27 (2), 93–115.

[4] Bal, F., and P. Nijkamp (1998). Exogenous and Endogenous Spatial Growth Models. *Annals of Regional Science* 32, 63–89.

[5] Barro, R. J., and X. Sala-i-Martin (1992). Convergence. *Journal of Political Economy* 100, 223–251.

[6] Barro, R. J., and X. Sala-i-Martin (1995). *Economic Growth.* McGraw-Hill, New York.

[7] Can, Y. (1992). Specification and estimation of hedonic housing price models. *Regional Science* 22, 453–474.

[8] Carrington, A. (2002). A Divided Europe? Regional Convergence and Neighborhood Spillover effects. *Mimeo, Bradford University.*

[9] Chen, B., and Y. Feng (2000). Determinants of economic growth in China: Private enterprise, education, and openness. *China Economic Review* 11, 1–15.

[10] Chen, J., and B. M. Fleisher (1996). Regional Income Inequality and Economic Growth in China. *Journal of Comparative Economics* 22, 141–164.

[11] Cliff, A., and K. Ord (1972). Testing for Spatial Autocorrelation Among Regression Residuals. *Geographical Analysis* 4, 267–284.

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

[13] Démurger, S. (2001). Infrastructure Development and Economic Growth: An Explanation for Regional Disparities in China? *Journal of Copmarative Economics* 29, 95–117.

[14] Démurger, S., and J. D. Sachs, W. T. Woo, S. Bao, G. Chang, and A. Mellinger (2002). Geography, Economic Policy and Regional Development in China. *Discussion Paper Number 1950, March 2002, Harvard Institute of Economic Research.*

[15] Graham, E. M., and E. Wada (2001). Foreign Direct Investment in China: Effects on Growth and Economic Performance. To appear in: Drysdale, P. (ed) *Achieving High Growth: Experience of Transitional Economies in East Asia. Oxford University Press, forthcoming 2001.*

[16] Hsueh, T., L. Qiang, and L. Shucheng (Eds) (1993). *China's Provincial Statistics, 1949–1989.* Westview Press, Boulder, San Fransisco, Oxford.

[17] Koopmans, T. C. (1965). *On the Concept of Optimal Economic Growth. The Econometric Approach to Development Planning.* North Holland, Amsterdam.

[18] Lucas, R. E. (1988). On the Mechanics of Economic Development. *Journal of Monetary Economics* 22, 3–42.

[19] National Bureau of Statistics of China. *China Statistical Yearbook 1986–2001.* China Statistical Press, China.

[20] Nijkamp, P., and J. Poot (1998). Spatial Perspectives on New Theories of Econoic Growth. *Annals of Regional Science* 32, 7–37.

[21] Oi, J. C. (1999). *Rural China Takes Off. Institutional Foundations of Economic Reforms.* University of California Press, Berkeley and Los Angeles.

[22] Ord, J. K., and A. Getis (1995). Local Spatial Autocorrelation Statistics: Distributional Issues and an Application. *Geographical Analysis* 27 (4), 286–306.

[23] Persson, J. (1997). Convergence Across the Swedish Counties, 1911–1993. *European Economic Review* 41, 1835–1852.

[24] Rey, S. J., and B. Montouri (1999). US Regional Income Convergence: A Spatial Econometric Perspective. *Regional Studies* 33 (2), 143–156.

[25] Romer, P. M. (1986). Increasing Returns and Long-Run Growth *Journal of Political Economy* 94, 1002-1037.

[26] Sala-i-Martin, X. (1996). Regional Cohesion: Evidence and Theories of Regional Growth and Convergence. *European Economic Review* 40, 1325–1352.

[27] Solow, R. M. (1956). A Contribution to the Theory of Economic Growth. *Quarterly Journal of Economics* 70, 65–94.

[28] Swan, T. V. (1956). Economic Growth and Capital Accumulation. *Economic Record* 32, 334–361.

[29] Tian, X. (1999). Market Orientation and Regional Economic Disparities in China. *Post–Communist Economies* 11 (2), 161–172.

[30] Vayá, E., E. López–Bazo, R. Moreno, and J. Surinach (2000). Growth and Externalities across Economies. an empirical Analysis using Spatial Econometrics. *Working Paper Economics and Social Sciences no. EOO/59, University of Barcelona.*

[31] Wheeler, C. H. (2001). A Note on the Spatial Correlation Structure of County–Level Growth in the US. *Journal of Regional Science* 41 (3), 433–449.

[32] Yao, S., and Z. Zhang (2001). Regional Growth in China Under Economic Reforms. *The Journal of Development Studies* 38 (2), 167–186.

[33] Ying, L. G. (2000). Measuring the spillover effects: Some Chinese evidence. *Papers in Regional Science* 79, 75–89.

[34] Zhang, J., and G. Kristensen (2001). The paradox of unequal regional investment and equal regional economic growth. *The Annals of Regional Science* 35, 637–655.

## Avhandlingar framlagda vid Institutionen för nationalekonomi, Umeå universitet

## List of dissertations at the Department of Economics, Umeå University

Holmström, Leif (1972) Teorin för företagens lokaliseringsval. UES 1. PhLic thesis

Löfgren, Karl-Gustaf (1972) Studier i teorin för prisdiskriminering. UES 2. PhLic thesis

Dahlberg, Åke (1972) Arbetsmarknadsutbildning - verkningar för den enskilde och samhället. UES 3. PhD thesis

Stage, Jørn (1973) Verklighetsuppfattning och ekonomisk teori. UES 4. PhLic thesis

Holmlund, Bertil (1976) Arbetslöshet och lönebildning - kvantitativa studier av svensk arbetsmarknad. UES 25. PhD thesis

Löfgren, Karl-Gustaf (1977) En studie i neokeynesiansk arbetslöshets- och inflationsteori. UES 34. PhD thesis

Lundberg, Lars (1976) Handelshinder och handelspolitik - Studier av verkningar på svensk ekonomi. Industriens Utredningsinstitut, Stockholm. PhD thesis

Johansson, Per-Olof (1978) Sysselsättning och samhällsekonomi - En studie av Algots etablering i Västerbotten. UES 53. PhD thesis

Wibe, Sören (1980) Teknik och aggregering i produktionsteorin. Svensk järnhantering 1850-1975; en branschanalys. UES 63. PhD thesis

Ivarson, Lars (1980) Bankers portföljvalsbeteende. En teoretisk studie. UES 64. PhD thesis

Batten, David (1981) Entropy, Information Theory and Spatial Input-output Analysis. UES 92. PhD thesis

Hårsman, Björn (1982) Housing Demand Models and Housing Market Models for Regional and Local Planning. Swedish Council for Building Research, D13:1981. PhD thesis

Holm, Magnus (1983) Regionalekonomiska modeller för planering och samordning i en decentraliserad ekonomi. Byggforskningsrådet, R118:1981 and R5:1983. PhD thesis

Ohlsson, Henry (1986) Cost-Benefit Analysis of Labor Market Programs - Applied to a Temporary Program in Northern Sweden. UES 167. PhLic thesis

Sarafoglou, Nikias (1987) A Contribution to Population Dynamics in Space. UES 179. PhD thesis

Ohlsson, Henry (1988) Cost-Benefit Analysis of Labor Market Programs - Applied to a Temporary Program in Northern Sweden. UES 182. PhD thesis

Anderstig, Christer (1988) Applied Methods for Analysis of Economic Structure and Change. CERUM 1988:2, Umeå University. PhD thesis

Karlsson, Charlie (1988) Innovation Adoption and a Product Life Cycle. UES 185. PhD thesis

Löfström, Åsa (1989) Diskriminering på svensk arbetsmarknad - En analys av löneskillnader mellan kvinnor och män. UES 196. PhD thesis

Axelsson, Roger (1989) Svensk arbetsmarknadsutbildning - En kvantitativ analys av dess effekter. UES 197. PhD thesis

Zhang, Wei-Bin (1989) Theory of Economic Development - Nonlinearity, Instability and Non-equilibrium. UES 198. PhD thesis

Hansson, Pär (1989) Intra-Industry Trade: Measurements, Determinants and Growth - A study of Swedish Foreign Trade. UES 205. PhD thesis

Kriström, Bengt (1990) Valuing Environmental Benefits Using the Contingent Valuation Method: An Econometric Analysis. UES 219. PhD thesis

Aronsson, Thomas (1990) The Short-Run Supply of Roundwood under Nonlinear Income Taxation - Theory, Estimation Methods and Empirical Results Based on Swedish Data. UES 220. PhD thesis

Westin, Lars (1990) Vintage Models of Spatial Structural Change. UES 227. PhD thesis

Wikström, Magnus (1992) Four Papers on Wage Formation in a Unionized Economy. UES 287. PhD thesis

Westerlund, Olle (1993) Internal Migration in Sweden - The Role of Fiscal Variables and Labor Market Conditions. UES 293. PhLic thesis

Bergman, Mats A. (1993) Market Structure and Market Power. The Case of the Swedish Forest Sector. UES 296. PhD thesis

Johansson, Per (1993) Count Data Models - Estimator Performance and Applications. UES 315. PhD thesis

Roson, Roberto (1994) Transport Networks and the Spatial Economy - A General Equilibrium Analysis. UES 340. PhD thesis

Li, Chuan-Zhong (1994) Welfare Evaluations in Contingent Valuation - An Econometric Analysis. UES 341. PhD thesis

Østbye, Stein (1994) Regional Labour and Capital Subsidies - Theory and Evidence of the Impact on Employment under Wage Bargaining. UES 344. PhLic thesis

Westerlund, Olle (1995) Economic Influences on Migration in Sweden. UES 379. PhD thesis

Mortazavi, Reza (1995) Three Papers on the Economics of Recreation, Tourism and Property Rights. UES 396. PhLic thesis

Østbye, Stein (1995) Regional Labour and Capital Subsidies. UES 397. PhD thesis

Hussain-Shahid, Imdad (1996) Benefits of Transport Infrastructure Investments: A Spatial Computable General Equilibrium Approach. UES 409. PhD thesis

Eriksson, Maria (1996) Selektion till arbetsmarknadsutbildning. UES 410. PhLic thesis

Karlsson, Niklas (1996) Testing and Estimation in Labour Supply and Duration Models. UES 413. PhD thesis

Olsson, Christina (1996) Chernobyl Effects and Dental Insurance. UES 428. PhLic thesis

Vredin, Maria (1997) The African Elephant - Existence Value and Determinants of Willingness to Pay. UES 441. PhLic thesis

Eriksson, Maria (1997) To Choose or not to Choose: Choice and Choice Set Models. UES 443. PhD thesis

Widerstedt, Barbro (1997) Employer Change and Migration. Two Papers on Labour Mobility in Sweden. UES 444. PhLic thesis

Lundberg, Sofia (1997) The Economics of Child Auctions in 19th Century Sweden. UES 445. PhLic thesis

Lundberg, Johan (1997) Two Papers on Revenue and Expenditure Decisions in the Swedish Local Public Sector. UES 454. PhLic thesis

Widerstedt, Barbro (1998) Moving or Staying? Job Mobility as a Sorting Process. UES 464. PhD thesis

Bask, Mikael (1998) Essays on Exchange Rates: Deterministic Chaos and Technical Analysis. UES 465. PhD thesis

Löfgren, Curt (1998) Time to Study Students: Two Essays on Student Achievement and Study Effort. UES 466. PhLic thesis

Sjögren, Tomas (1998) Union Wage Setting in a Dynamic Economy. UES 480. PhD thesis

Mortazavi, Reza (1999) Essays on Economic Problems in Recreation, Tourism and Transportation. UES 483. PhD thesis

Rudholm, Niklas (1999) Two Essays on Generic Competition in the Swedish Pharmaceuticals Market. UES 485. PhLic thesis

Olsson, Christina (1999) Essays in the Economics of Dental Insurance and Dental Health. UES 494. PhD thesis

Marklund, Per-Olov (1999) Environmental Regulation and Firm Efficiency. UES 504. PhLic thesis

Berglund, Elisabet (1999) Regional Entry and Exit of Firms. UES 506. PhD thesis

Hellström, Jörgen (1999) Count Data Autoregression Modelling. UES 507. PhLic thesis

Nordström, Jonas (1999) Tourism and Travel: Accounts, Demand and Forecasts. UES 509. PhD thesis

Johansson Vredin, Maria (1999) Economics Without Markets. Four papers on the Contingent Valuation and Stated Preference Methods. UES 517. PhD thesis

Schei, Torbjørn (2000) Natural recreation resources: production and a diversity of interests related to the management of grouse as an outfield resource in Finnmark, Norway, in the Euro-Arctic Barents region. UES 523. PhLic thesis

Backlund, Kenneth (2000) Welfare Measurement, Externalities and Pigouvian Taxation in Dynamic Economies. UES 527. PhD thesis

Andersson, Linda (2000) Job Turnover, Productivity and International Trade. UES 530. PhLic thesis

Ylvinger, Svante (2000) Essays on Production Performance Assessment. UES 531. PhD thesis

Bergkvist, Erik (2001) Freight Transportation. Valuation of Time and Forecasting of Flows. UES 549. PhD thesis

Rudholm, Niklas (2001) The Swedish Pharmaceuticals Market - Essays on Entry, Competition and Antibiotic Resistance. UES 552. PhD thesis

Lundberg, Johan (2001) Local Government Expenditures and Regional Growth in Sweden. UES 554. PhD thesis

Lundberg, Sofia (2001) Going Once, Going Twice, SOLD! The Economics of Past and Present Public Procurement in Sweden. UES 557. PhD thesis

Eliasson, Kent (2001) University Enrollment and Geographical Mobility: The Case of Sweden. UES 558. PhLic thesis

Samakovlis, Eva (2001) Economics of Paper Recycling. Efficiency, policies, and substitution possibilities. UES 563. PhD thesis

Daunfeldt, Sven-Olov (2001) Essays on Intra-Household Allocation and Policy Regime Shifts. UES 570. PhD thesis

Hellström, Jörgen (2002) Count Data Modelling and Tourism Demand. UES 584. PhD thesis

Andersson, Linda (2002) Essays on Job Turnover, Productivity and State-Local Finance. UES 586. PhD thesis

Rashid, Saman (2002) Invandrarinkomster, förvärvsdeltagande och familj. UES 588. PhLic thesis

Hanes, Niklas (2003) Empirical Studies in Local Public Finance: Spillovers, Amalgamations, and Tactical Redistributions UES 604. PhD thesis

Stenberg, Anders (2003) An Evaluation of the Adult Education Initiative Relative Labor Market Training. UES 609. PhD thesis

Stage, Jesper (2003) Mixing Oil and Water. Studies of the Namibian Economy. UES 611. PhD thesis

Marklund, Per-Olov (2004) Essays on Productive Efficiency, Shadow Prices , and Human Capital. UES 621. PhD thesis

Rashid, Saman (2004) Immigrants' Income and Family Migration. UES 625. PhD thesis

Sandberg, Krister (2004) Hedonic Prices, Economic Growth, and Spatial Dependence. UES 631. PhD thesis