Research on Finnish Society
Vol. 3 (2010), pp. 7–18

ISSN 1796-8739

An Econometric Examination on the Share of Land Value of Single-family Housing Prices in Helsinki

Elias Oikarinen
Turku School of Economics

This study brings empirical evidence on the importance of land value on housing prices in Helsinki Metropolitan Area (HMA). Utilizing econometric analysis and a quarterly dataset over 1988Q1-2008Q2, the results show that the value of land accounts for a significant fraction of single-family housing prices in HMA. In 2000-2007 the share of the land value component is estimated to be almost 50% of housing prices, on average. In line with prior expectations, the results also suggest that the land value component of housing has increased over time. The notable role and increase of the land value component has implications regarding housing price volatility. Since land prices appear to be more volatile than construction costs, it is anticipated that greater share of the land value component leads to more volatile housing prices. Given the significant role that housing wealth appears to play in the overall economy, this is of importance also for the economic policy makers.

Keywords: land price, housing price, construction costs

Introduction

The price of a house consists of the value of the physical structure together with the value of land upon which the house is built[1].The growth rate of the price of a house, in turn, is the weighted average of growth rates of the value of the structure and of the land upon which the house stands (Davis & Heathcote, 2007). The price of the structure is typically measured as the replacement cost of the physical building, after accounting for depreciation. Land, in turn, is the factor that makes a house worth more than the cost of putting up a new structure of similar size and quality on a vacant lot. In other words, land is the market value associated with the location, size and attractiveness of the site.

Even though a structure and the plot of land the structure stands upon are typically traded as a single bundle in the housing market, structures and land are quite different goods, whose prices should respond differently to shocks. Construction costs are generally relatively stable whereas land prices are much more volatile. The value of urban land is expected to be volatile, since land prices are typically driven by demand factors due to the inelasticity of the supply of land in desirable locations. Hence, the higher the land's share of the aggregate home value is, the more volatile housing prices are likely to be. Empirical evidence for the relatively high volatility of the value of land is presented by Somerville (1996) and Davis and Heathcote (2007) utilizing US data.

It is reasonable to believe that in large cities the share of land value of the total housing price level is significant, in general. Indeed, according to Rosenthal and Helsley (1999), Davis and Heathcote (2007) and Davis and Palumbo (2008) the value of land can account for over half of the price of housing in large cities. Furthermore, it is expected that the land value share increases as the population and real income level of the city grow. In line with this argument, Davis and Palumbo (2008) find the value of land to have accounted for about 50% of the total market value of housing in large US metro areas in 2004, while the share was only 32% in 1984.

Even though the relative share of the value of land of housing price level has implications regarding the magnitude of the impact of demand shocks on housing prices and thereby regarding housing price volatility, empirical research on the value of developed land is still scarce. While numerous studies such as Titman (1985), Capozza and Helsley (1989), Keushnigg and Nielsen (1996), Guntermann (1997), Rosenthal (1999) and Cunningham (2006), just to name a few, have analyzed price determination of land, only a few papers have investigated the value of developed land empirically.

The aim of this study is to bring empirical evidence on the share of the land value component of housing prices in Helsinki Metropolitan Area (HMA), the largest urban area in Finland. Utilizing econometric analysis and a quarterly dataset over 1988Q1-2008Q2, it is shown that the value of land accounts for a significant fraction of single-family hous-

The author works as a Postdoctoral Researcher at the Department of Economics, Turku School of Economics, University of Turku. Address: Turku School of Economics, Rehtorinpellonkatu 3, FI-20500 TURKU, Finland. Email: elias.oikarinen@tse.fi.

The author is thankful for the two anonymous referees and for the editor Jani Erola for valuable comments. Furthermore, the author gratefully acknowledges financial support from the OP-Pohjola Group Research Foundation.

[1] Due to weak demand for housing, in rapidly declining areas the price of a house can be less than the joint value of land and the replacement cost of the structure. At lowest the price can be the value of the land subtracted by demolition cost of the structure. In this study, however, it is only housing in non-declining metropolitan area that is considered.

ing prices in HMA. In 2000-2007 the share of the land value component is estimated to be almost 50%, on average. In line with prior expectations, the results also suggest that the land value component of housing has increased over time.

As explained above, it is anticipated that greater role of the land value component leads to more volatile housing prices. This should be of interest not only to those investors and households that own housing or are planning to buy housing but also to all the households that live in rental dwellings, since the growth of the share of land value component is likely to increase the volatility of rental prices as well. Furthermore, given the significant role that housing wealth appears to play in the overall economy, changes in the land value component are of importance also for the economic policy makers.[2] In particular, the policy makers should be aware of the fact that higher value of land is likely to strengthen the effects of changes in various economic policy instruments, such as alterations in tax rules or changes in the interest rates, on housing prices. Moreover, if expected housing price fluctuations affect policy decisions, the policy makers should understand that the influence of an economic shock on housing prices depends on the value of land in the area. That is, basing policy decisions on historical estimates of housing price sensitivity with respect to fundamentals in an area where real housing prices (and thereby land prices) have grown rapidly may notably increase the probability of misaligned policy decisions.

The paper proceeds as follows. The next section discusses the linkage between housing prices and land prices theoretically and sets the theoretical basis of the empirical examination. Furthermore, a review of the previous empirical literature on the value of urban developed land is presented. In the third section the data used in the empirical analysis is delineated. Then the findings from the econometric analysis are reported. In the end, the paper is summarized and some implications of the findings are discussed.

Theoretical considerations and previous literature

Typically, there are no data on the value of developed land. However, it is argued below that the price index of vacant lots can be employed to estimate the share of land value of the housing price level.

A simple way to look at the linkage between housing prices and prices of lots zoned for housing is to consider the price level for newly completed dwellings. The selling price of newly built housing can be presented by (1), where H denotes price of a unit of newly built housing, L is the unit price of land upon which the building is built and C signifies the unit cost of constructing the structure (including developers profit margin). Similarly, price of a unit of vacant land zoned for housing can be expressed by (2). This way of presenting the dependence between the prices of housing and of vacant land zoned for housing corresponds to the residual value of land view (see e.g. Somerville, 1996; Tse, 1998). Due to the tight linkages between the markets for housing and urban land, housing prices and the price of vacant land zoned for housing are dependent on each other and the prices are simultaneously determined. [3]

$$H = L + C \tag{1}$$

$$L = H - C \tag{2}$$

Equations (1) and (2) yield the following equations (where w denotes the share of land value component of the price of housing) for the housing and land price changes:

$$\triangle H = w * \triangle L + (1 - w) * \triangle C \tag{3}$$

$$\triangle L = \left[\triangle H - (1 - w) * \triangle C)\right] / w \tag{4}$$

Equation (3) states that the growth rate of housing prices is a simple weighted average of the growth rates of the land and structure components of housing. Naturally, (3) and (4) apply also for the old housing stock, H then being the unit price of old dwellings, L the unit value of developed residential land and C the value of the old structure. In empirical applications the evolution of C is typically estimated by movements in the construction costs.

As old dwellings and newly built homes within a metropolitan area can generally be considered to be close substitutes for each other, price growth of existing housing stock implies that higher prices can be charged also for newly built housing. In fact, substitutability between old and new housing implies that in (3) and (4) H can stand for the price of old housing while L can denote the price of vacant land zoned for housing. This is useful, since direct data on the value of developed land are not available, in general. Empirical support for the hypothesis that the price level of old housing stock is tightly linked to that of new housing construction in Finland is reported by Suoniemi (1990). Rosenthal (1999), in turn, finds different-vintage buildings to be close substitutes in Vancouver, Canada.

The framework considered above is, of course, a simplification of the reality. However, it illustrates the basic idea behind the linkage between housing prices and prices of lots. The price of vacant lots can be examined more rigorously by the real option pricing theory.

According to the real option pricing theory, owners of vacant land hold a call option that gives them the right, but not obligation, to develop the land. Due to the real option value, a lot can be more valuable as a potential site for development in the future than it is as an actual construction site at the present moment. It has been shown theoretically that the real option value of land is increased by greater uncertainty about future housing prices or rental prices (e.g. Titman, 1985; Capozza & Schwann, 1989; Capozza & Helsley,

[2] Significant housing wealth effect on consumption is reported e.g. by Case et al.(2005), Benjamin et al. (2004) and Campbell and Cocco (2007).

[3] See, for instance, the paper by Potepan (1996) where a simple theoretical model addressing the simultaneous determination of housing prices and undeveloped land prices is derived.

1989; Capozza & Sick, 1994; Capozza & Li, 2002). Empirical evidence by e.g. Quigg (1993) and Guntermann (1997) supports the existence of a notable real option value component of vacant land. Cunningham (2006), in turn, finds a standard-deviation rise in uncertainty to increase the price of vacant land by 1.6%.

The real option value capitalizes into the price of developed land and thereby into the price of housing. However, the relative impact of an increase in uncertainty regarding future rental price growth may not be the same on developed land as on vacant land situated outside the boundary of the city (see Capozza & Helsley, 1989). This may results in complications in an empirical analysis, such as the one conducted in this article. In general, a significant fraction of the vacant lots that are transacted in the market are located outside the boundary. Therefore, due to time variation in uncertainty, the perceived growth rate of price of vacant land zoned for housing may differ somewhat from the appreciation rate of the developed land. Hence, changes in uncertainty may have to catered for in an empirical analysis.

Previous published research on the share of land value of housing price level is limited. In an early study, Rosenthal and Helsley (1994) find the share of lot values of single-family housing prices to be as high as three quarters in the heavily developed areas adjacent to the downtown of Vancouver, Canada, in 1987. The computation is based on a comparison between the average selling prices of redeveloped properties and properties that were not redeveloped.

Also according to Davis and Heathcote (2008) and Davis and Palumbo (2008) the value of land can account for over half of the price of housing in large cities. Similarly to Rosenthal and Helsley, Davis and Heathcote as well as Davis and Palumbo infer the value of developed residential land indirectly. Employing a perpetual inventory system and using data on house prices and structure values, Davis and Heathcote construct quarterly time series for the aggregate value of housing stock. Then, by utilizing equation (4), they estimate quarterly time series for the value of aggregate stock of residential land. For the value of structures they use a price index for gross investment in new residential structures. They estimate that land accounted, on average, for 36% of the value of aggregate housing stock in the US between 1975-2006. Davis and Palumbo employ a methodology similar to that used by Davis and Heathcote. Their housing price and construction cost (structure value) variables differ somewhat from those utilized in Davis and Heathcote, however.

The method used by Davis and Heathcote and Davis and Palumbo requires time-series estimates of w. As Davis and Palumbo state, deriving these weights requires a good deal of work. This paper uses a different and more straightforward method, i.e. econometric analysis and the utilization of price index for vacant lots, to estimate the land value component. Also this method is indirect, though. The relative shares of the land and structure components are investigated by estimating the following regression model where H, L and C are in the natural log form::

$$H_t = \phi + \beta_1 L_t + \beta_2 C_t + \beta_3 U_t + \varepsilon_t \tag{5}$$

Due to the potential impact of uncertainty discussed above, a measure of uncertainty concerning rental price development (U) is included in the model. In (5) β_1, β_2 and β_3 are the estimated parameters on vacant lot price index (L), constructions cost index (C) and U. Furthermore, H stands for the price index of old housing stock. Since the housing price level consists of the two components (land and structure), the theory suggests that $\beta_1 > 0, \beta_2 > 0$ and $\beta_1 + \beta_2 = 1$. Moreover, since the growth rate of housing price is a simple weighted average of the growth rates of land and structures, $\beta_1(\beta_2)$ should equal the share of the land (structure) value component of housing prices. The coefficient on U, instead, is expected to be negative, since increase in uncertainty should result in greater relative rise in the value of vacant lots outside the boundary than in the value of developed land. The temporary deviation from the long-run relation is signified by ε. Finally, ϕ is a deterministic constant.

The relation in (5) is expected to be stationary[4] if the growth rate of L equals the growth rate of C over time. However, since desirable land is largely non-reproducible, changes in the demand for housing are likely to have substantial influence on the price of the land component of housing. By contrast, changes in the demand side are not expected to have much impact on the real price of structures, i.e. on construction costs, as noted by Davis and Heathcote (2007). Since, according the urban economic theory, growth in the demand factors such as population and income induce rise in the value of land, the increase in housing demand in a rapidly growing metropolitan area is expected to lead to faster growth in the value of land than in construction costs. That is, in a growing metropolitan area it is generally expected that L increases faster than C. Accordingly, in the HMA case β_1 is expected to be growing and β_2 decreasing in the long run. Therefore, recursive estimations are conducted in the empirical section to examine the time-variation in the parameters. As an additional specification check, model (6), including a time trend (t), is also estimated in the empirical analysis.

$$H_t = \phi + \beta_1 L_t + \beta_2 C_t + \beta_3 U_t + \delta t + \varepsilon_t \tag{6}$$

The trend may be needed to cater for the faster growth of land value than of construction costs. Because the share of the more rapidly growing component increases, δ is expected to be positive. In other words, housing appreciation is expected to be faster over the long horizon than implied by the parameters β_1 and β_2 alone. To preview, the findings reported in section 4 imply that relation (6) is, indeed, stationary over 1988Q1-2007Q4 whereas relation (5) is not. This supports the hypothesis that there are long-term trends in the shares of the land value and structure value components during the sample period.

[4] Covariance stationary, to be precise. Simply put, this means that the error term (ε_t) has a finite mean and variance and its mean and all autocovariances are unaffected by a change of time origin. For a more detailed discussion over stationarity and unit roots, see e.g. Enders (2004).

In line with the claim that the land price component of housing price level is likely to increase in growing areas over time, Davis and Heathcote (2007) find that land's share of aggregate home value has been trending upwards since the 1950s in the US. They show that the real price of residential land in the US rose 270% between 1975 and 2006, while the real price of housing structures increased only 33%. Based on their computations, land accounted for 46% of the value of housing stock in the US by the second quarter of 2006. According to Davis and Palumbo (2008), in turn, the value of land accounted for about 50% of the total market value of housing in large US metro areas in 2004, while the share was only 32% in 1984. Their analysis indicates that in 2004 Oklahoma City was the only large metropolitan area in which the share of land of housing price level was under one quarter. In San Francisco, the city with the highest land share figure, land accounted for nearly 90% of home value.

Finally, empirical evidence for the high volatility of the value of land is presented by Somerville (1996) and Davis and Heathcote (2007). Davis and Heathcote find land prices to be more than three times as volatile as the price of structures at business cycle frequencies. They bring evidence also for the demand driven nature of land prices and find the contemporaneous correlation between detrended real land and housing prices to be as high as .92.

The purpose of this paper is to bring empirical evidence on the magnitude of the land price component in HMA employing econometric analysis. The econometric analysis suggests that the share of land of the housing price level has increased over time and has been, on average, close to half during 2000-2007. These results are in line with the previous studies using US and Canadian data.

Data

The real housing price index (H) used in this study describes the price development of single-family detached houses in Helsinki Metropolitan Area.[5] The real land price index (L), in turn, depicts the evolution of price level of vacant land zoned for one-family houses in HMA. Both price indices are quality adjusted and published by Statistics Finland. The price indices are based on transactions in the privately financed market where prices are determined freely by demand and supply. Exceptional transactions have been excluded from the data based on which the indices are estimated.[6] The indices are quarterly and cover a period from 1985Q1 to 2008Q2. The quarterly values of the land price index are based on 80 transactions per quarter, on average. The corresponding figure for the housing price index is 101.

Since H relates to transactions taking place in the market for existing single-family houses, whereas L relates to the value of non-developed (but zoned) land, there is not perfect match between the two series. This may somewhat distort the empirical results. However, as noted in the theoretical considerations, substitutability between old and new housing implies that the prices of developed and non-developed land are tightly related to each other. The close linkage between developed and non-developed land prices should diminish the potential complications arising from the less than perfect match between the two price series.

Unlike typically in the empirical literature, the construction cost variable incorporates the developers' profit margin. Hence, C corresponds to the theoretical concept of construction costs in equations (1)-(4). The construction cost index is based on tender prices of new housing construction in HMA. The index is reported by Rapal Ltd and is available starting from 1988Q1. C approximates the development of the value of the physical structure.[7]

Furthermore, unemployment rate (U) reported by statistics Finland is incorporated into the dataset to approximate time variation in uncertainty regarding future housing (or rental) price development. It is assumed that people perceive the uncertainty regarding the future particularly high in times of high unemployment. On the other hand, greater uncertainty about the future development of demand side factors may induce higher unemployment rates (see Jellal et al., 2005). Therefore, the unemployment rate is likely to give information about the uncertainty concerning the future in the economy. The results from the econometric analysis are in line with this assumption. Obviously, unemployment is only a proxy for the uncertainty regarding future housing (or rental) price growth, however.[8]

Although hedonic indices are employed, there appear to be substantial short-run measurement error in the house and land price series. That is, the hedonic indices are probably

[5] HMA, as defined here, consists of Helsinki and the three nearest surrounding municipalities Espoo, Kauniainen and Vantaa.

[6] Regarding both houses and land, *Statistics Finland* excludes the transactions in which the size of the lot (empty or built) is exceptionally small or large or in which the transaction price per square meter is unusually small or large and the transactions that have been made between relatives. In the housing data, also transactions of exceptionally small or large houses are excluded. Moreover, the house price index is based on transactions where the structure and the plot of land the structure stands upon are traded as a single bundle. In other words, the index does not include the effect of houses that stand in rental land. The exclusion of the "exceptional" observations is conducted to diminish the heterogeneity complications regarding real estate units and to exclude potential errors in the transaction data. The share of the excluded observations is typically around one quarter of all the observations.

[7] The owner (instead of a construction company) is the "constructor" regarding a notable part of the new single-family housing production in Finland. Naturally, the time the owner devotes to building the structure is costly. If one would not build the house for own housing needs but to sell the property in the market, the value of the effort one devotes to building the house is the market value subtracted by the cost of the inputs other than labor. If the construction companies can sell their output at a price that includes notable profit margins, we can assume that the owner-builder could charge the same price if he/she did not occupy the house himself/herself. Therefore, even in the case where the owner is the "builder", movements in the index that includes profit margins work as a reasonable proxy for the changes in the value of the structure.

[8] Similar measures of housing price uncertainty to those employed by Cunningham (2006) were also tried in the analysis. Those measures did not yield sensible results.

Table 1
Descriptive statistics of differenced variables over 1988Q2-2007Q4.

Variable	mean (annualised)	Standard deviation (annualised)	Jarque-Bera (p-value)	Ljung-box test for auto-correlation (p-values, 4 lags)	Seasonal variation (p-value, F-test)
Housing price	0.021	0.078	0.00	0.00	0.57
Land price	0.030	0.127	0.00	0.00	0.02
Construction costs	-0.011	0.067	0.03	0.00	0.66
Unemployment	0.001	0.012	0.00	0.00	0.63

not able to perfectly track the quality variation of the transacted houses and lots. In particular, due to thin trading in the markets the price series are likely to include "noise" in the short-run, i.e. the short-run variation of the price series may well be overly large. Due to the apparently substantial noise especially in the land price series, the five-quarter centered moving averages of the H and L series are employed. The use of the moving averages is reasonable even though it reduces the number of observations by two, since the substantial noise in the "unsmoothed" price series might distort the results of the econometric analysis. Because of the use of moving averages and of the limits set by the construction cost index, the effective sample period is 1988Q1-2007Q4.

In the econometric analysis the series, except for U, are indexed and have the value of 100 in 1988Q1. Furthermore, natural logarithms are taken from all the indexed series. Only real values are employed in the study. Nominal values are deflated by the cost of living index to get real variables. Figure 1 exhibits the series included in the empirical analysis. Also the unsmoothed housing and land price series are shown in the Figure.

The dramatic rise since 1988 in housing and lot prices was largely a consequence of the financial market liberalization in the late 1980s that was followed by a boom in bank lending. The housing and lot markets finally collapsed at the beginning of the 1990s. The drop in housing and land prices as well as in the other asset prices was deepened by the severe recession in Finland in early and mid 1990s. Also C drastically decreased during 1990-1993. This was largely due to a drop in the profit margins. After the mid 1990s housing and lot prices have grown substantially faster than the general price level.

Expectedly, vacant land prices have grown more during the sample period than construction costs and housing prices. Also in line with the prior expectations, land prices have been more volatile than housing prices and construction costs. Figure 1 cannot be used to estimate the relative magnitudes of the land value and structure components of the total housing price level, however. Those magnitudes have to be evaluated by a proper econometric analysis. Note also that L and C series are only approximations of the true underlying factors (developed land value, structure value). This may be a partial reason behind the fact that H is above both L and C for several years during the late 1990s and in the early 21st century. Another reason may be the influence of uncertainty. Anyhow, it is reasonable to believe that in the long run the employed data describe well the underlying fundamentals.

Table 2
Contemporaneous quarterly correlations between differenced series.

	H	L	C	U
H	1			
L	.59**	1		
C	.57**	.38**	1	
U	-.58**	-.52**	-.50**	1

** and ** denote for statistical significance at the 5% and 1% level, respectively. The reported correlations are based on a sample period 1988Q2-2007Q4.*

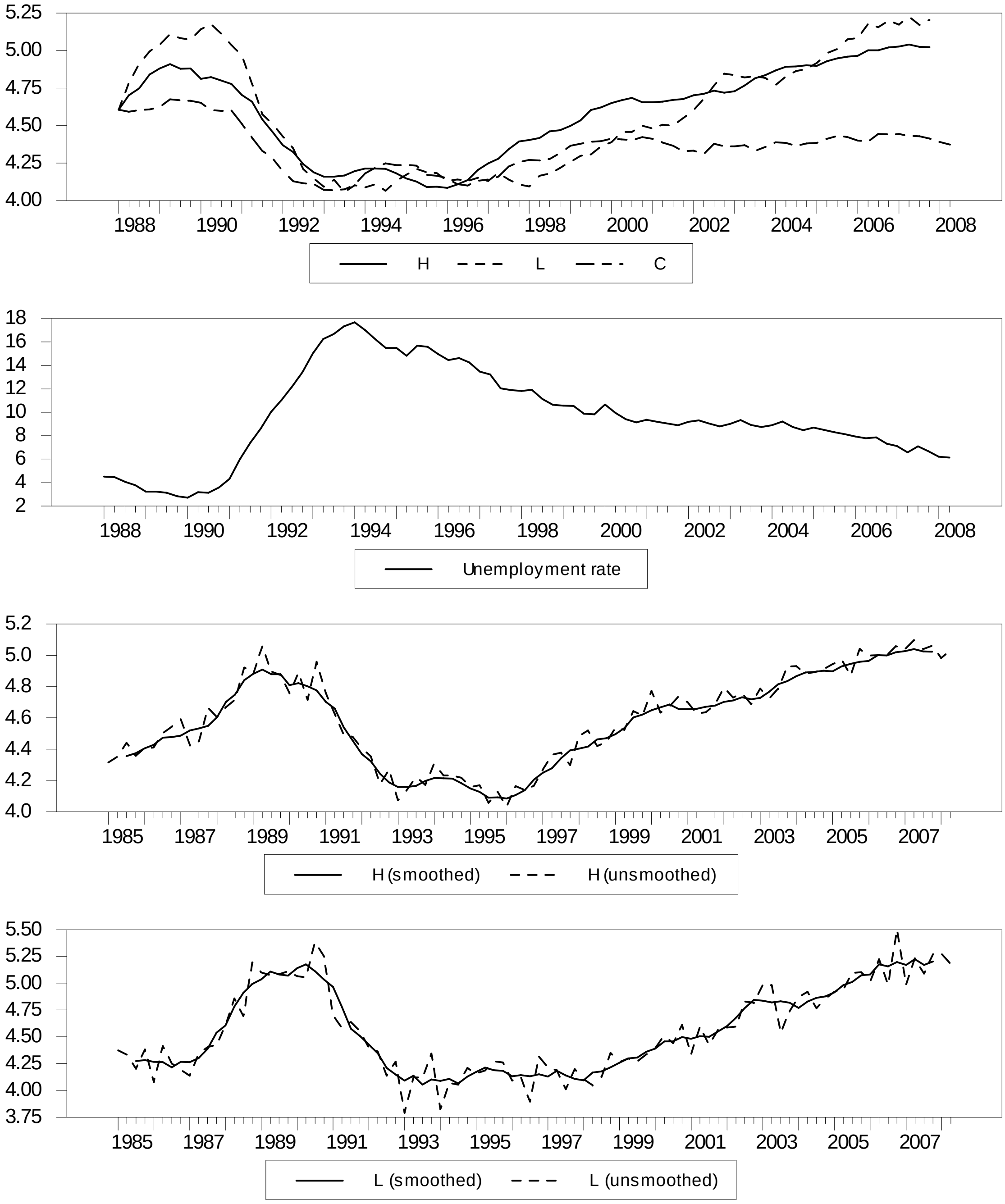
Table 1 reports some descriptive statistics of the differenced series. All the variables are highly autocorrelated.

Unsurprisingly, all the contemporaneous quarterly correlations between the differenced variables are statistically highly significant (see Table 2). The figure between lot and house appreciation is .59. The correlation between housing price growth and construction cost movements is of similar magnitude. The correlations based on the unsmoothed housing and land price series are considerably smaller than those reported in Table 2 because of the substantial noise in the original indices. The correlation between the unsmoothed changes in H and L is as low as .16.

Econometric evaluation

In this section, an econometric analysis is conducted to examine the magnitude of the shares of land and physical structure components of the total price of single-family housing in HMA. As explained in the theoretical section, housing prices, lot prices and construction costs are expected to form a stationary long-run relation towards which at least one of the variables adjusts and in which $\beta_1 + \beta_2$ is approximately one. In other words, the three variables are expected

Figure 1. Plot of the variables included in the empirical analysis.



to be cointegrated.[9] Time variation in uncertainty may have to be accounted for in order to find such a relation. Note that in none of the estimated long-run models the restriction $\beta_1 + \beta_2 = 1$ is imposed. This is because the prices of vacant lots are employed in the analysis, not the value of developed residential land. After all, old and newly completed houses might not be perfect substitutes. Besides, the data may well exhibit some measurement error. Anyhow, in none of the estimated regression models $\beta_1 + \beta_2$ differs from one statistically significantly.

In this analysis, the cointegration tests can be seen as tests for model misspecification. The existence of a stationary relationship is tested first by employing the Johansen (1996) Trace test for cointegration. However, due to the substantial difference between the growth rates of L and C, such stable long-run relation may not exist. Therefore, a model including a deterministic trend in the long-run relation (see equation 6 on page 9) is tested in addition to the conventionally used model excluding deterministic trend (corresponds to equation 6 on page 9). The trend term might cater for the fact that L has grown faster than C causing the actual coefficient on L to grow over time and thereby leading to faster growth of the equilibrium level for P than suggested by the time-invariant parameter estimates.

Indeed, the Johansen Trace test is able to find a cointegrating relation between H, L and C only if a trend is included in the long-term relation.[10] Furthermore, detection of a reasonable stationary relation requires the inclusion of unemployment rate as a proxy for uncertainty. The Trace test results are reported in Table 3.[11]

The theory indicates that if the relative shares of land and structure of house price level are constant over the long run, equation (5) should be stationary. The model based on equation (5) is misspecified, however, since it is not stationary. This, together with the stationarity of equation (6), is in line with the argument that the shares of land and structure are trending over the long horizon. The origin of the complication of the cointegration analysis without a trend term is further illustrated in Appendix A. In the model including the trend in the long-term relation, the long-run coefficients on L (.207) and C (.796) do not straightforwardly show the relative shares of land value and structure components of the housing price level during a particular time period. This is because the estimated equilibrium housing price level has grown by some .6% per quarter faster than implied by the coefficients on the stochastic variables only. In any case, $\beta_1 + \beta_2 \approx 1$ as anticipated. Furthermore, in line with the theory, increased uncertainty (unemployment rate) has lowered housing prices, i.e. the value of developed land, relative to vacant land prices.

The speed of adjustment of housing prices towards the relation is estimated to be 9.1% per quarter. The corresponding figures for L and C are 22% and 32%. The notable speed of adjustment parameter of C suggests that a positive housing price shock has a positive impact on the profit margins, i.e. that competition among developers is not perfect. This argument follows from the fact that the (national level) real construction costs without profit margins have been almost stable during the sample period whereas the costs including profit margins have varied substantially and simultaneously with housing prices (see Appendix B).

Table 3
Trace test statistics.

H_0 (rank)	*Trace test value (p-value) Model excluding trend*	*Trace test value (p-value) Model including trend*
$r = 0$	42.3 (.15)	76.7 (.00)
$r \leq 1$	21.7 (.33)	46.9 (.02)
$r \leq 2$	11.3 (.20)	19.2 (.28)
$r \leq 3$	.13 (.72)	7.51 (.30)

P-value in the test for weak exogeneity of U in the model including trend = .58
Estimated long-run relation based on the model including trend (standard errors in parenthesis):
$H = .207 * L + .796 * C - .023 * U + .006t$

The tested models include H, L, C and U, three lags in differences, and three seasonal dummies. The number of lags is selected based on the Hannan-Quinn information criteria and the LM(1) and LM(4) tests for residual serial correlation. The trace test values are based on the small-sample correction by Johansen (2002) and weak exogeneity is tested by the LR test by Johansen (1996). The parameter estimates are based on a model in which U is restricted to be weakly exogenous, i.e. U does not react to deviations from the estimated long-run relation.

The cointegration analysis shows that it is misleading to assume the relative shares of L and C to be time invariant. Therefore, recursive regression analyses are conducted. Stationarity of the estimated recursive regressions is examined employing the Engle-Granger (1987) methodology[12]. The more sophisticated Johansen Trace test is not employed, since the degrees of freedom in many of the tested models would be extremely small due to the several lags required in the tested models.

In the first recursive analysis the regression is run using the first ten quarters and then observations are added to the regression one at a time. Figure 2 exhibits the estimated

[9] Expectedly and in line with previous empirical evidence, the ADF unit root test indicates that all the variables are non-stationary in levels but stationary in differences. The unit root test results are reported in Appendix C.

[10] The p-value of .15 indicates that there might be, at least, a close to stationary relation between the variables in the model excluding the trend. However, it appears that the relatively low p-value is due to close to stationary C. The p-value in the test for exclusion of both H and L from this potentially stationary relation is .87.

[11] The Trace test statistics regarding the model with a time trend actually indicate that there are two stationary vectors between the variables. It appears that this is because C is close to stationary.

[12] The critical values are based on the response surface coefficients estimated by MacKinnon (1996). The detailed test results are available from the author on request.

coefficients together with their 10% confidence bands. As the residuals of the estimated models seem to exhibit heteroscedasticity and autocorrelation, all the reported confidence bands are based on heteroscedasticity and autocorrelation (allowing for one lag) consistent covariance matrices (see e.g. *RATS, version 6, user's guide*, 2006, 179-186). In general, the coefficient on L trends upwards, while that on C decreases over the sample period. The coefficient on L increases from about 40% to almost 60% as the sample period is extended. Occasionally, there is substantial deviation from the "trend", though, and the coefficient is at smallest (23%) when the employed sample period is 1988Q1-2001Q1. In Figures 2-3 the time period in the horizontal axis refers to the last quarter of the employed sample period.

A problem with the recursive analysis above is that the residual starts to trend upwards and the regressions are no longer stationary as the utilized sample period gets longer. Hence, the coefficient estimates and confidence bands shown in Figure 2 should be studied cautiously. If, instead, a rolling window of 30 observations is employed, the recursive regressions appear to be stationary. The estimates from this recursive analysis together with 10% confidence bands can be seen in Figure 3.

Also according to Figure 3 there is substantial time variation in the relative shares of the two components of housing prices. As anticipated, the late sub-sample (2000Q3-2007Q4) coefficient, 48%, on L is larger than the early sub-sample (1988Q1-1995Q2) one (39%).[13] The share of land value of housing price level substantially decreased during the recession and housing price bust of the 1990s. After the mid 1990s the share of land value has increased rapidly having more than doubled from the bottom of the bust. In 2007Q4 the land value share may well be somewhat greater than 48%, since L has grown substantially faster than C after 2000Q3. The findings according to which the share of land price component has grown over time and has accounted for about half of the single-family housing price level during the 2000s is not surprising in light of the previous literature.

Note that the recursive estimates in Figure 3 are based on regressions where U is incorporated in the model in the sub-samples from 1991Q1-1998Q2 to 1997Q2-2004Q3. This is because the regressions using sub-samples between 1991Q2-1998Q3 and 1995Q3-2002Q4 would be non-stationary without U in the model and the coefficient on U is significant and negative even longer, i.e. in all the sub-samples in which the starting date is between 1991Q1 and 1997Q2[14]. The absolute value of the coefficient on U is .04 at its greatest. Moreover, the standard errors of the coefficient estimates get notably smaller in the aforementioned sub-samples if U is present in the model.

It appears that the increase in the share of the land value component may have considerable effect on housing price volatility. However, the impact may be rather small on relatively short-run volatility. For instance, if the share of land value is 25%, the annual standard deviation of housing price growth is expected to be 11.7%, assuming that the annual volatilities of and annual correlation between land prices and construction costs remain the same as during the sample period on average. The expected volatility of housing appreciation would be only 17% larger, i.e. 13.8%, if the land value component was 50%. At business cycle frequencies the impact of the growth of the land value component is notably greater due to the greater long-horizon correlation between land price and construction cost changes and because of mean-averting land prices. At a ten-year horizon the corresponding expected (annualized) volatilities are 11.1% and 15.0%. That is, the increase of the share of land value from 25% to 50% would enlarge the expected volatility of housing price growth by 35%. A rise of this magnitude in the housing price volatility may have significant effects on the cyclicality of the overall economy.

Summary and discussion

The price of a house consists of the value of the physical structure together with the value of land upon which the house is built. The growth rate of the price of a house, in turn, is the weighted average of growth rates of the value of the structure and the land upon which the house stands. Previous empirical evidence by Rosenthal and Helsley (1994), Davis and Heathcote (2007) and Davis and Palumbo (2008) from North America suggests that the value of land can account for significant share of the price of housing in large cities. Furthermore, it is expected that the share of land value of total price of housing increases in rapidly growing areas. This follows from the scarcity of land in attractive locations.

The aim of this article is to estimate the share of the land value component of the total housing price level in Helsinki Metropolitan Area (HMA) employing econometric analysis. It is argued that the price data on vacant lots can be utilized to estimate the share of land value component of housing prices. Employing a quarterly dataset over 1988Q1-2007Q4, it is shown that the value of land accounts for a significant fraction of single-family housing prices in HMA. During 2000-2007 the share of the land value component is estimated to be 48%, on average. In the end of the sample period the land value share may well have been somewhat greater than 48%. The analysis also suggests that, as expected, the land value component of housing has increased over time. The recursive estimation proposes that the land value share was less than 20% during the great depression and housing market bust in the early and mid 1990s.

The impact of shocks in the demand factors (such as demographics, interest rates, or tax treatment of housing) on housing prices is likely to be the greater the larger is the land

[13] There is no readily available test to examine if the coefficients differ statistically significantly.

[14] If U is not included in the regressions, the standard deviation of the estimates is extremely large when the starting period of the sub-sample is between 1991Q1-1994Q4. At the same time, the estimated coefficient on L first drops dramatically and then increases rapidly. In fact, some of the point estimates for the coefficients on L and on C get implausible values (below zero or over one). However, in the early and late sub-samples U would be small and insignificant and in many cased would have the wrong (positive) sign. Therefore U is included only in the models that use the mentioned sub-periods.

Figure 2. Coefficients on land price (black curve) and construction costs (grey curve) based on the recursive analysis.

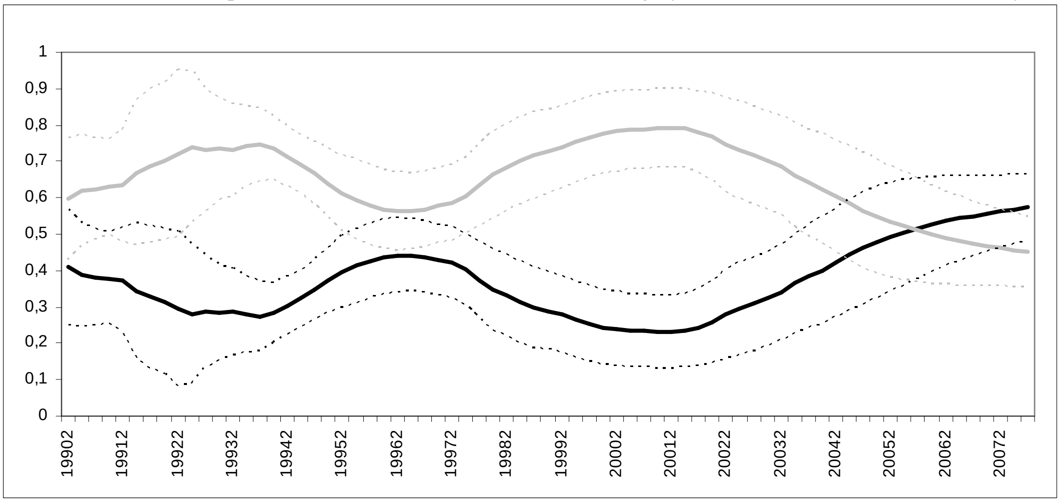
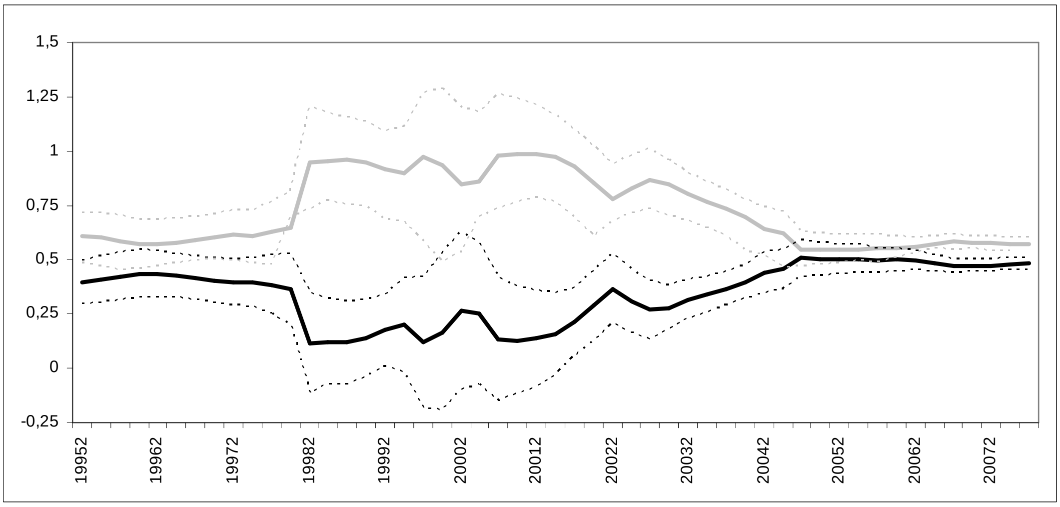


Figure 3. Coefficients on land price (black curve) and on construction costs (grey curve) based on the recursive analysis employing 30 quarter moving window.



value component. On the other hand, in sparsely populated areas where land is cheap housing price movements should be determined to a substantially greater extent by construction costs. The high value of land in HMA suggests that the demand side factors play a major role in the future evolution of housing prices in the area. Furthermore, as the value of land appears to be considerably more volatile than construction costs, greater land value component leads to higher expected housing price volatility[15]. This may have notable consequences not only in HMA but in the whole of Finland.

Because of the migration from periphery to the central cities, greater and greater share of dwellings is located in areas with relatively high value of land. Moreover, the value of land further rises in the central cities due to the increasing demand to housing (and thereby for developed land). That is, centralization may cause the national level housing prices to be more volatile in the future than in the past. For instance, doubling of the land value component from 25% to 50% is expected to increase the standard deviation of HMA single-family housing price growth by 35% at a ten-year horizon, assuming that the volatilities of and correlation between land prices and construction costs remain the same as during 1988-2007. This should be of interest not only to those investors and households that own housing or are planning to buy housing but also to all the households that live in rental dwellings, since the growth in the share of land value of housing prices is likely to increase the volatility of rental prices as well. Furthermore, given the significant role that housing wealth appears to play in the overall economy, the value of land is of importance also for the economic policy makers. Basing policy decisions on historical estimates of housing price sensitivity with respect to fundamentals in

[15] In other words, the fraction of home value that is accounted for by the value of land is a critical determinant of the elasticity of housing supply as emphasized by Glaeser and Gyourko (2005) and Gyourko and Saiz (2004).

an area where real housing prices (and thereby land prices) have grown rapidly may notably increase the probability of misaligned policy decisions.

One should also note that the increase in the value of land in growing areas works as a natural counterforce for centralization. As the value of land increases, the costs of living rise. Higher living costs lead to higher wage demands. This, together with the higher rental prices for the production facilities, raises production costs thereby lowering profitability. If the benefits from agglomeration are greater than the harms caused by land price growth and by the other counterforces for centralization, employment opportunities in the area are likely to grow. An important question is, to what extent the public sector (tax payers) should intervene this mechanism, i.e. enhance centralization, by supplying housing or land at lower than market prices.

References

Benjamin, J. D., Chinloy, P., & Jud, G. D. (2004). Why Do Households Concentrate Their Wealth in Housing? *Journal of Real Estate Research*, *26*, 329–344.

Campbell, J. Y., & Cocco, J. F. (2007, April). How do house prices affect consumption? Evidence from micro data. *Journal of Monetary Economics*, *54*(3), 591–621.

Capozza, D. R., & Helsley, R. W. (1989). The fundamentals of land prices and urban growth. *Journal of Urban Economics*, *26*(3), 295–306.

Capozza, D. R., & Li, Y. (2002). Optimal land development decisions. *Journal of Urban Economics*, *51*(1), 123–142.

Capozza, D. R., & Schwann, G. M. (1989). The asset approach to pricing urban land: Empirical evidence. *Real Estate Economics*, *17*(2), 161–174.

Capozza, D. R., & Sick, G. A. (1994). The risk structure of land markets. *Journal of urban Economics*, *35*, 297–297.

Case, K. E., Quigley, J. M., & Shiller, R. J. (2005). Comparing wealth effects: the stock market versus the housing market. *The B.E. Journal of Macroeconomics*, *5*(1).

Cunningham, C. R. (2006). House price uncertainty, timing of development, and vacant land prices: Evidence for real options in Seattle. *Journal of Urban Economics*, *59*(1), 1–31.

Davis, M. A., & Heathcote, J. (2007, November). The price and quantity of residential land in the united states. *Journal of Monetary Economics*, *54*(8), 2595–2620.

Davis, M. A., & Palumbo, M. G. (2008). The price of residential land in large US cities. *Journal of Urban Economics*, *63*(1), 352–384.

Enders, W. (2004). *Applied econometric time series* (2nd ed.). New York: John Wiley & Sons Inc.

Engle, R. F., & Granger, C. W. J. (1987, March). Co-Integration and error correction: Representation, estimation, and testing. *Econometrica*, *55*(2), 251–276.

Glaeser, E. L., & Gyourko, J. (2005, April). Urban Decline and Durable Housing. *Journal of Political Economy*, *113*(2), 345–375.

Guntermann, K. (1997). Residential land prices prior to development. *Journal of Real Estate Research*, *14*(1), 1–18.

Gyourko, J., & Saiz, A. (2004). Reinvestment in the housing stock: the role of construction costs and the supply side. *Journal of Urban Economics*, *55*(2), 238–256.

Jellal, M., Thisse, J., & Zenou, Y. (2005, November). Demand uncertainty, mismatch and (un)employment. *Economics Letters*, *89*(2), 248–254.

Johansen, S. (1996). *Likelihood-based inference in cointegrated vector autoregressive models*. Oxford: Oxford University Press.

Johansen, S. (2002). A Small Sample Correction for the Test of Cointegrating Rank in the Vector Autoregressive Model. *Econometrica*, *70*(5), 1929–1961.

Keuschnigg, C., & Nielsen, S. B. (1996, September). Housing markets and vacant land. *Journal of Economic Dynamics and Control*, *20*(9-10), 1731–1762.

MacKinnon, J. G. (1996). Numerical distribution functions for unit root and cointegration tests. *Journal of Applied Econometrics*, *11*(6), 601–18.

Potepan, M. J. (1996). Explaining Intermetropolitan Variation in Housing Prices, Rents and Land Prices. *Real Estate Economics*, *24*(2), 219–245.

Quigg, L. (1993). Empirical testing of real Option-Pricing models. *Journal of Finance*, *48*(2), 621–40.

RATS, version 6, user's guide. (2006). Evanston: Estima.

Rosenthal, S. S. (1999, May). Residential buildings and the cost of construction: New evidence on the efficiency of the housing market. *Review of Economics and Statistics*, *81*(2), 288–302.

Rosenthal, S. S., & Helsley, R. W. (1994). Redevelopment and the urban land price gradient. *Journal of Urban Economics*, *35*, 182–182.

Somerville, C. T. (1996). The contribution of land and structure to builder profits and house prices. *Journal of Housing Research*, *7*, 127–141.

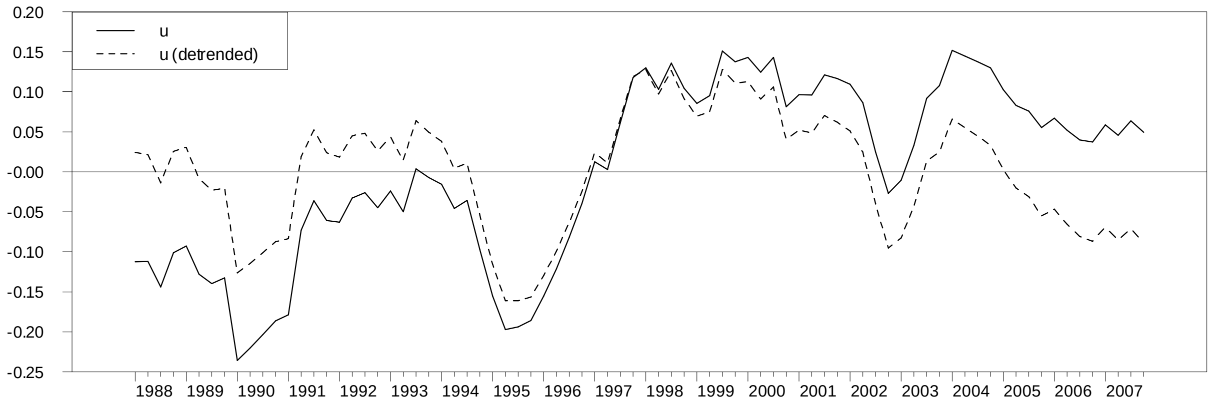
Suoniemi, I. (1990). *Uusien ja vanhojen asuntojen hinnat suomessa. hintasarjojen trendi- ja kausaaliominaisuudet* (Discussion papers No. 103). Helsinki: Finnish Labour Institute for Economic Research.

Titman, S. (1985, June). Urban land prices under uncertainty. *The American Economic Review*, *75*(3), 505–514.

Tse, R. Y. C. (1998). Housing price, land supply and revenue from land sales. *Urban Studies*, *35*(8), 1377–1392.

Appendix A

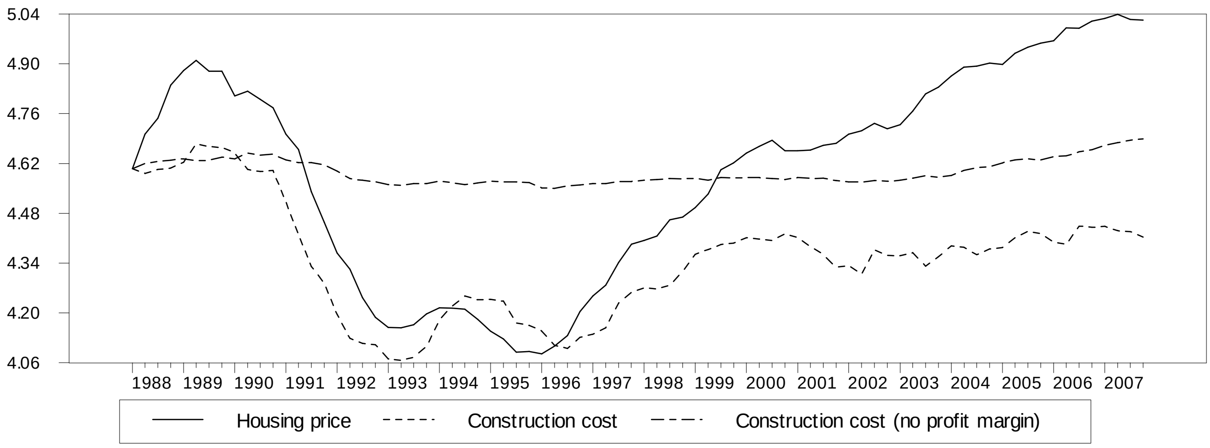
Plot of the residual from the whole sample OLS regression (u) and of the detrended residual.



u is the residual series from a conventional Ordinary Least Squares (OLS) regression without a trend: $H = .572 * L + .452 * C$. Since u trends clearly upwards, the model is misspecified. If u is regressed on a constant and trend, we get $u = -.182 + .003t$ with both of the estimated parameters being highly significant and the adjusted coefficient of determination being as large as .53. Residual from this regression is presented as a dash curve in the Figure. U is not included in the model since its coefficient is insignificantly different from zero and would have the wrong sign.

Appendix B

Housing price index and the construction cost indices both including profit margins and excluding profit margins.



Appendix C

Augmented Dickey-Fuller test results.

Variable	*Level (lags)*	*Difference (lags)*
House price	-1.65 (6) [c]	-2.34* (5)
Land price	-1.79 (6)[c,s]	-2.60** (5)[s]
Construction costs	-2.04 (1)[c]	-5.02** (0)
Unemployment rate	-0.67 (5)	-2.22* (4)

* and ** denote for statistical significance at the 5% and 1% level, respectively. Critical values at the 5% and 1% significance levels are -1.95 and -2.60 if constant is not included and -2.89 and -3.51 in the case where constant is present. The number of lags included in the ADF tests is decided based on the general-to-specific method. A constant term (c) is included in the tested model if the series clearly seems to be trending or if the ADF test without the constant term suggests that the series is exploding. In addition, three seasonal dummies (s) are added to the test if recommended by the F-test.