Workshop 22 - Housing Markets Dynamics

Dynamic linkages between prices of vacant land and housing – empirical evidence from helsinki

Elias Oikarinen
elias.oikarinen@tukkk.fi

Peltola Risto

Paper presented at the ENHR conference
"Housing in an expanding Europe: theory, policy, participation and implementation"
Ljubljana, Slovenia
2 - 5 July 2006
DYNAMIC LINKAGES BETWEEN PRICES OF VACANT LAND AND HOUSING – EMPIRICAL EVIDENCE FROM HELSINKI

By

Elias Oikarinen*

and

Risto Peltola†

Abstract

House price consists of the replacement cost of the physical structure together with the value of land upon which the house is build. Construction costs are, in general, relatively stable whereas land prices are much more volatile. Hence, volatility in house prices is typically due to changes in the value of constructed land. Movements in the value of constructed land, in turn, should have a direct impact on prices of vacant lots. In an efficient market the change in vacant lot prices should be immediate. Because of informational reasons due to factors such as thin trading in the land market, however, in reality housing appreciation may well have a leading role with respect to price movements of vacant lots. Employing data from the Helsinki Metropolitan Area, this study shows that housing price movements lead price changes in the market for vacant lots. Based on the empirical analysis it is suggested that house prices respond to shocks influencing the value of land first, after which the price level of vacant lots reacts to the information revealed by housing appreciation. Hence, the results indicate that the market for vacant lots is more inefficient than the housing market.

* Corresponding author. Contact details: Elias Oikarinen, Turku School of Economics, Rehtorinpellonkatu 3, 20500 Turku, Finland; elias.oikarinen@tse.fi (@tukkk.fi until 1.6.2006); tel. +358 40 5871943.

† National Land Survey of Finland
1 Introduction

It is obvious that housing price level is dependent on the price of land. House price consists of the replacement cost of the physical structure together with the value of land upon which the house is build. Yet, it is possible that housing appreciation leads the price changes perceived in the market for vacant land. Lead-lag relation between the housing market and the lot market would have implications for lot market efficiency and for the predictability of lot prices. Nevertheless, while a number of studies such as Titman (1985), Capozza (1989), Keushnigg and Nielsen (1996), Guntermann (1997), Rosenthal (1999) and Cunningham (2006), just to name a few, have analyzed price determination of vacant land, empirical research on the linkages between housing market and market for vacant land is still extremely scarce. The aim of this study is to bring evidence on this area by employing data from the Helsinki Metropolitan Area (HMA), the largest urban area in Finland.

The growth rate of price of a house is the weighted average of increase in the value of the structure and appreciation of land upon which the house stands. 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. The analyses by Rosenthal (1999) and Davis and Heathcote (2005) suggest that the value of land can account over half of the price of housing in large cities.

In general, housing price movements are likely to correlate strongly with land price changes. This is because construction costs are typically relatively stable whereas land prices are much more volatile. Thus, the higher the value of land is, the greater is correlation between housing appreciation and land price movements and the more volatile housing prices are likely to be. In fact, housing prices and land prices may well be cointegrated, i.e. there may be a long-run equilibrium relationship between land and housing prices in a given area. This could be the case if replacement costs of the physical structure are stationary. Reasoning behind the high volatility of land prices is that land prices are driven by demand, since the supply of land in desirable location is inelastic. Empirical evidence for the high volatility of the value of land is presented by Somerville (1999) and Davis and Heathcote (2005) utilizing US data. Davis and Heathcote 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 .94.

It is reasonable to assume that cross-regional differences in construction costs within a relatively homogenous and small country, such as Finland, are small relative to differences

---

1 Due to weak demand for housing, in rapidly declining areas 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 areas that is considered.
in land prices. Construction costs in HMA are the highest in Finland, approximately 20% higher than construction cost in some peripheral areas. In any case, as housing price level in HMA is substantially greater than in the other areas in Finland, it is reasonable to assume that also the value of land in HMA notably exceeds that in the other Finnish regions. The price per square meter of dwellings in HMA is approximately twice that of the price in the other parts of Finland. This suggests that the price of land may actually account for more than half of the price of a typical dwelling in HMA. Because of the high value of land, housing appreciation in HMA is likely to correlate strongly with land price changes.

At first thought there should be no lead-lag relation between the values of housing and land, because land prices form a major part of housing prices. In fact, according to the land pricing framework of Titman (1985) price movements of vacant lots should precede housing price movements. In practice, however, leading role of housing appreciation with respect to price movements of vacant lots may be caused by informational reasons due to factors such as thin trading in the lot market and lack of publicly available data concerning transactions in the land market. As explained above, housing appreciation in a metropolitan area is usually driven by changes in the value of land. Therefore, it is reasonable to assume that, in general, price changes perceived in the housing market reveal information concerning movements in the value of land. If the market for vacant lots is informationally efficient, price level of vacant lots should adjust immediately to the new information.

One option to look at the channel through which housing price movements might lead changes in the prices of vacant lots is to consider price level for newly completed dwellings. The selling price for new newly built housing can be presented by (1), where \( P^H \) denotes for price of a unit of newly built housing, \( P^L \) is the unit price of the 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 land unit can be expressed by (2).

\[
\begin{align*}
P^H &= P^L + C \\
P^L &= P^H - C
\end{align*}
\]

As existing dwellings and newly built homes within a metropolitan area can be considered to be relatively close substitutes for each other, growth of the price of existing housing implies that higher prices can be charged also for newly built dwellings, i.e. \( P^H \) increases. Assuming that \( C \) remains constant or at least grows less than \( P^H \), rise in \( P^H \) leads to an increase in the price of vacant lots. Hence, the assumption is that housing appreciation gives information regarding the value of land and that competition among housing developers, by restraining changes in \( C \), causes a lead-lag relation between housing and vacant lots. Again, in an efficient land market the duration of the lag would be negligible.

---

2 Within a larger and more heterogeneous country there may be more significant differences between regional construction costs (see e.g. Gyourko and Sainz, 2005, concerning differences in construction costs across US housing markets).

3 Titman (1985) points out that if the sizes of the buildings that can be built on the vacant lots are constrained by zoning regulation, the anticipatory feature of land price movements with respect to housing appreciation is undermined. In Finland the zoning regulations are typically tight.

4 Empirical evidence by Suoniemi (1990) supports the hypothesis that appreciation of existing housing stock Granger causes changes in the price level of new housing construction in Finland.
The aim of this paper is to study empirically the dynamic linkages between housing prices and prices of vacant lots employing quarterly data over 1985-2005 from HMA. The existence of a long-run relation between housing and land prices is tested by the Johansen cointegration test. Furthermore, cointegrated vector autoregressive model is estimated to study also the short-run dynamics. The implications of the empirical findings for lot market efficiency and lot price predictability are also evaluated.

The paper proceeds as follows. The next section presents the data used in the empirical analysis. In the third part the empirical methodology used in the study is described. The findings from the empirical analysis are presented in the fifth section. In the end, the paper is summarized and conclusions are derived.

2 Data

The housing price index \( H \) in this study describes the price development of single-family detached houses in HMA. The index of vacant land \( L \), in turn, depicts the evolution of price level of vacant land zoned for one-family houses. Both price indices are quality adjusted and received from National Land Survey of Finland. The data are quarterly and cover a period from 1985 to 2005. The land and house price indices are based on 7548 and 6597 observations, respectively. Transaction with extremely low or high price per square meter have been excluded from the data.

In addition, a number of control variables likely to affect housing and land prices significantly are incorporated in the empirical analysis:

- Inflation rate \( P \)
- Interest rate \( IR \)
- Construction cost index \( C \)
- Disposable income per household in HMA \( Y \)
- Number of households in the HMA \( HH \)

The series, except for \( P \) and \( IR \), are indexed and have the value of 100 in 1985Q1. 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.

The inflation rate variable, \( P \), is the difference of the cost of living index. Both cost of living index and the construction cost index are published by Statistics Finland. The market interest rate variable, \( IR \), in turn, is the twelve month Helibor until 1998Q4 and the corresponding Euribor after that. The interest rate data is only available from 1987Q1. The quarterly values of \( IR \) are the means of daily values during each quarter. Also the end of year figures for the number of households in HMA are published by Statistics Finland. The quarterly households statistics \( HH \) used in the analysis have been estimated based on the quarterly changes in the total population in HMA. Finally, \( Y \) is the taxed income of Helsinki residents in the state taxation deducted by the state, municipal and church taxes plus employees’ social security payments paid by Helsinki residents. The original income and taxation data are annual and provided by the City of Helsinki Urban Facts. For the needs of this study, quarterly variation
in $Y$ as well as the disposable income figures of 2005Q1-2005Q4 are estimated based on the nationwide income level index reported by Statistics Finland.

Figure 1 exhibits all the series included in the empirical analysis. 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. Housing and lot prices started to fall during 1989 and the markets finally collapsed at the beginning of the 1990s. Real house price level dropped by 50% from 1989Q3 to 1993Q1 and real lot prices were some 65% lower in 1993Q2 than in 1989Q3. 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. After mid 1990s housing and lot prices in HMA have grown substantially faster than disposable income and the number of households. This seems to be due to the low level of house and lot prices during the slump and because of the significant decrease in the interest rate. Expectedly, vacant land prices are substantially more volatile than housing prices and construction costs. Note also that, although hedonic indices are employed, there may be some measurement error in the house and land price series especially during 1988Q4-1990Q3 and the mid 1990s. At least the volatility in $L$ in 1995-1996 seems overly large. During that period there were somewhat less observations than during other parts of the sample. On the other hand, in both the periods mentioned, similar volatility occurs in both price series. This implies that major part of the high volatility of the indices during the turn of the decade and the mid 1990s is real. In any case, this is the best data available and it seems reasonable to believe that measurement errors in the data do not lead to fallacious conclusions in the empirical analysis.
Figure 1  The series

Table 1  Descriptive statistics

<table>
<thead>
<tr>
<th>Series</th>
<th>mean (annualised)</th>
<th>Standard deviation (annualised)</th>
<th>Jarque-Bera (p-value)</th>
<th>Ljung-box test for autocorrelation (p-values, 4 lags)</th>
<th>Seasonal variation (p-value, F-test)</th>
</tr>
</thead>
<tbody>
<tr>
<td>House appreciation</td>
<td>.026</td>
<td>.108</td>
<td>.00</td>
<td>.02</td>
<td>.75</td>
</tr>
<tr>
<td>Land appreciation</td>
<td>.032</td>
<td>.186</td>
<td>.00</td>
<td>.00</td>
<td>.06</td>
</tr>
<tr>
<td>Construction cost growth</td>
<td>.002</td>
<td>.013</td>
<td>.00</td>
<td>.00</td>
<td>.64</td>
</tr>
<tr>
<td>Inflation rate</td>
<td>.025</td>
<td>.012</td>
<td>.01</td>
<td>.00</td>
<td>.00</td>
</tr>
<tr>
<td>Interest rate</td>
<td>.042</td>
<td>.035</td>
<td>.05</td>
<td>.00</td>
<td>.00</td>
</tr>
<tr>
<td>Income growth</td>
<td>.020</td>
<td>.044</td>
<td>.00</td>
<td>.00</td>
<td>.03</td>
</tr>
<tr>
<td>Household growth</td>
<td>.017</td>
<td>.004</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
</tr>
</tbody>
</table>

Table 1 reports some descriptive statistics of the series. The standard deviation figures confirm the visual perception of the high volatility of land prices. There is also some evidence of seasonal variation in $\Delta L$. All the variables are highly autocorrelated and seasonal
variation in the inflation rate creates seasonality also in the real interest rate and in the real income growth.

Striking observation concerning the contemporaneous quarterly correlations, shown in Table 2, is that the figure between lot and house appreciation is only .2. Cross-autocorrelations show that interdependence between land and housing price movements is, expectedly, substantially stronger than implied by the contemporaneous correlation, however. In particular, current land appreciation correlates highly significantly with past housing price changes. Consequently, annual correlation between $\Delta L$ and $\Delta H$ is .77 and biannual correlation as high as .89. The finding that the correlation grows as the observation window is extended indicates that there are some form of lead-lag relations between house appreciation and price changes of vacant lots. It should be noted also that the housing and land series are likely to exhibit some noise due to the thin trading in the markets, and therefore the true contemporaneous quarterly correlation between $\Delta L$ and $\Delta H$ is likely to be somewhat greater than .2.

### Table 2 Contemporaneous quarterly correlations of differenced series

<table>
<thead>
<tr>
<th></th>
<th>$H$</th>
<th>$L$</th>
<th>$C$</th>
<th>$P$</th>
<th>$IR$</th>
<th>$Y$</th>
<th>$HH$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H$</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$L$</td>
<td>.21*</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$C$</td>
<td>.32**</td>
<td>.17</td>
<td>.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$P$</td>
<td>-.03</td>
<td>.07</td>
<td>-.29**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$IR$</td>
<td>-.02</td>
<td>.00</td>
<td>.21</td>
<td>-.60**</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Y$</td>
<td>.42**</td>
<td>.06</td>
<td>.40**</td>
<td>-.02</td>
<td>-.04</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>$HH$</td>
<td>.03</td>
<td>-.12</td>
<td>-.14</td>
<td>.02</td>
<td>.07</td>
<td>-.06</td>
<td>1</td>
</tr>
</tbody>
</table>

### 3 Methodology

In the econometric part of the paper, the order of integration of the series is tested, cointegration analysis is conducted and cointegrated vector autoregressive models are estimated. First, augmented Dickey-Fuller (ADF) test is used to study the order of integration of the variables. The number of lags included in the ADF tests is decided based on the general-to-specific method. A constant term is included in the ADF test 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 are added to the test if recommended by Akaike Information Criteria (AIC).

The existence of cointegration between variables which are integrated of order one [I(1)], i.e. whose levels are non-stationary but first differences are stationary, is tested employing the Johansen Trace test for cointegration. The lack of cointegration implies that dynamics between the series are only short-run in nature. The existence of one or more cointegrating

---

5 * and ** denote for statistical significance at the 5% and 1% level, respectively.
vectors between the variables, instead, indicates that also long-run interdependence exists. In other words, cointegration implies that there is at least one stationary long-run relation between the variables.

In the Johansen tests two possible CVAR models are considered:

Model 1:
\[ \Delta X_t = \mu + \Gamma_1 \Delta X_{t-1} + \ldots + \Gamma_{k-1} \Delta X_{t-k+1} + \alpha \beta' X_{t-1} + \Psi D_t + \varepsilon_t \]  
(1)

Model 2:
\[ \Delta X_t = \mu + \Gamma_1 \Delta X_{t-1} + \ldots + \Gamma_{k-1} \Delta X_{t-k+1} + \alpha (\beta', \beta_1)(X'_{t-1}, t)' + \Psi D_t + \varepsilon_t, \]  
(2)

where \( \Delta X_t \) is \( X_t - X_{t-1} \), \( X_t \) is a \( p \)-dimensional vector of the stochastic variables, \( t = 1, \ldots, T \), \( \mu \) is a vector of drift terms, \( \Gamma_i \) is \( p \times p \) matrix of coefficients for the lagged differences of the endogenous variables at lag \( i \), \( k \) is the maximum lag, i.e. the number of lags included in the corresponding vector autoregressive (VAR) model, \( \alpha \) is a \( p \times r \) (full rank) matrix of the speed of adjustment parameters, \( \beta' \) is a \( p \times r \) (full rank) matrix, \( \beta_1 \) is a \( r \)-dimensional coefficient vector and \( \varepsilon \) is a vector of error terms. Furthermore, three centered seasonal dummies are included in \( D \) if recommended by the Hannan-Quinn information criteria (HQ). The inclusion of centered seasonal dummies does not influence the critical values (Johansen, 1996). Also intervention dummies can be included in \( D \). \( \Psi \) is the coefficient vector for the dummy variables. Finally, \( r \) is the cointegration rank, i.e. the number of cointegrating vectors.

The difference between the models is that in Model 2 a deterministic time trend (\( t \)) is included in the cointegration space, i.e. in the long-run equilibrium relationship. Hence, Model 2 allows the cointegrating relation to be trend stationary. It is often difficult to know \textit{a priori} if trends cancel in the cointegrating relations. Therefore, it is sensible to start with Model 2 and test for the significance of \( t \). Furthermore, according to Doornik et al. (1998) adopting a model with a trend in the cointegration space has low cost even when the data generating process does not actually have one. The cost of excluding the trend term when there should be one is markedly larger. If, however, the trend term can clearly be excluded from the cointegration space according to Likelihood Ratio (LR) test (see Johansen, 1996), the number of cointegrating relations is tested based on Model 1.

The maximum lag is selected so that HQ is as small as possible and the residuals in the CVAR model do not exhibit significant serial correlation based on the LM(1) and LM(4) tests. The LM tests are conducted on residuals in the unrestricted model, i.e. in the model where \( r \) is set to equal \( p \).

The selection of the number of cointegrating vectors in a particular model is done by comparing the estimated Trace statistics with the quantiles approximated by the \( \Gamma \)-distribution (see Doornik, 1998).\(^6\) Because asymptotic distributions can be rather bad approximations to the finite sample distributions, also the Bartlett small sample corrected values, suggested by Johansen (2002) are reported.

\(^6\) If an intervention dummy is included in the tested model, a simulation program provided in CATS version 2 (see Dennis, 2006, pp. 142-145) is employed to estimate the quantiles. This is done although an intervention dummy taking a value of one only in one period and having a relatively small coefficient does, in general, not affect the relevant quantiles notably (see Doornik et al., 1998).
The LR test described in Johansen (1996) is employed to test if one or more variables can be excluded from the long-run relation, i.e. by testing if one or more of the coefficients in the $\beta$-vector can be set to equal zero. The LR test is also used to test the weak exogeneity of the variables. In these LR tests Bartlett small sample correction by Johansen (2000) is used.

The stability of the number of cointegrating relations and the stability of the cointegration space and the alphas are checked employing recursive analysis. All the cointegration tests are implemented by CATS in RATS, version 2, which also includes tools for the recursive estimation.

If cointegration between the variables is found, CVAR model is estimated to study the dynamics between the variables. Equilibrium-error (EQE), i.e. deviation from the long-run relation, is included in the model due to the fact that important information concerning long-run dynamics is lost if only differenced variables are used in the analysis. The inclusion of additional “control variables” in the CVAR models are decided based on the LR test using Sim’s correction for small samples together with economic arguments.

Granger non-causality (GNC) is tested by a standard F-test to further study the linkages between different variables. With cointegrated variables Granger causality can run also through the long-run equilibrium relation. Cointegration implies that the stationary EQE should Granger cause at least one of the cointegrated variables (Engle and Granger, 1987). Hence, also EQE must be included in the GNC test. GNC tests are done using the following model:

$$\Delta X_t = \mu + \sum_{j=1}^{n}\Delta X_{t-j} + \sum_{j=1}^{n}e_{t-j} + D_t + \varepsilon_t$$

If cointegration is not present, the lagged equilibrium-errors, $e_{t-j}$, are naturally not present in the GNC test. Furthermore, lagged EQEs are not included as explanatory factors for the variables that do not belong to the long-run relation.

Note that a finding that $x$ Granger-causes $y$ does not necessarily imply that $x$ causes $y$. It merely means that current and historical observations of $x$ are statistically significant in predicting future value of $y$. From now on, when causality is mentioned in the text it refers to Granger causality.

Finally, innovation accounting, i.e. impulse response analysis and variance decomposition, is performed based on the estimated models. The Choleski decomposition is utilized in the variance decomposition.

## 4 Empirical analysis

It this section, the order of integration of the variables is checked first. Then the existence of cointegrating relationships is tested and compositions of the cointegrating relations are analyzed. Finally, Granger causalities and impulse response functions are examined based on the estimated CVAR models.
The ADF test indicates that all the real variables as well as the inflation rate and the number of households are I(1). This finding is in line with majority of the previous related empirical work. The unit root test results are reported in Table 3. Note that the household series might actually be trend stationary. Only differenced $HH$ series is used in the forthcoming analysis, however.

Table 3 Augmented Dickey-Fuller test results

<table>
<thead>
<tr>
<th>Variable</th>
<th>Level (lags)</th>
<th>Difference (lags)</th>
</tr>
</thead>
<tbody>
<tr>
<td>House price</td>
<td>-1.62 (3)$^c$</td>
<td>-3.17** (2)</td>
</tr>
<tr>
<td>Land price</td>
<td>-1.76 (4)$^c$</td>
<td>-2.66* (3)</td>
</tr>
<tr>
<td>Construction cost</td>
<td>-1.49 (2)$^c$</td>
<td>-7.08** (0)</td>
</tr>
<tr>
<td>Inflation rate</td>
<td>-1.08 (3)$^c$</td>
<td>-10.2** (2)$^c$</td>
</tr>
<tr>
<td>Interest rate</td>
<td>-1.04 (4)</td>
<td>-10.7** (2)</td>
</tr>
<tr>
<td>Income</td>
<td>-0.75 (4)$^c$</td>
<td>-2.24* (3)</td>
</tr>
<tr>
<td>Households</td>
<td>-2.70 (4)$^c$</td>
<td>-3.93** (4)$^c$</td>
</tr>
</tbody>
</table>

The Johansen Trace test results support the existence of a cointegrating relation between $H$ and $L$, even though the non-stationarity of construction costs can not be rejected (see Table 4). The Bartlett small sample corrected value is not significant at the conventionally used 5% level. However, even the small sample corrected value is significant at the 10% level, and the Johansen test is also known for some power problems when small samples are employed. Furthermore, the speed of adjustment parameters of both of the variables are highly significant, visual inspection of the residual series from the long-run relation does not imply non-stationarity of the relation and both the alfas and the long-run relation are relatively stable. According to the pairwise model including three lags in differences, the speed of adjustment per quarter is 47% for housing and 31% for land. This implies that approximately 80% of the equilibrium-error vanishes within a quarter due to the adjustment of both variables towards the long-run relation.

Table 4 Trace test statistics

<table>
<thead>
<tr>
<th>H0 (rank)</th>
<th>Trace test value (p-value)</th>
<th>Small sample corrected Trace test value (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r = 0$</td>
<td>29.5 (.015)</td>
<td>24.3 (.076)</td>
</tr>
<tr>
<td>$r \leq 1$</td>
<td>3.77 (.772)</td>
<td>3.43 (.815)</td>
</tr>
<tr>
<td>$r \leq 2$</td>
<td>13.2 (.707)</td>
<td>11.2 (.845)</td>
</tr>
<tr>
<td>$r \leq 3$</td>
<td>3.51 (.788)</td>
<td>3.15 (.834)</td>
</tr>
</tbody>
</table>

* and ** denote for statistical significance at the 5% and 1% level, respectively. $^c$ indicates that constant was included in the test and $^s$ means that three seasonal dummies were included in the test. Critical values at the 5% and 1% significance levels, respectively, are -1.95 and -2.60 if constant is not included in the test and -2.89 and -3.51 in the case where constant is present.
As can be seen from Table 5, the pairwise CVAR model clearly indicates that $H$ Granger causes $L$ directly and through the long-run relation. Evidence on the existence of a feedback effect from land prices to house prices is inconclusive. The results suggest that there may be feedback through the long-run relation and possibly directly as well. In any case, based on previous changes in $H$ and $L$ better prediction can be made for land appreciation than for housing price movements.

Impulse responses from the pairwise model are graphed Figure 2. As land price level is more volatile than house price level, $L$ reacts more strongly to a shock in $H$ or $L$. It also seems that house prices stabilize more quickly after a shock.

Estimating a model that includes some control variables may give some further light to the housing and land price dynamics. The effect of interest rate and output shocks on housing and land prices is of a particular interest. Inclusion of additional control variables is not suggested by the LR test. Therefore, other variables are not added to the model to save degrees of freedom. Trace statistics in the multiple variable Johansen test are shown in Table 4. Sample starting from 1987Q1 is employed in the multiple variable analysis, since data on IR series do not exist prior to 1987. Note that a dummy variable (D92), which takes the value of one in 1992Q1 and is zero otherwise, is included in the eventual CVAR model that is tested. The dummy is statistically significant for $H$ and $Y$. The inclusion of D92 also leads to a substantial increase in the $R^2$ coefficients of both $H$ and $Y$ and to better normality of the residuals. The dummy is needed because of the extreme drops in the house price (20%) and income (8%) series from 1991Q4 to 1992Q1. The steep decreases in $H$ and $Y$ are largely due to the severe recession in Finland during the early and mid 1990s. From the beginning of 1991 to the beginning of 1992 the unemployment rate rose from 2.6% to 9.3% in HMA.

The cointegration analysis implies that there exists one long-run relation between the variables. The hypothesis that the cointegrating relation is the one earlier found between $H$ and $L$ cannot be rejected at the 5% level. The estimated alphas in the multiple variable model are 35% and 49% for $H$ and $L$, respectively. Hence, in contrast with the pairwise model, the model incorporating other factors affecting housing and land prices indicates that price of vacant land adjusts towards the long-run relation somewhat faster than house prices.

The relatively small number of degrees of freedom in the multiple variable CVAR may somewhat undermine the p-values in the GNC tests (Table 5). Also the multiple variable model shows strong evidence of housing appreciation Granger causing changes in vacant lot

---

8 According to the theory also $C$ should be included in the model. Because of the dynamics of the model the inclusion of $C$ would be problematic, however. $C$ would have to be included also in the long-run relation due to the long-term interdependence between $H$ and $L$ – otherwise growth in $H$ induced by an increase in $C$ would necessarily lead to an even larger growth in $L$. In theory, of course, appreciation of $H$ that is due to a rise in $C$ should not affect $L$. The problem is that $C$ is highly insignificant in the long-run relation and has got the wrong sign. One option to circumvent this problem would be to include $C$ as an exogenous variable (even though $C$ is not likely to be exogenous in reality). Nevertheless, the results of the CVAR model change only negligibly if $C$ is added as an exogenous variable and $C$ does not seem to bring significant new information to the model (this may be due to the stability of $C$ during the sample period). Hence, $C$ is excluded from the model in order to save degrees of freedom.

9 The small sample corrected p-value in the LR test for excluding $Y$ and IR $\beta$ is .06. Based on the results from the pairwise analysis and from the recursive analysis of the restricted multiple variable model, it seems justifiable to assume that $H$ and $L$ are pairwise cointegrated.
prices. Statistically significant feedback effect from $L$ to $H$ does not seem to exist. Instead, housing appreciation seems to be significantly affected by changes in the interest rate and in the household disposable income.

Impulse response analysis based on the multiple variable CVAR reveals some interesting observations. Most importantly, as can be seen in Figure 3, the model suggests that vacant land prices adjust substantially slower than house prices to interest and income shocks. For example, one quarter after a shock to $Y$ the change in $H$ is 45% of the eventual total effect of the shock, whereas the figure is 3% for $L$. Corresponding figures are 80% and 10% regarding a shock to $IR$. Still one year after a shock the differences are large – $H$ has adjusted 83% (income shock) and 90% (interest rate shock), while $L$ has adjusted only 58% and 72%. Although the results show that also the adjustment of the house price level is sluggish, it seems clear that informational inefficiencies are significantly greater in the market for vacant lots than in the housing market.

The impulses to shocks in $H$ and $L$ themselves are basically the same both based on the pairwise model and on the multiple variable model. The notable correction downwards in $L$ after a shock to itself is mainly due to the adjustment of $L$ towards the long-run relation. Generally, it can be assumed that a shock strikes the value of constructed land, i.e. housing prices, after which the price for vacant land adjusts to the new information. Hence, in reality it is unlikely that the initial shock occurs in the lot market. Instead, the findings suggest that $L$ just responds to changes in the other variables, especially to housing price movements. As suggested in the introductory section, housing price movements probably give information concerning the changes in the value of land. This information then influences demand in the market for vacant land with lag.

Expectedly, changes in the demand variables cause greater volatility in land prices than in house prices. The eventual increase in $L$ due to a one percent positive shock to $Y$ is 1.6% while $H$ grows by .9%. The influence of a change in $IR$ seems surprisingly small based on the graphed impulse responses. This is partly due to the endogeneity of $IR$. The initial shock to $IR$ is one percentage point, but as a consequence of the dynamics of the model the eventual rise in $IR$ is only .41 percentage points. A permanent one percentage point increase in $IR$ would naturally lead to larger negative effect on $H$ and $L$ than implied by the impulse response curves. This is not of great relevance in here though, since the main purpose of the impulse response analysis is to compare the adjustment speeds of $H$ and $L$ after a shock in the control variables. In any case, even taking account of the eventual smaller change in the interest rate, the response of house price level to an interest rate shock is smaller than that reported by Oikarinen (2005) concerning flats in HMA. This suggests that the price of land accounts for greater part of the flat prices than of the prices of singe-family detached housing, which could be explained by the fact that flats are typically located much closer to the centre of HMA and in more tightly build areas.

Note that the problem with $IR$ as an endogenous variable in the model is that nominal interest rate has, in practice, been exogenous since the late 1990s due to the Finnish membership in the European Monetary Union. As housing prices movements and income growth affect the inflation rate, the real interest rate can be seen as an endogenous variable, however.

Variance decomposition, exhibited in Table 6, confirms that $H$ is the principal factor driving $L$. In long horizon more than half of the forecast error in $L$ is due to innovations in $H$. By contrast, only a small part of housing price volatility can be explained by shocks to the other
three variables in the model. The effect of changes in $Y$ on land and housing prices is surprisingly small. The driving nature of $IR$ is somewhat stronger than that of $Y$.

Table 5  P-values in the Granger non-causality tests

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>∆H</th>
<th>∆L</th>
<th>∆Y</th>
<th>∆IR</th>
<th>eqe</th>
<th>$R^2$</th>
<th>Adj. $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>∆ Housing</td>
<td>.010</td>
<td>.148</td>
<td>.001</td>
<td>.324</td>
<td>.578</td>
<td>.446</td>
<td></td>
</tr>
<tr>
<td>∆ Land</td>
<td>.002</td>
<td>.001</td>
<td>.076</td>
<td>.455</td>
<td>.402</td>
<td></td>
<td></td>
</tr>
<tr>
<td>∆ Income</td>
<td>.546</td>
<td>.032</td>
<td>.315</td>
<td>.184</td>
<td>.721</td>
<td>.640</td>
<td></td>
</tr>
<tr>
<td>∆ Interest rate</td>
<td>.242</td>
<td>.144</td>
<td>.307</td>
<td>.001</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

MODEL INCLUDING CONTROL VARIABLES

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>∆H</th>
<th>∆L</th>
<th>∆Y</th>
<th>∆IR</th>
<th>eqe</th>
<th>$R^2$</th>
<th>Adj. $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>∆ Housing</td>
<td>.037</td>
<td>.133</td>
<td>.048</td>
<td>.039</td>
<td>.346</td>
<td>.578</td>
<td>.446</td>
</tr>
<tr>
<td>∆ Land</td>
<td>.002</td>
<td>.002</td>
<td>.311</td>
<td>.577</td>
<td>.026</td>
<td>.574</td>
<td>.440</td>
</tr>
<tr>
<td>∆ Income</td>
<td>.546</td>
<td>.032</td>
<td>.315</td>
<td>.184</td>
<td>.535</td>
<td>.400</td>
<td></td>
</tr>
<tr>
<td>∆ Interest rate</td>
<td>.242</td>
<td>.144</td>
<td>.307</td>
<td>.001</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 2  Impulses based on the pairwise CVAR model
Figure 3  Impulses based on the multiple variable CVAR model

Table 6  Variance decomposition of the multiple variable CVAR model

<table>
<thead>
<tr>
<th>Step (quarters)</th>
<th>House price level</th>
<th>Land price level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>∆Y</td>
<td>∆R</td>
</tr>
<tr>
<td>1</td>
<td>.002</td>
<td>.023</td>
</tr>
<tr>
<td>2</td>
<td>.064</td>
<td>.091</td>
</tr>
<tr>
<td>5</td>
<td>.066</td>
<td>.135</td>
</tr>
<tr>
<td>10</td>
<td>.059</td>
<td>.149</td>
</tr>
<tr>
<td>15</td>
<td>.057</td>
<td>.154</td>
</tr>
<tr>
<td>20</td>
<td>.056</td>
<td>.156</td>
</tr>
</tbody>
</table>

5  Summary and conclusions

Volatility in metropolitan housing prices is mainly attributable to demand driven changes in the value of land. Hence, housing appreciation typically reveals information regarding land price movements. In an informationally efficient market, both housing prices and price of vacant lots should adjust immediately to any new relevant information. Empirical evidence,
however, generally suggests that housing prices adjust sluggishly to changes in factors determining housing price level in the long run. As market for vacant lots is typically thinner than the housing market and the information concerning lot markets is usually more lacking than information regarding housing markets, it is probable that vacant lot prices adjust even more slowly to shocks than housing prices.

This paper suggests that housing prices respond to shocks influencing the value of land first, after which the price level of vacant lots reacts to the information revealed by housing appreciation. Thus, it is assumed that housing price movements lead changes in the market for vacant lots. Empirical analysis gives support to this hypothesis. Based on quarterly data over 1985Q1-2005Q4 from the Helsinki metropolitan area, the estimated cointegrated vector autoregressive models show that price movements of single-family houses Granger cause appreciation of vacant land zoned for one-family houses. According to the analysis, housing appreciation is clearly the principal factor driving vacant lot prices and there is only slight feedback from the vacant lot prices to the house prices.

Expectedly, lot price level is more volatile than housing price level and lot prices respond more strongly to interest rate and income shocks. It is evident, however, that it takes a long time before the vacant lot prices fully adjust to a shock. The adjustment speed of the land market seems to be notably slower than that of the housing market, even though also the adjustment of housing prices is sluggish. The findings suggest that the market for vacant lots is far from informationally efficient. In fact, it seems that future lot price movements can be predicted fairly accurately based on only current and lagged housing and lot appreciation.
References


