NATIONAL BANK OF POLAND
WORKING PAPER
No. 126

Econometric regime shifts and the US subprime bubble

André K. Anundsen

Warsaw 2012

André K. Anundsen – Department of Economics, University of Oslo

The paper has been presented at the Central Bank Macroeconomic Modeling Workshop: Modeling Imbalances in Warsaw, 13.–14. September 2012. I am indebted to Thomas von Brasch, Christian H. Christensen, Roger Hammersland, Steinar Holden, Tord S. H. Krogh, Oleg Kitov, Ragnar Nymoen, Max Roser, Bernt Stigum and Genaro Succarat for comments and discussions on earlier drafts of this paper. I would also like to thank Luke Van Cleve who was very helpful in providing me with the tax-data from the FRB-US model. Many of the ideas as well as the work with this paper were produced during my stay at the Institute for Economic Modelling, Oxford Martin School, University of Oxford. Their hospitality and stimulating research environment is greatly acknowledged. Norges Banks Fond til Økonomisk forskning and Professor Wilhelm Keilhaus Minnefond deserves a special thanks for funding the stay.

Design:

Oliwka s.c.

Layout and print:

NBP Printshop

Published by:

National Bank of Poland
Education and Publishing Department
00-919 Warszawa, 11/21 Świętokrzyska Street
phone: +48 22 653 23 35, fax +48 22 653 13 21

ISSN 2084–624X

http://www.nbp.pl

Contents

Abstract

Using aggregate quarterly data for the period 1975q1–2010q4, I find that the US housing market changed from a stable regime with prices determined by fundamentals, to a highly unstable regime at the beginning of the previous decade. My results indicate that these imbalances could have been detected with the aid of real time econometric modeling and that they were caused by the sharp rise in subprime lending in the early to mid 2000s. These results are based on the detection of huge parameter non-constancies and a loss of equilibrium correction in two theory derived cointegrating relationships shown to be very stable for earlier periods. Controlling for the increased subprime exposure during this period, enables me to reestablish the pre-break relationships also for the full sample. This suggests that the US housing bubble was caused by the increased borrowing to a more risky segment of the market, which may have allowed for a latent frenzy behavior that previously was constrained by the lack of financing. With reference to Stiglitz's general conception of a bubble, I use the econometric results to construct two bubble indicators, which clearly demonstrate the transition to an unstable regime. Such indicators can be part of an early warning system and are shown to Granger cause a set of coincident indicators and financial (in)stability measures.

Keywords: *Cointegration; Regime Shifts; US Housing Bubble; Subprime lending; Bubble Indicator*

JEL classification: *C22; C32; C51; C52; G01; R21*

1 Introduction

Starting in the late 1990s, the US housing market witnessed a tremendous and unprecedented boom. Real four quarter growth rates were positive for ten consecutive years between 1997q2–2007q1. Much of this increase was subsequently reversed, and by 2011 real housing prices were back at their 2001 level. The repercussions of the housing collapse have been enormous and it was one of the causes of the recession that still impairs the global economy. There is a great need to understand US housing price formation and dynamics, in order to develop an "early warning system", to robustify the institutional framework and to prevent such events from repeating in the future.

Furthermore, housing prices play a key role in transmitting shocks to the real economy. Mortgage equity withdrawal (MEW) represents a channel in which gains from soaring housing prices may be capitalized through an increase in private consumption, see Aron et al. (2011) for emprical evidence of how it contributed to the US consumption boom of the early 2000s. Leamer (2007) argues that housing starts and the change in housing starts are the best leading business cycle indicators. The evolution of housing prices may be one important factor that influences the activity in the building and construction sector, i.e. by increasing the profitability of new construction projects through a Tobin-Q effect (Tobin, 1969). Quantitative models of housing prices are therefore highly relevant for a model based forecasting system of the US economy.

The surge in housing prices over the previous decade was parallelled by dramatic changes in banks' lending practices and securitization of questionable loans increased substantially. Before 2003, most mortgage originations were prime conforming loans, while the share of subprime and Alt-A mortgages increased steadily after this. At the same time, the share of subprime mortgages and Alt-A mortgages that were repacked and sold as private label asset backed securities (ABS)[1] rose from 45% of a total value of about 215 billion dollars in 2001 to 80% of 2 trillion(!) in 2005/2006 (Hendershott et al., 2010). The enormous increase in lending to more risky borrowers may have caused US housing prices to shoot away trajectories consistent with underlying fundamentals. Subprime borrowers typically have very high LTV ratios and given the non-recourse option in many US states, the downside risk of taking up a mortgage is very low for this group of borrowers. In addition, as pointed out by Hendershott et al. (2010), since the foreclosure process typically takes between 6 and 18 months, a household can live rent free over this period. In combination with very low interest rates, so called teaser rates, the first couple of years, there was not much to stop people from taking on excessive debt.

With this background, it is interesting to note that already for some time, there has been a discussion in the academic literature about the econometric modeling of US housing prices. Much of this debate has been concerned with the question of whether US housing prices are determined by so called fundamentals or not, where typical fundamentals are thought to be variables such as housing rents, income, the cost of financing or owning a property, along with a supply side measure. In addition to being an interesting and challenging econometric question, the role of fundamentals in determining housing prices may be relevant for the bubble debate. As my econometric results demonstrate,

[1]Loans satisfying the conforming loan limits of the GSEs are eligible for GSE securitization, while subprime and Alt-A mortgages are not. If resold, these loans are repacked into ABSs by private label securitizers. For more details on securitization, see the discussion in Hendershott et al. (2010)

the conflicting results in the literature may be explained by the transition from a stable to an unstable (bubble) regime in the early 2000s, and thus the diverging results may be ascribed to the different sample periods considered.

According to the definition in Stiglitz (1990), a bubble exists "if the reason why the price is high today is *only* because investors believe that the selling price will be high tomorrow – when "fundamental"factors do not seem to justify such a price"(Stiglitz, 1990, p.13). In this paper, I combine this definition with the modeling assumption that fundamental factors – if they exist – are non-stationary economic time series. Given this assumption, housing prices are determined by fundamentals *if and only if* there exists a cointegrating relationship between housing prices and these non-stationary economic variables. Since cointegration and equilibrium correction are non-trivial statistical properties, this approach opens for several insights that are relevant for discussing whether or not – in the Stiglitz (1990) sense – the evolution of US housing prices over the previous decade is best characterized as a bubble. First, if cointegration can be established over the full sample period as well as for different sub-samples, the bubble hypothesis is clearly rejected. At the other extreme, if no evidence for cointegration can be found, we cannot reject that hypothesis. That said, this may also be an indication that our information set does not include the relevant economic variables, i.e. the fundamental determinants of housing prices. The intermediate case may be even more relevant: If a cointegrating relationship can be established early in the sample but is lost subsequently, we may suspect a structural break. Even more interesting: If cointegration disappears a number of periods before the onset of a wider financial crisis, the results can be used to test if the transition from a stable market with equilibrium correction (no bubble) dynamics to an unstable market (a bubble) have predictive power for the wider crisis.

Several researchers have estimated equilibrium correction models for US housing prices, but without necessarily drawing the implications for whether or not there is – or has been – a bubble in the housing market. As the literature review in Section 2 will reveal, the results are diverging, which by itself calls for further research in an attempt to consolidate the evidence. My results, based on a system based as well as a single equation cointegration analysis, demonstrate that a structural break took place in US housing price formation in the early 2000s. While real housing prices are shown to follow fundamentals both in a price-to-rent framework and in an inverted demand equation prior to this, there is no evidence of such a relationship after the break. That said, I show that including a measure for the share of subprime loans relative to total loans explains much of this breakdown. This suggests that the econometric breakdown, interpreted as a bubble, was caused by the increased lending to subprime borrowers.

The results from the econometric models are used to construct two regime shift indicators that may be interpreted as "bubble indicators". Mikhed and Zemcik (2009b) constructed a similar indicator, where they defined a bubble as a situation where either housing prices are non-stationary and rents are stationary, or where both series are non-stationary and the price-to-rent ratio is non-stationary as well. Compared to that approach, the indicators presented in this paper has the advantage of being directly derived from an econometric model linking housing prices to economic fundamentals. I show that these indicators – which could have been calculated in real time – are able to detect the transition to a bubble regime early in the 2000s.

Finally, tests for Granger non-causality (GNC) show that these indicators Granger cause delinquency rates and non-performing loans, the unemployment rate and industrial production. My results therefore suggest that the expansion of subprime lending caused the housing bubble, which again was an important factor leading up to the wider financial crisis.

As already mentioned, the paper starts with a review of the existing literature on the econometric modeling of US housing prices. The literature review is followed by a discussion of how a traditional life-cycle model for housing may be interpreted within an equilibrium correction framework. In Section 4, I turn to a description of the data and their temporal properties. The succeeding section, Section 5, documents a structural break in US housing price formation in the early 2000s. Including a measure for the number of subprime loans as a share of total loans enables me to model this structural break in Section 6. The "bubble indicators"are presented in Section 7. In the same section, I report results from tests for GNC between the "bubble indicators"and a set of financial (in)stability measures and coincident indicators. The paper completes with some concluding remarks.

2 Cointegration or not: An unsettled debate

There is no consensus in the literature on the question of whether US housing prices and fundamentals are cointegrated. Some papers have found evidence of cointegration, while others have reached the opposite conclusion. In broad terms, the literature can be divided into two groups: Those who consider local differences and large panels and those who look at aggregate time series data. Given the level of aggregation, there are two theoretical approaches that are commonly considered when the relationship between housing prices and fundamentals is studied. The first takes as a starting point an inverted demand equation linking housing prices to income, a measure of the cost of housing and a supply measure. The second approach looks at the relationship between housing prices and rents. The present study is an aggregate study of both the inverted demand approach and the price-to-rent approach, but a brief – though non-exhaustive – summary of the findings from both aggregate and regional analyses seems relevant. Table 1 gives a summary of the main results as well as the sample periods used in the papers reviewed in this section.

Meen (2002) adopts a single equation approach to estimate the fundamental determinants of real housing prices at the national level. Based on a sample covering the period 1981q3–1998q2, he reports evidence of cointegration between real housing prices, real personal disposable income, real net financial wealth, the real interest rate and the housing stock. The author demonstrates that the estimated elasticities are sensitive to the inclusion of the housing stock variable. In fact, the income elasticity turns negative if the housing stock is omitted from the cointegrating relation.

Based on the Johansen (1988) approach, McCarthy and Peach (2004) estimate a stock-flow model for the US housing market. They find the long run determinants of housing prices to be the stock of dwellings, non-durables and services consumption – which is used as a proxy for permanent income – as well as the user cost of housing. The variables are all measured in real terms. McCarthy and Peach (2004) conclude that there is no evidence of a bubble in the US housing market when the model is estimated over the sample 1981q1–2003q3, but conclude that housing prices have risen as a result of higher incomes and low interest rates.

An early contribution to the panel data literature is Abraham and Hendershott (1996), who estimate an equilibrium correction type of model for 30 Main Statistical Areas (MSAs) using annual data for the 1977–1992 period . They find that housing prices depend on construction costs, disposable income and the real interest rate in the long run, which supports the main conclusions of the aforementioned papers.

Though several authors have found that US housing prices are determined by fundamentals, Gallin (2006) argues that US housing prices cannot be modeled in an equilibrium correction framework. First, he looks at national housing price data over the sample 1975q1–2002q2 using a two-step Engle and Granger (1987) procedure. Then, the author considers a panel of annual data covering 95 cities over the period 1978–2002. In neither case does he find evidence of cointegration. The findings of Gallin (2006) contradicts the results of Malpezzi (1999) who considered a similar panel and found evidence of cointegration on the sample 1979–1996. The same author (see Gallin (2008)) looks at the relationship between housing prices, rents and the direct user cost of housing for a sample covering the period 1970q1-2005q4. Estimating a conditional equilibrium correction model, he shows that there is no evidence of cointegration between housing prices

Table 1: Results from previous studies regarding cointegration between housing prices and fundamentals

Author(s)	Linear	No evidence	Non-linear	Regional	National	Sample
Abraham and Hendershott (1996)	x			x		1977–1992
Malpezzi (1999)	x			x		1979–1996
Meen (2002)	x				x	1981q3–1998q2
McCarthy and Peach (2004)	x				x	1981q1–2003q3
Gallin (2006)		x		x	x	1975q1–2002q2/1978–2002
Gallin (2008)		x		x		1970q1–2005q4
Mikhed and Zemcik (2009a)†		x		x	x	1980q2–2008q2/1978–2007
Mikhed and Zemcik (2009b)		x		x		1978h1–2006h2
Zhou (2010)			x	x	x	1978q1–2007q4
Clark and Coggin (2011)		x			x	1975q1–2005q2
Duca et al. (2011a)	x				x	1981q1–2007q2
Duca et al. (2011b)	x				x	1981q2–2009q3

Notes: The table gives a summary of the main conclusions in the literature on whether US housing prices and fundamentals are cointegrated or not. Strictly speaking, Abraham and Hendershott (1996) do not test for cointegration, but the model they derive may be interpreted within an equilibrium correction framework.

† For samples ending in 2006q4 and 2008q2, Mikhed and Zemcik (2009a) find evidence of cointegration between housing prices and construction wages, while housing prices and fundamentals are not found to be cointegrated for samples ending before this.

2

and these fundamentals for the full sample.

The main conclusions of Gallin (2006) are supported by Clark and Coggin (2011) and Mikhed and Zemcik (2009a), who both study the long run determinants of real housing prices at the national and at the regional level. Mikhed and Zemcik do however find that a cointegrating relationship may be established if the sample ends in 2006 or later, while no such relationship exists in earlier periods.

Mikhed and Zemcik (2009b) use semi-annual data on housing prices and rents for 23 MSAs over the period 1978-2006 and find similar results as Gallin (2008). Considering the full sample, they do not find evidence of cointegration between housing prices and rents and conclude that there is a bubble. The authors go further and construct a "bubble indicator" based on the relationship between housing prices and rents using 10-year rolling windows. It is assumed that the indicator takes the value one if prices are I(1) and rents are I(0) over a given time interval, while it is equal to zero for stationary housing prices and either stationary or non-stationary rents. If both housing prices and rents are I(1), the value of the indicator is equal to the p-value from the panel unit root test of Pesaran (2007) on the price-to-rent ratio. In other words, they implicitly assume that – if there is cointegration – the CI-vector is $(1, -1)$ between prices and rents. For most of the rolling windows considered, this indicator provides no evidence of cointegration and takes a value well above 0.20, which strictly speaking should be interpreted as a bubble using their methodology. An alternative approach to constructing such a "bubble indicator" will be discussed later in this paper.

Contrary to the many recent papers finding no evidence of a cointegrating relationship between housing prices and fundamentals, Duca et al. (2011a,b) argue that the reason why most models of US housing prices break down in the 2000s is the exclusion of a measure of exogenous changes in credit availability. In Duca et al. (2011b), it is shown that adding a measure of the loan-to-value (LTV) ratio of first time home buyers in a model linking housing prices to income, the housing stock and the user cost outperform non-LTV models judged by interpretation of the estimated elasticities as well as the numerical size of the equilibrium adjustment coefficient. Similar conclusions are reached in Duca et al. (2011a), where the relationship between the rent-to-price ratio and the user cost is considered.

Finally, Zhou (2010) uses data for the period between 1978q1 and 2007q4 to test for linear, and if that is not found, non-linear cointegration between housing prices, income, the mortgage interest rate and construction costs. To determine whether the variables in the information set are linearly cointegrated, both the Engle and Granger (1987) and Johansen (1988) procedures are employed. Only for the case of Cleveland does the author find evidence of linear cointegration, which is also the case when the Johansen procedure is considered. For the country and six cities, he finds evidence of non-linear cointegration using the two-step procedure of Granger and Hallman (1991) and Granger (1991), which transforms the non-linear relationship to a linear one and then cointegration tests for the linear case may be applied.

3 A conceptual framework for equilibrium correcting housing prices

As mentioned in the literature review, there are generally two different theoretical approaches that are considered when looking at the relationship between housing prices and fundamentals; the inverted demand approach and the price-to-rent approach. To be clear about the origin of these relationships, I will briefly discuss their relation to the life-cycle model of housing, see e.g. Meen (2001, 2002) or Muellbauer and Murphy (1997).

Based on the life-cycle model, the following condition must be satisfied in equilibrium:

$$\frac{U_H}{U_C} = PH\left[(1-\tau^y)(i+\tau^p) - \pi + \delta - \frac{\dot{PH}}{PH}\right] \tag{1}$$

The condition in (1) follows from the representative consumer's maximization problem, where $\frac{U_H}{U_C}$ is the marginal rate of substitution between housing, H, and a composite consumption good, C. The condition states that that the consumers marginal willingness to pay for housing services in terms of other consumption goods should in optimum be equal to the cost in terms of forgone consumption. The term in brackets is usually labeled the real user cost of housing, which can be split into three different components. The first is the sum of the nominal interest rate, i, and the property tax, τ^p, less tax deductions at a rate τ^y, and corrected for an increase in the overall price level, π. The second component is the housing depreciation rate, δ. The final component is the expected real housing price inflation, $\frac{\dot{PH}}{PH}$, with PH denoting real housing prices. The sum of the first two components is often referred to as the direct user cost of housing, which will be my operational measure of the user cost in the econometric analysis.[2]

Market efficiency requires the following no-arbitrage condition to be satisfied:

$$Q = PH\left[(1-\tau^y)(i+\tau^p) - \pi + \delta - \frac{\dot{PH}}{PH}\right] \tag{2}$$

The expression in (2) states that the user cost of housing should in equilibrium be equal to the real imputed rent on housing services, Q. That is, the user cost of a given dwelling should be equal to what it would have costed to rent a dwelling of similar quality (the value of living in the property). Rearranging equation (2) slightly, gives the following equilibrium relationship:

$$\frac{PH}{Q} = \frac{1}{(1-\tau^y)(i+\tau^p) - \pi + \delta - \frac{\dot{PH}}{PH}} \tag{3}$$

[2]It should be noted that I have experimented with alternative measures of the user cost, where I also included expected capital gains as a moving average of the housing price growth over previous years or simply as the last period four quarter growth (static expectations). What I found was that the results were sensitive to the number of lags I included in the moving average process. For that reason, and because I have no a priori reason to assume a given structure on the moving average process, I decided to use the real direct user cost instead. Note that this implies that expectations about future price changes are captured by the lags included in the econometric models.

The real imputed rent is unobservable, and two approximations are custom in the empirical literature. The first approximation is to assume that the real imputed rent can be proxied by the observed rent, i.e. the unobservable Q is replaced by an observable R in equation (3). Since the user cost takes negative values over the sample period considered in this paper, I shall consider (3) on a semi-logarithmic form in the empirical analysis. The expression based on the price-to-rent approach therefore reads:

$$ph = \gamma_r r + \gamma_{UC} UC \tag{4}$$

where lower case letters indicate that the variables are measured on a log scale. In contrast to Gallin (2006), Mikhed and Zemcik (2009b) and Duca et al. (2011a), I do not impose a unitary coefficient between housing prices and rents from the outset. One reason for this is due to measurement issues, since the observable R is only an approximation of the theoretical Q. Furthermore, the implied unitary elasticity between housing prices and rents is a testable restriction. Finally, it is not clear a priori whether rents can be considered weakly exogenous with respect to the long run parameters, which is another testable restriction. That said, using the price-to-rent ratio instead (imposing $\gamma_r = 1$ in equation (4) from the outset) does not affect the results in this paper. The equilibrium correction representation of the price-to-rent model can be expressed in the following way:

$$\begin{aligned}\Delta ph_t = {} & \mu + \alpha_{ph} \left(ph - \gamma_r r - \gamma_{UC} UC\right)_{t-1} \\ & + \sum_{i=1}^{p} \rho_{ph,i} \Delta ph_{t-i} + \sum_{i=0}^{p} \rho_{r,i} \Delta r_{t-i} + \sum_{i=0}^{p} \rho_{UC,i} \Delta UC_{t-i} + \varepsilon_t\end{aligned} \tag{5}$$

The second approach followed in the literature is to assume that the imputed rent is a function of variables such as income, Y, and the housing stock, in which case we have:[3]

$$Q = g(Y, H) \tag{6}$$

Inserting for equation (6) in equation (3), a log-linear approximation becomes:

$$ph = \gamma_y y + \gamma_h h + \gamma_{UC} UC \tag{7}$$

where lower-case letters again indicate that the variables are measured in logs. The transformations and approximations imply that equation (7) may not be very different from the demand part of a reduced form demand and supply model (see Meen (2002) for more discussion), but it helps for interpretation to be clear about the origin of this equation in the life-cycle model.

Since the housing stock evolves slowly, it is assumed to be fixed in the short run, i.e. it is assumed that the short run supply schedule is vertical. In the short run, it is

[3] I have also tested for population and financial wealth effects, but none of these variables entered significantly in an inverted demand equation.

therefore assumed that prices clear the market, which again implies that short run price movements reflect changes in demand. The equilibrium correction representation of (7) can be formulated in the following way:

$$\Delta ph_t = \mu + \alpha_{ph}\left(ph - \gamma_y y - \gamma_h h - \gamma_{UC} UC\right)_{t-1} + \sum_{i=1}^{p} \rho_{ph,i} \Delta ph_{t-i} + \sum_{i=0}^{p} \rho_{y,i} \Delta y_{t-i} + \sum_{i=0}^{p} \rho_{UC,i} \Delta UC_{t-i} + \varepsilon_t \tag{8}$$

Whether the underlying theories represented by equation (4) and equation (7) are sufficient to explain US housing price formation may be judged by the significance of the estimated long run elasticities and – in particular – the significance and numerical size of the equilibrium correction coefficient, α_{ph}, in equation (5) and equation (8).

From a theoretical point of view, we expect γ_r in equation (5) to be positive. In equation (8), we expect γ_y to be positive and γ_h to be negative. In both (5) and (8), we expect γ_{UC} to be negative. Further, we expect α_{ph} to be negative and significantly different from zero if housing prices are determined by fundamentals. In the case of a bubble, one would not expect α_{ph} to be significantly different from zero – or at least that it would change markedly towards zero relative to the value it takes during a period of equilibrium correction (no bubble) dynamics. If that is the case, deviations from an estimated equilibrium would be restored very slowly – or not at all. Thus, with reference to Stiglitz definition of a bubble, I will think of a bubble as a situation in which housing prices and fundamentals are not cointegrated.

4

4 Data description and temporal properties

4.1 Data description

As the operational measure of housing prices, I use the housing price index of the Federal Housing Finance Agency (FHFA), which is available from 1975q1.[4] To measure housing rents, I use the rent component of CPI as reported by the Bureau of Labor Statistics (BLS).

My operationalization of the user cost uses the effective interest rate measured as a weighted average of the effective fixed and flexible mortgage interest rates. These data are based on the Monthly Interest Rate Survey Data as reported by FHFA. The weights are determined by the origination shares of the different mortgages. This detail is important in order to get a precise measure of the financing cost at an aggregate level, since – as Figure 1 demonstrates – the share of fixed and flexible rate mortgages have changed quite substantially over the time period I consider.[5]

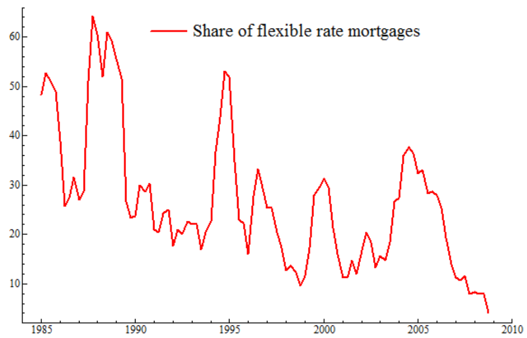


Figure 1: The share of mortgages that have flexible rates in the US, 1985q1–2008q4

The sum of the property tax rate and the interest rate is corrected for tax deductions using the marginal personal income tax rate (at twice median family income). Both tax rates are from the database of the FRB-US model. The final component in the direct user cost is the depreciation rate, which is from the National Income and Product Accounts.[6] The real direct user cost is constructed by subtracting the inflation rate measured by CPI for all items.

The income series is the disposable personal income series collected from the St. Louis Fed's database FRED. The housing stock series is from Moody's analytics and is interpolated from the annual data published by the Census Bureau.[7]

[4]This housing price index is calculated according to the weighted repeat sales method of Case and Shiller (1987) and is the longest time series available for US housing prices. For further documentation on how the index is calculated, the reader is referred to Calhoun (1996).

[5]The weighted interest rates are available all the way back to 1973, while I was able to track down data on the shares of the different mortgages from 1985 to 2008 only.

[6]I was only able to collect data for the depreciation rate until 2007q3. After this, I have assumed that the depreciation rate remains unchanged.

[7]In an earlier version of this paper, I constructed a quarterly series using annual housing stock data from Census Bureau that I was able to collect from 1980. Together with both annual and quarterly data on housing completions, I then used a law of motion of capital motion equation to calibrate the implied

All data are seasonally *unadjusted* except the disposable income and housing stock series, which were only available seasonally *adjusted*. In the econometric analysis, I used the unadjusted series and included seasonal dummies in the usual way. Housing prices, rents and disposable income are measured in real terms, where the nominal to real transformations have been achieved by deflating with the CPI for all urban consumers, less shelter.

To control for the interest rate uncertainty during the inflation period of the late 1970s, I include a dummy, MT, that is equal to one between 1975q1 and 1982q3. Without this dummy, the user cost effect is estimated less precisely. In fact, it is insignificant in some inverted demand equations, which does not seem reasonable from a theoretical point of view. That said, this adjustment does not materially affect the other coefficients and helps to get more precise estimates of the user cost effect. Duca et al. (2011a,b) used a similar dummy for a sample starting in 1979q4 to control for the monetary targeting period between 1979q4 and 1982q3. Finally, I follow Duca et al. (2011a,b) and include a dummy for the Tax Reform Act of 1997, which is not properly accounted for by the user cost (see Duca et al. (2011a,b) and Cunningham and Engelhardt (2008) for more discussion). This dummy, CGT, is set equal to one from 1997q3.

4.2 Temporal properties

It is well known that standard inference theory in general ceases to be valid if the data are non-stationary (see Granger and Newbold (1974)). One solution is to use cointegration methods, which takes as a starting point that even though economic data display individual stochastic non-stationarities, there may exist linear (or even non-linear) combinations that are stationary.

Because of this, I started by testing for unit roots using both the Augmented Dickey-Fuller (ADF) test (Dickey and Fuller (1979) and Dickey and Fuller (1981)) and the Phillips-Perron (PP) test (Phillips (1987) and Phillips and Perron (1988)). The results from these tests are conveniently summarized in Table B.1 in Appendix B, while Figure A.1 and Figure A.2 in Appendix A display the series in levels and first differences.

Based on the unit root tests, it is clear that all series are non-stationary. With the exception of the housing stock, which according to the tests has an I(2) component, all series are found to be integrated of first order. That said, if I include six lags in the ADF-regression initially, where the sixth lag is found significant, the test suggest that also this series is integrated of first order. With this small caveat in mind, I continue the analysis under the assumption that all series are at most integrated of order one.

scrapping rate. This gave me a series that is similar to the series from Moodys, but the latter has the advantage of covering 5 more years (20 observations) of data. That said, similar conclusions were reached in that version of the paper.

5 The recent regime shift in US housing price formation

5.1 Methodological approach

In this section, I present the results obtained when the two theoretical models are confronted with the data. To test for cointegration, I have used the system based approach due to Johansen (1988, 1991, 1995). As a robustness check, I have also considered a single equation test. The Johansen method relies on a reparameterization of a vector autoregressive (VAR) model. In the case where we consider a p'th order VAR, the vector equilibrium correction model (VECM) – which forms the basis for inference in the cointegrated VAR (CVAR) – takes the following form.

$$\Delta \boldsymbol{y}_t = \boldsymbol{\Pi}\boldsymbol{y}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\Gamma}_i \Delta \boldsymbol{y}_{t-i} + \boldsymbol{\Phi}\boldsymbol{D}_t + \boldsymbol{\varepsilon}_t \tag{9}$$

where $\boldsymbol{y}_t$ is a $k \times 1$ vector of endogenous variables, $\boldsymbol{D}_t$ is a vector of deterministic terms (including a constant) and $\boldsymbol{\varepsilon}_t \sim IIN(0, \boldsymbol{\Omega})$. With reference to a VAR model, we have that $\boldsymbol{\Pi} = \sum_{i=1}^{p} \boldsymbol{\Pi}_i - \boldsymbol{I}$ and $\boldsymbol{\Gamma}_i = -\sum_{j=i+1}^{p} \boldsymbol{\Pi}_j$, with $\boldsymbol{\Pi}_i$ referring to the coefficient matrix attached to lag number i of the vector $\boldsymbol{y}_t$.

A test for cointegration is then to test for the number of independent linear combinations of the variables in $\boldsymbol{y}_t$ that are stationary, which amounts to testing the rank, r, of the matrix $\boldsymbol{\Pi}$. If $\boldsymbol{\Pi}$ has reduced rank, it can be decomposed in the following way $\boldsymbol{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}'$, where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are matrices of dimension $k \times r$ representing the loading factors and the long run coefficients, respectively.[8] I follow the standard practice and let a deterministic trend enter the space spanned by the matrix $\boldsymbol{\alpha}$.

When considering the price-to-rent based model, the vector $\boldsymbol{y}_t$ is a 3×1 vector containing real housing prices, real rents and the real direct user cost. The inverted demand equation, is tested based on a slightly modified version of equation (9), since I condition on the housing stock in the cointegration space. To illustrate what this implies in terms of the VECM representation, it is convenient to partition $\boldsymbol{y}_t$ into a vector of endogenous variables, $\boldsymbol{x}_t$, and a vector of exogenous variables, $\boldsymbol{z}_t$. The VECM can then be written in the following way

$$\Delta \boldsymbol{x}_t = \boldsymbol{\Pi}\boldsymbol{y}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\Gamma}_{\boldsymbol{x},i} \Delta \boldsymbol{x}_{t-i} + \sum_{i=0}^{p-1} \boldsymbol{\Gamma}_{\boldsymbol{z},i} \Delta \boldsymbol{z}_{t-i} + \boldsymbol{\Phi}\boldsymbol{D}_t + \boldsymbol{\varepsilon}_t \tag{10}$$

where $\boldsymbol{y}_t = (\boldsymbol{x}_t', \boldsymbol{z}_t')'$. Thus, when I consider the inverted demand equation, the vector $\boldsymbol{x}_t$ will contain real housing prices, real disposable income and the real user cost, while $\boldsymbol{z}_t$ is a scalar containing the housing stock only. Since the housing stock is assumed constant in the short run, I impose the additional restriction that $\boldsymbol{\Gamma}_{\boldsymbol{z},i} = \Gamma_{h,i} = 0\ \forall i$.

[8]An additional assumption is needed to rule out the possibility of I(2). More precisely, with reference to the second differenced VAR, we can write $\boldsymbol{\alpha}_\perp' \boldsymbol{\Gamma}\boldsymbol{\beta}_\perp = \boldsymbol{\xi}\boldsymbol{\eta}'$, where $\boldsymbol{\Gamma} = \sum_{i=1}^{p-1} \boldsymbol{\Gamma}_i - \boldsymbol{I}$, while $\boldsymbol{\alpha}_\perp$ and $\boldsymbol{\beta}_\perp$ are the orthogonal complements of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ (i.e $\boldsymbol{\alpha}_\perp\boldsymbol{\alpha}' = \boldsymbol{\beta}_\perp\boldsymbol{\beta}' = 0$) with dimension $(k-r) \times s$. In general, if $s < (k-r)$ then there are $k-r-s$ I(2) trends in the data, so under the assumption of no I(2) trends, we must have that $s = k - r$, i.e. there are $k-r$ common stochastic I(1) trends.

5.2 Results from the system based approach

Given the conflicting results in the literature, I started by exploring the stability of the two theoretical relationships for housing price determination described by equation (4) and equation (7). Relying on the statistical framework described in the previous section, I first estimated the VECM representation (equation (9) and (10), respectively) of the two models for a sample ending in 1995q4. Then, I sequentially added four new observations until both models were estimated over the full sample period, 1975q1–2010q4.

I started with a VAR of fifth order, then I tested down the lag length using a series of Wald F-tests. In both models and for all end points, the appropriate lag length was found to be five.[9] After this, I tested for cointegration using the trace test of Johansen (1988). Finally, I tested the joint restriction of excluding the trend from the cointegration space and whether weak exogeneity of the other variables in the VAR could be supported. More precisely, when looking at the long run relationship between housing prices, rents and the user cost (see equation (4) and equation (9)), I tested whether rents and the user cost could be considered weakly exogenous with respect to the long run coefficients, while the same test was done with respect to disposable income and the user cost when I tested the inverted demand equation (confer equation (7) and equation (10)).

In Table 2 and Table 3, I have summarized the main results from these recursive theory-data confrontations. Column 1-2 report the estimation end point and the rank of the $\mathbf{\Pi}$-matrix. Conditional on a non-zero rank[10], the next column reports the p-value from the likelihood ratio test for overidentifying restrictions. The final three (four) columns report the estimated adjustment coefficient (α_{ph}) and the long run elasticities, with standard errors shown below the point estimates.

There are several noteworthy results in Table 2 and Table 3. Most clear are the results from the price-to-rent approach, but they are confirmed by the results from the inverted demand approach.

Looking first at the results from the price-to-rent approach (Table 2), it is seen that there is strong evidence for one cointegrating vector (rank = 1) until 2001. Also, the overidentifying restrictions are accepted and the estimated coefficients do not change notably as the estimation end point is extended gradually from 1995q4 to 2000q4. However, when 2001q4 is included in the sample, that relationship can no longer be supported (rank = 0). At the end of the sample, there are evidence of a return of equilibrium correction (rank = 1). That said, the adjustment coefficient is much lower and that the other coefficient estimates have changed substantially relative to their pre-break values.

An inspection of the results from the inverted demand approach (see Table 3), gives a similar impression. Though the rank of $\mathbf{\Pi}$ does not drop to zero, it is clearly seen that the equilibrium correction coefficient is reduced substantially when the sample is extended to cover the early 2000s and that it changes towards zero around 2002/2003. In addition, the estimated coefficients change markedly and the overidentifying restrictions are no longer supported.

It is worth noting that the estimated long run elasticities in the inverted demand

[9]With four lags used to construct the inflation rate used in the user cost expression and five lags in the econometric model, the full effective sample covers the period 1977q2–2010q4.

[10]I have used small sample adjusted test statistics, and – for the inverted demand approach – I have used consistent critical values from Table 13 in Doornik (2003) for the case of one exogenous variable. A 5% significance level was used as a cut-off.

Table 2: Results from recursive CVAR analysis using the price-to-rent approach (confer equation (4) and equation (9)), 1977q2–T

End point (T)	Rank	Test for restrictions	α_{ph}	β_r	β_{UC}
1995q4	1	0.1720	−0.232	0.998	−1.319
			0.043	0.155	0.379
1996q4	1	0.1721	−0.233	1.064	−1.307
			0.042	0.150	0.374
1997q4	1	0.3590	−0.227	1.070	−1.367
			0.041	0.153	0.379
1998q4	1	0.2881	−0.229	1.062	−1.334
			0.040	0.148	0.369
1999q4	1	0.1346	−0.225	1.075	−1.249
			0.039	0.145	0.365
2000q4	1	0.2576	−0.199	1.152	−1.176
			0.037	0.164	0.409
2001q4	1	*	*	*	*
2002q4	0	*	*	*	*
2003q4	0	*	*	*	*
2004q4	0	*	*	*	*
2005q4	0	*	*	*	*
2006q4	0	*	*	*	*
2007q4	0	*	*	*	*
2008q4	0	*	*	*	*
2009q4	0	*	*	*	*
2010q4	1	0.3175	−0.060	2.184	0.059
			0.012	0.348	1.270

Notes: This table reports a summary of the main results when the system based approach of Johansen (1988) is implemented by sequentially adding four new observations to the sample. The first end point is 1995q4, while the last is 2010q4. The endogenous variables in the system are real housing prices, ph, real rents, r and the real direct user cost, UC. A deterministic trend is restricted to enter the cointegration space, while a constant, three centered seasonal dummies and the MT and CGT dummies enter unrestrictedly.

Table 3: Results from recursive CVAR analysis based on inverted demand approach (confer equation (7) and equation (10)), 1977q2–T

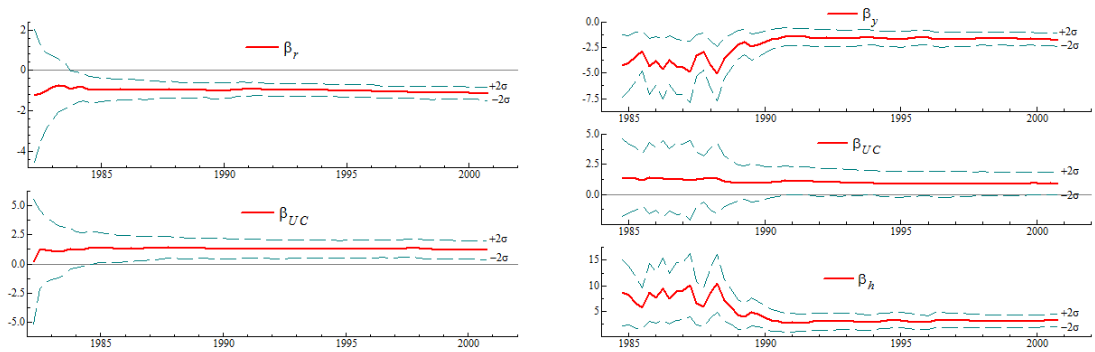
End point (T)	Rank	Restrictions supported	α_{ph}	β_y	β_{UC}	β_h
1995q4	1	0.2602	−0.187 0.033	1.500 0.344	−0.893 0.496	−2.794 0.693
1996q4	1	0.3914	−0.175 0.030	1.730 0.356	−0.831 0.532	−3.301 0.705
1997q4	1	0.3664	−0.181 0.031	1.693 0.343	−0.855 0.515	−3.174 0.676
1998q4	1	0.3012	−0.184 0.031	1.663 0.329	−0.841 0.494	−3.119 0.648
1999q4	1	0.4507	−0.186 0.031	1.580 0.306	−0.956 0.462	−2.957 0.605
2000q4	1	0.4639	−0.174 0.029	1.762 0.312	−0.903 0.485	−3.307 0.619
2001q4	1	0.0399	−0.151 0.028	1.950 0.370	−0.893 0.76	−3.695 0.735
2002q4	1	0.0035	−0.106 0.021	2.549 0.538	−0.743 0.837	−4.865 1.069
2003q4	1	0.0002	−0.057 0.014	4.416 1.056	0.312 1.617	−8.523 2.101
2004q4	1	0.0000	−0.026 0.008	8.286 2.363	0.904 3.556	−16.161 4.692
2005q4	1	0.0000	−0.006 0.002	30.104 10.819	7.121 16.876	−60.078 21.472
2006q4	1	0.0000	−0.011 0.003	17.540 5.475	4.634 8.729	−34.728 10.823
2007q4	1	0.0000	−0.029 0.007	5.836 1.967	−0.785 3.097	−11.573 3.919
2008q4	1	0.0000	−0.035 0.008	5.245 1.708	−0.496 2.655	−10.438 3.440
2009q4	1	0.0000	−0.033 0.007	5.815 1.746	0.027 2.725	−11.628 3.550
2010q2	1	0.0000	−0.033 0.008	5.505 1.758	0.758 2.635	−10.865 3.559

Notes: This table reports a summary of the main results when the system based approach of Johansen (1988) is implemented by sequentially adding four new observations to the sample. The first end point is 1995q4, while the last is 2010q2. The endogenous variables in the system are real housing prices, ph, real disposable income, y and the real direct user cost, UC. A deterministic trend and the housing stock, h, are restricted to enter the cointegration space. A constant, three centered seasonal dummies and the MT and CGT dummies enter unrestrictedly. Consistent critical values for one exogenous variable are tabulated in Doornik (2003).

model are interpretable and in accordance with the international literature when the estimation end point is set to 2000q4 or earlier, see Girouard et al. (2006) for an overview of results from international studies. I also find that the coefficient on housing rents in the price-to-rent model is close to one and that it is weakly exogenous, which justifies the a priori restriction made by Gallin (2006), Mikhed and Zemcik (2009b) and Duca et al. (2011a). Figure 2 displays the recursively estimated coefficients from both models when the end point is set to 2000q4.

From Table B.2 and Table B.3 in Appendix B, it can be seen that the models are mostly well specified over the stable period. That said, there are some minor evidence of autocorrelation in the inverted demand model. I find that excluding the trend from the model (a restriction that is supported), removes this autocorrelation and the model is well specified over the entire stable period in that case (see Table B.4 in Appendix B).

With reference to my earlier claim that the two dummies included in the analysis mainly helps to more sharply estimate the effect of the user cost, it is reassuring to take a look at the results in Table B.5 and B.6 of Appendix B, where I have redone the recursive analysis without the two dummies in the models. It is clear that excluding these dummies mainly affect the user cost estimates, as all other coefficients and findings are largely unaltered.



(a) (b)

Figure 2: Panel a) Recursively estimated coefficients for the rent and the user cost in the price-to-rent model, 1984q1–2000q4 Panel b) Recursively estimated coefficients for disposable income, the user cost and the housing stock from the inverted demand approach, 1984q1–2000q4

The results from the system based cointegration analysis strongly suggest a breakdown of both the price-to-rent model and the inverted demand model in the early 2000s. In the next section, I will shed some more light on this breakdown resorting to a single equation analysis.

5.3 Results from a single equation cointegration analysis

An alternative approach to testing for cointegration is to estimate equation (5) and equation (8) directly, and then test the significance of the adjustment coefficient. This follows from the Engle-Granger representation theorem (see Engle and Granger (1987)) that states that equilibrium correction implies cointegration and *vice versa*. Ordinary critical values for the t-distribution can however not be used under the null of no cointegration as the distribution of α_{ph} is non-standard and skewed to the left.[11]

Since the theoretical models tell us little about the dynamics of housing prices, I have estimated equation (5) and equation (8) following a *general-to-specific* (Gets) procedure. I used the automatic model selection algorithm *Autometrics* implemented within PcGive (see Doornik (2009) and Doornik and Hendry (2009)).[12] The lagged levels were restricted to enter the final specification, which ensures theory consistency.

Table 4 and Table 5 report the long run elasticities and the adjustment coefficients along with their finite sample p-values, when I sequentially add four more observations to the sample and use Autometrics to select the relevant variables.

Table 4: Recursive coefficients for price-to-rent model using a single equation approach (confer equation (4) and equation (5)), 1977q2–T

End Point (T)	β_r	β_{UC}	α_{ph}	p-value
1995q4	1.164	-0.816	-0.224	0.0007
1996q4	1.177	-0.796	-0.228	0.0004
1997q4	1.206	-0.816	-0.219	0.0006
1998q4	1.200	-0.819	-0.223	0.0003
1999q4	1.202	-0.819	-0.222	0.0002
2000q4	1.266	-0.828	-0.203	0.0005
2001q4	1.409	-1.001	-0.161	0.0027
2002q4	1.630	-0.909	-0.130	0.0050
2003q4	1.900	-0.726	-0.105	0.0379
2004q4	3.528	-0.488	-0.048	0.5892
2005q4	4.072	-1.456	-0.022	0.8479
2006q4	4.764	-0.733	-0.026	0.6603
2007q4	2.175	-1.306	-0.041	0.1285
2008q4	1.919	1.004	-0.046	0.0607
2009q4	1.935	-1.470	-0.056	0.0131
2010q4	2.095	-0.922	-0.061	0.0022

Notes: This table reports the estimated cointegrating vector along with the loading factor and the corresponding p-value when the price-to-rent model is estimated using a single equation approach.

It is reassuring that these results mimic those I find in the system based analysis and the results strongly suggests that the two models for US housing price formation broke down early in the previous decade. The estimated coefficients for the stable period are also close to those I find from the system based analysis. Furthermore, the same results regarding equilibrium correction are obtained, though this alternative approach seem to support cointegration in the rent-to-price model for a longer period than the system based

[11] A program for calculating finite sample critical values for the conditional equilibrium correction model accompanies the paper by Ericsson and MacKinnon (2002) and is available on http://qed.econ.queensu.ca/pub/faculty/mackinnon/.

[12] This algorithm automatizes the Gets approach and can also handle cases where regressors are not mutually orthogonal. A recent evaluation of the search algorithm is given in Castle et al. (2011).

Table 5: Recursive coefficients for inverted demand equation using a single equation approach (confer equation (7) and equation (8)), 1977q2–T

End Point (T)	β_y	β_h	β_{UC}	α_{ph}	p-value
1995q4	1.414	-2.579	-0.626	-0.145	0.0417
1996q4	1.799	-3.381	-0.965	-0.138	0.0420
1997q4	1.805	-3.378	-1.191	-0.155	0.0137
1998q4	1.498	-2.768	-0.885	-0.168	0.0051
1999q4	1.697	-3.134	-0.926	-0.145	0.0123
2000q4	1.835	-3.396	-0.922	-0.139	0.0129
2001q4	2.205	-4.138	-1.049	-0.107	0.1002
2002q4	2.837	-5.366	-0.912	-0.081	0.1529
2003q4	8.832	-17.067	0.890	-0.035	0.7560
2004q4	15.484	-30.004	0.186	-0.015	0.9295
2005q4	-15.022	30.025	-6.927	0.011	0.9976
2006q4	29.649	-58.872	-1.292	-0.007	0.9593
2007q4	5.355	-10.562	-2.053	-0.030	0.3936
2008q4	4.154	-8.161	0.547	-0.034	0.2855
2009q4	4.417	-8.960	1.052	-0.033	0.3413
2010q2	5.720	-11.248	-1.053	-0.034	0.3180

Notes: This table reports the estimated cointegrating vector along with the loading factor and the corresponding p-value when the inverted demand model is estimated using a single equation approach.

approach does. That said, the estimated loading factor changes towards zero already in 2001/2002, which closely resembles the results from the system based analysis.

5.4 Encompassing previous findings

As I discussed in Section 2, the results in the literature show no consensus about the issue of whether an equilibrium correction model can capture the dynamics of US housing prices well or not. There may be several reasons for the divergence of results and my results indicate that the different sample periods used can be one explanation.

In that respect, the results reported in Table 2–5 tell an intriguing story[13]: As long as the estimation end point is set to 2000q4 or earlier, my results suggest that considering an inverted demand model, housing prices and fundamentals are cointegrated. Interestingly, both Meen (2002), Abraham and Hendershott (1996) and Malpezzi (1999) whose samples end prior to this all reach that conclusion.

However, a researcher estimating the same model for a sample ending in any period between 2001 and 2010 would have been lead to the conclusion that an equilibrium correction model cannot possibly explain the fluctuations in US housing prices. That is the case for both Gallin (2006), Clark and Coggin (2011) and Zhou (2010) whose sample

[13] I compare to both studies that have employed national data and studies that have considered large panels. Though the comparison is not meant to be exact in the sense that start years, operationalizations of the data and test procedures may differ across the studies, it is still interesting to observe that parts of the diverging results in the literature may be attributed to different sample periods.

ends in 2002q2, 2005q2 and 2007q4, respectively. It is interesting to note that while Mikhed and Zemcik (2009a) find evidence of cointegration between housing prices and construction wages for a sample ending in 2006q4 but not in 1996q4, my results – using a slightly different information set – suggest the opposite.

Also based on the results from the rent-to-price approach am I able to encompass previous findings in the literature. Neither Gallin (2008) nor Mikhed and Zemcik (2009b) find evidence for cointegration when looking at the relationship between housing prices and rents for samples ending in 2005 and 2006, respectively. This corroborates the findings reported in Table 2 and Table 5.

The above discussion indicates that – to a large extent – the diverging results in the literature can be ascribed to the use of different estimation end points. The two studies that stand out from the rest are Duca et al. (2011a,b), who document that there is evidence of cointegration in both a price-to-rent model and an inverted demand equation for samples ending in 2007q2 and 2009q3, respectively. They include a measure of the loan-to-value ratio for first time home buyers in their analysis, which may explain why they find cointegration for the period as a whole. Nevertheless, as Figure 2 shows, the cointegrating relations I am able to establish prior to 2001 are very stable when estimated recursively and there is strong evidence of cointegration also prior to this, confer Table 2–5. With that in mind, another interpretation of the results in Duca et al. (2011a,b) is that by conditioning on the LTV ratio, they are able to model a structural break. The next section provides additional evidence to the claim by Duca et al. (2011a,b) that it is the major changes in the credit market that caused the breakdown of these models in the early 2000s and therefore was an important factor causing the US housing bubble.

6 The increased subprime exposure as a cause of the breakdown

One possible cause of the econometric breakdown documented in the previous section is that the substantial changes in the subprime market allowed previously constrained and risky borrowers to finance the housing bubble. If that was the case, we should not expect housing prices and fundamentals to be cointegrated. In that respect, Figure 3 tells an interesting story. The graph displays the number of subprime loans as a share of total loans serviced by the participants in the mortgage delinquency survey over the period 1998q1 to 2010q4.

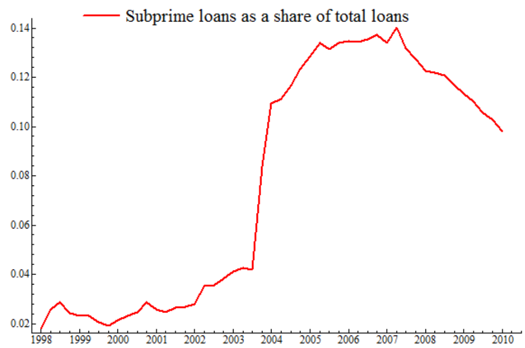


Figure 3: The number of subprime loans as a share of total loans, 1998q1–2010q4 (Source: Moody's

6

It is clear from that figure that the explosion in subprime lending comes very close in date to the equilibrium correction breakdown I documented in the previous section, with the ratio of subprime loans as a share of total loans going from only 2% in 1998q1 to 14% at its peak in 2007.

To investigate the role played by the increased lending to a more risky segment of the market a little further, I have included this ratio, *sp*, as a variable in the VECMs of the previous section.[14] The sudden jump in this series in 2003 leads to some mis-specification in the VARs that was not present earlier, but it is nevertheless interesting to see what happens when this variable is included in the VARs. I have summarized these findings in Table 6 and Table 7.

It can clearly be seen from the results in Table 6 and Table 7 that by including this variable in the two VARs, I find evidence for one cointegrating vector over the full sample. In addition, I find that the trend can be excluded and weak exogeneity of all the variables in the VAR (including the new variable) is supported. Most striking is the fact that including this variable, which is positive and highly significant, changes the estimates of the other coefficients for the full sample analysis in such a way that they move very close to their pre-break values, compare to the results in Table 2 and 3. Furthermore, the loading factor is also increased substantially and now has a more reasonable numerical size.

[14]Due to the lack of data, I have set this series to zero prior to 1998q1. That said, since subprime lending is a relatively new phenomena, this approximation should not be very important for my results.

Table 6: CVAR analysis for the rent-to-price approach with subprime share in VAR, 1977q2-2010q4

$Eigenvalue: \lambda_i$	H_0	H_A	λ_{trace}	5%-critical value[b]
0.281	$r = 0$	$r \geq 1$	67.81	62.66
0.126	$r \leq 1$	$r \geq 2$	29.94	42.77
0.082	$r \leq 2$	$r \geq 3$	14.41	25.73
0.039	$r \leq 3$	$r = 4$	4.57	12.45

Results when trend is excluded and weak exogeneity of user cost, rents and subprime share is imposed (standard errors below point estimates):

$ph + \underset{0.486}{1.201}UC - \underset{0.167}{1.219}r - \underset{0.179}{1.419}sp$
$\alpha_{ph} = \underset{0.023}{-0.143}, \alpha_{UC} = 0, \alpha_r = 0, \alpha_{sp} = 0$

Log likelihood: 2110.57

Likelihood ratio test for overidentifying restrictions:

$\chi^2(4) = 4.7267[0.3165]$

Estimation period: 1977q2–2010q4

Table 7: CVAR analysis for the inverted demand approach with subprime share in VAR, 1977q2-2010q2

$Eigenvalue: \lambda_i$	H_0	H_A	λ_{trace}	5%-critical value[b]
0.340	$r = 0$	$r \geq 1$	94.14	73.13
0.226	$r \leq 1$	$r \geq 2$	47.27	50.08
0.110	$r \leq 2$	$r \geq 3$	18.38	30.91
0.046	$r \leq 3$	$r = 4$	5.26	15.33

Results when trend is excluded and weak exogeneity of user cost, rents and sp is imposed (standard errors below point estimates):

$ph + \underset{0.588}{0.672}UC - \underset{0.378}{2.054}y + \underset{0.765}{3.921}h - \underset{0.194}{2.045}sp$
$\alpha_{ph} = \underset{0.020}{-0.136}, \alpha_{UC} = 0, \alpha_y = 0, \alpha_{sp} = 0$

Log likelihood: 2110.57

Likelihood ratio test for overidentifying restrictions:

$\chi^2(4) = 11.201[0.0244]$

Estimation period: 1977q2–2010q2

6

Though the sample period is small, it is interesting to see how these findings are affected if we instead set the estimation end point to 2000q4 (just before the break). As seen from Table 8 and Table 9, I now find evidence of two cointegrating vectors, which suggest that the share of subprime loans is stationary over this period, given the finding of one cointegrating vector when this variable was not included in the VARs. Furthermore, imposing the restrictions that the share of subprime loans has no effect on housing prices (i.e. testing for the same relationships as found in Section 5) and at the same time testing for the stationarity of the share of subprime loans is easily accepted by the likelihood ratio tests. Though the trend is insignificant in the *sp* equations, I have left it there to have an exactly identified system and to make a comparison to the results I obtained when *sp* was not included in the VAR and the end point was set to 2000q4. Looking at the line reading 2000q4 in Table 2 and Table 3 and comparing to the results in Table 8 and Table 9, we see that the coefficient estimates are practically identical. This is a reassuring finding, given the mis-specification that is induced in the models when *sp* is included as an additional variable.

These results suggest that the we can, without loss of generality, exclude this variable from the model in the pre-break period. It further suggests that the breakdown of the stable relationship between housing prices, the user cost and rents as well as the inverted demand equation was caused by the increased exposure to the more risky segment of the market. In the next section, I shall explore whether we can find formal statistical evidence that the equilibrium correction breakdown – which I have interpreted as a bubble – have any predictive power for the wider financial crisis and to what extent the econometric models presented in this paper could have been used to monitor the housing market in real time.

Table 8: CVAR analysis for the rent-to-price approach with subprime share in VAR, 1977q2-2000q4

$Eigenvalue : \lambda_i$	H_0	H_A	λ_{trace}	5%-critical value[b]
0.379	$r = 0$	$r \geq 1$	82.99	62.66
0.323	$r \leq 1$	$r \geq 2$	47.30	42.77
0.206	$r \leq 2$	$r \geq 3$	18.00	25.73
0.010	$r \leq 3$	$r = 4$	0.77	12.45

Results when sp is excluded from the first vector and weak exogeneity is imposed (standard errors below point estimates):

Vector 1: $ph + \underset{0.400}{1.338}UC - \underset{0.159}{1.128}r$

Vector 2: $sp - \underset{0.00004}{0.000009}t$

$\alpha_{ph,1} = \underset{0.037}{-0.206}, \alpha_{sp,2} = \underset{0.058}{-0.367}$

Restrictions on loadings:

$\alpha_{ph,2} = \alpha_{sp,1} = \alpha_{r,1} = \alpha_{r,2} = \alpha_{UC,1} = \alpha_{UC,2} = 0$

Log likelihood: 1650.063

Likelihood ratio test for overidentifying restrictions:

$\chi^2(9) = 7.8184[0.5525]$

Estimation period: 1977q2–2010q2

Table 9: CVAR analysis for the inverted demand approach with subprime share in VAR, 1977q2-2000q4

$Eigenvalue : \lambda_i$	H_0	H_A	λ_{trace}	5%-critical value[b]
0.390	$r = 0$	$r \geq 1$	96.57	73.13
0.355	$r \leq 1$	$r \geq 2$	59.54	50.08
0.231	$r \leq 2$	$r \geq 3$	26.62	30.91
0.088	$r \leq 3$	$r = 4$	6.90	15.33

Results when sp is excluded from the first vector and weak exogeneity is imposed (standard errors below point estimates):

Vector 1: $ph + \underset{0.490}{0.933}UC - \underset{0.314}{1.691}y + \underset{0.623}{3.170}h$

Vector 2: $sp - \underset{0.00004}{0.00002}t$

$\alpha_{ph,1} = \underset{0.020}{-0.179}, \alpha_{sp,2} = \underset{0.058}{-0.373}$

Restrictions on loadings:

$\alpha_{ph,2} = \alpha_{sp,1} = \alpha_{y,1} = \alpha_{y,2} = \alpha_{UC,1} = \alpha_{UC,2} = 0$

Log likelihood: 1570.356

Likelihood ratio test for overidentifying restrictions:

$\chi^2(10) = 8.8845[0.5431]$

Estimation period: 1977q2–2010q2

7 Equilibrium correction breakdown as a bubble indicator and a predictor for the wider financial crisis

I have constructed two "bubble indicators"(BI's) in the spirit of Mikhed and Zemcik (2009a), but my indicators are based on the relationship between housing prices and fundamentals from recursively estimating and respecifying the models represented by equation (5) and equation (8) using Autometrics.

I have let my indicators take the values of the finite sample p-values calculated when the variable selection is done recursively quarter-by-quarter all the way back to 1995q4.[15] This means that the derived bubble measure is dependent on the extent to which housing prices and fundamentals are cointegrated at different points in time, which can be seen as an operationalization of Stiglitz (1990) definition of a bubble. Thus, if we believe that the lack of cointegration corresponds to a bubble (or at least that prices are not responding to deviations from fundamentals in a "normal" way), then any p-value in excess of, say 10%, may indicate a major distortion in the housing market.

Given the data sources and methodology proposed in this paper, my indicators could have been constructed already in 2000 (or earlier) and used to say something about the temperature in the US housing market and to assert the role of fundamentals in real time. The two indicators are plotted along with a straight line indicating a 10% (no bubble) significance level in Figure 4.

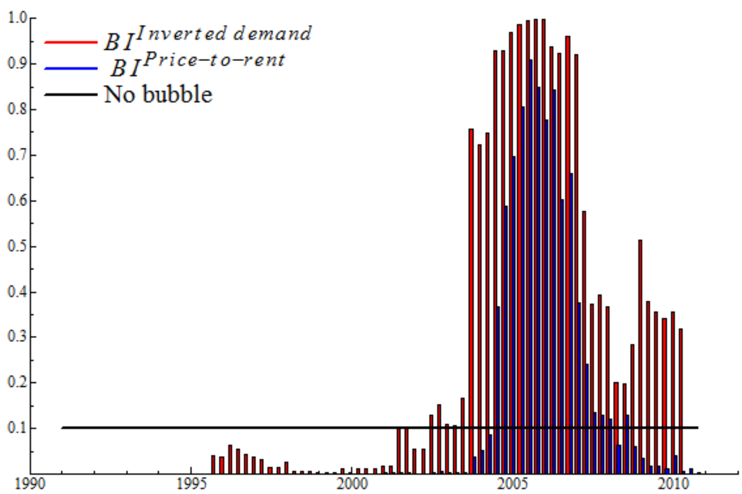


Figure 4: Bubble indictor from price-to-rent approach (blue) and inverted demand approach (red), 1995q1–2010q4

Although the two indicators are not identical, they both send a quite clear signal already in the early 2000s. In 2004, it is evident that both indicators suggest a bubble in the US housing market. They also stay at a high level until 2006, where both start

[15]The calculation of finite sample critical values was done using the program accompanying Ericsson and MacKinnon (2002). As they emphasize, the critical values for the conditional equilibrium correction model depends on a number of features such as the sample size, the number of variables in the hypothesized cointegrating vector, what deterministic terms are included as well as the number of estimated coefficients.

dropping (the price-to-rent based indicator more so). While the price-to-rent indicator hits the no bubble line in 2009, that is not the case for the one derived from the inverted demand equation. That may either reflect the notion of a negative bubble or simply be the result of the fact that this alternative approach requires more observations to reestablish cointegration.

As a first step to investigate the relevance of these bubble indicators a little further, I have addressed two additional questions: Are the BI's leading 1) Financial (in)stability measures, such as delinquency rates and nonperforming loans? and 2) Coincident indicators such as the unemployment rate and industrial production? In an attempt to answer these questions, I have tested for Granger non-causality (see Granger (1969)).

To represent measures of financial (in)stability, I have used the delinquency rates on loans secured by real estate,Del, as well non-performing loans as a share of total loans, NPL. These series were collected from FRED. The unemployment rate, U, which is the civilian unemployment rate from BLS and industrial production, IP, which is measured by the industrial production index collected from FRED are used as coincident indicators. All the series used in the GNC tests are plotted in Figure A.3 in Appendix A.

The standard setup to test for Granger causality is to consider a bi-variate VAR of the following form:

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \sum_{i=1}^{p} \begin{pmatrix} a_{11,i} & a_{12,i} \\ a_{21,i} & a_{22,i} \end{pmatrix} \begin{pmatrix} y_{1,t-i} \\ y_{2,t-i} \end{pmatrix} \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix} \tag{11}$$

Then the appropriate lag length,l, may be determined by a sequence of F-tests. A test for Granger non-causality of y_1 (y_2) with respect to y_2 (y_1) is then a test on whether $a_{12,i}$ (a_{21_i}) $= 0 \ \forall \ i = 1, ..., l$. However, since several of the variables considered here appear to be non-stationary over the sample period considered, I adopt a slightly different procedure. I start, in the usual way, by determining the optimal lag length, l, by a sequence of F-tests. Then, I test for cointegration between the variables in the VAR. If there is no evidence of cointegration, I consider the variables in first differences. However, if there is evidence of cointegration, I consider the bi-variate VAR on VECM form, i.e:

$$\begin{pmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{pmatrix} = \begin{pmatrix} \alpha_1\beta_1 & \alpha_1\beta_2 \\ \alpha_2\beta_1 & \alpha_2\beta_2 \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} + \sum_{i=1}^{l-1} \begin{pmatrix} \gamma_{11,i} & \gamma_{12,i} \\ \gamma_{21,i} & \gamma_{22,i} \end{pmatrix} \begin{pmatrix} \Delta y_{1,t-i} \\ \Delta y_{2,t-i} \end{pmatrix} \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix} \tag{12}$$

Thus, if y_2 is Granger non-causal for y_1, it must be the case that $\alpha_1 = 0$ and $\gamma_{12,i} = 0 \ \forall \ i = 1, ..., l-1$. Given that we find cointegration, either $\alpha_1 \neq 0$, $\alpha_2 \neq 0$, or both. Hence, cointegration implies Granger causality in at least one direction (Granger, 1986).

Initially, I started with a generous lag length of 8. Then I decided the optimal lag truncation using ordinary F-tests. The results from these tests for GNC are displayed in Table 10 and Table 11 using the bubble indicator implied by the price-to-rent and inverted demand approach, respectively. The first column lists the variable that is used to test for GNC. The next two columns report the chosen lag length along with the number of cointegrating relationships I find support for. The final two columns show the p-value from the tests for GNC from BI to the variable considered, and *vice versa*.

7

Table 10: Tests for Granger non-causality using BI from price-to-rent model

Variable tested (x)	*Lags*	Rank	$BI^{\text{Price-to-rent}} \rightarrow x$	$x \rightarrow BI^{\text{Price-to-rent}}$
Del	8	1	0.0000	0.2699
NPL	8	1	0.0000	0.0120
U	5	1	0.0000	0.5102
IP	3	0	0.4323	0.9432
(Sample: 1997q4–2010q4)				

Notes: The table reports the p-values from standard F-tests for Granger non-causality between the the bubble indicator derived from the price-to-rent model ($BI^{\text{Price-to-rent}}$) and a set of financial (in)stability measures and coincident indicators. The financial (in)stability measures comprise delinquency rates (Del) and non-performing loans as a fraction of total loans (NPL). The coincident indicators are made up by the unemployment rate (U) and an industrial production index (IP).

Table 11: Tests for Granger non-causality using BI from inverted demand equation

Variable tested (x)	*Lags*	Rank	$BI^{\text{Inverted demand}} \rightarrow x$	$x \rightarrow BI^{\text{Inverted demand}}$
Del	5	1	0.0000	0.3016
NPL	4	1	0.0000	0.0812
U	5	1	0.0000	0.6863
IP	2	1	0.0000	0.8084
(Sample: 1997q2–2010q2)				

Notes: The table reports the p-values from standard F-tests for Granger non-causality between the the bubble indicator derived from the inverted demand model ($BI^{\text{Inverted demand}}$) and a set of financial (in)stability measures and coincident indicators. The financial (in)stability measures comprise delinquency rates (Del) and non-performing loans as a fraction of total loans (NPL). The coincident indicators are made up by the unemployment rate (U) and an industrial production index (IP).

The results from the GNC tests suggest that the BI's have some predictive power for the two financial (in)stability measures and the coincident indicators. There is however little evidence of a causal relationship going in the other direction. A strict interpretation of the combined results from the previous and the current section is that there exists formal statistical evidence implying that the extension of subprime lending caused the breakdown (the bubble) and that this contributed to the instability in the banking sector and the wider financial crisis. The most interesting finding with regard to the bubble indicators is, however, that they clearly warn of the imbalances in the US housing market at a quite early stage. The relevance of such indicators for monitoring the housing market should, however, be assessed by looking at more countries or possibly by disaggregating to a state or MSA level in the US.

8 Conclusion

Based on both system based tests and single equation test for the absence of cointegration, this paper has documented how two stable equilibrium relationships linking real US housing prices to real rents and the real direct user cost and another one linking real housing prices to the real direct user cost,real disposable income and the housing stock breaks down in the early 2000s. Though there are some evidence of restored equilibrium correction at the end of the sample, the adjustment coefficient and the long run elasticities are diametrically different in the post-break period. Including a measure for the number of subprime mortgages as a share of total mortgages, I am able to model this structural break. Further, I show that this variable is stationary for the pre-break period. These findings suggest that it was the expansion of subprime borrowing that caused the breakdown.

Because cointegration is a non-trival finding, the breakdown of a cointegrating relationship can often be interpreted as a result of a far-reaching or fundamental change in an interwoven system like the US housing and credit market. It can also be interpreted as a passage from a regime where fundamentals drive housing prices, to a regime dominated by bubble dynamics. In that perspective, I developed two regime switch indicators, which can be interpreted as a "bubble indicators". According to these indicators, the US housing bubble started in the early 2000s, was pricked in 2007 and by the end of 2010 housing prices were more closely in line with the pre-break fundamentals.

Tests for Granger non-causality showed that the indicators have predictive power for financial (in)stability measures such as delinquency rates and non-performing loans as well as coincident indicators represented by the unemployment rate and an industrial production index.

References

Abraham, J. M. and P. H. Hendershott (1996). Bubbles in metropolitan housing markets. *Journal of Housing Research 7*(2), 191–207.

Aron, J., J. V. Duca, J. Muellbauer, K. Murata, and A. Murphy (2011). Credit, housing collateral and consumption: Evidence from the uk, japan and the us. *Review of Income and Wealth*.

Calhoun, C. A. (1996). OFHEO house price indexes: HPI technical description. Unpublished note, Office of Federal Housing Enterprise Oversight, Washington D.C.

Case, K. E. and R. J. Shiller (1987). Prices of single-family homes since 1970: New indexes for four cities. *New England Economic Review*, 45–56.

Castle, J. L., J. A. Doornik, and D. F. Hendry (2011). Evaluating automatic model selection. *Journal of Time Series Econometrics 3*.

Clark, S. P. and T. D. Coggin (2011). Was there a U.S. house price bubble? an econometric analysis using national and regional panel data. *The Quarterly Review of Economics and Finance 51*(2), 189–200.

Cunningham, C. R. and G. V. Engelhardt (2008). Housing capital-gains taxation and homeowner mobility: Evidence from the taxpayer relief act of 1997. *Journal of Urban Economics 63*(3), 803–815.

Dickey, D. A. and W. A. Fuller (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association 74*(366), 427–431.

Dickey, D. A. and W. A. Fuller (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica 49*(4), 1057–1072.

Doornik, J. (2009). Autometrics. In J. L. Castle and N. Shepard (Eds.), *The Methodology and Practice of Econometrics*, pp. 88–121. Oxford University Press.

Doornik, J. A. (2003). Asymptotic tables for cointegration tests based on the gamma-distribution approximation. Unpublished note, Nuffield College, Oxford.

Doornik, J. A. and D. F. Hendry (2009). *Empirical Econometric Modelling: PcGive 13*, Volume 1. Timberlake.

Duca, J., J. Muellbauer, and A. Murphy (2011a). House prices and credit constraints: Making sense of the us experience. *Economic Journal 121*, 533–551.

Duca, J., J. Muellbauer, and A. Murphy (2011b). Shifting credit standards and the boom and bust in U.S. home prices. Technical Report 1104, Federal Reserve Bank of Dallas.

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

Ericsson, N. R. and J. G. MacKinnon (2002). Distributions of error correction tests for cointegration. *Econometric Journal 5*, 285–318.

Gallin, J. (2006). The long-run relationship between house prices and income: Evidence from local housing markets. *Real Estate Economics 34*(3), 417–438.

Gallin, J. (2008). The long-run relationship between house prices and rents. *Real Estate Economics 36*(4), 635–658.

Girouard, N., M. Kennedy, P. van den Noord, and C. Andre (2006). Recent house price developments: The role of fundamentals. OECD Economics Department Working Papers 475.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica 37*(3), 424–38.

Granger, C. W. J. (1986). Developments in the study of cointegrated economic variables. *Oxford Bulletin of Economics and Statistics 48*(3), 213–228.

Granger, C. W. J. (1991). Some recent generalizations of cointegration and the analysis of long-run relationships. In R. F. Engle and C. W. J. Granger (Eds.), *The Methodology and Practice of Econometrics*. Oxford University Press.

Granger, C. W. J. and J. J. Hallman (1991). Long memory series with attractors. *Oxford Bulletin of Economics and Statistics 53*(1), 11–26.

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

Hendershott, P., R. Hendershott, and J. Shilling (2010). The mortgage finance bubble: Causes and corrections. *Journal of Housing Research 19*(1), 1–16.

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

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica 59*(6), 1551–1580.

Johansen, S. (1995). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press.

Leamer, E. E. (2007). Housing is the business cycle. Working paper, NBER.

Malpezzi, S. (1999). A simple error correction model of house prices. *Journal of Housing Economics 8*(1), 27–62.

McCarthy, J. and W. Peach (2004). Are home prices the next "bubble"? Economic policy review, Federak Reserve Bank of New York.

Meen, G. (2001). *Modelling Spatial Housing Markets: Theory, Analysis and Policy.* Kluwer Academic Publishers, Boston.

Meen, G. (2002). The time-series behavior of house prices: A transatlantic divide? *Journal of Housing Economics 11*(1), 1–23.

Mikhed, V. and P. Zemcik (2009a). Do house prices reflect fundamentals? Aggregate and panel data evidence. *Journal of Housing Economics 18*(2), 140–149.

Mikhed, V. and P. Zemcik (2009b). Testing for bubbles in housing markets: A panel data approach. *The Journal of Real Estate Finance and Economics 38*, 366–386.

Muellbauer, J. and A. Murphy (1997). Booms and busts in the UK housing market. *Economic Journal 107*(445), 1701–1727.

Pesaran, M. H. (2007). A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics 22*(2), 265–312.

Phillips, P. C. B. (1987). Time series regression with a unit root. *Econometrica 55*(2), 277–301.

Phillips, P. C. B. and P. Perron (1988). Testing for a unit root in time series regression. *Biometrika 75*(2), 335–346.

Stiglitz, J. E. (1990). Symposium on bubbles. *The Journal of Economic Perspectives 4*(2), 13–18.

Tobin, J. (1969). A general equilibrium approach to monetary theory. *Journal of Money, Credit and Banking 1*(1), pp. 15–29.

Zhou, J. (2010). Testing for cointegration between house prices and economic fundamentals. *Real Estate Economics 38*(4), 599–632.

A Figures

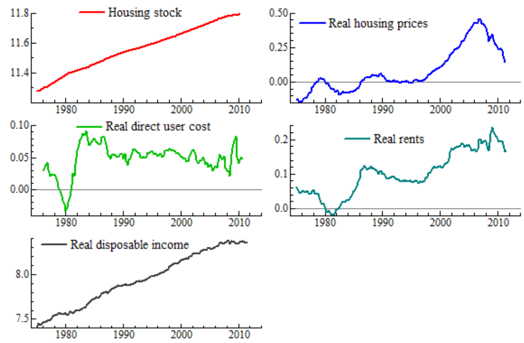


Figure A.1: The data series in levels, 1975q1–2010q4

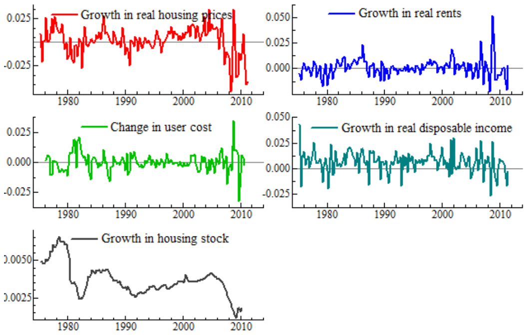


Figure A.2: The data series in first difference, 1975q1–2010q4

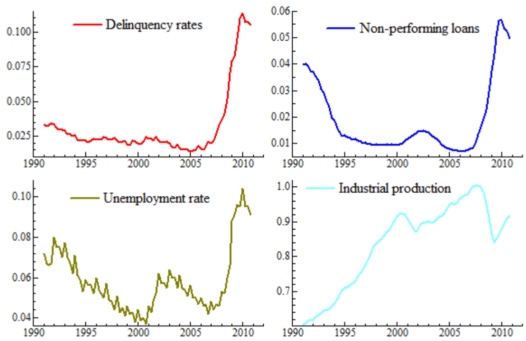


Figure A.3: The data series used for Granger non-causality tests, 1995q1–2010q4

B Tables

Table B.1: Tests for the order of integration

	ADF			PP			
Variable	t-ADF	5%	k	Adj. t-stat	5%	BW	Characteristics
ph	−2.218	−3.44	3	−1.736	−3.44	8	t
h	−3.442	−3.44	2	−2.728	−3.44	9	t
y	−1.933	−3.44	4	−1.996	−3.44	5	t
UC	−2.817	−3.44	5	−2.386	−3.44	2	t
r	−2.586	−3.44	3	−2.549	−3.44	6	t
Δph	−3.459	−2.88	2	−8.232	−2.88	8	i
Δh	−1.576	−2.88	1	−1.738	−2.88	7	i
Δy	−11.940	−2.88	0	−13.454	2.88	5	i
ΔUC	−4.479	−2.88	4	−8.705	−2.88	12	i
Δr	−5.502	−2.88	2	−8.962	−2.88	6	i
$\Delta^2 h$	−8.693	−2.88	0	−9.311	−2.88	6	i
(Sample: 1975q1–2010q4)							

Notes: The table reports the results from two different unit root tests. ADF refers to the Augmented Dickey-Fuller test and PP is the Phillips-Perron test. k denotes the optimal lag truncation for the ADF-test and BW is the bandwidth selected for the PP-test. For the ADF tests, I started with 5 lags and tested down the lag length according to an ordinary t-test. Under the column heading *Characteristics*, t denotes a trend and an intercept in the test regression, while i refers to the case where only an intercept was included.

Table B.2: Vector diagnostics from CVAR based on price-to-rent approach (confer equation (4) and equation (9)), 1977q2–T

End point (T)	Autocorrelation	Non-normality	Heteroskedasticity
1995q4	0.4026	0.0423	0.8108
1996q4	0.4471	0.0239	0.6919
1997q4	0.2715	0.0439	0.7280
1998q4	0.3804	0.0261	0.6084
1999q4	0.2328	0.0621	0.6694
2000q4	0.2989	0.0318	0.6142
2001q4	0.2281	0.1822	0.4962
2002q4	0.1704	0.2110	0.4025
2003q4	0.3091	0.0100	0.1081
2004q4	0.3747	0.0105	0.0345
2005q4	0.5299	0.0047	0.0497
2006q4	0.2880	0.0210	0.0189
2007q4	0.1437	0.0059	0.0034
2008q4	0.1113	0.0763	0.0000
2009q4	0.0486	0.0743	0.0000
2010q4	0.0266	0.0163	0.0000

Notes: This table reports the diagnostics from the recursively estimated price-to-rent VAR. The rest of the results from this analysis are reported in Table 2.

Table B.3: Vector diagnostics from CVAR based on inverted demand approach (confer equation (7) and equation (10)), 1977q2–T

End point (T)	Autocorrelation	Non-normality	Heteroskedasticity
1995q4	0.0244	0.7224	0.6940
1996q4	0.0229	0.9137	0.6254
1997q4	0.0114	0.8311	0.5237
1998q4	0.0199	0.8200	0.3885
1999q4	0.0177	0.8972	0.4197
2000q4	0.0169	0.9430	0.5939
2001q4	0.0686	0.8884	0.3084
2002q4	0.0300	0.8603	0.2436
2003q4	0.1685	0.5440	0.1508
2004q4	0.2078	0.3930	0.1935
2005q4	0.1555	0.5202	0.3101
2006q4	0.1448	0.6997	0.4420
2007q4	0.2031	0.5177	0.4306
2008q4	0.1188	0.6179	0.0001
2009q4	0.0429	0.3750	0.0002
2010q2	0.0777	0.1875	0.0001

Notes: This table reports the diagnostics from the recursively estimated inverted demand VAR. The rest of the results from this analysis are reported in Table 2.

Table B.4: Vector diagnostics from CVAR based on inverted demand approach *excluding* the trend (confer equation (7) and equation (10)), 1977q2–T

End point (T)	Autocorrelation	Non-normality	Heteroskedasticity
1995q4	0.1825	0.1907	0.7713
1996q4	0.2193	0.3020	0.6664
1997q4	0.1623	0.3016	0.5315
1998q4	0.1790	0.3695	0.4894
1999q4	0.1072	0.4930	0.4660
2000q4	0.1916	0.5727	0.4776
2001q4	0.3038	0.5255	0.2795
2002q4	0.2289	0.5954	0.1808
2003q4	0.5100	0.3471	0.0641
2004q4	0.4902	0.2278	0.0529
2005q4	0.3384	0.4397	0.0676
2006q4	0.3360	0.6365	0.0247
2007q4	0.5055	0.4217	0.0390
2008q4	0.2709	0.6048	0.0000
2009q4	0.0485	0.3583	0.0000
2010q4	0.1038	0.1071	0.0000

Notes: This table reports the diagnostics from the recursively estimated inverted demand VAR when the trend is excluded from the model.

Table B.5: Results from recursive CVAR analysis based on price to rent approach *without* dummies, (confer equation (4) and equation (9)), 1977q2–T

End point (T)	Lags	Autocorrelation	Non-normality	Heteroskedasticity	Rank	P-value restrictions	α_{ph}	β_r	β_{UC}
1995q4	5	0.3493	0.0425	0.8216	1	0.1901	−0.211 0.040	1.161 0.121	−0.998 0.276
1996q4	5	0.3963	0.0211	0.6869	1	0.1774	−0.217 0.039	1.176 0.117	1.077 0.262
1997q4	5	0.3308	0.0376	0.6558	1	0.2472	−0.194 0.038	1.238 0.131	−1.152 0.294
1998q4	5	0.2318	0.0890	0.5904	1	0.2998	−0.170 0.034	1.321 0.143	−1.300 0.325
1999q4	5	0.1512	0.1032	0.6969	1	0.2692	−0.142 0.031	1.426 0.165	−1.421 0.381
2000q4	5	0.0714	0.0451	0.4982	1	0.0191	−0.103 0.026	1.670 0.224	−1.719 0.532
2001q4	5	0.0218	0.1689	0.3259	0	*	*	*	*
2002q4	5	0.0150	0.1569	0.3374	1	0.0011	−0.080 0.021	1.960 0.231	−2.310 0.580
2003q4	5	0.0415	0.0099	0.0589	0	*	*	*	*
2004q4	5	0.1820	0.0084	0.0271	0	*	*	*	*
2005q4	5	0.4350	0.0050	0.0440	0	*	*	*	*
2006q4	5	0.2676	0.0132	0.0142	1	0.0000	−0.033 0.012	2.892 0.503	−2.230 1.470
2007q4	5	0.0294	0.0093	0.0015	0	*	*	*	*
2008q4	5	0.0421	0.0460	0.0000	0	*	*	*	*
2009q4	5	0.0154	0.0412	0.0000	0	*	*	*	*
2010q2	5	0.0180	0.0082	0.0000	1	0.1735	−0.054 0.012	2.433 0.201	−1.096 0.762

Notes: This table reports a summary of the main results when the system based approach of Johansen (1988) is implemented by sequentially adding fourth new observations to the sample. The first end point is 1995q4, while the last is 2010q4. The endogenous variables in the system are real housing prices, ph, real rents, r and the real direct user cost, UC. A deterministic trend is restricted to enter the cointegration space, while a constant amd three centered seasonal dummies enter unrestrictedly.

Table B.6: Results from recursive CVAR analysis based on inverted demand approach *without* dummies (confer equation (7) and equation (10)), 1977q2–T

End point (T)	Lags	Autocorrelation	Non-normality	Heteroskedasticity	Rank	P-value restrictions	α_{ph}	β_y	β_{UC}	β_h
1995q4	5	0.0338	0.8148	0.4726	1	0.2672	$-0.155_{0.029}$	$1.795_{0.424}$	$0.149_{0.316}$	$-3.212_{0.856}$
1996q4	5	0.0412	0.7389	0.4275	1	0.5493	$-0.143_{0.027}$	$2.111_{0.442}$	$0.281_{0.335}$	$-3.900_{0.882}$
1997q4	5	0.0224	0.6157	0.2233	1	0.4025	$-0.150_{0.028}$	$2.015_{0.417}$	$0.108_{0.306}$	$-3.644_{0.828}$
1998q4	5	0.0437	0.5176	0.1974	1	0.5524	$-0.147_{0.028}$	$2.054_{0.415}$	$0.041_{0.281}$	$-3.689_{0.823}$
1999q4	5	0.0319	0.6550	0.1634	1	0.7941	$-0.149_{0.027}$	$1.958_{0.385}$	$0.035_{0.268}$	$-3.515_{0.769}$
2000q4	5	0.0204	0.6947	0.3491	1	0.7618	$-0.140_{0.025}$	$2.188_{0.384}$	$0.011_{0.278}$	$-3.947_{0.775}$
2001q4	5	0.0853	0.6757	0.0962	1	0.1006	$-0.113_{0.022}$	$2.624_{0.487}$	$0.149_{0.353}$	$-4.764_{0.986}$
2002q4	5	0.0714	0.8215	0.0554	1	0.0194	$-0.086_{0.017}$	$3.242_{0.645}$	$0.281_{0.466}$	$-5.959_{1.307}$
2003q4	5	0.1753	0.4731	0.0612	1	0.0018	$-0.050_{0.011}$	$5.214_{1.150}$	$0.978_{0.829}$	$-9.838_{2.335}$
2004q4	5	0.2312	0.4798	0.0290	1	0.0004	$-0.024_{0.006}$	$9.518_{2.519}$	$2.194_{1.769}$	$-18.246_{5.103}$
2005q4	5	0.1717	0.6157	0.0828	1	0.0000	$-0.003_{0.001}$	$53.161_{17.755}$	$21.500_{12.903}$	$-103.95_{35.840}$
2006q4	5	0.1559	0.7780	0.2607	1	0.0000	$-0.009_{0.002}$	$21.047_{6.113}$	$7.642_{4.477}$	$-40.912_{12.287}$
2007q4	5	0.1365	0.7872	0.1178	1	0.0000	$-0.028_{0.007}$	$6.212_{1.936}$	$1.902_{1.418}$	$-11.695_{3.905}$
2008q4	5	0.0752	0.2581	0.0000	1	0.0004	$-0.036_{0.007}$	$4.976_{1.548}$	$2.244_{1.201}$	$-9.333_{3.152}$
2009q4	5	0.0130	0.1343	0.0000	1	0.0001	$-0.034_{0.007}$	$5.584_{1.552}$	$2.607_{1.280}$	$-10.545_{3.179}$
2010q2	5	0.0199	0.0661	0.0000	1	0.0001	$-0.033_{0.007}$	$5.571_{1.592}$	$2.482_{1.312}$	$-10.484_{3.260}$

Notes: This table reports a summary of the main results when the system based approach of Johansen (1988) is implemented by sequentially adding fourth new observations to the sample. The first end point is 1995q4, while the last is 2010q4. The endogenous variables in the system are real housing prices, ph, real disposable income, y and the real direct user cost, UC. A deterministic trend and the housing stock, h, are restricted to enter the cointegration space. A constant and three centered seasonal dummies enter unrestrictedly.