Urban Studies, Vol. 44, No. 8, 1517–1537, July 2007

Routledge
Taylor & Francis Group

The Dynamics of Intraurban Quantile House Price Indexes

N. Edward Coulson and Daniel P. McMillen

[Paper first received, March 2006; in final form, November 2006]

Summary. Estimating price indexes for different quantiles shows how prices vary for homes that are in different stages of the filtering process. The paper describes a filtering model and its implication for the dynamic interaction of housing prices at these various stages. A time-series analysis of quantile price indexes for three municipalities near Chicago supports the predictions of the filtering model.

1. Introduction

Despite an enormous literature on price index construction and a host of studies comparing appreciation rates across urban areas, few studies have estimated house price indexes for different locations *within* an urban area. Prominent exceptions include Meese and Wallace (1991) and Case and Mayer (1996), who estimate house price indexes for the San Francisco Bay and Boston areas respectively. Other studies, such as Poterba (1991), Smith and Tesarek (1991), Mayer (1993) and Case and Shiller (1994) find that appreciation rates vary depending on a home's place in the sales price distribution. Attempts to explain differences in house price appreciation rates within urban areas are less common. Here, the prominent exception is the study by Case and Mayer (1996), who argue that home price appreciation rates depend on a number of spatially fixed amenities and locational characteristics that vary across sub-markets within an urban area. They found that house prices rose less rapidly in Boston area municipalities that had a high percentage of manufacturing employment or middle-aged residents in 1980, and prices grew more rapidly in towns closer to Boston. Similarly, McMillen (2003) found that house prices within the City of Chicago rose more rapidly during the 1990s in locations closer to the city centre and in census tracts that in 1990 had high proportions of residents who were African American and had completed college. McMillen (2003) also found that appreciation rates were related to the characteristics of the housing stock: appreciation rates were higher in census tracts with a high proportion of vacant housing in 1990, with a housing stock of more recent vintage and with a higher proportion of owner-occupied homes.

In section 2, we describe quantile estimation and, in section 3, we estimate price indexes for three suburban municipalities in the Chicago area—Arlington Heights, Evanston and Oak Park. Our data include sales of single-family homes for 1983–2003. We use a hedonic price function to control for structural characteristics; the estimated price index is the series

N. Edward Coulson is in the Department of Economics, Penn State University, University Park, Pennsylvania, PA 16802, USA. Fax: 814 863 4775. E-mail: fyj@psu.edu. Daniel P. McMillen is in the Department of Economics (MC 144), University of Illinois at Chicago, 601 S. Morgan Street, Chicago, IL 60607, USA. Fax: 312 996 3344. E-mail: mcmillen@uic.edu.

0042-0980 Print/1360-063X Online/07/081517–21
DOI: 10.1080/00420980701373446

of coefficients on dummy variables indicating the quarter of sale. We use a quantile estimator to construct our price indexes for housing at various quality levels. The quantile approach has a distinct advantage over standard regression procedures because it is less sensitive to outliers that are common in hedonic housing studies. Importantly for our purposes, the quantile approach also provides a means of identifying a home's stage in the filtering process. For example, estimates based on the 90 per cent quantile show how prices vary over time for homes with errors that imply significantly higher prices than would be expected given their structural characteristics. Such homes can be expected to be those that have not depreciated significantly. In contrast, the estimates at the 10 per cent quantile show how prices vary over time for homes with prices that are significantly lower than expected—i.e. those that have depreciated significantly.

In section 4, we outline a filtering model mostly due to O'Flaherty (1995, 1996) and discuss the implications of that model for the dynamic interaction. We find that the theory has fairly tight implications for the causal relations between the price series, both in bivariate and multivariate settings, as well as for the cointegrating relationships between the series. In section 5, we estimate these time-series models and find a surprising congruence between theory and empirics. Section 6 concludes.

2. Quantile Hedonic Price Indexes

Examples of the hedonic approach to price index construction include Kiel and Zabel (1997), Mark and Goldberg (1984), Palmquist (1980) and Thibodeau (1989). The hedonic approach is typified by the following equation

$$y_{it} = \alpha + \beta'\mathbf{x}_i + \delta_2 D_{2,it} + \cdots + \delta_T D_{T,it} + u_{it} \quad (1)$$

where, y_{it} is the natural logarithm of the price of home i at time t; $\mathbf{x}_i$ is a vector of housing characteristics such as square footage and the number of bedrooms; and u_{it} is an error term. Sales dates range from 1 to T and the variables $D_{2,it} \ldots D_{T,it}$ indicate that the home sold during the period represented by the first subscript.

The estimated price index will be biased if missing variables are correlated with the time dummy variables. An obvious example of a missing variable is house quality. If homes selling at later dates tend to be of higher quality than those from early sales, the δs from later periods will be biased upward and will overstate the rate of price appreciation. Moreover, the estimates of equation (1) will be biased if coefficients are changing over time. To account for missing variable and time-varying coefficients, we can add a new variable to equation (1), z, with values that change over time

$$y_{it} = \alpha + \beta'\mathbf{x}_i + \delta_2 D_{2,it} + \cdots + \delta_T D_{T,it} + z_{it} + u_{it} \quad (2)$$

The variable z accounts for typical missing variables as well as changes in the coefficients and changes in $\mathbf{x}$ over time. Equation (2) simplifies to (1) if z_{it} always equals zero. We can rewrite equation (2) as

$$y_{it} = \alpha + \beta'\mathbf{x}_i + \delta_2 D_{2,it} + \cdots + \delta_T D_{T,it} + (z_i + \lambda_2 z_i D_{2,it} + \cdots + \lambda_T z_i D_{T,\,it} + u_{it}) \quad (3)$$

The new variables measure changes in z between time t and the base period. For example, since $D_{2,it} = 1$ while $D_{3,it} \ldots D_{T,it} = 0$ if the home sold during the second period, we have $z_{i2} = z_i + \lambda_2 z_i$ while $z_{i3} \cdots z_{iT} = 0$. Thus, z_{it} in equation (1) is the same as $z_i + \lambda_t z_i$ in equation (2). The bracketed terms in equation (3) are the error terms when z is unobserved. The missing variables are correlated with the time variables, which leads to biased estimates of the price index.

As with any mean-based procedure, the ordinary regression model is sensitive to outliers. An obvious example in the case of house price models is depreciation, which can differ across households with the same set of X's, and is likely to produce a relatively low value for the error term when a substantial amount of depreciation occurs (but is not observed in the data). The 'quality' variable may also be the source of outliers: given observed housing characteristics, unusually high-quality homes will tend to have high prices and large values for the error term. Either variable is missing from most datasets.

The quantile approach, which was originally proposed by Koenker and Bassett (1978), is less sensitive to outliers than the ordinary regression model.[1] Quantile regressions produce different coefficients for each pre-specified per centile of the error distribution. Let q represent the target quantile, which is specified before estimation, and let e_{it} be the residual implied by the econometric model. Quantile parameter estimates are the coefficients that minimise the following objective function

$$\sum_{e_{it}>0} 2q|e_{it}| + \sum_{e_{it}\leq 0} 2(1-q)|e_{it}| \tag{4}$$

At the median, $q = 0.5$, which implies that equal weight is given to positive and negative residuals. At the 90th per centile, $2q = 1.8$ and $2(1 - q) = 0.2$, which implies that more weight is given to positive residuals—observations with high values for the dependent variable, given the values of the explanatory variables. Equation (4) will be minimised at a set of parameter values where $100q$ per cent of the residuals are positive. This result differs from ordinary least squares, in which the *sum* of the residuals equals zero and otherwise there is no constraint on the number of positive residuals.

Quantile effects have a straightforward missing variables interpretation that follows directly from the hedonic and repeat sales price index estimators. The typical study presents estimated equations with the general form $y_i = \beta_q' x_i + u_{qi}$, which implies that coefficients differ by quantile. For example, the marginal effect of x at the median is $\beta_{0.5}$ while the marginal effect at the 90th per centile is $\beta_{0.9}$. The error term u_q will also vary across quantiles. In our example, the contribution of a sale at time $t = 2$ to the price index—the term analogous to β_q—can be found by taking the derivative of equation (3) with respect to $D_{2,it}$. The result, $\delta_2^* = \delta_2 + \lambda_2 z_i$, varies with the missing variable z. If $\lambda_2 > 0$, then higher values of z lead to higher values for δ_2^*. But z is part of the error term. Thus, high values of the error term imply high values for δ_2^*– a quantile effect. Unobserved depreciation is an excellent example of a missing variable that produces a quantile effect. A lack of maintenance shows up as an outlier in a standard regression model. Such outliers are drawn from the lower tails of the error distribution. Standard regression estimates of δ_t^* are biased downward in times with large numbers of homes that have such an unobservable characteristic.

The missing-variables interpretation of the quantile estimator has implications for filtering models of the housing market (Sweeney, 1974; Coulson and Bond, 1990). Homes of unusually high quality will be concentrated in the upper ends of the error distributions. After controlling for such variables as square footage and lot size, these homes show up as observations with large, positive errors. As a home depreciates and falls in quality, the error term turns eventually to negative values. Thus, we can observe how prices vary as homes progress through the filtering process by comparing estimated coefficients across regression quantiles.

3. Data and Hedonic Regression Results

The dataset for the empirical application of the quantile regression estimator was drawn from two sources, the Illinois Department of Revenue (IDOR) and the Cook County Assessor's Office. IDOR conducts reviews of assessment practices for all counties in Illinois, including Cook County. IDOR provided data on all sales of single-family homes in Cook County for 1983–2003 with the exception of 1992. Important variables include the sales price, date of sale and the parcel identification number, which matches the IDOR data with the 1997 Cook County file of assessments. The assessment file provides standard housing characteristics as of 1997. Although housing characteristics are available for only one date, a median-based estimator will provide accurate estimates of constant-quality appreciation rates if most homes are not remodelled during the sample period.

We chose three suburbs of Chicago for the analysis. Evanston and Oak Park are contiguous to Chicago. They have an older housing stock and a combination of high- and low-income households. With populations of 74 360 and 50 824 respectively, both suburbs are relatively large compared with the average in the Chicago area. The third suburb, Arlington Heights, is comparable with Evanston in population (76 031). Lying more than 20 miles north-west of downtown Chicago, Arlington Heights is newer than the other two suburbs. Whereas Evanston and Oak Park combine wealthy areas with pockets of poverty, Arlington Heights' income distribution is more uniform.[2] Each suburb can reasonably be considered to form a separate sub-market of the overall Chicago area housing market. Importantly, a filtering model is likely to be relevant in these suburbs, all of which are large and have ample variation in the age and price of the housing stock. In addition, spatial autocorrelation is unlikely to be a problem within individual suburbs, which is an important benefit as quantile estimators have not yet been developed for standard spatial models.

Table 1 provides descriptive statistics. Over the full 1983-2003 period, home prices averaged \$257 877 in Evanston, compared with slightly under \$200 000 in both Arlington Heights and Oak Park. The natural logarithm of the nominal sales price forms the dependent variable for our hedonic models. In addition to 84 variables indicating the quarter of sale, other variables are standard controls for structural characteristics. The explanatory variables include building area, lot size, the age of the structure, the total number of rooms, number of bedrooms and number of bathrooms. The remaining, discrete variables are fairly standard. The wide variation in these variables shows that the housing stock varies greatly across the three municipalities.

Table 2 presents standard OLS regression results along with quantile estimates for the 10 per cent, 25 per cent, 50 per cent, 75 per cent, and 90 per cent quantiles. The regressions are estimated with the full sample of sales across all three suburbs. Although our subsequent analysis is based on separate estimates for each town, the estimates for the pooled sample are sufficient to illustrate our approach. The OLS results are similar to those of other studies. Nearly all the variables are statistically significant with the expected signs. One apparently surprising result—the positive coefficient for age—is explained by the tendency in these municipalities for older homes to be better constructed than newer houses in these suburbs. In addition, older homes tend to be located in good neighbourhoods, close to schools, stores and commuter train lines.

The results in the last five columns of Table 2 show that there are significant quantile effects in the regressions. An additional square foot of land area adds much more to sales prices at higher quantiles. Controlling for building area and the total number of rooms, adding an additional bedroom (or equivalently, a smaller average bedroom size) reduces prices at the 50 per cent, 75 per cent and 90 per cent quantiles, but increases prices at the lower quantiles. Variables such as a finished basement, central air conditioning, masonry or frame/masonry construction and the presence of a garage have much more of an effect on sales prices at the lowest quantiles, a result that is likely to be due to their being relatively less common in homes in this group. The tendency for the equations to fit better at higher quantiles is probably due to the heterogeneity of the housing stock at lower quantiles, which combines new but low-quality housing with high-quality but poorly maintained older homes.

Figure 1 shows the estimated price indexes for the three regions based on the 50 per cent quantiles. Prices generally moved together throughout this time, with particularly high appreciation rates in the mid 1980s and since 2000. Prices rose more rapidly in Oak Park after 2000 than in Arlington Heights or

Table 1. Descriptive statistics

Variable	Arlington Heights	Evanston	Oak Park
Price	198 529	257 877	196 104
	(191 427)	(193 357)	(122 424)
ln price	12.086	12.237	12.014
	(0.497)	(0.722)	(0.656)
Building area (square feet)	1 727.263	1 752.049	1 696.177
	(720.822)	(780.392)	(636.308)
Ln building area	7.385	7.385	7.374
	(0.358)	(0.398)	(0.346)
Land area (square feet)	9 472.959	6 308.634	5 625.446
	(4 442.941)	(2 953.013)	(2 499.744)
Ln land area	9.086	8.647	8.543
	(0.358)	(0.470)	(0.448)
Age	32.308	63.887	73.849
	(16.121)	(24.619)	(17.943)
Rooms	6.731	6.742	6.648
	(1.350)	(1.680)	(1.415)
Bedrooms	3.501	3.426	3.389
	(0.753)	(0.977)	(0.833)
Bathrooms	1.921	1.791	1.656
	(0.664)	(0.747)	(0.640)
Discrete variables: percentage that equal one			
Basement	40.30	75.89	92.18
Basement is finished	37.91	23.27	13.63
Water heating	6.32	36.36	53.15
Central air conditioning	59.87	30.19	19.53
Fireplace	48.18	55.79	32.88
Porch	8.68	24.80	42.90
Masonry construction	17.38	40.80	26.40
Frame and masonry construction	57.70	16.36	5.17
Attic	28.29	51.12	54.54
Attic is finished	6.54	20.83	20.71
One-car garage	24.32	23.49	13.81
Two-car garage (or more)	63.55	46.10	67.13
Garage is attached	72.70	15.93	6.41
N	14 233	9 403	9 942

Note: Standard deviations are in parentheses for the continuous variables.

Evanston. Figures 2–4 display the estimated price indexes for the 10 per cent, 50 per cent, and 90 per cent quantiles for the three different suburbs. The prices generally move together across quantiles in all three suburbs. However, the tendency for prices to appreciate more slowly in the lowest quantiles is shared only by Arlington Heights and Oak Park. In Evanston, the price index for homes at the 10 per cent quantile index was higher than the index for the other quantiles over much of the sample period.

4. The Filtering Model and Its Implications

The goal of our empirical model is to capture the interrelationships of price movements across the quality spectrum within a neighbourhood housing market. In order to motivate our investigation, in this section we sketch a filtering model based on work by Sweeney (1974) and O'Flaherty (1996) and examine its implications for dynamic models of housing prices and their interactions.

Table 2. Regression results for pooled sample

Variable	OLS	10 per cent	25 per cent	50 per cent	75 per cent	90 per cent
Ln building area	0.175	0.157	0.176	0.189	0.190	0.202
	(29.30)***	(20.19)***	(34.99)***	(52.32)***	(50.99)***	(37.54)***
Ln land area	0.423	0.278	0.399	0.468	0.514	0.520
	(33.28)***	(17.68)***	(38.98)***	(61.07)***	(65.42)***	(46.78)***
Age	0.002	0.002	0.002	0.002	0.001	0.001
	(13.33)***	(11.45)***	(13.38)***	(17.39)***	(15.67)***	(8.23)***
Rooms	0.012	0.009	0.013	0.014	0.012	0.014
	(3.80)***	(2.59)***	(5.49)***	(7.43)***	(6.72)***	(5.45)***
Bedrooms	0.008	0.038	0.005	−0.011	−0.015	−0.018
	(1.71)***	(7.19)***	(1.49)	(4.05)***	(5.49)***	(5.03)***
Bathrooms	0.056	0.028	0.048	0.071	0.074	0.088
	(11.13)***	(4.74)***	(12.11)***	(23.44)***	(24.75)***	(21.59)***
Basement	0.007	0.007	0.031	0.034	0.033	0.029
	(1.35)	(1.15)	(7.47)***	(10.75)***	(10.39)***	(6.63)***
Basement is finished	0.051	0.083	0.066	0.061	0.045	0.035
	(9.87)***	(14.32)***	(16.08)***	(19.46)***	(14.91)***	(8.77)***
Water heating	0.071	0.068	0.064	0.054	0.048	0.044
	(12.13)***	(10.27)***	(13.91)***	(15.32)***	(13.63)***	(9.37)***
Central air conditioning	0.037	0.042	0.037	0.026	0.023	0.024
	(7.09)***	(7.68)***	(9.27)***	(8.14)***	(7.19)***	(5.54)***
Fireplace	0.147	0.131	0.130	0.129	0.112	0.092
	(28.83)***	(22.55)***	(31.52)***	(41.76)***	(37.17)***	(22.97)***
Porch	0.004	−0.010	−0.002	0.002	0.001	0.004
	(0.82)	(1.73)	(0.53)	(0.47)	(0.27)	(0.82)
Masonry construction	0.043	0.080	0.032	0.030	0.035	0.027
	(7.61)***	(13.07)***	(7.33)***	(8.71)***	(9.98)***	(5.80)***
Frame and masonry construction	0.055	0.114	0.050	0.033	0.030	0.019
	(8.80)***	(17.05)***	(10.59)***	(8.81)***	(7.82)***	(3.71)***
Attic	0.018	0.014	0.012	0.012	0.012	0.019
	(3.53)***	(2.38)	(2.89)***	(3.90)***	(3.74)***	(4.53)***
Attic is finished	−0.050	−0.078	−0.042	−0.031	−0.017	−0.006
	(7.32)***	(10.61)***	(7.93)***	(7.45)***	(4.14)***	(1.11)
One-car garage	0.064	0.069	0.065	0.060	0.042	0.025
	(9.06)***	(9.15)***	(11.99)***	(14.02)***	(9.78)***	(4.41)***
Two-car garage (or more)	0.069	0.084	0.082	0.071	0.055	0.038
	(11.51)***	(13.00)***	(17.90)***	(19.66)***	(15.09)***	(7.77)***
Garage is attached	0.028	0.048	0.042	0.017	0.011	0.012
	(4.19)***	(7.08)***	(8.38)***	(4.26)***	(2.71)***	(2.08)
Evanston	0.159	0.029	0.105	0.170	0.224	0.282
	(20.42)***	(3.40)***	(16.97)***	(36.19)***	(47.65)***	(41.75)***
Oak Park	−0.009	−0.045	−0.036	−0.020	0.011	0.053
	(1.01)	(4.89)***	(5.40)***	(3.69)***	(1.90)	(6.66)***
R^2 (OLS) or pseudo-R^2 (quantile)	0.618	0.460	0.534	0.588	0.621	0.640

Notes: Absolute t-values are in parentheses. The regressions include 84 variables representing the quarter of sale. The number of observations is 33 578. ***probability value <0.01.

The housing market is characterised by a discrete set of housing qualities. Households are characterised by income and each household has a bid-rent function that determines willingness-to-pay for any quality level. Each quality of housing is allocated to the household with the highest willingness to pay and the equilibrium allocation can be represented by a function that monotonically maps quality to household income (as in the hedonic equilibrium models discussed in Epple, 1987). Each of

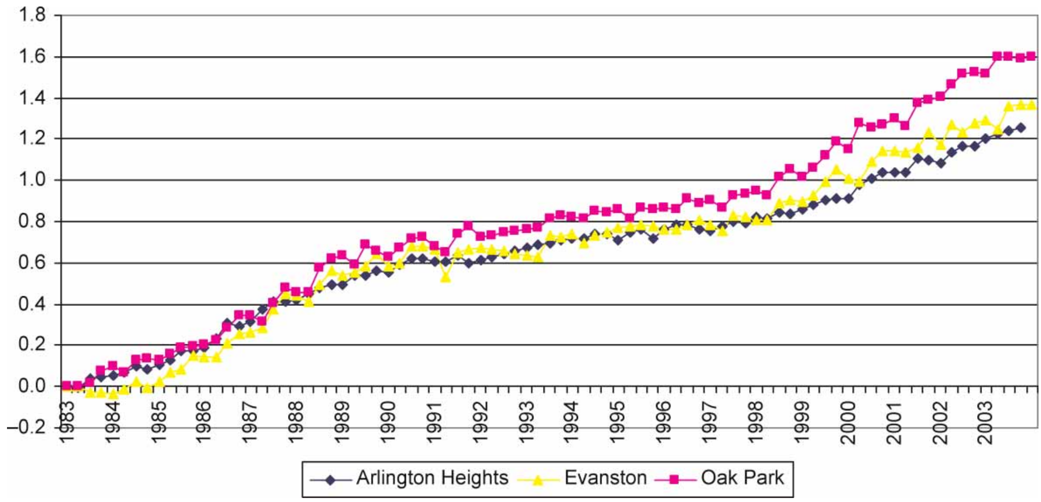

Figure 1. Price index for 50 per cent quantiles, by municipality.

the discrete quality levels is occupied by a convex set of incomes—i.e. there is an upper and lower bound of income that occupies housing of each quality level.[3] Households on the boundary are, in equilibrium, indifferent between the adjoining quality levels.

The supply side of the model can be specified either by the existing stock of housing at each quality level, or by the cost of construction at each level. In the latter case, as described by O'Flaherty (1996), the marginal cost of quality is increasing, and steeply so at higher-quality levels. The low marginal cost at low-quality levels implies that housing of low quality (for the housing market) is never newly constructed and is only created by the depreciation of higher-quality units. However, in O'Flaherty's model the very highest qualities of housing are not only constructed but also maintained and not allowed to depreciate.

The model implies constraints on both the source of supply shocks and the causal

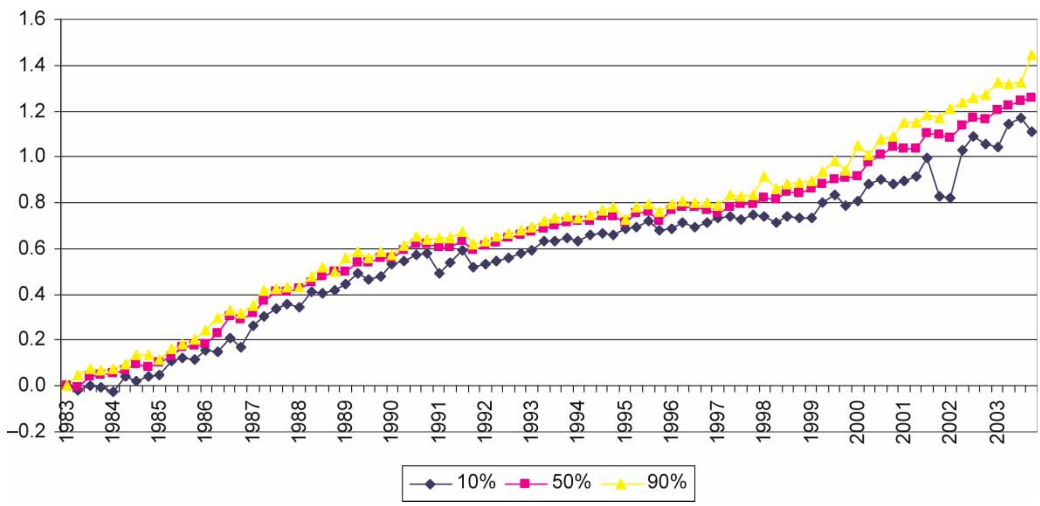

Figure 2. Price index by quantile for Arlington Heights.

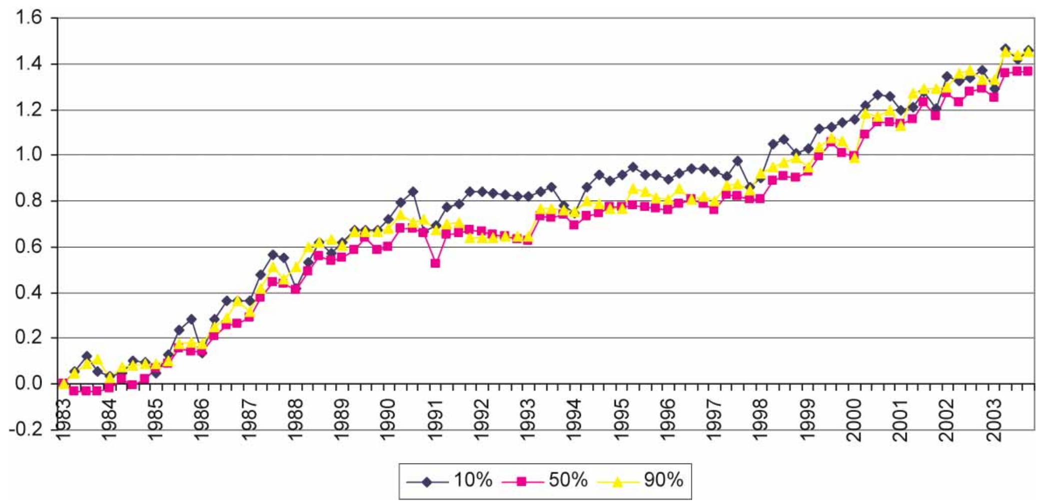

Figure 3. Price index by quantile for Evanston.

connections that arise from them and thus on the dynamic relationship between the price of housing at the various levels. First, we take new construction to be the source of shocks to the equilibrium that arise from the supply side and, as noted, such supply shocks can only arise at the top-quality levels because only there will new construction be brought to market. It is straightforward to trace out the effects of a supply shock on prices; for specificity, let this shock come at the very highest quality level. The immediate impact is of course at that highest level, as the expanding supply lowers the relative price, all the while increasing the range of households who obtain the highest quality of housing. In order to maintain the equilibrium, as some consumers move from the second-highest level to the highest level, the price of the second-highest level will fall. Thus the causality runs from the highest-quality price to the second-highest price. A ripple effect ensues, where price movements at one level cause price movements at the next level down, to

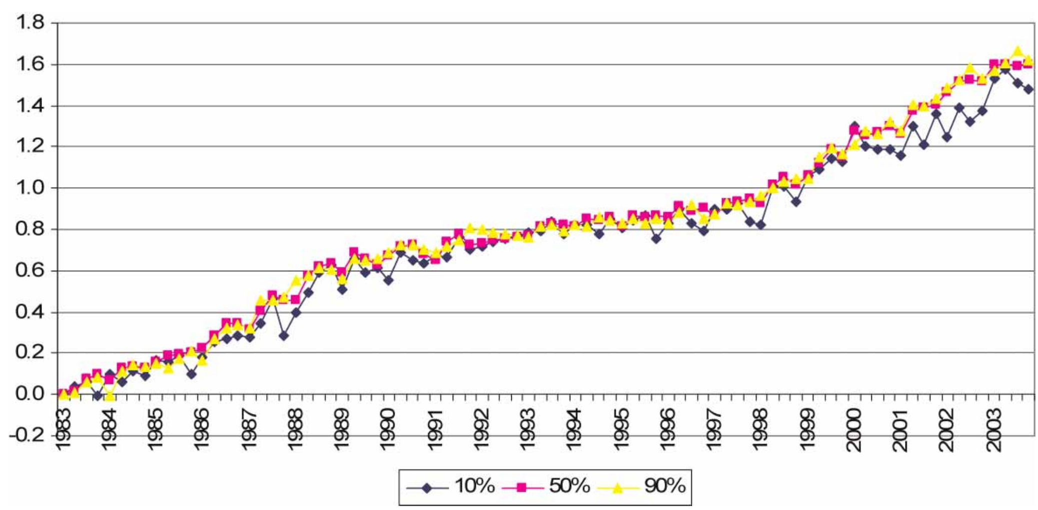

Figure 4. Price index by quantile for Oak Park.

the lowest-quality level, where the price becomes low enough to cause demolition.

In our empirical work, we make these concepts operational in the following way. In the first instance, we use Granger-causality tests to analyse the causal relations of prices at different levels. The tests arise out of a regression of the form

$$Y_{i,t} = \alpha + \sum_k \beta_k Y_{i,t-k} + \sum_k \gamma_k Y_{j,t-k} + e_t \tag{5}$$

where $Y_{i,t}$ refers to the price of housing of quality i at time t; $Y_{j,t}$ refers to housing of quality j at time t; $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are parameter vectors to be estimated; k is a lag index that is greater than zero; and e is a serially uncorrelated error term. Y_j is said to Granger cause Y_i if the γ_ks are jointly statistically significant. The first implication that we draw from the theory is that if Y_i is of relatively high quality compared with Y_j then $\gamma = 0$. And vice versa: if Y_i is of relatively low quality, then we can reject the hypothesis that $\gamma = 0$.

Recall that we specify five quality levels in our quantile regression specification. Unfortunately, it is not possible to specify precisely the quality levels which correspond to high and low quality in the filtering model. It may be the case, for example, that housing from both the 90th and 75th percentiles is high quality and thus experiences supply shocks. Tracing out the implications of such a supply shock is slightly more complex. The increase in supply to 75th percentile housing will lower its relative price and draw buyers from not only the 50th percentile stratum but also the 90th. The latter is necessary in order to maintain the indifference of the household on the border of the two upper quality levels. In this case, there will be two-way causality between the two highest price levels.

We can also analyse the dynamics of these price movements in a multivariate setting. We specify a vector autoregression of the form

$$[I - A(B)]\mathbf{Y}_t = \mathbf{e}_t \tag{6}$$

where, $\mathbf{Y}_t$ is a 5×1 vector of values of the house price index at the five different per centiles for a given neighbourhood. The lag polynomial (with the normalisation $A(0) = 0$) governs the dynamic interaction of the five price series.

Following standard procedures, the 5×1 vector of VAR residuals, e_t, has the representation

$$\mathbf{e}_t = B\mathbf{u}_t \tag{7}$$

where, the 5×1 vector $\mathbf{u}_t$ is composed of five independent shocks to each of the price levels. Thus, $\text{cov}(u) = \mathbf{D}$, where $\mathbf{D}$ is a diagonal matrix with the variances of $\mathbf{u}$ along the diagonal and $\text{cov}(\mathbf{e}) = \mathbf{W} = \mathbf{BDB}'$. The 5×5 covariance matrix $\mathbf{W}$ has 15 distinct parameters. Since each element of $\mathbf{W}$ is a function of the parameters of $\mathbf{B}$ and $\mathbf{D}$, a necessary condition for (just) identification is that the number of distinct parameters in the two matrices is 15. Five of these distinct parameters are the diagonal elements of $\mathbf{D}$, so $\mathbf{B}$ should have no more than 10 parameters. Noting that $\mathbf{B}$ is not symmetrical, this implies 15 restrictions, of which five of are the normalisation of the diagonal elements at unity. A convenient method of imposing the remaining restrictions is to develop a contemporaneous causal ordering on the elements of $\mathbf{e}$. Such an ordering implies a lower triangular structure for $\mathbf{B}$. The theory implies, at least to a first order (see the discussion above) that the contemporaneous causal structure runs from high-quality to low-quality prices. This implies

$$\begin{aligned}
e_{90,t} &= u_{90,t} \\
e_{75,t} &= b_{21}u_{90,t} + u_{75,t} \\
e_{50,t} &= b_{31}u_{90,t} + b_{32}u_{75,t} + u_{50,t} \\
e_{25,t} &= b_{41}u_{90,t} + b_{42}u_{75,t} + b_{43}u_{50,t} + u_{25,t} \\
e_{10,t} &= b_{51}u_{90,t} + b_{51}u_{75,t} + b_{53}u_{50,t} \\
&\quad + b_{54}u_{25,t} + u_{10,t}
\end{aligned} \tag{8}$$

with the proviso that this may unduly restrict the reverse contemporaneous causality of the 75th percentile to the 90th. However, loosening this restriction would leave the model unidentified.

In the context of the Chicago-area housing market, we discuss the influence of particular quality sub-markets on other sub-markets in the following way. The influence of Y_{90} on Y_{25}, for instance, is measured by the way the orthogonal shock of the high-quality housing, u_{90}, affects Y_{25}. A positive shock has a positive effect on e_{50} (assuming $b_{31} > 0$) which in turn has a direct positive effect on the actual value of the price index, Y_{50}. In addition the impact of u_{90} on each of the other residuals will have a contemporaneous value on Y_{50}. As time passes, the changes in each of the indexes will have additional effects through the workings of the VAR itself. Tracing out the dynamic effects of these original shocks creates the impulse response function, which will be one tool in assessing the results of the estimation. Additionally, we will present the variance decomposition which measures the contribution of each u_{it} to the variation of each Y_{it}. What the filtering model suggests is that, even if we find causality running from (say) the 25th percentile to the 10th in a bivariate setting, this influence should disappear in the multivariate setting because ultimately the bulk of the original shocks to the system (now specified as u_{it}s) are from the top (perhaps top two) quality levels. Thus, if the filtering model has empirical support, we will find that the impulse response functions and variance decompositions for all of the price series will display dependence on the shock to the top (or top two) quality levels.

We turn now to the joint questions of the persistence of the data and the functional form of the VAR. As is well-known, the VAR can take three functional forms. If the data are stationary—which is to say that there are no unit root processes in the data—then the VAR should be specified in level form. If the data do contain such unit roots, then it is possible that the data should be entered as first differences. However, if the various price series in the VAR contain common unit processes, then the data are said to be cointegrated and the VAR should have an error-correction mechanism as part of the model (along with the differenced data) (Hamilton, 1994).

We will find in the next section that the data do indeed have unit roots. Roughly speaking, the economic import of this finding is that the presence of unit roots implies that the shocks to the VAR system have permanent impact on the levels of the series. We will then test for cointegration. Cointegration implies that, while the impact on the levels is permanent, the number of permanent shocks is less than the number of series and that as a consequence the relationship *between* some of the series remains the same in the long run and, indeed, these relationships characterise the long-run equilibrium.

Traditional (but by no means unanimous) economic theorising (beginning with Nelson and Plosser, 1982; see also, for example, Shapiro and Watson, 1988) suggests that permanent shocks are those that emanate from the supply side, because they evolve from (permanent) shifts in technology. Thus the supply shocks in our framework, which are also described as new construction, must be small in number (one or two) and on that account the model predicts that the five price series will be cointegrated; moreover, the number of cointegrating relationships will be three or four, since there are only the one or two permanent shocks among the five series.

To summarise, our characterisation of the filtering model implies three sets of predictions for the time-series behaviour of our price series

(1) Bivariate causality runs from high-quality to low-quality prices, with the possible exception of some two-way causality in the upper quality levels.
(2) In a multivariate setting, we predict that the bulk of the explanatory power for all price series will come from the top one (or two) quality levels.
(3) We predict that the five data series will contain one or two unit roots, congruent with the idea that supply shocks occur only in the highest-quality segments of

the market. Thus, there will be three or four cointegrating relationships.

5. Empirical Results

The results are presented in Table 3, which contains F-tests for each of the possible causality tests. The asterisks indicate rejection at the 1 per cent level. The results are striking. Of the 30 cases in which we test for Granger causality of a higher-percentile index on a lower-percentile index, 27 of them reject the null of no causality. Higher-quality house prices cause lower-quality house prices. The three non-rejections are all from the same municipality (Evanston) and all from the same three highest-quality pairings. Conversely, of the 30 tests of causality running from lower per centiles to higher per centiles, there are only 3 rejections. If the null hypothesis were true, we would expect 5 per cent of the tests to exhibit rejections anyway (if the tests are independent). Thus, 3 rejections do no represent a particularly significant proportion of 30 tests, and we may safely conclude that lower-quality house prices do not cause higher-quality prices. This result is quite congruent with our theory of housing market filtering described in the previous section, of shocks to the equilibrium price schedule starting at the top of the quality spectrum and working their way down. It is not congruent with our theory in one small sense: we had speculated that any causal power that provided an exception to this general causality scheme would arise in the top levels. Instead, that is the one part of the spectrum (in Evanston, anyway) where causality is absent. The exceptions actually occur in the lower strata, but as noted this may merely be Type 1 error.

In preparation for our VAR construction, Table 4 presents the results of unit root tests for each of the series. The results are as expected; in each of the series, we are unable to reject the null hypothesis of a unit root. Each of the series has a permanent component, which we identify with supply shocks caused by new construction.

Table 3. Granger causality tests

Null Hypothesis	Arlington Heights	Evanston	Oak Park
25 per cent does not Granger Cause 10 per cent	22.8861***	5.60073***	7.02459***
10 per cent does not Granger Cause 25 per cent	7.92408***	1.40521	2.31793
50 per cent does not Granger Cause 10 per cent	17.8642***	7.94138***	9.26203***
10 per cent does not Granger Cause 50 per cent	3.95101***	1.87924	2.7579
75 per cent does not Granger Cause 10 per cent	15.2246***	6.22891***	10.3926***
10 per cent does not Granger Cause 75 per cent	2.31141	1.13122	4.87112***
90 per cent does not Granger Cause 10 per cent	10.5843***	7.42436***	11.9733***
10 per cent does not Granger Cause 90 per cent	0.584	0.24741	2.5299
50 per cent does not Granger Cause 25 per cent	7.20772***	7.4986***	6.40702***
25 per cent does not Granger Cause 50 per cent	2.14574	1.04074	2.85611
75 per cent does not Granger Cause 25 per cent	11.4979***	4.23935***	8.73524***
25 per cent does not Granger Cause 75 per cent	3.01426	0.87238	3.11488
90 per cent does not Granger Cause 25 per cent	9.5422***	5.78371***	9.80485***
25 per cent does not Granger Cause 90 per cent	0.6239	0.58519	2.30384
75 per cent does not Granger Cause 50 per cent	9.41921***	1.12186	4.60719***
50 per cent does not Granger Cause 75 per cent	3.00401	2.50845	2.15546
90 per cent does not Granger Cause 50 per cent	9.08053***	3.12177	6.78356***
50 per cent does not Granger Cause 90 per cent	0.19151	2.58148	1.65838
90 per cent does not Granger Cause 75 per cent	4.45491***	1.86756	3.6828***
75 per cent does not Granger Cause 90 per cent	1.3281	3.06644	0.74648

Notes: A small value of the F-test statistics indicates rejection of the null hypothesis that the first percentile does not Granger cause the second percentile in a bivariate Granger causal model.
***probability value less than 1 per cent.

Table 4. Augmented Dickey–Fuller tests for unit roots

	10 per cent	25 per cent	50 per cent	75 per cent	90 per cent
Arlington Heights	−0.52	−0.58	−0.62	−0.73	0.27
Evanston	−0.57	−1.17	−1.08	−0.53	−0.63
Oak Park	−0.06	0.06	−0.12	−0.31	0.37

Notes: The table entries are derived from Augmented Dickey–Fuller tests for the presence of a unit root in the indicated series. The Dickey–Fuller regressions contain lags as selected by the Schwarz information criterion. A test-statistic below the critical value of −2.58 would entail (at the 10 per cent level) rejection of the null hypothesis that a unit root is present. No such rejection occurs.

Following the results in Table 4, in Table 5 we present cointegration tests. These tests are trace tests as discussed in Johansen (1995). They take as the null hypothesis the restriction that there are no more than a specified number of cointegrating relationships in the data. In Table 5, we present for each housing market the value of the test statistic, the corresponding critical value and the probability value of the statistic. In Oak Park, we find that the number of cointegrating vectors is 2 or 3, depending on the level of Type I error used. This corresponds to 2 or 3 sources of long-run supply shocks. Qualitatively identical results were obtained for Evanston. In Oak Park, the results are somewhat different. There are apparently 4 cointegrating relationships in this municipality and the results are invariant to the choice between 5 and 10 per cent Type I errors.

Recall that the theory specifies a small number of supply shocks and thus a large number of cointegrating vectors. Only in Oak Park do the cointegration results match the theory in all respects, but the other two municipalities at least provide some evidence that there are at most two supply shocks in the system, which would correspond with the

Table 5. Tests for cointegration

Hypothesised number of CE(s)	Trace statistic	0.05 critical value	Probability***
Arlington Heights			
None***	130.0287	69.81889	0.0000
At most 1***	57.10420	47.85613	0.0053
At most 2**	28.15588	29.79707	0.0764
At most 3	12.63688	15.49471	0.1288
At most 4	0.227035	3.841466	0.6337
Evanston			
None***	138.8765	69.81889	0.0000
At most 1***	80.43839	47.85613	0.0000
At most 2**	27.30383	29.79707	0.0944
At most 3	8.776189	15.49471	0.3865
At most 4	0.206985	3.841466	0.6491
Oak Park			
None***	116.5319	69.81889	0
At most 1***	77.96896	47.85613	0
At most 2***	42.596	29.79707	0.001
At most 3***	15.81148	15.49471	0.0448
At most 4	0.021059	3.841466	0.8845

Notes: Cell entries are test statistics for the null hypothesis of the given null hypothesis on the extent of cointegration of the five quantiles in the indicated community.
**rejection at the 5 per cent level.
***rejection at the 1 per cent level.

notion discussed in the previous section that both of the top two quality levels experience supply shocks.

One thing that is clear from Table 5 is that cointegration at some order is definitely present in each of the three VARs. Therefore, we use vector error correction models in the estimation of the VAR. We present the results using results from the cointegration tests using 10 per cent critical values—3 in Arlington Heights and Evanston, and 4 in Oak Park (as opposed in particular to specifying 2 in Arlington Heights and Evanston). The impulse response functions and variance decompositions do not in fact depend on this difference in specification in any material way. We forebear from displaying the numerous coefficients of the (two-lag) VAR itself; they are available on request.

Recall that our theory predicts that the only shocks that should have a substantive influence on any of the prices are the shocks at the top of the quality spectrum, that is u_{90} in particular, and possibly u_{75} if housing at the 75th per centile is 'high-quality' in the sense described by the model above. In the case of Arlington Heights, this latter possibility fits the data moderately well. It can be observed in Figure 5 that the 90th and 75th percentile shocks are always the most important shocks for prices at every quality level. An interesting finding, however, is that for the lower-quality levels the 75th per centile shocks have a bigger impact than those from the 90th. This finding makes sense if both quality levels are the source of supply shocks; the lower of the two qualities will plausibly have more impact on lower-quality markets than the more 'removed' housing of the very highest quality.

A somewhat different picture emerges from Figure 6, which displays the impulse response functions for Evanston. In this case, the 90th percentile is quite important: it is always the first or second-largest impulse response function. However, the other important shock to the system comes not from the 75th percentile but from the 50th percentile. Indeed, for the lowest levels this can be the most important source of fluctuations. The 75th percentile shock is in fact not at all quantitatively interesting in those impulse response functions. This pattern of impulse responses is somewhat contradictory to our theory.

Naturally enough, the set of impulse responses that comes the closest to matching our theory model is that which also came the closest when we tested for cointegration. In Oak Park, the 90th percentile shock for the most part quantitatively dominates the other shocks, as displayed in Figure 7.[4] Clearly in that municipality there is one supply shock, it appears to derive from construction in the highest part of the quality spectrum and then filters down from there.

In Tables 6, 7 and 8, we display the variance decompositions at a variety of forecast horizons. The results there are quite congruent with those of the impulse response graphs. In Table 6, the sources of variation in the various quantile prices of Arlington Heights are decomposed into percentiles attributable to the various shocks to the system. Note that for the highest-quality housing the 'own-shock' has the most importance. However, for lower-quality housing, the 75th percentile shock has much more importance, particularly (as might be expected given our supply-shock interpretation) in the long run. In light of our theory, both types are high quality, but the latter has much more impact on lower-quality levels, as indicated as well in the impulse response functions.

In Evanston (Table 7), we again observe that the 90th percentile shock has a big impact on 90th percentile housing, and both it and the shock to the 75th have an impact on the 75th percentile. However, the 75th percentile loses its influence in lower percentile housing and is surpassed by the 90th and 50th percentile shocks. This, as noted above, does not exactly coincide with our theory-model. While there is no 'reverse-causation' evident in this variance decomposition, in light of both the theory sketched above and the results for Arlington Heights, one could conceive that 50th percentile housing could be in the 'high-quality' group and have a filtering-type effect on the 25th and 10th percentile

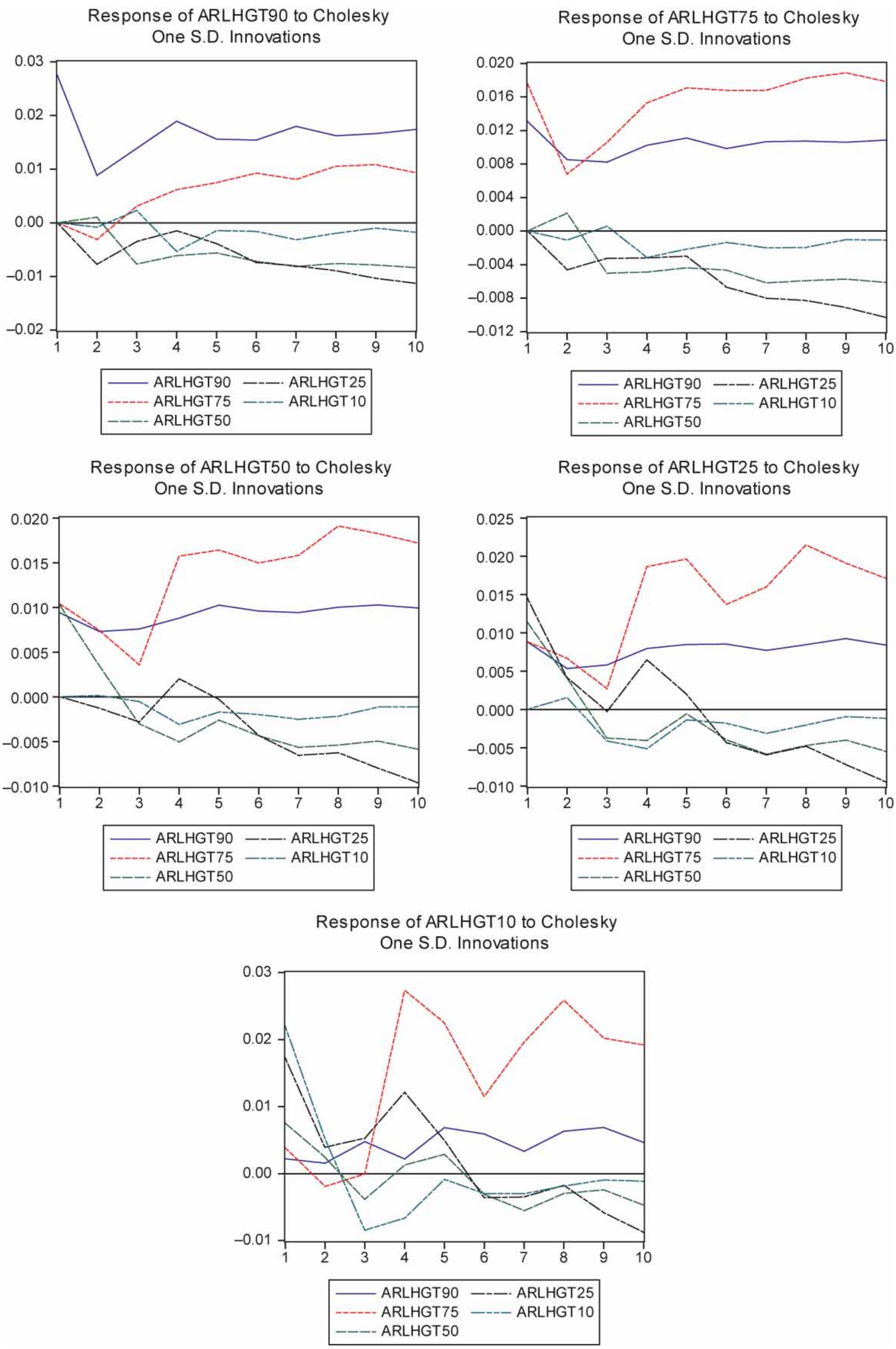

Figure 5. Impulse response functions: Arlington Heights.

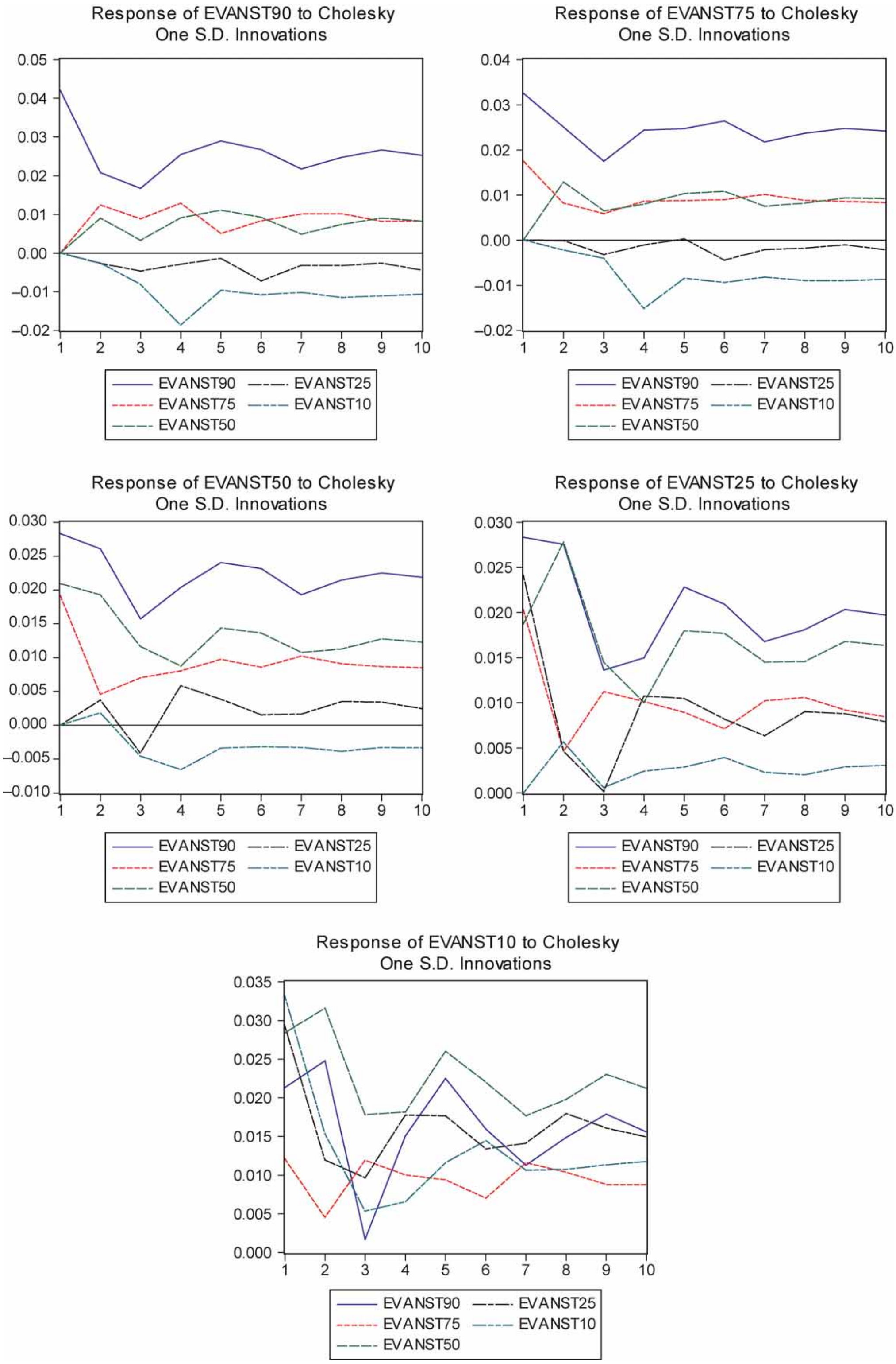

Figure 6. Impulse response functions: Evanston.

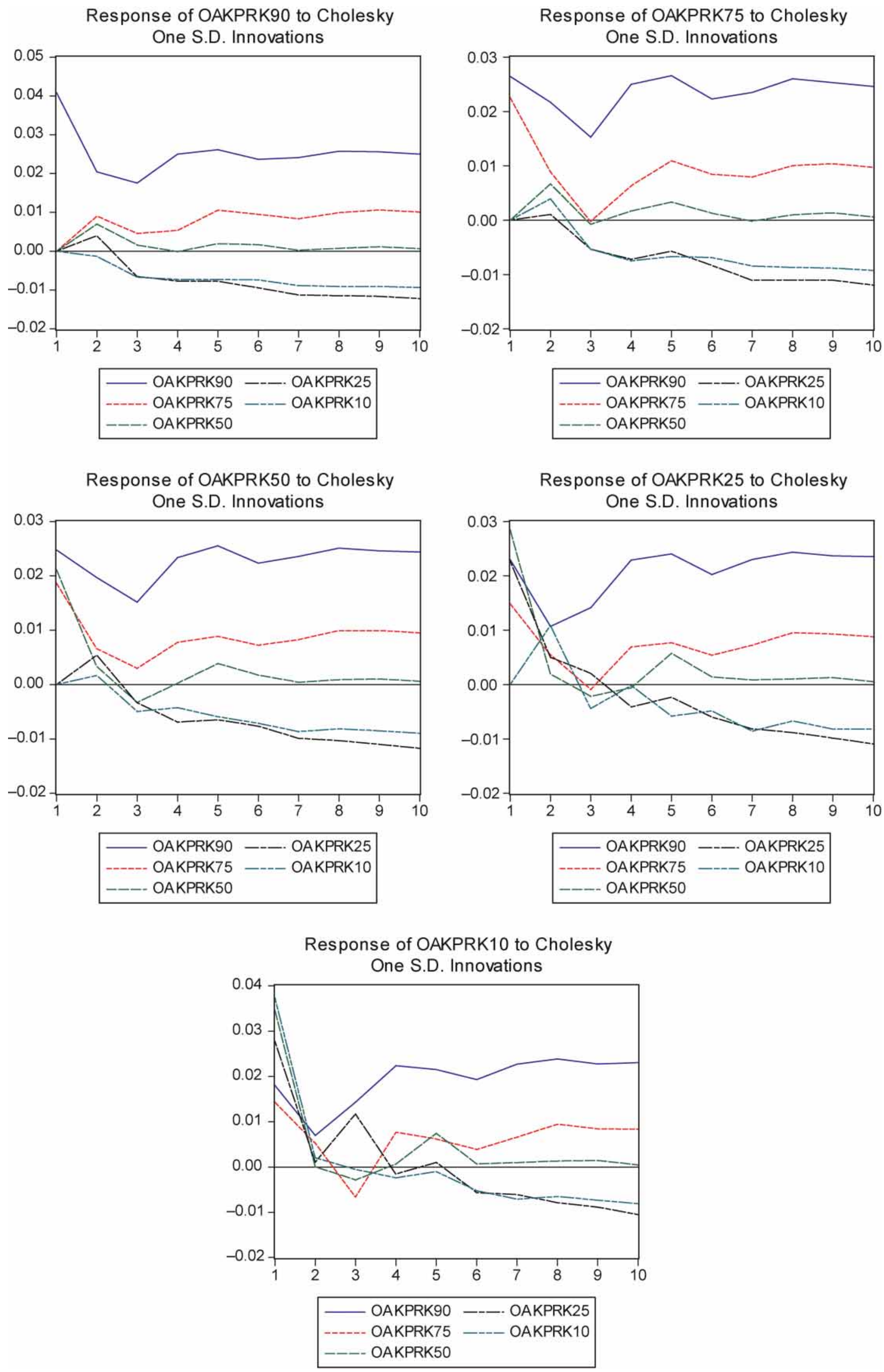

Figure 7. Impulse response functions: Oak Park.

Table 6. Variance decompositions, Arlington Heights

Period	90 per cent	75 per cent	50 per cent	25 per cent	10 per cent
90 per cent					
1.00	100.00	0.00	0.00	0.00	0.00
2.00	92.08	1.11	0.11	6.62	0.08
3.00	86.64	1.65	5.09	6.11	0.51
4.00	83.93	3.48	5.94	4.53	2.12
8.00	70.52	10.77	8.81	8.36	1.55
12.00	61.45	13.85	9.93	13.72	1.05
16.00	56.81	14.97	10.39	17.03	0.80
20.00	54.08	15.61	10.63	19.02	0.65
75 per cent					
1.00	35.63	64.37	0.00	0.00	0.00
2.00	38.93	56.73	0.73	3.43	0.19
3.00	36.98	55.48	3.55	3.82	0.18
4.00	33.96	57.23	4.39	3.47	0.95
8.00	27.23	59.43	5.30	7.23	0.82
12.00	23.79	58.33	5.74	11.58	0.56
16.00	22.19	57.28	5.95	14.16	0.42
20.00	21.29	56.62	6.06	15.67	0.35
50 per cent					
1.00	29.30	35.68	35.01	0.00	0.00
2.00	33.48	38.35	27.77	0.40	0.00
3.00	38.95	34.26	24.83	1.90	0.06
4.00	31.60	48.27	17.42	1.57	1.14
8.00	25.73	59.39	9.32	4.47	1.09
12.00	22.39	59.45	7.78	9.68	0.71
16.00	20.73	58.70	7.22	12.83	0.52
20.00	19.82	58.17	6.95	14.64	0.42
25 per cent					
1.00	15.61	15.77	26.07	42.56	0.00
2.00	17.52	20.27	24.11	37.70	0.40
3.00	20.65	19.20	23.59	33.71	2.86
4.00	17.34	40.70	15.03	23.04	3.89
8.00	16.48	60.70	8.54	12.06	2.22
12.00	15.51	63.20	7.00	12.91	1.38
16.00	14.79	63.42	6.41	14.39	0.99
20.00	14.35	63.41	6.11	15.36	0.77
10 per cent					
1.00	0.56	1.72	6.59	34.82	56.31
2.00	0.79	2.02	6.85	34.39	55.95
3.00	2.87	1.76	7.36	32.56	55.45
4.00	1.75	38.37	3.96	24.52	31.41
8.00	4.25	62.17	3.45	13.72	16.41
12.00	5.00	68.96	3.24	12.27	10.52
16.00	5.13	71.47	3.19	12.64	7.56
20.00	5.21	72.72	3.18	13.02	5.87

Notes: Table entries display the percentage of the forecast error variance which is attributable to the shock to the indicated price quantile. The first column gives the forecast horizon.

groups. What is harder to understand is how the 90th percentile housing retains its influence at the expense of the 75th.

Finally in Oak Park (Table 8), we have (again) a nice congruence between theory and evidence. The 90th percentile shock is the shock of primary importance throughout the quality spectrum and that importance increases as the time-horizon increases, just as expected.

Table 7. Variance decompositions, Evanston

Period	90 per cent	75 per cent	50 per cent	25 per cent	10 per cent
90 per cent					
1	100	0	0	0	0
2	89.81422	6.286859	3.314792	0.301525	0.2826
3	85.35234	7.971195	3.160555	1.003136	2.512778
4	75.24593	9.575314	4.212125	0.900569	10.06606
8	72.96943	8.817786	5.829208	1.424179	10.9594
12	71.97935	8.71928	6.190473	1.428305	11.68259
16	71.42562	8.731683	6.33002	1.435244	12.07744
20	71.06732	8.760943	6.404006	1.44024	12.32749
75 per cent					
1	77.51239	22.48761	0	0	0
2	75.59415	16.75521	7.422405	0.000688	0.227546
3	75.48237	15.46181	7.862908	0.390182	0.802735
4	71.78051	13.39009	7.516656	0.319447	6.993298
8	70.79019	11.7587	8.874425	0.560451	8.016239
12	70.68768	11.07457	9.140194	0.513963	8.583592
16	70.649	10.73534	9.258882	0.495773	8.860999
20	70.60386	10.54836	9.326174	0.485364	9.036241
50 per cent					
1	49.88121	22.89826	27.22053	0	0
2	54.93511	14.42786	29.99797	0.514097	0.12496
3	54.6083	13.83736	29.82246	0.964805	0.767072
4	56.43756	13.22495	26.85582	1.705731	1.775932
8	60.03154	12.57696	24.28728	1.418783	1.685437
12	61.64776	12.20498	23.12252	1.338801	1.685938
16	62.52906	12.02784	22.46975	1.28423	1.689113
20	63.06773	11.93683	22.05209	1.248993	1.694356
25 per cent					
1	37.3604	19.2004	16.33707	27.10213	0
2	41.58743	11.55235	29.92625	16.08472	0.849245
3	40.85248	13.0939	31.172	14.12797	0.753659
4	40.85864	13.72797	29.70728	14.91638	0.789733
8	43.55016	12.39927	30.65641	12.52108	0.873083
12	44.59761	11.94693	31.09013	11.45614	0.909196
16	45.12847	11.77112	31.31449	10.86587	0.920048
20	45.46119	11.67374	31.44682	10.49209	0.926159
10 per cent					
1	13.51464	4.365704	23.97498	25.52075	32.62392
2	19.88826	3.118422	33.55443	18.62914	24.80975
3	17.992	5.186648	35.59446	18.36491	22.86198
4	18.62874	5.86791	35.13892	20.21138	20.15304
8	20.22791	6.614705	36.2245	20.33575	16.59714
12	20.56265	6.937172	36.88415	20.55335	15.06268
16	20.64639	7.151865	37.1738	20.77827	14.24967
20	20.68128	7.310266	37.34183	20.93812	13.7285

Notes: see Table 6.

6. Conclusions

We use quantile regressions to create price indexes for various housing qualities. We then investigate the dynamic interaction of these indexes through the lens of the filtering model. The theory supposes that the most important shocks are those on the supply side, which are permanent and directly experienced only by the high-quality types, so that the data are cointegrated. This cointegration

Table 8. . Variance decompositions, Oak Park

Period	90 per cent	75 per cent	50 per cent	25 per cent	10 per cent
90 per cent					
2	93.36334	3.642598	2.195179	0.716541	0.08234
3	90.22409	3.848976	1.938338	2.212427	1.776168
4	88.30091	3.815224	1.504024	3.466507	2.91333
8	79.11525	7.176087	0.842151	7.563114	5.303394
12	73.94802	8.506064	0.55573	10.35425	6.635941
16	71.0324	9.177411	0.411071	12.00412	7.375001
20	69.2543	9.575588	0.326907	13.02242	7.820782
75 per cent					
1	57.89816	42.10184	0	0	0
2	64.30835	32.31613	2.462848	0.0622	0.85047
3	66.47431	27.85134	2.15084	1.428204	2.095304
4	70.25286	21.74676	1.67176	2.852008	3.476607
8	71.09199	15.63822	0.986923	6.847462	5.435401
12	68.94539	13.756	0.63713	9.904514	6.756964
16	67.38055	12.93847	0.463068	11.73148	7.486429
20	66.37257	12.49836	0.364081	12.84326	7.92173
50 per cent					
1	43.62529	24.65184	31.72287	0	0
2	53.25218	20.74158	24.29097	1.56834	0.146924
3	56.85832	18.40044	21.58328	1.879029	1.278932
4	62.63347	16.16137	16.46167	3.12279	1.620691
8	68.2911	12.53749	8.060859	6.55172	4.558826
12	67.32751	11.78879	4.995536	9.70926	6.178905
16	66.24717	11.52983	3.561095	11.58865	7.073261
20	65.49841	11.40656	2.756975	12.73418	7.603874
25 per cent					
1	25.45552	10.64132	39.0786	24.82455	0
2	27.17908	10.59928	34.41994	22.8176	4.984103
3	32.48813	9.695274	31.56007	20.95678	5.299741
4	42.96381	9.390578	25.7187	17.60908	4.317835
8	58.71516	8.939918	14.42688	12.61215	5.305903
12	61.8077	9.449298	9.146023	13.08571	6.511272
16	62.42966	9.840295	6.53807	13.90741	7.284567
20	62.57651	10.09578	5.058507	14.50705	7.762163
10 per cent					
1	8.348811	5.266615	30.7042	19.77744	35.90293
2	9.399177	5.854643	30.07559	19.40051	35.27008
3	13.25819	6.325391	27.54267	20.78964	32.08411
4	21.89718	6.783062	24.37583	18.4407	28.50323
8	40.85729	7.069816	17.13097	14.16886	20.77306
12	49.29743	7.976418	11.92079	13.85794	16.94742
16	53.13568	8.68018	8.889921	14.33807	14.95615
20	55.23471	9.153018	7.03887	14.78502	13.78838

Notes: see Table 6.

is manifested by the fact that the supply shocks set off a chain of unidirectional causal relations running from high- to low-quality prices. We perform the analysis for three municipalities in Chicago and find a high degree of congruence between the theory and the evidence. The causality does run from high to low quality; the series do have permanent components; and the data are cointegrated, suggesting a limited number of permanent (supply) shocks. In the multivariate analysis, most of the explanatory power for price changes comes from the highest qualities, as the theory would also predict.

Notes

1. The quantile estimator has been used in such studies as Albrecht *et al.* (2003); Bassett and Chen (2001); Buchinsky (1994, 1998a, 2001); Dimelis and Louri (2002); Garcia *et al.* (2001); Hartog *et al.* (2001); Levin (2001); Martins and Pereira (2004); and Thorsen (1994). Buchinsky (1998b), Koenker (2005) and Koenker and Hallock (2001) present useful surveys, while Gyourko and Tracy (1999) and McMillen and Thorsnes (forthcoming) present housing price examples.
2. According to the 2000 US Census, the median household income is \$67 807 in Arlington Heights, \$56 335 in Evanston, and \$59 183 in Oak Park. Poverty rates are 2.5 per cent, 11.1 per cent and 5.6 per cent respectively.
3. The stratification can be according to taste, rather than income, as in Epple (1987).
4. Skaburskis (2006) finds convergence between high- and low-quality housing prices in Canadian cities. This phenomenon, while by no means universal, is admitted in our framework. For an example from the short run, view the impulse responses in Oak Park, where the 90th percentile shock has a bigger impact on lower quantiles than on the 75th quantile housing.

References

Albrecht, J., Bjorklund, A. and Vroman, S. (2003) Is there a glass ceiling in Sweden?, *Journal of Labor Economics*, 21, pp. 145–177.

Bassett, G. W. jr. and Chen, H.-L. (2001) Portfolio style, return-based attribution using quantile regression, *Empirical Economics*, 26, pp. 293–305.

Buchinsky, M. (1994) Changes in the US wage structure 1963–1987: application of quantile regression, *Econometrica*, 62, pp. 405–458.

Buchinsky, M. (1998a) The dynamics of changes in the female wage distribution in the USA: a quantile regression approach, *Journal of Applied Econometrics*, 13, pp. 1–30.

Buchinsky, M. (1998b) Recent advances in quantile regression models: a practical guideline for empirical research, *Journal of Human Resources*, 33, pp. 88–126.

Buchinsky, M. (2001) Quantile regression with sample selection: estimating women's return to education in the US, *Empirical Economics*, 26, pp. 87–113.

Case, K. E. and Mayer, C. J. (1996) Housing price dynamics within a metropolitan area, *Regional Science and Urban Economics*, 26, pp. 387–407.

Case, K. E. and Shiller, R. J. (1994) A decade of boom and bust in the prices of single-family homes: Boston and Los Angeles, 1983 to 1993, *New England Economic Review*, March, pp. 40–51.

Coulson, N. E. and Bond, E. W. (1990) A hedonic approach to residential succession, *Review of Economics and Statistics*, 72, pp. 433–444.

Dimelis, S. and Louri, H. (2002) Foreign ownership and production efficiency: a quantile regression analysis, *Oxford Economic Papers*, 54, pp. 449–469.

Epple, D. (1987) Hedonic prices in implicit markets: estimating supply and demand parameters for differentiated products, *Journal of Political Economy*, 95, pp. 59–80.

Garcia, J., Hernandez, P. J. and Lopez-Nicolas, A. (2001) How wide is the gap? An investigation of gender wage differences using quantile regression, *Empirical Economics*, 26, pp. 149–167.

Gyourko, J. and Tracy, J. (1999) A look at real housing prices and incomes: some implications for housing affordability and quality, *Federal Reserve Bank of New York Policy Review*, September, pp. 63–77.

Hamilton, J. (1994) *Time Series Analysis*. Princeton, NJ: Princeton University Press.

Hartog, J., Pereira, P. T. and Jose, A. C. (2001) Changing returns to education in Portugal during the 1980s and early 1990s: OLS and quantile regression estimators, *Applied Economics*, 33, pp. 1027–1037.

Johansen, S. (1995) *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*. New York: Cambridge University Press.

Kiel, K. A. and Zabel, J. E. (1997) Evaluating the usefulness of the American housing survey for creating housing price indices, *Journal of Real Estate Finance and Economics*, 14, pp. 189–202.

Koenker, R. (2005) *Quantile Regression*. New York: Cambridge University Press.

Koenker, R. and Bassett, G. W. jr (1978) Regression quantiles, *Econometrica*, 46, pp. 33–50.

Koenker, R. and Hallock, K. F. (2001) Quantile regression, *Journal of Economic Perspectives*, 15, pp. 143–156.

Levin, J. (2001) For whom the reductions count: a quantile regression analysis of class size and peer effects on scholastic achievement, *Empirical Economics*, 26, pp. 221–246.

MARK, J. H. and GOLDBERG, M. A. (1984) Alternative housing price indices: an evaluation, *Journal of the American Real Estate and Urban Economics Association*, 12, pp. 30–49.

MARTINS, P. S. and PEREIRA, P. T. (2004) Does education reduce wage inequality? Quantile regression evidence from 16 countries, *Labour Economics*, 11, pp. 355–371.

MAYER, C. J. (1993) Taxes, income distribution, and the real estate cycle: why all houses do not appreciate at the same rate, *New England Economic Review*, May, pp. 39–50.

MCMILLEN, D. P. (2003) Neighborhood house price indexes in Chicago: a Fourier repeat sales approach, *Journal of Economic Geography*, 3, pp. 57–73.

MCMILLEN, D. P. and THORSNES, P. (forthcoming) Housing renovations and the quantile repeat sales price index, *Real Estate Economics*.

MEESE, R. and WALLACE, N. (1991) Nonparametric estimation of dynamic hedonic price models and the construction of residential housing price indices, *Journal of the American Real Estate and Urban Economics Association*, 19, pp. 308–332.

NELSON, C. and PLOSSER, C. (1982) Trends and random walks in macroeconomic time series: some evidence and implications, *Journal of Monetary Economics*, 10, pp. 139–169.

O'FLAHERTY, B. (1995) An economic theory of homelessness and housing, *Journal of Housing Economics*, 4, pp. 13–49.

O'FLAHERTY, B. (1996) *Making Room: The Economics of Homelessness*. Cambridge, MA: Harvard University Press.

PALMQUIST, R. B. (1980) Alternative techniques for developing real estate price indexes, *Review of Economics and Statistics*, 66, pp. 394–404.

POTERBA, J. M. (1991) House price dynamics: the role of tax policy and demography, *Brookings Papers on Economic Activity*, 2, pp. 143–183.

SHAPIRO, M. and WATSON, M. (1988) Sources of business cycle fluctuations, *NBER Macroeconomics Annual*, 11, pp. 1–48.

SKABURSKIS, A. (2006) Filtering, city change and the supply of low-priced housing in Canada, *Urban Studies*, 43, pp. 533–558.

SMITH, B. A. and TESAREK, W. P. (1991) House prices and regional real estate cycles: market adjustments in Houston, *Journal of the American Real Estate and Urban Economics Association*, 19, pp. 396–416.

SWEENEY, J. L. (1974) A commodity hierarchy model of the rental housing market, *Journal of Urban Economics*, 1, pp. 288–293.

THIBODEAU, T. G. (1989) Housing price indexes from the 1973–83 SMSA Annual Housing Survey, *Journal of the American Real Estate and Urban Economics Association*, 17, pp. 110–117.

THORSEN, J. A. (1994) The use of least median of squares in the estimation of land value equations, *Journal of Real Estate Finance and Economics*, 8, pp. 183–190.