The Interaction of Land Markets and Housing Markets in a Spatial Context: A Case Study of Helsinki

Risto PELTOLA, Finland

Key words: Spatial analysis, housing land, housing policy, valuation, development land

SUMMARY

The purpose of the paper is to increase our understanding on how land markets and housing markets interact in a spatial context. The term housing land supply chain is used to comprise development land, housing lot and house markets. Hedonic models are estimated for each based on a large good quality dataset.

Two concepts are used to measure price ratios in the housing land supply chain: (1) the land share of a house price and (2) a ratio of development land price to lot price.

The paper combines housing economics and spatial analysis. Hedonic models produce a trend surface, and residuals are mapped to reveal the local effects. Thematic maps are used to visualize the spatial structure of error terms. Two scales are used: grid level to get an overview, and transaction level for exact local effects.

The paper tries to offer a broad, deep and transparent view of the housing market. The data consists of more than 45.000 transactions during the last 21 years in Helsinki metropolitan area.

The results may be valuable in property valuation and management of housing policy.

The Interaction of Land Markets and Housing Markets in a Spatial Context: A Case Study of Helsinki

Risto PELTOLA, Finland

1. INTRODUCTION

The purpose of the paper is to increase our understanding on how land markets and housing markets interact in a spatial context. The paper combines housing economics and spatial analysis.

1.1 Housing economics

The valuation of houses is important for household, business and public finance interests. From a household point of view homes are not just places to live, they are also expensive items to buy, their value constitutes the main part of most households assets, and on top of all, house prices are unconveniently cyclical bringing an unwelcome element of risk to the life of a usually risk-aversive household.

A house can be divided in two components: land and structures. Sometimes the value of land is more than half of the value of a home, and land is the main cyclical component in the price of a home.

Valuing land and housing is difficult due to spatial and other variation of the properties. This is one of reasons that transaction costs for houses are so high, and search costs in particular.

Apart from valuation of residential property, there are several other housing issues that are relevant in the context of my study. The affordability of housing is one of them. Near important centres of employment housing is usually expensive and is getting more expensive. In some city-regions a home may appreciate faster than median incomes grow. In my study area, Helsinki metropolitan region, this race seems to be pretty even in a 30-year period to this day.

Is there a way to increase production of new homes to meet the demand? Intuitively it looks as if the inadequate production of new houses is connected to a rather inelastic supply of housing lots. There is some evidence to this, too. In the British context, where supply of housing lots is one of most inelastic in the world, this issue has often been pointed out by Evans (most recently Evans and Hartwich 2005).

While Britain may be a case in the extreme end, it is not the only country where rapidly rising house prices are not met with increasing production of homes. The same problems of inadequate production of homes and inelastic supply of housing lots are clearly present in land-rich Finland, too, at least in Helsinki metropolitan region.

My contribution to the issue is to study how land and house prices interact in a spatial context. I try to offer a simultaneous view of the housing and land markets, including all the land transactions that precede the production of new homes, namely: market for development land and market for serviced housing lots. Together with markets of homes they constitute, as I call it, the **supply chain of housing land.** Also the term **land development chain** could be applied**.**

The studies that compare prices of serviced lots and houses are surprisingly rare. Studies of their interaction in a spatial context are even rarer. When studies are made, they are usually at a very aggregate level (e.g. Cheshire and Sheppard). International comparisons are rare and anecdotal. (See Evans 1999 for a survey of studies in the field.)

The studies that include development land in the analysis of the supply chain are unknown to me at least, even if this is critical given the low price elasticity of supply of housing lots. Costs of servicing land are also a neglected issue in economics, and that is not part of my study either.

The lack of studies is probably due to poor availability of data. Given the importance of the housing affordability problem and its connection to inelastic supply of housing land, this clearly is a neglected issue in housing economics.

Development land , lot and house markets are subject to analysis in this paper. I try to offer some benchmark to help to evaluate how the housing land supply chain works and, when problems are identified, how to manage it better.

1.2 Spatial analysis

My approach consists of two parts:

1) I fit a conventional hedonic price model to a large data set.
2) I examine the spatial structure of error terms manually, in a visual way.

The first part reveals the general trends. Given the quality, quantity and timeliness of the data, it gives some valuable information about factors affecting the prices.

Hedonic analysis is a standard property valuation tool. Spatial models have also been used in property market analysis, although they do not belong to an ordinary real estate economists toolbox. They have been developed and used vastly more in fields other than real estate, most notably earth sciences. However, it’s not unusual to combine hedonics and spatial models such as kriging in real estate, too.

My approach to spatial analysis is different and more intuitive. Exploring microspatial variation visually – instead of modelling it – has some advantages:

- it is computationally fast and simple (as opposed to sophisticated modelling techniques)
- software needed is easily available

- visualization makes the results easy to interpret (again, as opposed to some model results)
- spatial effects are usually very strong and easily understood even by a layman
- even if more sophisticated spatial models are to be used, a visual exploration is useful as a preliminary, hypothesis generating stage

The following sections of this paper are: data, models, visualization of local effects, and finally concluding remarks.

2. DATA

2.1 Definitions

The housing land supply chain consists of three parts:

1) development land
2) serviced housing lots
3) houses

Development land is land without a design plan or planning permission, usually near urban fringes. Serviced housing lots are made of development land through planning and investment in services, such as roads, pipelines and parks, and education, health and nursery services. Houses are residential structures built on serviced lots, so house values include the land component.

The definitions and operational criteria applied to the data are following:

Table 1. Definitions and operational criteria for research data

	area	planning	location	project size	buildings
houses	300-10000 m2	design plan	urban, suburban	1-2 homes	yes
housing lots	300-10000 m2	design plan	urban, suburban	1-2 homes	no
development land	1-100 hectares	without a design plan or planning permission	potential demand for urban land		no valuable buildings

The sale must also qualify as an arms-length transaction, and only sales between private persons or companies were accepted. E.g. all 8500 lots sold by local governments were excluded. Some of the prices in those sales are subsidized and do not represent the market value. Time period is 1985-2005 and the area of study is larger Helsinki metropolitan area (Uusimaa region).

For data availability reasons, the analysis is restricted to single-family houses and lots. There is plenty of data of those sales available. The data, although heterogenous, is more homogenous then any other type of residential property. Given the large amount of observations, this is an ideal data to analyze spatial effects.

Table 2. Some descriptive statistic of the research data

		development land	housing lot	house
N		2681	12683	30290
price (euro) (constant 2004 value)	mean	141717	47524	137260
	std dev	606409	72509	130301
land area (m2)	mean	58880	1156	1187
	std dev	161117	688	661
price (euro/m2) (constant 2004 value)	mean	3,2	47,0	
	std dev	8,5	57,0	
distance to Helsinki (km)	mean	37,2	29,3	33,3
	std dev	17,8	17,7	21,8

3. MODEL

3.1 Price variables

The variable in the models is either the total price or the unit price of the property (table 3). When lot value has been estimated by the model, the price ratios are calculated, too.

Table 3. The dependent price variables in the models and calculation of price ratios

	total price (euros)	unit price (euros/m2)
houses	x	
housing lots	x	x
development land		x
(estimated) lot price / house price	x	
development land price / (estimated) lot price		x

3.2 Price factors

The demand for housing land is derived from the demand for housing. Hence, factors affecting the value of housing should affect the value of residential lots and development land. (In practice the impact of many of these factors is impossible to measure in the context of development land.)

There are factors that affect the value in all stages of the production chain. Some factors, on the other hand, are specific to certain stage. The common factors are:

- time (trend and cycle)
- macro location (access to centres, administrative subdivision: local services and local taxation)
- micro location (e.g. sea, lakes and road network)

If these general price factors, location and time, are dominant enough and if they affect similarly in all stages of the land development chain, valuing development land in particular becomes easier, since lack of comparables of development land can partly be substituted by sales in the higher stages of the land development chain and applying the relevant price ratio.

3.3 Model specification

Conventional hedonic models are used to reveal the price effects and to produce the price estimates. The price model matrix is as follows:

$$p = a_0 + a_{i1} S + a_{i2} T + a_{i3} F + a_{i4} 0 + \varepsilon_i$$

p = price
S = spatial trend variable matrix, such as distances to city center
T = temporal variable matrix, such as time trend and cyclical variation
F = physical variable matrix, such as size and age of building, size of lot
O= other variable matrix

a_i = parameter vectors
ε = error term vector

Three price models were specified:

1. house price model
2. lot price model
3. development land price model

Price ratio models were specified, too:

4. ratio of development land price to lot price
5. ratio of lot price to house price

Models are not presented due to lack of space, but some model results are illustrated (figures 1-4).

The models have several functions:

- Models make it possible to compare factors affecting house, lot and development land prices to each other.
- Model 2 produces an estimate of the lot price in any time and place, such as the time and place of sale of house or development land.

- Models produce a trend surface, against which it's meaningful and illustrative to study the micro spatial variation in prices.

The main price factors where chosen by stepwise analysis, where variables were added one by one in order of their ability to increase the model efficiency. Table 4 shows, which continuous variables were of prime importance:

Table 4. Model specification for house and housing lot price models: relative importance of variables (stepwise procedure)

variable		price model: houses	price model: housing lots
temporal	time trend	4	6
	business cycle (*)	2	2
spatial	distance to Helsinki CBD	3	1
	distance to large town	13	5
	distance to small town	7	7
	distance to shopping center	17	18
	distance to seashore	8	11
	distance to lakeshore	14	13
	adjacent to lake or sea	16	21
	accessibility to main road	20	12
	proximity to main road	19	16
	400 ha grid: error correction	6	3
lot specific	lot size	10	4
	building density in lot	15	9
house specific	house size	1	
	house age	5	
	number of variables	24	23
	R2	0,69	0,69
	R2 for 5 most important	0,63	0,60
	(*) business cycle = MPA -flat price index (deviation from trend)		

The table reveals two important spatio-temporal variables: distance to Helsinki city centre and business cycle. They are the two most important determinants of housing lot price, and among the three most important in case of house prices.

The table also tells how only less than 70 % of the price variation can be explained by a simple global model. When price indices and some other dummy variables are included, the fit is less than 75 %. How severe are the limitations of the models?

There certainly is a fair amount of random variation (noise) in real estate prices. Given the quantity of the data, the noise is not a problem here. More important, there probably is a lot of systematic spatial variation which the global model can't explain. This is exactly is the problem we are interested in.

Finally, as to the prices of houses, some of the most important price characteristics are not available in the data. Age and size of house, though important, are the only ones. This is a source of variation of unknown amount.

The model results reveal the global trends. A GIS tool is now applied to explore the local effects. This is based on visualization of the spatially autocorrelated error term. According to O'Sullivan and Unwin, map of residuals is a useful way of exploring the data to suggest local factors that are not included in the trend surface (O'Sullivan and Unwin, p 261).

A large scale map is used to show the error term of any of the 45.000+ sales. A medium scale map is used for 1500 grids. The grid-error-correction variable (table 4) is based on these grids.

3.4 Price trends

The results of the global model show general price trends. Only a few examples of the price trends are given, since the main emphasis is not in model results as such, but in the error term. The pictures show, how the effect of a particular price factor is usually, but not always, very similar in all stages of the supply chain.

The pictures also give examples of what is controlled in the models and how it is done. Of course, the effects of these controlled variables, such as access to the amenities of the sea, are not expected to be seen in the error-visualization maps any more. Figures 1-4 show the price indices and the impact of three spatial trend variables.

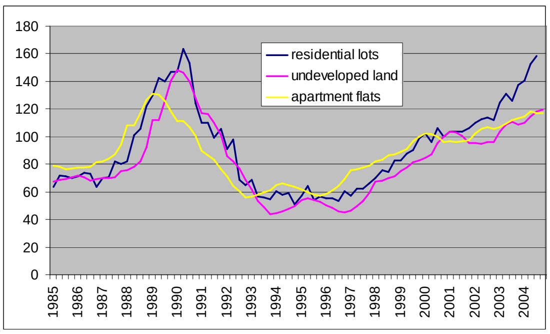


Figure 1. Real price indices

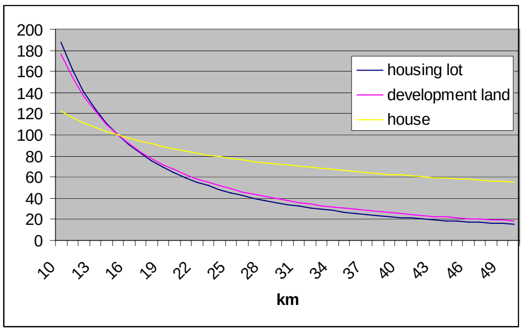
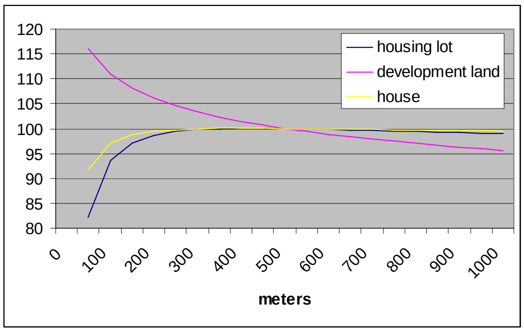


Figure 2. Distance to Helsinki city centre

3.4.1 Temporal effects

Figure 1 reveals strong cycles in all markets (especially the 1988-1992 period) and larger volatility of land prices compared to housing cycles. The volatility of land prices is approximately 1½ that of housing, and there is a lag of 9-12 months.

There are no clear differences in trends. However, the trend of lot prices is steeper than that of both housing and development land prices. The reasons are outside of the scope of this paper, so is the analysis of temporal variation in general.



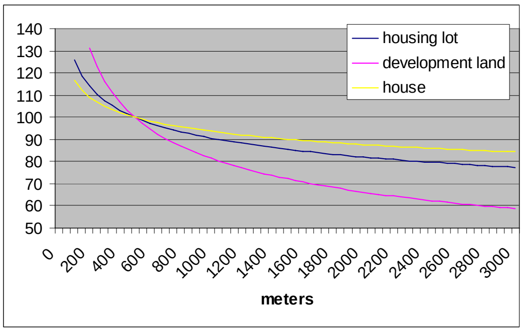


Figure 3. Distance to highway

Figure 4. Distance to Baltic Sea

3.4.2 Spatial effects

Distance to Helsinki has the same impact in the price of development land and residential lots (figure 2).

The price of development land is sensitive to access to main road (figure 3). Because land adjacent to main road is often developed to commercial or industrial use, close proximity to main road affects in opposite way to development land and residential lots.

Distance to small residential town has the larger impact in the price of development land than in the price of residential lots, and so is the impact of closeness to Baltic Sea (figure 4):

4. VISUALIZING SPATIAL VARIATION

The microspatial price variation is almost always poorly captured by a standard model where spatial autocorrelation is not controlled.

This section gives examples on how maps of different scales can be used to give an overview or a detailed look at the local effects. The examples cover all residential property in the supply chain. It also covers two price ratios between those three stages in the chain.

The use of a desktop -GIS to visualize property market characteristics is straitforward. Variables such as transaction prices, lot or floor areas, house age and other characteristics, and transaction volumes are routinely mapped.

Table 5 shows some typical attributes to visualize.

Table 5. Some typical thematic maps for exploring housing markets

type of a thematic map		**scale**	
		grid level	**transaction level**
volumes	number of transactions	x	
	sums (prices, lot areas, floor areas)	x	
relative prices	houses	x	x
	housing lots	x	x
	development land		x
price ratios	lot price / house price	x	x
	development land price / lot price		x
other transaction chasteristics	means for:		
	price (per unit or total)	x	
	area (lot or floor area)	x	
	age of house etc	x	

Some attributes are only suitable at the aggregate level, while most price information is valuable both at the gird level and at the observation level. However, if the markets are thin, there is no sense in aggregating to grid level, as is the case of development land sales.

Examples of thematic maps are given (figures 5-11).

4.1 An overview: grid level

Figure 5 illustrates the number of house and lot sales in a 400 hectare grid.

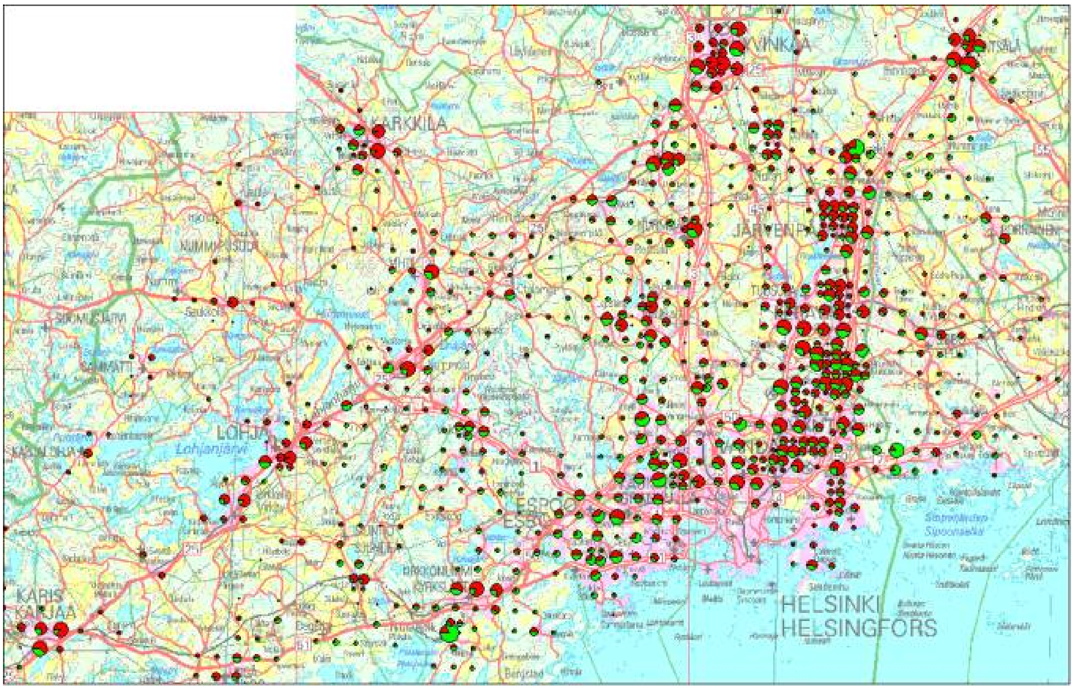


Figure 5. The number of house and lot sales in a 400 hectare grid.

Only up to 70-80 % of the variation of prices could be explained by conventional hedonic models. Of special concern here is the unexplainable part of the variation, the error term. It's easy to see intuitively, that a large part of this variation is due to spatial effects, which the hedonic model is unable to capture.

Figures 6 and 7 reveal strong spatial continuation, as was expected. Relatively expensive sales tend to lie in clusters, and so do relatively inexpensive sales. Houses and lots have also a very similar price pattern, as can be seen when figures 6 and 7 are united. Houses and lots are expensive in the same areas, of course. This should be clear on theoretical grounds only. However, the evidence confirms the theory in a nice way.

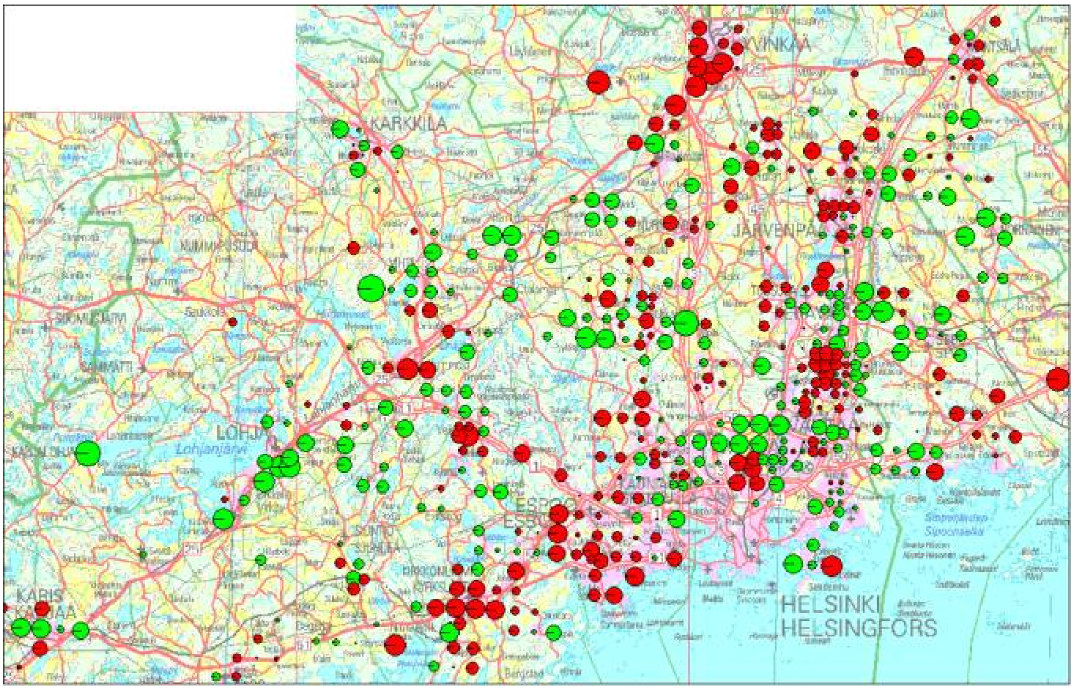


Figure 6. The error term of lot prices in a 400 hectare grid.

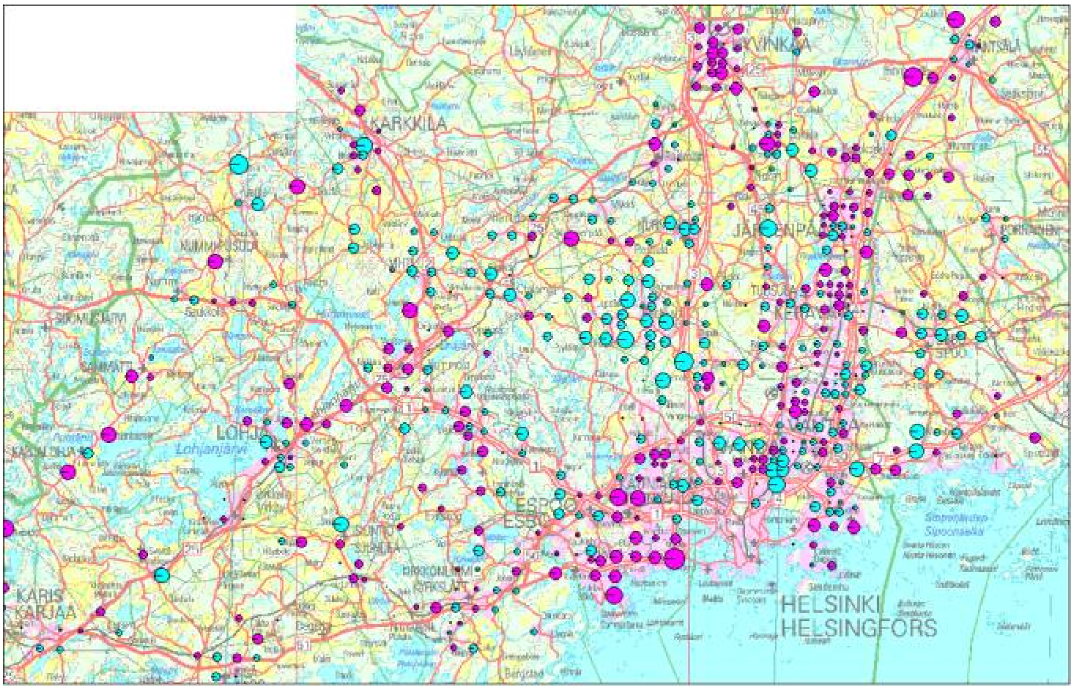


Figure 7. The error term of house prices in a 400 hectare grid.

4.2 A detailed view

The spatial variation occurs not only between the grids, but also within the grid. It may occur between any individual pair of sales. (This is one of the basic ideas of spatial analysis.) In the following two examples a large scale map is used to illustrate the error term of individual observations (figures 8-9).

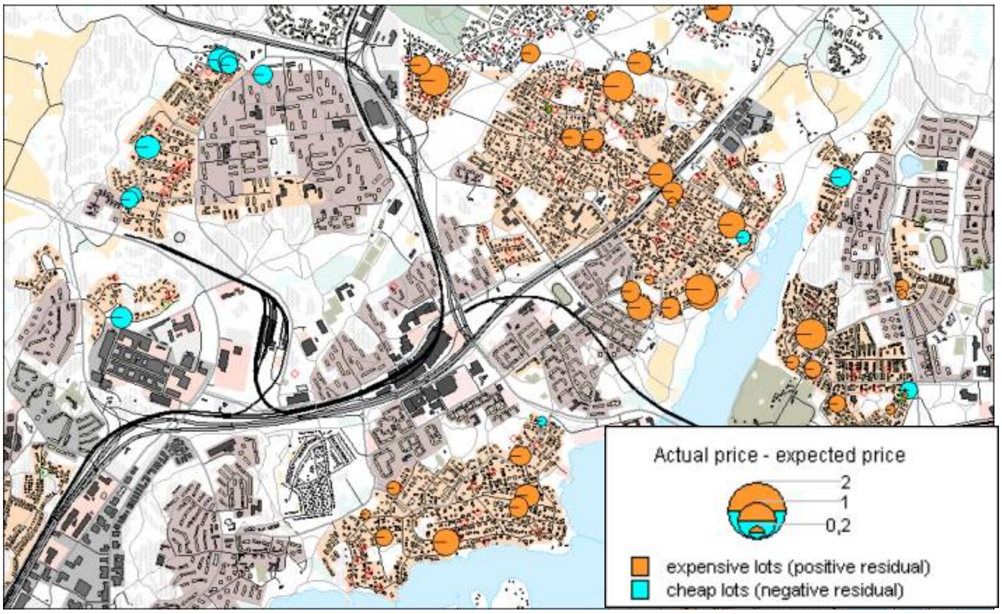


Figure 8. The error term of individual sales of housing lots.

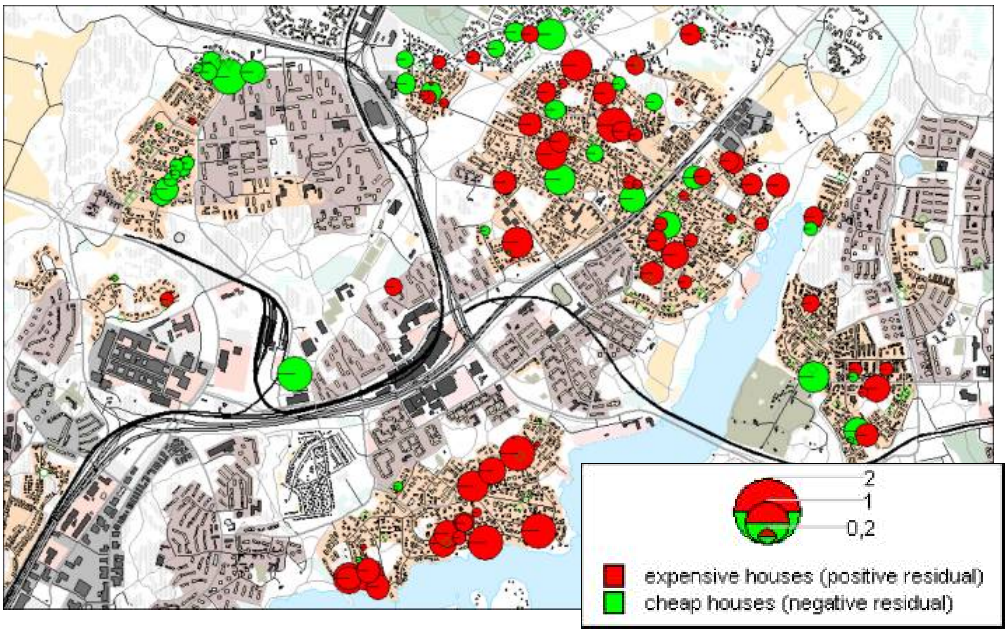


Figure 9. The error term of individual sales of houses.

4.3 A view at prices ratios

Finally, the prices ratios can be presented on a map, either in a grid level or in an individual sales level. Only two examples are given:

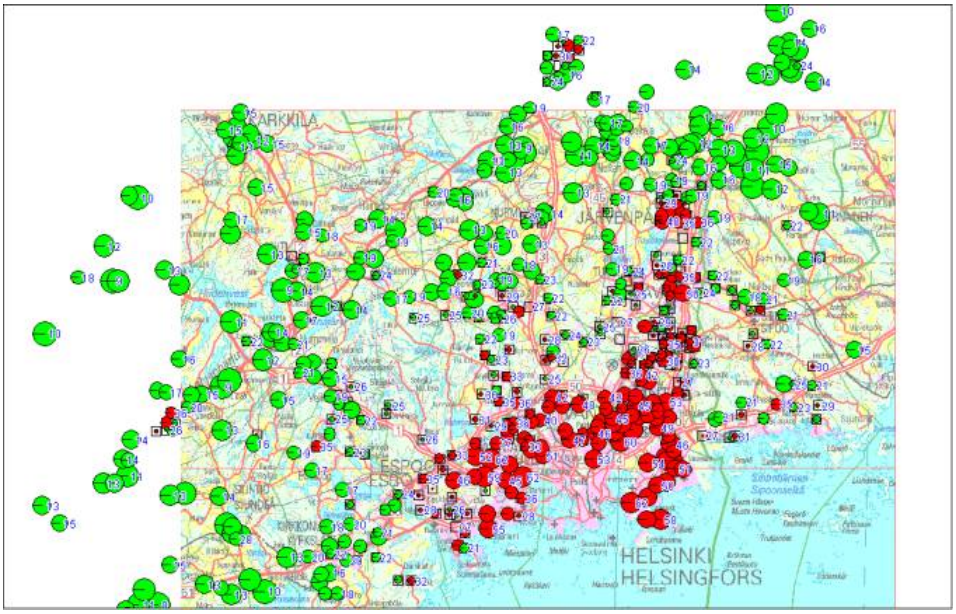


Figure 10. The land share of house price in a grid

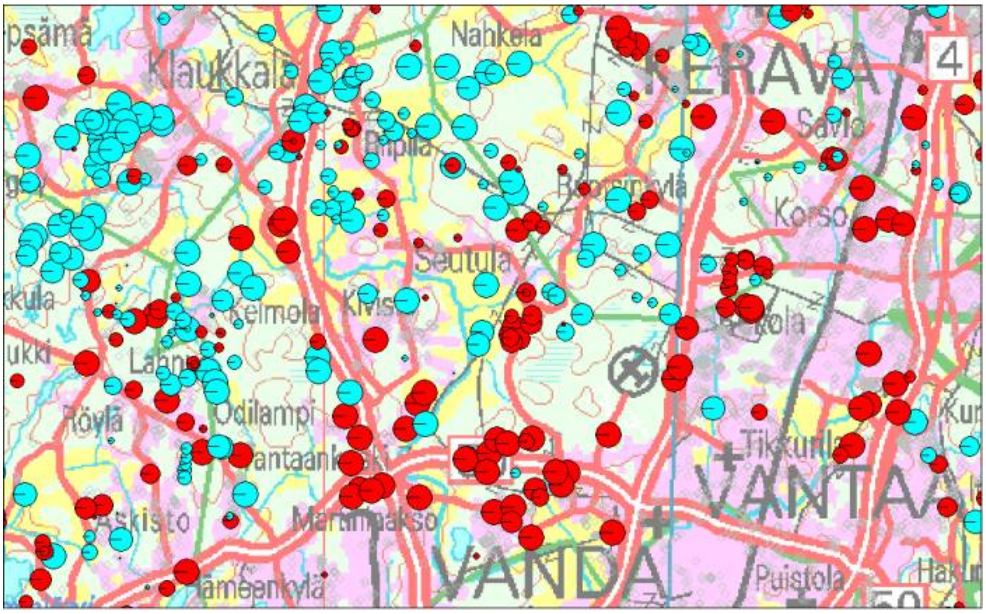


Figure 11. The ratio of development land to lot price. (red circles: ratio 5-25 %, blue circles: 1-5 %)

Figure 10 illustrates the land share of house price in a grid. The lot price has been estimated by model in time and space of the actual house transaction. An alternative way would have been to use estimated prices on both sides of the calculation to produce a quasi-constant-quality price ratio.

Figure 11 illustrates the ratio of development land price to lot price. Again, the lot price has been estimated by model in time and space of the actual development land transaction.

5. CONCLUSIONS

5.1 Spatial analysis implications

The toolbox for spatial analysis consists of two parts: hedonic modelling to produce a trend surface and thematic maps to explore the local effects.

The hedonic models reveal similar effects for spatio-temporal price factors along the housing land supply chain, but also some important differences. For houses the curves are flatter than for land. Proximity to main road has a special impact to development land.

The thematic maps reveal distinct spatial patterns in the error term indicating clearly, that global hedonic models are inadequate to capture spatial effects properly. It’s easy, even for a layman, to interprete the maps in a visual way. The maps convey effectively an overview of the price landscape. A closer view, to the level of individual transactions, is also possible, which makes the tool transparent, and spatial effects are possible to examine between any pair or group of observations.

The tool offers an easy way to produce visualizations of property markets. If a user has this toolbox at his disposal, he can easily produce any number thematic maps and zoomings in a short time, and check the underlying attribute data when necessary.

The explored spatial patterns are of varying type and importance. Recognizing the pattern help make decisions:

- A visible pattern can be used for hypotheses generating purposes, like any explorative method.
- The pattern may imply a way to improve the underlying hedonic model.
- The pattern may reveal a spatial autocorrelation, which cannot be regressed to any missing variable. Perhaps more sophisticated spatial statistics, such as kriging, is then needed to control the autocorrelation.
- Sometimes (quite seldom) no clear pattern emerges, so the underlying trend surface explains the spatial variation.
- A visible pattern is detected, and this is just enough for certain practical purposes.
- The pattern may reveal outliers or even a crude error in the data.

5.2 Housing policy implications

I have introduced (or probably more correctly: reinvented) two benchmarks for house price studies: (1) the land share of a house price and (2) a ratio of development land to a serviced lot price.

These ratios have some direct policy or business implications:

(1) if the land share of a house price is low, compared to near-by areas, it may be profitable to increase the supply of houses

(2) if the ratio of development land price to lot price is low, compared to near-by areas, it may be profitable to undertake planning and increase the supply of serviced housing lots.

I have demonstrated the spatial distribution of those two ratios in the Helsinki metropolitan area based on a transaction data of over 21 years and more than 45.000 sales. A clear spatial pattern emerges. The analysis of that pattern is as yet indefinite, but it is safe to say something:

The land share of a house price reveals an expected pattern: in central locations, where houses are most expensive, the land share of a house price is also highest.

The ratio of development land price to lot price reveal a similar, but more complex pattern. Development land price is more sensitive to location than lot price, so the ratio is highest in prime locations. A special case is land near main roads, where development land sales are plenty, land is expensive and location for residential purposes is not ideal.

However, the ratio is unusually low in subprime locations. In locations just 20-30 kilometres from Helsinki centre, or half an hours drive time, land may cost only 1-3 % of the estimated lot price in the same location. Why is this land not brought for residential construction but is kept on agricultural use?

Planning restrictions clearly play a role, but only in a short run. More important are costs to service those potential lots with infrastructure, not just roads, pipelines and parks, but also the education, health and nursery services that local governments are responsible for. But even these costs, or some substantial part of them, should be easy to finance given the appreciation in the value brought about by development.

There should be a way to finance those services and bring about the much needed construction of new homes. Obviously, the high price of housing lots should be considered as a solution to service finance problem. Property transaction data and a toolbox used here should help identify profitable areas for development. It should also help find ways to finance these development projects.

REFERENCES

Cheshire, P.C. and S. Sheppard (2003) The introduction of price signals into land use planning decision-making: a proposal. Research papers in Environmental and Spatial Analysis Series 89

Evans, Alan W. The Land Market and Government Intervention. Handbook of Urban Economics, Volume 3, Applied Urban Economics (p 1637-669). North-Holland 1999.

Evans, Alan W. and Hartwich, Oliver Marc: Unaffordable Housing, Fables and Myths, Policy Exchange 2005

O'Sullivan, David and Unwin, Dadid J.: Geographic Information Analysis, John Wiley & Sons, 2003.

CONTACTS:

Risto Peltola
National Land Survey of Finland
Development Centre
P.O.Box 84 00521 Helsinki
Finland

Telephone: +358 205 41 5628
Fax: +358 205 41 5700
Email: risto.peltola@nls.fi