J Real Estate Finan Econ (2016) 53:1–28
DOI 10.1007/s11146-015-9523-2

Long-run Equilibrium Shift and Short-run Dynamics of U.S. Home Price Tiers During the Housing Bubble

Damian S. Damianov[1] **· Diego Escobari**[2]

Published online: 30 August 2015

Abstract We use vector error correction models to examine the interdependence between the high and the low price tiers during the latest housing market boom and bust. For 118 of the 364 US statistical areas analyzed, the tiered price indexes are bound by a long-run relationship. In general, low tier homes appreciated more than high tier homes in the past two decades. In contrast to previous periods of high volatility, however, low tier homes appreciated more during the boom and lost more value during the bust of the market. We find a shift in the long-run equilibrium during the bubble —the cointegration parameter that ties the tiers together is greater in absolute value during the bubble period compared to the periods of more moderate appreciation and depreciation rates. Moreover, the shift in the long-run equilibrium can be explained by differences in subprime originations across housing markets. We also find that short run price dynamics is driven by momentum in both segments of the market.

Keywords Residential real estate markets · Housing price tiers

We thank Anurag Banerjee, Sergio García, Erling Røed Larsen, Robert Wassmer, Julian Williams and two anonymous referees for very helpful comments.

✉ Diego Escobari
diego.escobari@utrgv.edu

Damian S. Damianov
damian.damianov@durham.ac.uk

1 Department of Economics and Finance, Durham University, Mill Hill Lane, Durham DH1 3LB, UK

2 Department of Economics and Finance, The University of Texas Rio Grande Valley, Edinburg, TX 78539, USA

JEL Classification R3

Introduction

The dynamics of U.S. housing prices has been extensively studied in the literature since the real estate market boom and bust of the past decade. Most of the analyses examine the dynamics of price indices for the nation and for different geographic regions. As a result, we now have a better understanding of the contributing forces to the bubble and the relationship between housing prices and various fundamental factors, e.g. construction costs,[1] loan-to-value and price-to-rent ratios,[2] interest rates, lending standards,[3] demographic variables, etc. (see, e.g. (Mayer 2011) for a survey).

Most studies are based on aggregate indices that condense the appreciation and depreciation rates of all houses in a metropolitan area into a single measure. While this approach provides insights into the relationship between aggregate prices and fundamentals, it masks the effects arising from the interactions among different segments of the housing market. Recent life cycle models, however, demonstrate that these interactions are important as they have implications not only for the relationship between various house price tiers, but, through their effects on credit constraints, on the price dynamics of the entire housing market (Stein (1995) and Ortalo-Magné and Rady (1999) and (2006)). Thus, identifying trends and relationships between segments of local housing markets facilitates not only a comparison across time periods and geographic areas, but also affords tests of market efficiency and informs future theoretical research on housing markets.

In this paper we study the housing price dynamics at a disaggregated level in 364 U.S. statistical areas using the Zillow Home Value tiered indices during the time period from April 1996 through November 2014. These indices are constructed with a hedonic adjustment methodology intended to capture effects that are due to market trends only. The entire market is stratified in three tiers according to market value —bottom, middle, and top— and the indices track the appreciation and depreciation rates of these three segments over time.

We estimate a vector error correction model which accounts not only for the long run relationship between price tiers, but also for their short run dynamics. We then allow the cointegrating parameters to vary over time by using a rolling regression specification. We thus create a panel of estimated cointegrating parameters and

[1] Glaeser et al. (2008) discuss the role of supply elasticity for the severity and duration of housing bubble periods.

[2] Gallin (2006) shows that for the time period 1970-2005 prices and rents are cointegrated. Price-rent ratios tend to predict future changes of both prices and rents whereby the corrective response of prices is greater than that of rents.

[3] See e.g. Mian and Sufi (2009) and Pavlov and Wachter (2011).

examine how the cointegrating relationship between tiers differs during the period before and during the housing bubble.

Our analysis generates a number of new insights. First, in agreement with studies of previous time periods (see, e.g. Poterba (1991) and Mayer (1993) for the period 1970-1990 in a number of metropolitan areas) we find that, for the entire data panel,[4] and for the time series of about half of the metropolitan areas,[5] the low and the high price tiers are non-stationary in levels but are stationary in first differences (i.e. are integrated of order one).

Our second and probably most unexpected finding is the tendency of low tier homes to appreciate faster than properties at the high end of the market prior to the housing market bust. After the burst of the bubble, however, low tier homes depreciate faster than high tier properties. As we discuss in more detail in the next sections, this pattern stands in contrast to previous periods of market volatility. We quantify the extent of this phenomenon for all metropolitan areas using cointegration tests between the price tiers. We find that price tiers are cointegrated, i.e. they are bound by a long run relationship in 118 statistical areas. This suggests that the price tiers are driven by the same common factors that pull the series together in the long run (Mayer 1993). Cointegration in our context can also be interpreted as evidence in support of the efficient market hypothesis in the long run. Meese and Wallace (1994) find that house prices are cointegrated with home owner cost of capital and rents for the time period 1970-1988, a finding that they similarly interpret as evidence for the efficient market condition in real estate. They find, however, that the present value relation does not hold in the short run.

Third, the short run dynamics between the tiers exhibits strong positive correlation between current appreciation rates and lagged appreciation rates for both high and low tier homes. As suggested by Case and Shiller (1990), this is evidence for momentum and information inefficiency in the housing market. While Case and Shiller (1990) report autoregressive coefficients on the magnitude of 0.3, when broker down in tiers, we find coefficients for the majority of the metropolitan areas in excess of 0.5. Because we find that tiers are cointegrated, commonly used univariate models or vector autoregressive models to assess momentum (as the ones presented in Case and Shiller 1990) would be subject to an omitted variable bias. Our assessment of momentum relies on estimates from vector-error correction models thus taking into account the long-run relationship between tiers. Our finding that dynamics of real estate prices is driven by momentum in both market segments also accords well with the recent empirical evidence on social influences in the housing market. Pan and Pirinsky (2013), for instance, show that the probability of a household to purchase a home is positively correlated with the home ownership rate of ethnic peers

[4] Panel augmented Dickey-Fuller unit root tests as proposed by Pesaran (2007) reject the null of a unit root in the first difference both for the low and the high tier at the 1 % significance level.

[5] Augmented Dickey-Fuller unit root test cannot reject the null of a unit root in levels and reject the null of a unit root in first differences at conventional statistical levels for 176 out of the 364 statistical areas.

residing in the same area and report that this effect is particularly strong among younger, less educated, and poorer individuals. As these individuals are likely to be the buyers of starter homes, we also conduct formal tests and find that momentum is significantly stronger for the low tier market segment in about half of the geographic areas.

Fourth, the cointegrating parameter β is smaller than negative one. That is, for each dollar of price appreciation in the high tier, the low tier appreciates by a greater amount in the long run. This result is consistent with recent life cycle models of housing markets. Ortalo-Magné and Rady (2006) show that, after a positive income shock or shock to credit constraints, "starter" homes appreciate proportionally and "trade up" homes appreciate less than proportionally to the magnitude of the shock in the new steady state. However, Ortalo-Magné and Rady (2006) also find that the high tier is more volatile in the transition to the new steady state —a feature that does not accord to our results. Using vector error correction models, Gallin (2006) studies the time series of prices, rents, and the rent-to-price ratios for the period 1994-2005. He finds that this ratio helps predict changes in real prices (but not real rents) over a 4-year period. To our knowledge, we are the first to study the short run and long run dynamics of the time series of different segments of the market during the recent market bubble.

Finally, we estimate a rolling regression specification of a vector-error correction model by estimating the cointegrating factor β for windows of 100 months thus obtaining a time series of realizations of this parameter for the 118 statistical areas in which the price tiers are cointegrated. With the resulting panel of metropolitan areas we test for differences between the cointegrating parameters for the periods prior to and during the bubble. We find a statistically and economically significant shift in the long run equilibrium. As statistical areas enter into the bubble period, the cointegration coefficient increases in absolute value reaching a peak around the time when the local housing market bubble bursts. That is, prices of low tier homes are relatively more volatile compared to the prices of high tier homes during the bubble period. We also find that the cointegration relationship between the tiers is sensitive to percentage of subprime originations in local areas while the dependence on demographic factors such as population and income is not statistically significant.

The rest of the paper is organized as follows. Section "Data" presents a description of the dataset and provides details on the Zillow tiered home price indices. Sections "Existence of a Long-run Equilibrium and Price Convergence", "Estimating the Long-run Equilibrium and Short-run Dynamics" and "Stability of the Long-run Equilibrium" present the empirical results, and deal with stationarity of the time series and unit root tests, long run and short run dynamics, and equilibrium stability, respectively. In Section "Comparison to Previous Bubble Periods" we review the similarities and differences between the latest and previous housing market cycles. Section "Explaining the Tiered Price Dynamics" explores to what extent demographic variables, lending practices and autocorrelation in appreciation rates (momentum) can explain the observed price dynamics during the bubble period. Section "Conclusion" summarizes the main findings and implications.

Data

This study uses the seasonally adjusted monthly low and high tier segments of the Zillow Home Value Index (ZHVI).[6] The ZHVI as well as the Case-Shiller index (CSI) capture changes in single-family home prices over time. We opted for the ZHVI over the CSI for two key reasons. First, the ZHVI is a hedonic price index that, as explained in Dorsey et al. (2010), overcomes limitations of repeat-sales indexes such as the CSI. The ZHVI is constructed from the same deed records as the CSI but includes also the properties that sold just once during the relevant period in addition to the repeat sales transactions used for the construction of the CSI index. Further, ZHVI utilizes attributes of individual houses such as size and the number of bedroom and bathrooms thus capturing their changes over time. Guerrieri et al. (2013) provide a detailed discussion on the similarities and differences between the two indices and report a correlation coefficient of 0.96 between their appreciation rates during the time period from 2000 to 2006. Second, while the CSI is constructed only for twenty metropolitan statistical areas, the ZHVI covers a much larger number of areas. Our sample includes the low tier and high tier segments of the ZHVI for 364 statistical areas and spans the period from April 1996 through November 2014.[7] This allows us not only to study the dynamics for a large number of areas, but also to exploit existing cross-sectional differences.[8]

The ZHVI and CSI are monthly indexes designed to reflect changes due to market trends only and not to physical changes in individual houses or neighborhoods. The tiered indexes are obtained by assigning properties into one of three price tiers. The thresholds for the price tiers are determined from the distribution of home values within each statistical areas in such a way that an equal number of homes appear in each tier. The repeat sales methodology used in the CSI entails forming pairs of recorded prices from arms-length transactions of the same property. While repeat-sales indexes contain valuable information on the dynamics of prices, there are some concerns when using repeat sales indices to estimate autoregressive model specifications as the ones studied here. They are related to the potential bias and efficiency loss due to the inclusion only of properties which sold more than once in a given period. These houses constitute only a relatively small subsample of the entire market. The ZHVI that we use is created from estimated sales prices of every home, not just the ones that have recently been sold. Hedonic measures of house prices more accurately reflect changes in the characteristics of individual houses, but requires high quality data on a number of additional characteristics of houses. The ZHVI uses home

[6] The data is obtained from the online real estate database Zillow, available at http://www.zillow.com/research/

[7] The statistical areas include both metropolitan and micropolitan statistical areas.

[8] An alternative index that can be used for an analysis of house price dynamics is the repeat sales index constructed by the Office of Federal Housing Enterprise Oversight (OFHEO). This index, however, is limited to houses for which mortgages were provided by Fannie Mae and Freddy Mac thus excluding the subprime segment of the market. For a discussion on the use of indexes in studying house price dynamics see Miao et al. (2011).

attributes that include physical facts about the home and land, prior sale transactions, tax assessment information and geographic location.

In the Appendix we present the list of the 364 statistical areas used in this study. For practical reasons, in Table 1 we present the summary statistics of the low tier (LowTier) and the high tier (HighTier) indexes for only a selected group of areas. In particular, following an alphabetical order, we include the first four and the last four areas as well as the descriptive statistics for the whole sample. We have 224 geographic area time series in the sample comprising a total of 81,536 tiered index observations. All indices are equal to 100 on April 1996.

Existence of a Long-run Equilibrium and Price Convergence

We denote by $\text{LowTier}_t^{\text{SA}}$ and $\text{HighTier}_t^{\text{SA}}$ the low and high price tiers, respectively, for a given statistical area (SA) during month t. With the following linear combination of these two price tiers,

$$\text{LowTier}_t^{\text{SA}} + \beta \text{HighTier}_t^{\text{SA}} = d_t^{\text{SA}}, \tag{1}$$

we say that a long-run equilibrium exists if for a given constant β the difference d_t^{SA} is stationary. That is, short-run deviations away from the long-run equilibrium $\text{LowTier}_t^{\text{SA}} = -\beta \text{HighTier}_t^{\text{SA}}$ are observed when $d_t^{\text{SA}} \neq 0$. However, the

Table 1 Summary statistics

	(1)	(2) LowTier	(3)	(4)	(5)	(6) HighTier	(7)	(8)	(9)
Statistical Area	Obs.	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max
Aberdeen, WA	224	92.22	15.19	69.32	121.1	132.5	28.81	96.38	183.5
Adrian, MI	224	117.1	20.60	86.42	153.8	131.4	18.96	100	162.8
Akron, OH	224	124.1	13.99	100	144.8	120.8	10.07	99.65	136.0
Albany, GA	224	124.6	14.32	100	147.9	134.8	17.80	100	157.1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
York, PA	224	135.9	24.98	100	175.4	138.0	27.72	98.83	184.0
Youngstown, OH	224	103.9	10.18	88.52	123.2	123.1	10.60	100	139.2
Yuba City, CA	224	170.6	72.06	100	332.7	166.6	55.65	90.68	282.6
Yuma, AZ	224	139.7	38.09	98.68	219.0	154.8	41.70	99.84	237.0
Overall (# SAs=364)	81,536	141.6	43.09	68.43	404.2	153.2	45.74	79.90	500.3

Notes: The sample contains monthly observations from 1996m4 through 2014m11. There are 224 times-series observations of LowTier and HighTier for each of the 364 statistical areas stationarity of d_t^{SA} means that these deviations are by definition only transitory. If a long-run equilibrium exists, we say that LowTier_t^{SA} and HighTier_t^{SA} are cointegrated with vector $[1, \beta]$; moreover, the price tiers are integrated of order one, $I(1)$, while d_t^{SA} is integrated of order zero, $I(0)$.

We also define convergence between price tiers if the long-term forecast of both price tiers is the same:

$$\lim_{k\to\infty} E_t\left(\text{LowTier}_{t+k}^{SA} \middle| \Omega_t\right) = \lim_{k\to\infty} E_t\left(\text{HighTier}_{t+k}^{SA} | \Omega_t\right), \tag{2}$$

where Ω_t denotes the information set at time t. Price convergence is first tested by restricting β in Eq. 1 to be equal to negative one. There is a number of unit root tests that can be used to analyze the stationarity of d_t^{SA} in Eq. 1 and convergence in Eq. 2. In this paper we employ the GLS augmented DF tests as originally proposed by Dickey and Fuller (1979) along with the more recent panel unit root test proposed by Pesaran (2007).[9]

We initially test for the stationarity for the variables in levels (LowTier_t^{SA} and HighTier_t^{SA}), the variables in first-differences ($\Delta\text{LowTier}_t^{SA}$ and $\Delta\text{HighTier}_t^{SA}$), and the difference in price tiers ($\text{LowTier}_t^{SA} - \text{HighTier}_t^{SA}$). This is done for every time series in the sample as well as for the pooled panel. Table 2 presents the unit root test results for the first four and the last four statistical areas —organized in alphabetical order— as well as the panel unit root test for all the 364 SAs in our panel. The first and fifth columns report the ADF t-statistics, τ_{df}, for the tiers in levels showing very little evidence of stationarity. We reject the unit root null at at least 10 % level for $\text{LowTier}^{\text{York}}$, $\text{LowTier}^{\text{Yuba City}}$, and $\text{LowTier}^{\text{Yuma}}$. All the other SA high and low tiers reported in the table have a unit root, which means that shocks are permanent. Overall at a 10 % significance level 306 (364−58) of the 364 low tiers and 320 (364 − 44) of the 364 high tiers have a unit root. Columns three and seven show that for the series in differences the low tiers in 244 SAs and the high tiers in 255 SAs are stationary. Combining these results with the results in levels we find that at 10 % significance level, 176 of the series are integrated of order one, $I(1)$. Keeping track of these 176 SAs is important because this is the set of SAs for which we will test for a cointegrating relationship. The bottom part of Table 2 shows the Pesaran (2007) panel results. Both levels are non-stationary, while both first-differences are stationary. The null hypothesis in this test is that all series are non-stationary and the alternative is that only a fraction of the series are stationary.

[9]The Pesaran (2007) test uses a system of ADF regressions:

$$\begin{aligned}\Delta\left(\text{LowTier}_t^{SA} - \text{HighTier}_t^{SA}\right) = {} & \rho^{SA}\left(\text{LowTier}_{t-1}^{SA} - \text{HighTier}_{t-1}^{SA}\right) \\ & + \sum_{j=1}^{k_j} \phi^{SA}(j)\Delta\left(\text{LowTier}_{t-j}^{SA} - \text{HighTier}_{t-j}^{SA}\right) + \varepsilon_t^{SA},\end{aligned}$$

where the statistic of interest ($\bar{Z}_{\text{adf}}$) is distributed standard normal standard normal. It is based on means of individual ADF t-statistic (τ_{df}) for each SA.

Table 2 Unit Root tests

Statistical Area	(1) LowTier Levels τ_{df}	(2) Lags	(3) Differences τ_{df}	(4) Lags	(5) HighTier Levels τ_{df}	(6) Lags	(7) Differences τ_{df}	(8) Lags	(9) LowTier−HighTier Levels τ_{df}	(10) Lags
Aberdeen, WA	−1.209	13	−8.871*	14	−2.226	14	−3.230†	12	−1.120	13
Adrian, MI	−1.445	8	−1.811	7	−1.304	14	−2.555‡	7	−1.103	8
Akron, OH	−0.713	10	−2.097	8	−1.788	7	−1.780	5	−0.178	9
Albany, GA	0.488	13	−10.08*	14	−0.642	11	−5.490*	14	−0.909	9
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
York, PA	−2.630‡	3	−2.986†	13	−2.307	4	−1.465	7	−2.647‡	4
Youngstown, OH	−1.007	8	−3.763†	13	−1.481	14	−4.569*	5	−1.492	7
Yuba City, CA	−2.599‡	5	−1.923	5	−2.073	14	−2.162	13	−1.903	5
Yuma, AZ	−2.730‡	3	−1.460	7	−1.599	11	−2.449	7	−3.044‡	10
Overall (# SAs=364):[a]										
10 %	58		244		44		255		97	
5 %	34		218		26		235		70	
1 %	15		171		7		192		23	
	$\bar{Z}_{adf}$		$\bar{Z}_{adf}$		$\bar{Z}_{adf}$		$\bar{Z}_{adf}$		$\bar{Z}_{adf}$	
Panel	0.956	12	−36.262*	11	0.734	12	−39.720*	11	−1.176	14
	[0.830]		[0.000]		[0.769]		[0.000]		[0.120]	

Notes: τ_{df} is the test-statistics of the modified Dickey-Fuller Generalized Least Squares. $\bar{Z}_{adf}$ are the augmented ADF z-statistic as proposed in Pesaran (2007). Figures in brackets are p-values. * significant at 1 %; † significant at 5 %; ‡ significant at 10 %. Critical values are −3.480 (1 %); −2.811 (5 %); −2.531 (10 %). [a] number of SAs that number of SA for which the corresponding null hypothesis is rejected at 10 %, 5 % and 1 % significance levels

Column 9 in Table 2 presents the tests for the existence of a long-run equilibrium while imposing the restriction that β in Eq. 1 is equal to one. This test for the stationarity of $\text{LowTier}_t^{\text{SA}} - \text{HighTier}_t^{\text{SA}}$ can also be viewed as a test for the existence of price convergence between tiers as defined in Eq. 2. The results show important evidence of price convergence (and long-run equilibrium) between price tiers. From the eight SAs we report, only for York and Yuma there is no evidence of convergence at a 10 % significance level. Overall at a 10 % level there is no evidence of convergence for 97 of the 364 SAs in the sample (i.e., there is evidence of convergence for about 3/4 of the areas). In line with these findings the $\bar{Z}_{\text{adf}}$ statistic for the panel fails to reject the null that price convergence (long-run equilibrium) exists. The constraint in these tests is that it is assumed that $\beta = -1$. In the next section we provide cointegration tests that analyze the existence of a long-run equilibrium with an unrestricted β.

Estimating the Long-run Equilibrium and Short-run Dynamics

While Eq. 1 characterizes the long-run equilibrium, this simple specification is unlikely to capture the true data generating process. We therefore estimate a more flexible vector-error correction model which accounts both for the long-run equilibrium and the short-run dynamics around the long-run equilibrium. The model is described by the following equations:

$$\begin{aligned}
\Delta\text{LowTier}_t^{\text{SA}} &= a_L + \alpha_L\Big(\text{LowTier}_{t-1}^{\text{SA}} + \beta\text{HighTier}_{t-1}^{\text{SA}}\Big) \\
&\quad + \sum_{j=1}^{k} a_{LL}(j)\Delta\text{LowTier}_{t-j}^{\text{SA}} + \sum_{j=1}^{k} a_{LH}(j)\Delta\text{HighTier}_{t-j}^{\text{SA}} + \varepsilon_{Lt}, \quad (3) \\
\Delta\text{HighTier}_t^{\text{SA}} &= a_H + \alpha_H\Big(\text{LowTier}_{t-1}^{\text{SA}} + \beta\text{HighTier}_{t-1}^{\text{SA}}\Big) \\
&\quad + \sum_{j=1}^{k} a_{HL}(j)\Delta\text{LowTier}_{t-j}^{\text{SA}} + \sum_{j=1}^{k} a_{HH}(j)\Delta\text{HighTier}_{t-j}^{\text{SA}} + \varepsilon_{Ht}, \quad (4)
\end{aligned}$$

where ε_{Lt} and ε_{Ht} are white-noise disturbance terms that may be correlated. The term in parenthesis, which is the same as d_t^{SA} in Eq. 1, captures the long-run equilibrium dynamics. The left-hand side variables capture the short-run dynamics, while α_L and α_H are the speed-of-adjustment coefficients. If deviations are positive (i.e., $d_t^{\text{SA}} > 0$) and we have a negative β, the low tier index is relatively larger than the high tier index and it is reasonable to believe that we will have a decrease in the low tier index (i.e., $\Delta\text{LowTier}_t^{\text{SA}} < 0$) and an increase in the high tier index (i.e., $\Delta\text{HighTier}_t^{\text{SA}} > 0$). Hence, we would expect α_L to be negative and α_H to be positive. The lag dependent variables in this system of equations are included to control for serial correlation and to test for momentum.

Note that if there is no cointegrating relationship between the price tiers, that is, if we restrict $\alpha_L = \alpha_H = 0$, Eqs. 3 and 4 define a Vector Autoregressive model.

Our vector-error correction model thus is a generalized version of an autoregressive specification that allows us to test for momentum in the dynamics of price tier. From Eqs. 3 and 4 we can see, for example with $k = 1$, that a positive coefficient a_{LL} can be interpreted as a momentum effect in LowTier. That is, an increase (decrease) in the low tier price in the last period, $\Delta\text{LowTier}^{SA}_{t-1}$, is associated with an increase (decrease) in the price in the current period, $\Delta\text{LowTier}^{SA}_{t}$. Likewise a_{HH} would capture the momentum in HighTier. While each of these coefficients a_{LL} and a_{HH} come from separate equations, the estimation of the vector-error correction model as a system of equations allows us to test whether momentum is particularly greater for a specific tier.[10]

The estimation of the vector-error correction model follows the methods in Johansen (1988) and Johansen and Juselius (1990). While Engle and Granger's (1987) two-step error-correction model may also be used in this context, the Johansen's error-correction model yields more efficient estimates of the cointegrating coefficients. This is because it is a full information maximum likelihood estimator which allows testing for cointegration in the system of equations in a single step without requiring a specific variable to be normalized. Moreover, it avoids carrying over the errors from the first step into the second, and does not require any prior assumptions regarding the causality of the variables.

In Table 3 we present tests for the existence of a cointegration relationship between the high and the low tier in each statistical area. As we are testing for cointegration between two time series in each statistical area, the number of cointegrating equations can be at most one, which occurs when a coefficient β exists such that the error terms ε_{Lt} and ε_{Ht} defined in Eqs. 3 and 4 are stationary.[11]

The maximum eigenvalue statistic reported in columns 1 and 2 presents a likelihood-ratio test of the null hypothesis that there are exactly r cointegrating equations against the alternative that there are $r + 1$ cointegrating equations. Based on this test, the null hypothesis of no cointegration between the price tiers can be rejected for Aberdeen WA, Albany GA, Winston-Salem NC, and Wooster OH at the 5 % significance level. The null hypothesis of the trace statistic reported in columns 4 and 5 is that there are no more than r cointegrating equations. Consistent with the findings in Lütkepohl et al. (2001), the maximum eigenvalue in columns 1 through 3 and the trace statistic in columns 4 through 6 yield similar results. The number of cointegrating equations reported in columns 3 and 6 is equal to one in 63 and 62 of the 176 statistical areas, respectively.[12]

We interpret this as a mixed evidence of a cointegrating relationship between tiers when no breaks are allowed based on the maximum eigenvalue and trace statistic.

[10]Note that, under the existence of a long-run equilibrium between price tiers, simpler autoregressive (AR) or vector-autoregressive (VAR) specifications of the first-differences of the price tiers would provide biased estimates of the momentum coefficients. An AR would not control for the link between price tiers, while a VAR would be missing the dynamics around the long-run equilibrium.

[11]In general, with n time series, there might exist up to $n - 1$ cointegrating vectors. Johansen's (1988) approach can be used to estimate these distinct relationships.

[12]Notice that we only work with the 176 SAs that have both tiers integrated of order one, to follow the definition of cointegrating relationship.

Table 3 Cointegration tests

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Maximum Eigenvalue			Trace Statistic			Minimum Hannan-Quinn			
Statistical Area	$r = 0$	$r = 1$	# Eq.	$r = 0$	$r \leq 1$	# Eq.	$r = 0$	$r = 1$	$r = 2$	# Eq.
Aberdeen, WA	33.25	2.176	1	41.43	2.176	1	6.392	6.288	6.297	1
Albany, GA	20.89	2.200	1	23.09	2.200	1	4.537	4.488	4.494	1
Albertville, AL	4.035	0.242	0	4.278	0.242	0	6.091	6.118	6.132	0
Albuquerque, NM	3.990	0.0167	0	4.007	0.0167	0	4.871	4.898	4.913	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Winston-Salem, NC	67.30	0.0168	1	67.32	0.0168	1	1.911	1.655	1.670	1
Wooster, OH	16.48	2.979	1	19.46	2.979	1	3.626	3.597	3.599	1
Yakima, WA	5.925	0.952	0	6.877	0.952	0	3.150	3.169	3.180	0
Youngstown, OH	13.55	0.00201	0	13.55	0.00201	0	3.094	3.079	3.094	1
Overall (# SAs=176):[a]			63			62				118

Notes: Critical values for the maximum eigenvalue for $r = 0$ are 14.07 (5 %), 18.63 (1 %); and for $r = 1$ are 3.76 (5 %), 6.65 (1 %). Critical values for the trace statistic for $r = 0$ are 15.41 (5 %); 20.04 (1 %). For $r \leq 1$ are 3.76 (5 %); 6.65 (1 %). [a] number of SAs in which there is a cointegrating equation

The number of cointegrating equations in a vector error correction model can also be determined by minimizing the Hannan and Quinn information criterion (see e.g. Gonzalo and Pitakaris (1998), Aznar and Salvador (2002)). For every SA in which the maximum eigenvalue and the trace statistic found a single cointegrating equation, the minimum Hannan-Quinn statistic also found a single cointegrating equation. However, the minimum Hannan-Quinn finds evidence of cointegration in more statistical areas. The last column reports that a cointegrating relationship exists for Aberdeen WA, Albany GA, Winston-Salem NC, Wooster OH, and Youngstown OH. Overall based on this criteria there exists a cointegrating relationship in 118 of the areas. We will use this set of 118 SAs when estimating the vector-error correction model.

It is possible that the lack of cointegration between some of the housing tiers is due to the fact that the Johansen methods do not allow for shifts. Table 4 presents the cointegration tests proposed by Gregory and Hansen (1996a) and Gregory and Hansen (1996b) that include regime and trend shifts. The test's null hypothesis is no cointegration against the alternative of cointegration with changes in level and trend. We run this test only for the 58 statistical areas in which the minimum Hannan-Quinn statistics did not find a cointegrating equation. The different columns in Table 4 present three test statistics, the ADF-type test, the Z_α-type test and the Z_t-type test for eight of the SAs and a summary of the overall results. These tests

Table 4 Cointegration tests with shifts

	(1) ADF-type test	(2)	(3)	(4) Z_α-type test	(5)	(6) Z_t-type test	(7)
Statistical Area	t-stat	Break	Lags	Z_α-stat	Break	Z_t-stat	Break
Albertville AL	−6.376*	2000m5	1	−6.195*	1996m12	−42.66†	1996m6
Albuquerque NM	−4.896	1998m6	1	−5.758†	1998m3	−36.00	1998m4
Astoria OR	−4.606	1994m7	3	−7.084*	1993m4	−23.90	2000m10
Athens TN	−5.785†	2001m6	3	−7.188*	2001m11	−36.80	1996m2
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Wapakoneta OH	−5.954†	1997m10	3	−6.550*	1997m10	−41.72†	1998m2
Wausau WI	−4.922	1993m6	1	−5.073‡	1994m4	−32.99	1993m5
Whitewater WI	−6.602*	1994m5	1	−6.147*	1994m1	−38.03‡	1994m8
Yakima WA	−4.235	1999m9	1	−5.633†	2001m2	−27.22	2000m8
Overall (# SAs=58):[a]							
10 %	27			51		18	
5 %	16			46		11	
1 %	6			32		2	

Notes: Optimal lag for ADF-type test chosen by Akaike criterion. All specifications model a change in regime and trend. * significant at 1%; † significant at 5 %; ‡ significant at 10 %. Critical values for the ADF-type test are −6.02 (1 %); −5.50 (5 %); −5.24 (10 %). Critical levels for the Z_α-type test and the Z_t-type test are different for different SAs. [a] number of SAs in which there is a cointegrating equation at 10 %, 5 % and 1 % significance levels

can be viewed as multivariate extensions of Perron (1989), Banerjee et al. (1992), Perron and Vogelsang (1992), and Zivot and Andrews (1992). The Z_α-type statistic is the one that finds the most evidence supporting the existence of cointegrating relationships. Overall at a 10 % significance level there is a cointegrating relationship for 51 of the 58 SAs. Columns 2, 5, and 7 estimate the dates of the shifts.

Table 5 presents the maximum likelihood estimation of the system of Eqs. 3 and 4 for each of the 118 SAs that have both tiers integrated of order one and for which a cointegration equation exists.[13] The lag length is selected based on the Akaike Information criterion. The long-run equilibrium is captured by the cointegrating vector $[1, \beta]$, with the estimates of β being reported in column 9. The results show that there

[13]This estimation does not take into account shifts. However, later on we estimate a rolling regression that allows for a time-varying β, which is more in line with the evidence of shifts in the cointegrating relationships found in Table 4.

Table 5 Vector error correction

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	ΔLowTier				ΔHighTier					H_0 :		
Statistical Area	α_L	a_L	a_{LL}	a_{LH}	α_H	a_H	a_{HL}	a_{HH}	β	$a_{LL} = a_{HH}{}^a$	$\alpha_L = \alpha_H = 0$	Log Lik.
Aberdeen, WA	−0.0292*	−0.00183	0.670*	0.0599	−0.00191	0.0279	0.0158	0.901*	−0.525*	[7.80E-05]	[0.000353]	−432.7
	(0.00739)	(0.0558)	(0.0492)	(0.0490)	(0.00485)	(0.0366)	(0.0323)	(0.0322)	(0.0653)			
Adrian, MI	0.00024	−0.0386	0.597*	0.272†	0.000192	0.0483†	−0.00438	0.775*	-10.40‡	[0.00452]	[0.260]	−329.1
	(0.000331)	(0.0606)	(0.0536)	(0.107)	(0.00013)	(0.0238)	(0.0211)	(0.0419)	(5.926)			
Allegan, MI	0.000256	−0.00229	0.737*	0.0561	0.000364†	0.00161	−0.0327	0.831*	−11.73†	[0.072]	[0.091]	−345.7
	(0.000191)	(0.0484)	(0.047)	(0.0423)	(0.000178)	(0.045)	(0.0438)	(0.0394)	(5.095)			
Anderson, IN	0.00401	−0.0237	0.766*	−0.0987	0.00801*	0.0119	0.114*	0.482*	−2.459*	[0.000134]	[0.0162]	−533.6
	(0.00394)	(0.0698)	(0.049)	(0.079)	(0.00281)	(0.0497)	(0.0349)	(0.0563)	(0.526)			
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Waterloo, IA	−0.00635†	−0.0371	0.656*	0.0712	−0.00246	0.0959‡	−0.0454	0.768*	−0.323	[0.0735]	[0.0759]	−432.1
	(0.00286)	(0.0826)	(0.0531)	(0.071)	(0.00197)	(0.0568)	(0.0365)	(0.0489)	(0.271)			
Wausau, WI	0.00485*	−0.015	0.665*	0.0991	0.00233†	0.0312‡	−0.00273	0.772*	−1.569*	[0.077]	[0.0097]	−11.52
	(0.00175)	(0.0251)	(0.0541)	(0.0713)	(0.00114)	(0.0163)	(0.0351)	(0.0463)	(0.276)			
Wooster, OH	4.19E-06	0.0104	0.707*	0.0679	0.00765†	−5.71E−06	0.0699	0.686*	−1.196*	[0.398]	[0.044]	−203.2
	(0.00377)	(0.0393)	(0.0512)	(0.0554)	(0.00353)	(0.0369)	(0.0479)	(0.0519)	(0.267)			

Table 5 (continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	ΔLowTier				ΔHighTier					H_0 :		
Statistical Area	α_L	a_L	a_{LL}	a_{LH}	α_H	a_H	a_{HL}	a_{HH}	β	$a_{LL} = a_{HH}$[a]	$\alpha_L = \alpha_H = 0$	Log Lik.
Youngstown,OH	−0.000728‡	−0.0326	0.629*	0.182†	−0.000506‡	0.0469†	−0.0675†	0.753*	6.996	[0.0391]	[0.0625]	−180.7
	(0.000438)	(0.0355)	(0.0511)	(0.0768)	(0.000264)	(0.0214)	(0.0308)	(0.0463)	(3.034)			
Overall (# SAs=118):[b]												
10 %	101	6	118	52	50	41	32	118	104	63	112	
5 %	87	4	118	44	38	31	26	118	102	44	98	
1 %	57	0	118	20	18	12	14	118	97	25	60	

Notes: Figures in parentheses are standard errors. Figures in square brackets are p-values. * significant at 1 %; † significant at 5 %; ‡ significant at 10 %. [a] the alternative hypothesis is $H_a : a_{LL} > a_{HH}$. [b] number of SAs in which the null hypothesis (of statistical significance for columns 1 through 9) is rejected at 10 %, 5 % and 1 % significance levels exists important heterogeneity in the point estimates of β across SAs with the average β being -1.121. This long-run equilibrium has an interesting interpretation. On average we have $\text{LowTier}_t = 1.121 \cdot \text{HighTier}_t$. That is, a one point increase in the high tier price index is linked with a 1.121 points increase in the low tier index. Without any assumptions of the causality between LowTier and HighTier, a β coefficient that is greater than one (in absolute value) means that fluctuations of the low tier are larger than fluctuations in the high tier. In the next section we analyze the role of the housing bubble for the stability of β.

From the estimates of a_{LL} and a_{HH} reported in Table 5 we can see that there is strong evidence for momentum in both tiers. At the 1 % level a_{LL} and a_{HH} are both statistically significant for all the 118 SAs in this sample. The momentum in the low tier ranges between 0.464 for Provo, UT to 0.994 for Key West, FL and a similar range is observed for the high tier. Given the estimate of β that indicates that fluctuations in the low tier are larger than in the high tier, we also explore whether the momentum effect is stronger in the low segment of the market. Column 10 provides the p-values of the null of $a_{LL} = a_{HH}$ versus the alternative $a_{LL} > a_{HH}$. The results show that at a 10 % significance level we reject the null for 63 of the 118 areas in this sample. One of the interpretations for the observed momentum effect is buyer irrationality in the real estate market (Pan and Pirinsky, 2013). Enders (2010) p. 367 explains that at least one of the speed of adjustment terms in Eqs. 3 and 4 must be nonzero to establish that the short run dynamics responds to deviations from long run equilibrium. Column 11 reports the p-values of testing the null hypothesis $\alpha_L = \alpha_H = 0$. We find that at 10 % significance level, the null is rejected for 112 of the 118 statistical areas in the sample. We will next use these 112 SAs in our analysis of the stability of the long-run equilibrium.

Stability of the Long-run Equilibrium

The estimation of the system of Eqs. 3 and 4 assumes that β is time invariant. This is potentially a strong assumption given the woes faced by the housing markets since the onset of the subprime crisis. To assess the stability of the long-run equilibrium relationship ($\text{LowTier}^{SA}_{t-1} + \beta \text{HighTier}^{SA}_{t-1}$) we use a rolling regression approach in which we allow β to change over time.[14] For a given window of size w we estimate the system of Eqs. 3 and 4 using the first w observations, $[1, w]$. Then we estimate the system again using the sample $[2, w+1]$ and so forth until we exhaust the data using the subsequent samples $[2, w+1], [3, w+2], \ldots, [N-w+1, N]$. This procedure generates a sequence $\{\beta_t\}$ of coefficients that characterize the dynamics of the cointegrating equation. Because we do this for every statistical area i we actually obtain a panel of β_{it} estimates.

[14] In the context of bubbles, Driffill and Sola (1998) use similar rolling regression approach in a random walk with drift regression to provide evidence of instability and regime-switching. Swanson (1998) uses rolling regressions in vector error correction models similar to the ones we estimate here.

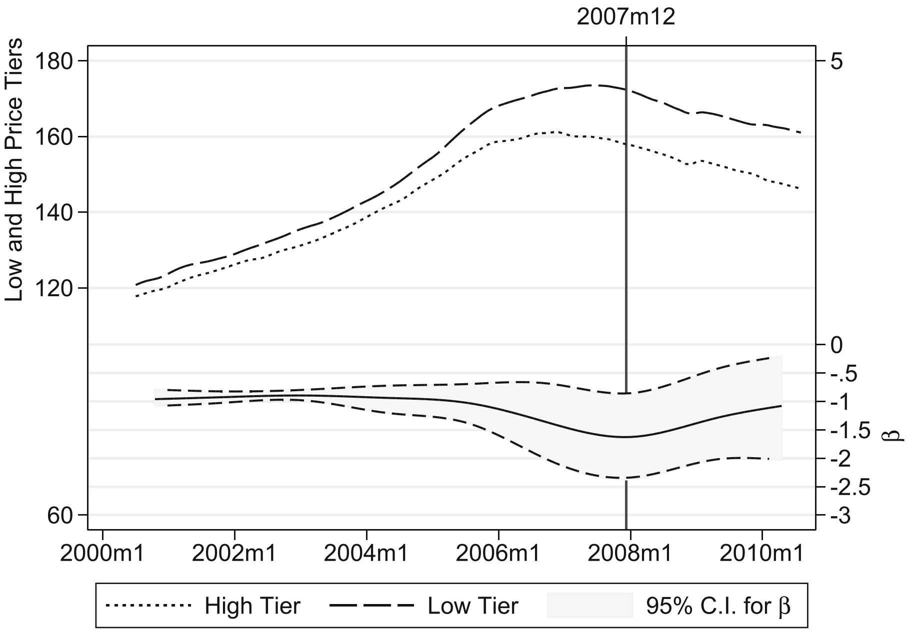

Fig. 1 Tiers and Rolling Regressions Estimates of β. Notes: The upper section of the figure presents the average TIERLOW and TIERHIGH across all statistical areas. The lower section presents the average rolling regressions estimates of the βs along with their 95 % confidence intervals

Using this panel, we analyze how the housing bubble has affected the long-run relationship between price tiers. The model we estimate is the following:

$$\beta_{it} = \delta \cdot \mathbb{I}_{|\tau_i - t| \leq \theta} + \mu_i + \eta_{it}, \tag{5}$$

where $\mathbb{I}$ is an indicator function capturing the period of bubble formation and burst. It assumes the value of one during the time of the bubble, i.e. when $|\tau_i - t| \leq \theta$, where τ_i denotes the time at which the housing bubble in metropolitan area i bursts and θ is a positive integer that captures the time distance away from the burst.[15] The coefficient of interest δ captures any shift in the cointegrating coefficient β_{it} in the neighborhood around the burst of the bubble. μ_i is the statistical area time-invariant specific effect, and η_{it} is the remaining stochastic term. We use $\theta = (3, 6, 12)$ as robustness checks to account for the uncertainty regarding the beginning and end dates of the bubble.

Figure 1 illustrates the shift in the long-run equilibrium during the bubble for the sample of 112 SAs that comply with the conditions for a vector error correction model. In the upper part of the figure we have the average low and average high

[15]The time of the burst τ_i is obtained as the date in which LOWTIER_t^{SA} for the corresponding SA reaches its maximum.

price tiers. Averages are obtained at every point in time across all SAs. The average burst date across statistical areas (shown as the vertical line on December 2007) is computed as $\tau = \frac{1}{N}\sum_{i=1}^{N}\tau_i$. In the lower part we plot the average βs across SAs along with their 95 % confidence intervals.[16]

The salient feature of Fig. 1 is that during the early years of the sample β appears very stable and close to -1. This means that the long-run relationship between LowTier and HighTier was very close to a one-to-one relationship —for every point in the appreciation of the LowTier index, the HighTier index would also appreciate by one point. However, around the date of the burst during the bubble years β appears to be larger in absolute value suggesting a higher appreciation and depreciation of the low tier. The formal analysis of any possible shift in the long-run relationship between LowTier and HighTier follows the estimation of Eq. 5. The results are presented in Table 6. Different columns report different sets of estimates of β prior to the bubble burst (β_{PRE}), during the bubble (β_{DUR}), and the difference between the two ($\beta_{\text{DUR}} - \beta_{\text{PRE}}$). The odd-numbered columns present pooled regression estimates while the even-numbered columns report fixed effects estimates that control for any time-invariant SA specific characteristic. For example with $\theta = 3$ the estimates in the second column indicate that during the six months period around the bubble burst, a one point increase (decrease) in the HighTier index was associated with a 1.238 points increase (decrease) in the LowTier index. This implies that the rate of appreciation (depreciation) of the low tier housing prices was greater than the rate of appreciation (depreciation) of the high tier housing prices.

Column 2 also shows that the point estimate of β_{PRE} is closer to -1, but the p-value associated with the null of $\beta_{\text{PRE}} = -1$ shows strong evidence against the null. The difference between β_{PRE} and β_{DUR} is statistically significant at at least 1 %. Consistent with Fig. 1, differences in the appreciation/depreciation rates between LowTier and HighTier are greater around the bubble burst. Columns 3 through 6 present robustness checks to account for the uncertainty about the beginning and end of the bubble periods. Columns 3 and 4 consider a six-month window ($\theta = 6$), while columns 5 and 6 a one-year window ($\theta = 12$). All the point estimates are statistically significant at at least 1 % level. The last row in Table 6 (above the number of observations) presents the p-values of the null hypothesis that $\beta_{\text{PRE}} = \beta_{\text{DUR}}$ with the alternative that $\beta_{\text{PRE}} > \beta_{\text{DUR}}$. The results across columns indicate strong evidence against the null hypothesis —the β coefficient shifted away from one around the bubble periods. The point estimate in the last column reads that the difference in the appreciation/depreciation rates between the two price tiers is 0.166 smaller in the years leading to the bubble period than during the bubble.

As a robustness check and to see if a larger number of statistical areas meet the requirements for the estimation of vector error correction models, we take the natural logarithms of HighTier and LowTier and follow the same steps as before for the transformed variables. The unit root tests, cointegration tests, and the estimation of these additional vector error correction models with the variables resulted in a set of

[16] The estimation uses a window $w = 100$. The lines were smoothed using the fitted cubic splines.

Table 6 Long run equilibrium

	(1) Within $\theta = 3$	(2)	(3) Within $\theta = 6$	(4)	(5) Within $\theta = 12$	(6)
Variables	Pooled	FE	Pooled	FE	Pooled	FE
β_{PRE}	−1.075*	−1.077*	−1.067*	−1.070*	−1.048*	−1.055*
	(0.00870)	(0.00685)	(0.00890)	(0.00701)	(0.00934)	(0.00737)
$H_0 : \beta_{\text{PRE}} = 0$	[0]	[0]	[0]	[0]	[0]	[0]
$H_0 : \beta_{\text{PRE}} = -1$	[0]	[0]	[0]	[0]	[3.29e-07]	[0]
β_{DUR}	−1.281*	−1.238*	−1.265*	−1.227*	−1.257*	−1.221*
	(0.0408)	(0.0322)	(0.0287)	(0.0227)	(0.0204)	(0.0161)
$H_0 : \beta_{\text{DUR}} = 0$	[0]	[0]	[0]	[0]	[0]	[0]
$H_0 : \beta_{\text{DUR}} = -1$	[0]	[0]	[0]	[0]	[0]	[0]
$\beta_{\text{DUR}} - \beta_{\text{PRE}}$	−0.206*	−0.161*	−0.199*	−0.157*	−0.209*	−0.166*
	(0.0417)	(0.0329)	(0.0301)	(0.0238)	(0.0224)	(0.0178)
$H_0 : \beta_{\text{DUR}} = \beta_{\text{PRE}}{}^a$	[7.84e-07]	[1.07e-06]	[0]	[0]	[0]	[0]
Obs.	13,888	13,888	13,888	13,888	13,888	13,888

Notes: Figures in parentheses are standard errors. Figures in brackets are p-values. * significant at 1 %; † significant at 5 %; ‡ significant at 10 %. [a] the alternative hypothesis is $H_a : \beta_{\text{DUR}} < \beta_{\text{PRE}}$.

123 statistical areas. We obtain new βs from rolling regression of the variables in logarithms and then estimated Eq. 5 again. The results are reported in Table 7. As the new variables are logarithm transformations, the interpretation of the βs is in terms of percentage changes. Column 6 indicates that a one percentage increase (decrease) in HighTier is associates with a 1.106 % increase (decrease) in LowTier in the period prior to the bubble. The elasticity increases to 1.289 during the bubble period with the difference between the period prior to the bubble and the bubble period being statistically significant. The results are robust to changes in the length of the bubble period θ as observed across different columns.

One concern when studying the link between market segments is the existence of a spurious correlation if any of the segments has a trend (deterministic or stochastic). Granger and Newbold (1974) explain that spurious correlation may still exist even after detrending. In the estimation of the vector error correction models we need the variables to be stationary, in addition to having a cointegrating relationship. This approach means that we only estimate the βs for the statistical areas for a reduced sample where there is a genuine relationship and where a long-run relationship exists. If the focus is on the appreciation and depreciation rates and we are not interested in distinguishing between long- and short-run dynamics, we can implement a much

Table 7 Long run equilibrium (in logs)

	(1) Within $\theta = 3$	(2)	(3) Within $\theta = 6$	(4)	(5) Within $\theta = 12$	(6)
Variables	Pooled	FE	Pooled	FE	Pooled	FE
β_{PRE}	−1.127*	−1.129*	−1.118*	−1.122*	−1.099*	−1.106*
	(0.00904)	(0.00725)	(0.00925)	(0.00742)	(0.00972)	(0.00780)
$H_0 : \beta_{\text{PRE}} = 0$	[0]	[0]	[0]	[0]	[0]	[0]
$H_0 : \beta_{\text{PRE}} = -1$	[0]	[0]	[0]	[0]	[0]	[0]
β_{DUR}	−1.378*	−1.339*	−1.342*	−1.307*	−1.322*	−1.289*
	(0.0420)	(0.0338)	(0.0297)	(0.0239)	(0.0210)	(0.0169)
$H_0 : \beta_{\text{DUR}} = 0$	[0]	[0]	[0]	[0]	[0]	[0]
$H_0 : \beta_{\text{DUR}} = -1$	[0]	[0]	[0]	[0]	[0]	[0]
$\beta_{\text{DUR}} - \beta_{\text{PRE}}$	−0.251*	−0.210*	−0.224*	−0.185*	−0.224*	−0.183*
	(0.0430)	(0.0345)	(0.0311)	(0.0250)	(0.0232)	(0.0187)
$H_0 : \beta_{\text{DUR}} = \beta_{\text{PRE}}{}^a$	[5.52e-09]	[1.21e-09]	[0]	[0]	[0]	[0]
Obs.	15,252	15,252	15,252	15,252	15,252	15,252

Notes: For 123 SAs. Figures in parentheses are standard errors. Figures in brackets are p-values. * significant at 1 %; † significant at 5 %; ‡ significant at 10 %. [a] the alternative hypothesis is $H_a : \beta_{\text{DUR}} < \beta_{\text{PRE}}$

simpler analysis by using the whole sample.[17] Figure 2 presents the scatterplots and the histograms of the returns of the low and high segments. The histograms and data points on the scatterplot with black border correspond to the observations prior to the bubble, while the gray histograms and data points correspond to the periods during the bubble.

Figure 2 shows a relationship between the returns of the high and low segments that is close to a one-to-one as the data points appear to be scattered around the 45^{o} line. Moreover the returns for both segments appear to be lower prior to the bubble. A Wilcoxon test of the null hypothesis that the distributions are the same prior and during the bubble finds strong evidence against the null for both price segments. In addition, simple t-tests with the null of equality between means prior and during the bubble also find strong support against the null. This evidence using the whole sample is consistent with the shifts in the $beta$s. In the next two sections we compare our results to the literature on the price dynamics during previous bubble periods.

[17]One key benefit from separating the dynamics into a long-run and a short-run dynamics is that momentum is a short-run concept while the long-run shift that we find is interpreted as a shift in the permanent components of the price indexes.

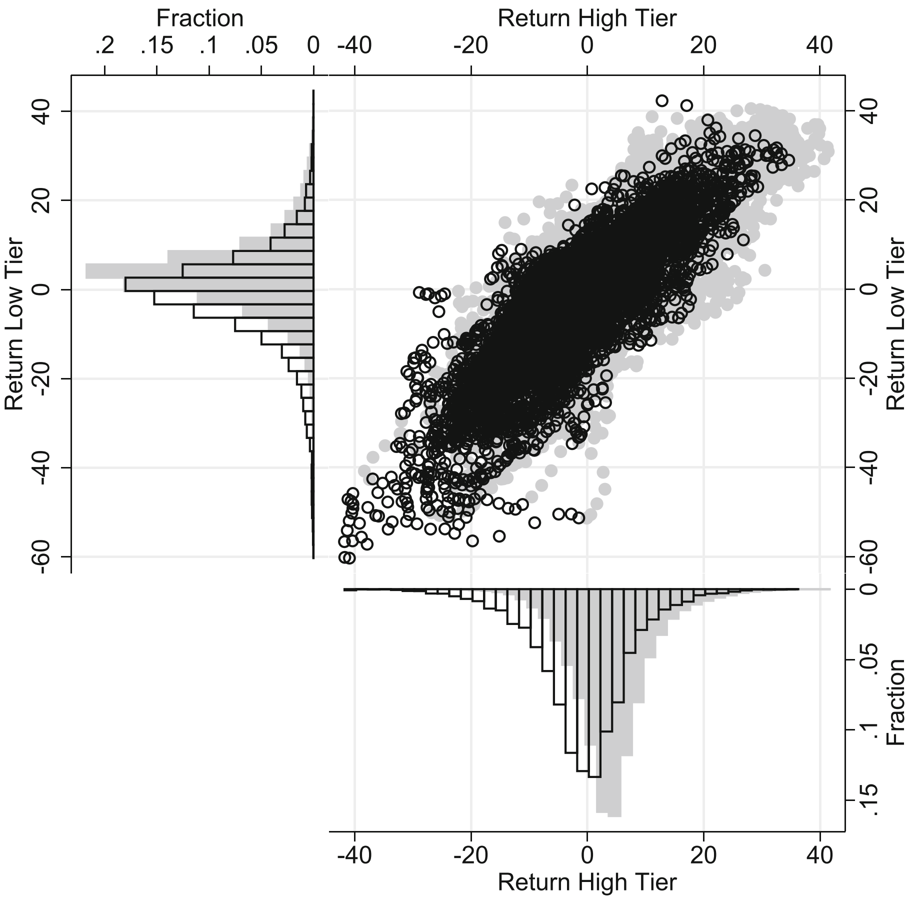

Fig. 2 LowTier and HighTier returns. Notes. This figure presents the scatter plot of the returns of LowTier and HighTier along with the corresponding histograms. The periods prior to the bubble are illustrated in the scatter plot and histogram in grey, while the observations that correspond to the bubble periods are in black

Moreover, we present some cross-sectional characteristics that might help explain the observed variation in the βs.

Comparison to Previous Bubble Periods

One of the unique aspects of the recent housing bubble, as supported by our estimates of β, was the tendency of low tier homes to appreciate faster during the boom and to fall more precipitously during the bust of the market. This feature stands in stark contrast to previous housing market cycles. Smith and Tesarek (1991) find that during the boom period between 1970 and 1985 in Houston, high value homes appreciated at an annual rate of 9 % while low value homes appreciated by only 8.3 %. In the subsequent downturn period 1985-1987, however, high-quality houses lost nearly 30 % of their value while low quality houses depreciated by only 18 %. Seward et al. (1992) document similar price dynamics during the same time period for St. Petersburg, Florida. The only exception known to us is the study by Case and Shiller (1994) who compare Los Angeles and Boston during the period 1983-1993. For the most part the price tiers move closely together in both cities, yet in Boston the low tier continued to rise as the upper tier flattened out and then fell more sharply during the downturn of the market. The Boston housing market exhibited a similar trend during the recent housing bubble as well.

Poterba (1991) presents a more comprehensive analysis of house prices over the time period 1970-1982 in four cities: Atlanta, Chicago, Dallas, and Oakland. He reports that houses in the top quartile appreciate faster than houses in the bottom quartile. Probably most closely related to the present paper is the study by Mayer (1993). Analyzing the same four cities, he concludes that "...homes do not appreciate at the same rate, but that more of the volatility occurs among high-priced homes." He offers two explanations for his findings. The one is that periods of rising prices mitigate the down payment constraints for trade up homes (see, e.g. Stein (1995)). The other is that periods of home appreciations improve the "balance sheet" of households leading to an increased demand for luxury goods, in particular more expensive houses. Our rolling regression approach leads us to conclude that low tier homes in the entire panel appreciated faster during the price run-up and depreciated more precipitously after the bubble burst after controlling for metropolitan area-specific fixed effects (see Table 6).

Explaining the Tiered Price Dynamics

In this section we explore to what extent demographic variables, lending practices, and behavioural factors (momentum) can explain the observed price dynamics during the bubble period.

Demographic Variables and Lending Practices

When viewed as entirely investment assets, according to the efficient market hypothesis, different houses should experience roughly the same rates of appreciation if we assume that they are linked to the same market risk factors. Further, according to the construction cost viewpoint, if house prices merely reflect the variation in construction inputs, in markets without substantial frictions, price tiers should exhibit roughly the same patterns over time (Poterba (1991)). If this pattern holds in the long-run we should observe a $\beta = -1$. The dynamics of our rolling estimates of β over time presented in Fig. 1 suggests, however, that beta drops below -1 and is generally lower during the period 2006-2010 reaching a minimum around the time when the local housing market bubbles bursted during the period 2007-2008. The

Table 8 Explaining long run equilibrium shifts

Variables	(1) OLS	(2) OLS	(3) OLS	(4) FE	(5) FE	(6) FE
HIGHERPRICED	−0.338*	−0.338*	−0.242†	−0.355*	−0.371*	−0.383*
	(0.0971)	(0.0973)	(0.0972)	(0.0690)	(0.0690)	(0.0782)
$\mathbb{I}_{\lvert\tau_i - t\rvert \leq \theta}$		0.00470	0.0479		−0.211*	−0.211*
		(0.0724)	(0.0715)		(0.0538)	(0.0538)
POPULATION			−5.69e-07*			5.87e-07
			(6.98e-08)			(8.58e-07)
INCOME			0.0132*			−0.00197
			(0.00233)			(0.00562)
Constant	−1.145*	−1.146*	−2.107*	−1.131*	−1.083*	−1.100†
	(0.0867)	(0.0882)	(0.217)	(0.0616)	(0.0626)	(0.553)
Obs.	3,015	3,015	3,015	3,015	3,015	3,015

Notes: The dependent variable is β_{it}. The sample is from January 2007 through December 2013. Figures in parentheses are standard errors. $\theta = 3$. * significant at 1 %; † significant at 5 %; ‡ significant at 10 %

more formal results presented in Tables 6 and 7 also support this conclusion. We find strong evidence that β deviates away from -1 during the bubble period. These observed equilibrium shifts may be the result of demographic factors and behavioral forces related to demand. We therefore further investigate the impact of subprime lending practices as well as the effect of population and income on the β coefficients.

Following Mayer and Pence (2009) we use "high-priced" loans as a proxy for subprime loans. Under the Home Mortgage Disclosure Act (HMDA), most originators must report attributes of mortgage applications. Avery et al. (2007) explain that HMDA covers over 80 percent of all home loans and it is considered the most comprehensive source of mortgage data. As in Mayer and Pence (2009), we classify a loan as "high-priced" if the annual percentage rate (APR) is at least three percentage points above the Treasury benchmark for first-lien mortgages and if the APR is five percentage points over that benchmark for junior liens. We construct the variable HIGHERPRICED $\in [0, 1]$ that measures the proportion of higher-priced loans in the statistical area. We then augment the regression Eq. 5 to include this variable along with the statistical area characteristics POPULATION (measured in number of individuals), and INCOME (measured in thousands of dollars per year). The results from this regression are reported in Table 8. Consistent across all columns we find that HIGHERPRICED has a negative and statistically significant effect on β.[18] Higher

[18]The number of observations is smaller in this table as the sample starts in 2007 and was not available for all statistical areas.

proportion of high-priced loans is associated with a greater divergence between price segments. There results provide further evidence for the role of the mortgage lending channel for the housing market boom and bust. Mian and Sufi (2009) show that areas with high latent demand for credit (as measured by the percentage of mortgage applications rejected in 1996) saw the sharpest mortgage credit expansion and experienced the fastest increase in house prices in the 2001-2005 period. Credit was extended to borrowers with low credit records, who presumably constituted over proportionally buyers of low tier properties. Further evidence for this conjecture is provided by Landvoigt et al. (2012). In their analysis of the San Diego housing market they conclude that "cheaper credit for poor households was a major driver of prices, especially at the low end of the market." Along similar lines, Pavlov and Wachter (2011) show that financial deregulation, in particular the use of interest-only and subprime mortgages accentuates the real estate cycle and the price response to demand shocks. Using a novel measure of a housing bubble, Berkovec et al. (2012) further show that Alt-A and subprime market shares are correlated with the bubble.

The contribution of our analysis is in providing further evidence that the equilibrium shift between the high and the low price tiers that we observe during the bubble period is related to subprime lending. The estimated effects of the bubble indicator variable $\mathbb{I}_{|\tau_i - t| \leq \theta}$ on our cointegration estimates β is consistent with the estimate of $\beta_{\text{DUR}} - \beta_{\text{PRE}}$ reported in Tables 6 and 7, while POPULATION and INCOME are not statistically significant when controlling for statistical area specific characteristics in the fixed effects estimates in column 6.

Momentum Effects

In his Journal of Economic Perspectives summary article to the symposium on bubbles, Stiglitz emphasized the central role of momentum and positive feedback for the formation of bubbles in asset markets. He wrote that a bubble occurs "if the reason that the price is high today is only because investors believe that the selling price will be high tomorrow." (Stiglitz (1990), p. 13). The literature on housing market has since then taken two distinct approaches to identifying momentum and positive feedback.

The one approach relies on surveys of households regarding their expectations about future prices. Piazzesi and Schneider (2009) analyze responses from the Michigan Survey of Consumers and report that in the early 2000s an increasing number of households have become optimistic about future expected price appreciation rates after observing steady increase in prices for several years. They dub these optimistic households "momentum traders" and document that their percentage peaked to 20.2 percent in 2005 reaching a 25 year high. Using a search model to capture market frictions in the housing market, they show that that even a small number of momentum traders can drive up the prices significantly without increasing trading volume and market shares.

The other approach examines the correlation between lagged and current appreciation rates in housing price time series exploiting the idea that in the presence of momentum overly optimistic expectations transform into self-fulfilling prophecies during the buildup of the bubble. These correlations are commonly taken as an indication that housing markets are not weak form efficient.

Case and Shiller (1989, 1990) study the housing markets in Atlanta, Chicago, Dallas, and San Francisco/Oakland for the years 1970 to 1986 and find that appreciation rates of the previous year enter with an estimated coefficient of 0.3 in their autoregressive model specifications. Lags of more than one period are negatively related to current appreciation rates, although the coefficient is not statistically significant. More recently, Beracha and Skiba (2011) study autoregressive models of price appreciation in 381 metropolitan areas in the US during the 1983-2008 sample period and quarterly lags. While the one period lag appears to be negatively correlated —an indication of a slight mean reversion— the two and three period lags are positive and significant with coefficients on the magnitude of 15 % and 24 %, respectively. Our estimates of a_{LL} and a_{HH} in Table 5 are in line with a positive momentum and go beyond Case and Schiller (1989, 1990) by providing separate estimates for different market segments. Our empirical specification is more detailed in the sense that we study momentum for each price tier taking into account that the price tiers are bound by a long-run relationship. We find a strong evidence of momentum in both market segments.

Conclusion

It is now generally accepted that that the bust of the US residential housing market was a major cause for the deepest financial crisis in the US since the Great Depression (Mian and Sufi 2009). The events leading to the financial turmoil in the US and globally raised awareness about the systemic links between the housing market, the financial sector and the broader economy. Thus, understanding the price dynamics of residential real estate in the US during this period has been front and center in the recent academic literature on real estate. Most of the studies focus on aggregate indexes to identify global trends and information transmission mechanisms nationally and between various metropolitan areas (Miao et al. 2011).

This article broadens the scope of the existing literature by studying the links between market segments within statistical areas. Taking advantage of a large sample of US statistical areas, we analyze the conditions for the existence of a long-run relationship between the price tiers for each area. We estimate vector-error correction models with rolling windows and create a panel of cointegration coefficients. Using this approach, we establish that the cointegration parameter that bounds the tiers together is greater during the bubble period compared to the period of more moderate price movements. That is, according to the equilibrium relationship, low tier homes appreciate and depreciate at greater rates during the bubble. We also find evidence that the shift is linked to the percentage of subprime loan originations. While controlling for the existence of a long-run relationship, we find strong evidence of positive momentum in both tiers. Our analysis carries implications for a variety of participants in the real estate market. To current homeowners and prospective homebuyers it may suggest strategies regarding the timing of their decision to get into the market, sell their homes, or move up the property ladder (Seward, Delaney, and Smith 1992). To real estate investors and property developers it may shed light on the sources of financial risks. Given that mortgage default rates are linked not only to the characteristics of the borrowers, but also to home prices, our findings bear implications also to mortgage lenders and mortgage insurers. Probably the most important message is that mortgage insurance premiums have to take into account not only the risk profile of borrowers but also the price dynamics of the relevant housing market segment.

Appendix

The following is the list of all the SAs employed in the analysis. [a] denotes that the SA meets only the stationarity requirements. [b] denotes that the SA additionally meets the cointegration requirements. [c] denotes that the SA additionally meets the requirement that at least one of the speed of adjustment coefficients is nonzero in the vector error correction model: Aberdeen, WA[c]; Adrian, MI; Akron, OH; Albany, GA[b]; Albany, NY; Albany, OR; Albertville, AL[a]; Albuquerque, NM[a]; Alexandria, LA; Allegan, MI[c]; Allentown, PA; Anderson, IN[c]; Anderson, SC; Ann Arbor, MI[c]; Appleton, WI[b]; Asheville, NC; Ashtabula, OH; Astoria, OR[a]; Athens, GA[c]; Athens, TN[a]; Atlanta, GA; Atlantic City, NJ; Auburn, IN[c]; Auburn, NY[a]; Bainbridge, GA[a]; Bakersfield, CA; Baltimore, MD; Baraboo, WI[a]; Batavia, NY; Battle Creek, MI[c]; Bay City, MI[c]; Beaver Dam, WI[c]; Bedford, IN[a]; Bellingham, WA; Bend, OR; Big Rapids, MI[a]; Binghamton, NY[c]; Birmingham, AL[c]; Bloomington, IL[c]; Bloomington, IN; Bloomsburg, PA[a]; Boise City, ID; Boston, MA; Boulder, CO; Bremerton, WA; Burlington, NC[c]; Calhoun, GA[a]; Cambridge, MD[c]; Canton, OH; Cape Cod, MA; Carson City, NV; Cedar Rapids, IA[c]; Cedartown, GA[a]; Centralia, WA[c]; Champaign-Urbana, IL[c]; Charleston, SC; Charlotte, NC; Chattanooga, TN[c]; Chicago, IL; Chico, CA; Cincinnati, OH; Clarksville, TN; Cleveland, OH; Cleveland, TN[c]; Clewiston, FL[a]; College Station, TX[c]; Colorado Springs, CO; Columbia, MO[a]; Columbia, SC; Columbus, GA[c]; Columbus, OH; Concord, NH; Connersville, IN[c]; Coos Bay, OR[c]; Corning, NY[a]; Cortland, NY[c]; Corvallis, OR; Crawfordsville, IN[c]; Cumberland, MD; Dalton, GA; Davenport, IL[c]; Dayton, OH[c]; Daytona Beach, FL; Decatur, IN[a]; Defiance, OH; Denver, CO; Des Moines, IA[c]; Destin, FL; Dover, DE; DuBois, PA[c]; Dubuque, IA[c]; Duluth, MN[c]; Dunn, NC[c]; Durango, CO; Durham, NC; Dyersburg, TN[a]; Eagle Pass, TX; East Liverpool, OH[a]; East Stroudsburg, PA; Easton, MD[c]; Eau Claire, WI[c]; El Centro, CA; Elizabeth City, NC[c]; Elkhart, IN[c]; Ellensburg, WA[c]; Elmira, NY[a]; Erie, PA[a]; Eugene, OR; Eureka, CA[c]; Evansville, IN[c]; Fallon, NV[a]; Fargo, ND[a]; Fayetteville, AR; Fayetteville, NC; Flagstaff, AZ; Florence, SC; Fond du Lac, WI[c]; Fort Collins, CO[a]; Fort Myers, FL; Fort Smith, AR[a]; Fort Valley, GA[a]; Frankfort, IN[c]; Fremont, NE[c]; Fremont, OH[a]; Fresno, CA; Gainesville, FL; Gainesville, GA; Georgetown, SC[c]; Glens Falls, NY; Gloversville, NY[a]; Grand Island, NE[c]; Grand Junction, CO; Grand Rapids, MI; Grants Pass, OR[c]; Greeley, CO; Greensboro, NC; Greensburg, IN; Greenville, OH[c]; Greenville, SC[c]; Hanford, CA; Harrisburg, PA; Hartford, CT; Heber, UT[c]; Hickory, NC[c]; Hinesville, GA[a]; Holland, MI[b]; Honolulu, HI;

Hot Springs, AR[a]; Houma, LA[c]; Humboldt, TN; Huntington, IN[a]; Indianapolis, IN; Ithaca, NY[c]; Jackson, MS; Jackson, TN[c]; Jacksonville, FL; Jacksonville, NC[c]; Jasper, IN; Jesup, GA[c]; Johnson City, TN[a]; Kankakee, IL; Kapaa, HI; Keene, NH[a]; Kendallville, IN[c]; Kennewick, WA[c]; Key West, FL[c]; Kingsport, TN[a]; Kingston, NY[c]; Klamath Falls, OR; Knoxville, TN; La Crosse, WI[a]; La Follette, TN[a]; Lafayette, IN[b]; Lake Charles, LA; Lake Havasu City, AZ; Lakeland, FL; Lancaster, PA; Lancaster, SC[c]; Lansing, MI; Las Vegas, NV; Lawrenceburg, TN; Lebanon, PA; Lewisburg, TN[c]; Lewiston, ID[c]; Lexington Park, MD; Lexington, KY[c]; Lima, OH; Lincoln, NE[c]; Lincolnton, NC[c]; Little Rock, AR[c]; Logansport, IN[c]; Longview, WA[c]; Los Angeles, CA; Louisville-Jefferson County, KY[c]; Macon, GA; Madera, CA; Madison, IN; Madison, WI[c]; Manitowoc, WI[c]; Marion, IN[c]; Martin, TN; McMinnville, TN[c]; Medford, OR; Melbourne, FL; Memphis, TN[c]; Merced, CA; Miami-Fort Lauderdale, FL; Minneapolis-St Paul, MN; Mobile, AL; Modesto, CA; Morristown, TN[a]; Moses Lake, WA[c]; Mount Vernon, WA[c]; Muskegon, MI; Muskogee, OK[c]; Napa, CA; Naples, FL; Nashville, TN; New Haven, CT; New London, CT; New Philadelphia, OH[a]; New York, NY; Nogales, AZ; Norwalk, OH[b]; Oak Harbor, WA[a]; Ocala, FL; Ocean City, NJ; Ocean Pines, MD; Ogden, UT[c]; Ogdensburg, NY[a]; Oil City, PA; Okeechobee, FL[c]; Oklahoma City, OK[c]; Olean, NY[c]; Olympia, WA; Omaha, NE[c]; Orlando, FL; Oshkosh, WI[a]; Oxford, MS[c]; Pahrump, NV; Palatka, FL[a]; Palm Coast, FL; Panama City, FL; Paris, TN; Payson, AZ[c]; Pendleton, OR; Pensacola, FL; Peru, IN[a]; Philadelphia, PA; Phoenix Lake, CA[a]; Phoenix, AZ; Pittsburgh, PA[a]; Plymouth, IN[c]; Port St. Lucie, FL; Portland, OR; Portsmouth, OH[c]; Pottsville, PA; Poughkeepsie, NY; Prescott, AZ; Prineville, OR; Providence, RI; Provo, UT[c]; Pueblo, CO[c]; Punta Gorda, FL; Racine, WI[a]; Raleigh, NC; Reading, PA; Redding, CA[c]; Reno, NV; Richmond, IN[c]; Richmond, VA; Riverside, CA; Roanoke, VA[c]; Rochester, NY[c]; Rocky Mount, NC[c]; Rome, GA[b]; Roseburg, OR[c]; Sacramento, CA; Saginaw, MI; Salem, OR; Salinas, CA; Salisbury, MD; Salisbury, NC; Salt Lake City, UT[a]; San Diego, CA; San Francisco, CA; San Jose, CA; San Luis Obispo, CA; Sandusky, OH[c]; Sanford, NC[c]; Santa Cruz, CA; Santa Fe, NM; Santa Rosa, CA; Sarasota, FL; Savannah, GA[a]; Searcy, AR; Seattle, WA; Sebring, FL; Seneca Falls, NY; Sevierville, TN; Sheboygan, WI[c]; Shelbyville, TN[c]; Shelton, WA; Show Low, AZ; Sidney, OH; Sierra Vista, AZ; Spartanburg, SC[c]; Spokane, WA; Springfield, IL; Springfield, MA; Springfield, MO[c]; Springfield, OH[c]; St. George, UT[a]; St. Louis, MO[c]; St. Marys, GA; Stamford, CT; State College, PA[a]; Statesville, NC[a]; Stockton, CA; Syracuse, NY; Tallahassee, FL; Tampa, FL; Terre Haute, IN; Thomasville, NC[c]; Tiffin, OH; Toledo, OH; Torrington, CT; Trenton, NJ; Truckee, CA[a]; Tucson, AZ; Tullahoma, TN[c]; Tulsa, OK; Tuscaloosa, AL[a]; Union City, TN[a]; Urbana, OH[c]; Utica, NY[c]; Vallejo, CA; Ventura, CA; Vero Beach, FL[c]; Vineland, NJ; Virginia Beach, VA; Visalia, CA; Wabash, IN[c]; Walla Walla, WA[a]; Wapakoneta, OH[a]; Warner Robins, GA; Warsaw, IN[c]; Washington, DC; Waterloo, IA[c]; Watertown, NY; Watertown, WI; Wausau, WI[c]; Whitewater, WI[a]; Willimantic, CT; Wilmington, NC; Winchester, VA; Winston-Salem, NC[a]; Wooster, OH[c]; Worcester, MA; Yakima, WA[a]; York, PA; Youngstown, OH[c]; Yuba City, CA; Yuma, AZ.

References

Avery, R., Brevoort, K., & Canner, G. (2007). Opportunities and Issues In Using HMDA Data. *Journal of Real Estate Research*, *29*, 351–379.

Aznar, A., & Salvador, M. (2002). Selecting the rank of the cointegration space and the form of the intercept using an information criterion. *Econometric Theory*, *18*, 926–947.

Banerjee, A., Lumsdaine, R., & Stock, J. (1992). Recursive and sequential tests of the unit-root and trend-break hypotheses: Theory and international evidence. *Journal of Business and Economic Statistics*, *10*, 271–287.

Beracha, E., & Skiba, H. (2011). Momentum in residential real estate. *Journal of Real Estate Finance and Economics*, *43*, 299–320.

Berkovec, J., Chang, Y., & McManus, D.A. (2012). Alternative lending channels and the crisis in U.S. housing markets. *Real Estate Economics*, *40*, 8–31.

Case, K.E., & Shiller, R. (1989). The efficiency of the market for single family homes. *American Economic Review*, *79*, 125–37.

Case, K.E., & Shiller, R. (1990). Forecasting prices and excess returns in the housing market. *AREUEA Journal*, *18*, 253–273.

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

Delaney, C.J., Seward, J.A., & Smith, M.T. (1992). An empirical analysis of housing price appreciation in a market stratified by size and value of the housing stock. *Journal of Real Estate Research*, *7*, 195–205.

Dickey, D.A., & Fuller, W.A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, *74*, 427–431.

Dorsey, R.E., Hu, H., Mayer, W.J., & Wang, H. (2010). Hedonic versus repeat-sales housing price indexes for measuring the recent boom-bust cycle. *Journal of Housing Economics*, *19*, 87–105.

Driffill, J., & Sola, M. (1998). Intrinsic bubbles and regime-switching. *Journal of Monetary Economics*, *42*, 375–373.

Enders, W. (2010). *Applied Econometric Time Series*, 3rd ed. Hoboken, NJ: Wiley.

Engle, R.E., & Granger, C. (1987). Cointegration and error-correction: Representation, estimation and testing. *Econometrica*, *55*, 251–276.

Gallin, J. (2006). The long-run relationship between house prices and income: evidence from local housing markets. *Real Estate Economics*, *34*, 417–438.

Glaeser, E., Gyourko, J., & Saiz, A. (2008). Housing supply and housing bubbles. *Journal of Urban Economics*, *64*, 198–217.

Gonzalo, J., & Pitarakis, J.Y. (1998). Specification via model selection in vector error correction models. *Economics Letters*, *60*, 321–328.

Granger, C.W.J., & Newbold, P. (1974). Spurious regressions in econometrics. *Journal of Econometrics*, *2*, 111–120.

Gregory, A.W., & Hansen, B.E. (1996a). Residual-based tests for cointegration in models with regime shifts. *Journal of Econometrics*, *70*, 99–126.

Gregory, A.W., & Hansen, B.E. (1996b). Tests for cointegration in models with regime and trend shifts. *Oxford Bulletin of Economics and Statistics*, *58*, 555–560.

Guerrieri, V., Hartley, D., & Hurst, E. (2013). Endogenous gentrification and housing price dynamics. *Journal of Public Economics*, *100*, 45–60.

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, *12*, 231–254.

Johansen, S., & Juselius, K. (1990). Maximum likelihood estimation and inference on cointegration-with applications to the demand for money. *Oxford Bulletin of Economics and statistics*, *52*, 169–210.

Landvoigt, T., Piazzesi, M., & Schneider, M. (2012). The housing market(s) of San Diego (January 2012), NBER Working Paper No. w17723.

Lütkepohl, H., Saikkonen, P., & Trenkler, C. (2001). Maximum eigenvalue versus trace tests for the cointegrating rank of a VAR process. *The Econometrics Journal*, *4*, 287–310.

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

Mayer, C. (2011). Housing bubbles: a survey. *Annual Review of Economics*, *3*, 559–577.

Mayer, C., & Pence, K. (2009). *Subprime Mortgages: What, Where, and To Whom? In Housing markets and the economy: risk, regulation, and policy: essays in honor of Karl E. Case ed. by Edward Glaeser and John Quigley*. Cambridge, MA: Lincoln Land Institute.

Meese, R., & Wallace, N. (1994). Testing the present value relation for housing: Should I leave my house in San Francisco. *Journal of Urban Economics*, *35*, 245–266.

Mian, A., & Sufi, A. (2009). The consequences of mortgage credit expansion: Evidence from the U.S. mortgage default crisis. *Quarterly Journal of Economics*, *124*, 1449–1496.

Miao, H., Ramchander, S., & Simpson, M.W. (2011). Return and volatility transmission in U.S. housing markets. *Real Estate Economics*, *39*, 701–741.

Ortalo-Magné, F., & Rady, S. (2006). Housing market dynamics: On the contribution of income shock and credit constraints. *Review of Economic Studies*, *73*, 459–485.

Ortalo-Magné, F., & Rady, S. (1999). Boom in, bust out: Young households and the housing market cycle. *European Economic Review*, *43*, 755–766.

Pavlov, A.D., & Wachter, S. (2011). Subprime lending and real estate prices. *Real Estate Economics*, *39*, 1–17.

Perron, P. (1989). The great crash, the oil price shock, and the unit root hypothesis. *Econometrica*, *57*, 1361–1401.

Perron, P., & Vogelsang, T.J. (1992). Nonstationarity and level shifts with an application to purchasing power parity. *Journal of Business and Economic Statistics*, *10*, 301–320.

Pesaran, M.H. (2007). A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics*, *22*, 265–312.

Piazzesi, M., & Schneider, M. (2009). Momentum traders in the housing market: survey evidence and a search model. *American Economic Review P&P*, *99*, 406–411.

Pirinsky, C., & Pan, C. (2013). Social influence in the housing market. *Journal of Financial and Quantitative Analysis,* forthcoming.

Poterba, J. (1991). House price dynamics: the role of tax policy and demography. *Brookings Papers on Economic Activity*, *2*, 143–203.

Seward, J.A., Delaney, C.J., & Smith M.T. (1992). An empirical analysis of housing price appreciation in a market stratified by size and value of the housing stock. *The Journal of Real Estate Research*, *7*, 195–205.

Stein, J.C. (1995). Prices and trading volume in the housing market: A model with down-payment effects. *Quarterly Journal of Economics*, *110*, 379–406.

Smith, B.A., & Tesarek, W.P. (1991). House prices and regional real estate cycles: Market adjustments in Houston. *AREUEA Journal*, *19*, 396–416.

Stiglitz, J. (1990). Symposium on bubbles. *Journal of Economic Perspectives*, *4*, 13–18.

Swanson, N.R. (1998). Money and output viewed through a rolling window. *Journal of Monetary Economics*, *41*, 455–473.

Zivot, E., & Andrews, D.W.K. (1992). Further evidence to the great crash, the oil price shock, and the unit-root hypothesis. *Journal of Business and Economic Statistics*, *10*, 251–270.