# Bubbles and House Price Dispersion in the United States during 1975-2007


Yang Tang
Nanyang Technological University

Ting Zeng
Southwestern University of Finance and Economics

Shenghao Zhu
National University of Singapore


January 13, 2017


**Abstract**

We investigate the rapid growth in the dispersion of housing prices across metropolitan statistical areas (MSAs) in the United States during 1975-2007. We first examine several explanations for this pattern, and find that it is difficult to fully explain it. Our econometric analyses show that the Log of price-to-rent ratios follows a random walk process. We then set up a parsimonious asset-pricing island model. We find that the dispersion of fundamental housing prices grows too slow relative to that in data. Incorporating rational bubble solutions, our calibrated model can match the rapid growth in the dispersion of housing prices.

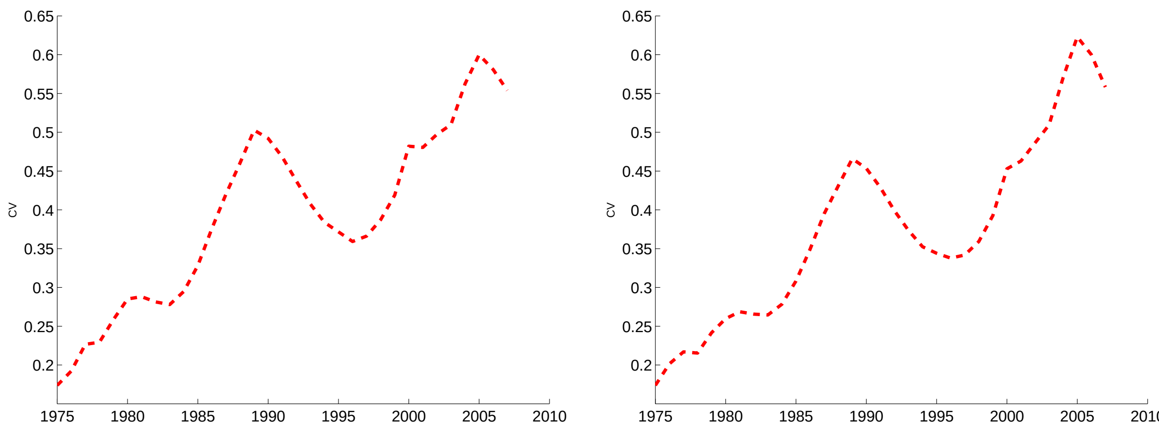


**JEL classification:** E30, G12, R30
**Keywords:** The cross-sectional dispersion of housing prices, excessive dispersion, bubbles

# 1 Introduction

The housing market in the United States during 1975-2007 is featured by a rapid growth in the dispersion of housing prices across metropolitan statistical areas (MSAs). Since housing typically takes up a major proportion of household net worth, fluctuations in housing prices can thus exert significant impacts on the macro-economy.[1] Hence, it is important to understand what drives the rapid rise in the dispersion of housing prices. In this paper we attempt to investigate this issue.

In panel (a) of Figure 1, we plot the cross-sectional coefficient of variation (CV) for housing prices in the United States during 1975-2007.[2] Our balanced panel of housing prices contains 81 major MSAs. In panel (b) of Figure 1, we also present the CV of housing prices for an unbalanced panel of 330 MSAs. Figure 1 shows that there is a rapid increase in the dispersion of housing prices. The CV of housing prices in 1975 is 0.17, while this number reaches 0.55 in 2007.[3]




(a) The CV of housing prices (balanced panel) (b) The CV of housing prices (unbalanced panel)

Figure 1: The Dispersion of Housing Prices: 1975-2007

Notes: The left panel is based on data from a balanced panel of 81 MSAs. The right panel is based on data from an unbalanced panel of 330 MSAs. Housing prices are deflated by the consumer price index (excluding shelters). The original data source is Van Nieuwerburgh and Weill (2010). See Appendix A for details.

[1] Wolff (2006), using the 2001 Survey of Consumer Finances (SCF) data, finds that housing (principal residence and other real estate) accounts for 38% of household assets.

[2] The coefficient of variation (CV) is defined as the ratio of the standard deviation to the mean.

[3] The CV computed in Figure 1 is unweighted. We also compute the weighted CV using housing units in each MSA as the weight. The pattern does not change.

We first examine several intuitively plausible explanations for patterns observed in Figure 1. These explanations emphasize the population concentration, the divergence in income growth rates across MSAs, or the divergence in income growth rates of top income groups across MSAs. Through investigations of these explanations we find that it is difficult to fully explain the rapid rise in the dispersion of housing prices during 1975-2007.

We then conduct econometric analyses of panel data. We empirically show that housing price processes are non-stationary and housing prices grow at different rates across MSAs. These findings are consistent with the pattern in Figure 1 since a unit-root process and different growth rates of housing prices can cause the rapid rise in the dispersion of housing prices. We find that rental growth rates are stationary. While different MSAs have different growth rates of housing prices, they have the same average growth rate of rental prices. This comparison hints that rentals are not the main reason of the rapid rise in the dispersion of housing prices. This comparison also encourages us to further investigate house price-to-rent ratios. Through a panel-data unit-root test we find that house price-to-rent ratios are also non-stationary.

To investigate further patterns in Figure 1 we set up a parsimonious asset-pricing island model. Each island corresponds to an MSA. We first study the fundamental solution of the asset pricing model. The fundamental housing price can be completely supported by rents. Our calculations show that the cross-sectional CV of housing prices is larger than that of housing prices implied by the fundamental solution for each year during 1975-2007. Housing prices in the United States display excessive dispersion. Also we find that the growth in the dispersion of fundamental housing prices is too slow relative to the pattern in data.

Inspired by our empirical finding that the Log of price-to-rent ratios follows a random walk process, we then investigate rational bubbles in our asset pricing model. Following Froot and Obstfeld (1991) and Lansing (2010), we generate bubble components of asset prices in the Lucas asset-pricing model by removing the transversality condition. Specifically, the stochastic growth component of the price-to-rent ratio causes housing price bubbles in our model. Bubble components can only be supported by speculations. Furthermore, our model can deliver explicit expressions for both fundamental and bubble components of the price-to-rent ratio. Our calibrated model with rational bubbles can simultaneously match four stylized facts in the United States during 1975-2007, the rapid growth in the dispersion of housing prices, a moderate increase in the dispersion of rental prices, the rising mean of housing prices, and the rising mean of rental prices. The stochastic growth component of the price-to-rent ratio in the bubble solution is the key mechanism, through which our model could match the rapid growth in the dispersion of housing prices. We also perform several robustness checks. Our model still successfully match housing price moments.

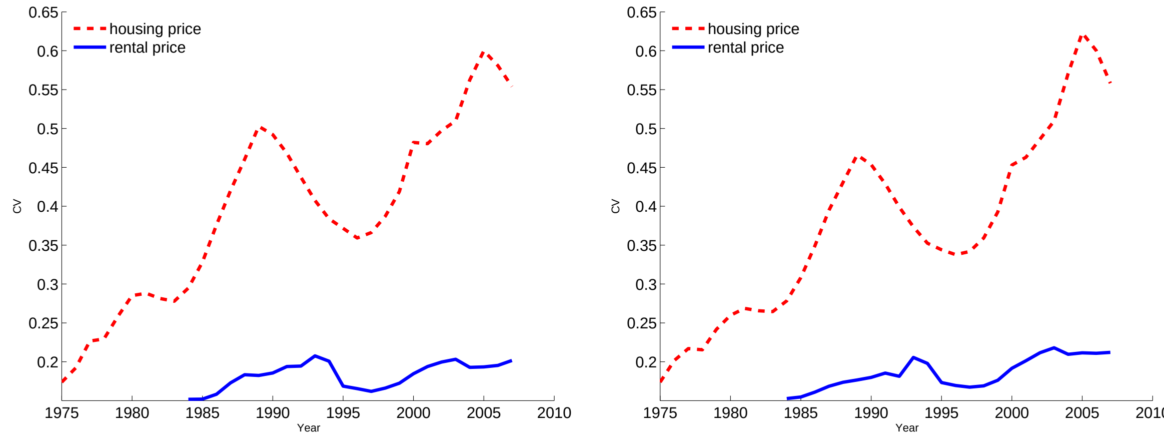
One may think that the stochastic growth component of the price-to-rent ratio eventually leads to explosive dispersion of housing prices. In an extension, we introduce an extrinsic uncertainty, which represents the confidence state, into the benchmark model. We construct a sunspot equilibrium in which bubbles eventually burst in the long-run. But before bubbles burst, there is a rapid growth in the dispersion of housing prices. Thus our paper also contributes to the literature of sunspot equilibria and asset pricing.

## 1.1 Related literatures

Our research is mainly related to two strands of literature: the housing price dispersion and asset bubbles.

Van Nieuwerburgh and Weill (2010) develops a dynamic spatial equilibrium model and calibrate productivity dispersion to match the increase of cross-sectional earnings dispersion across MSAs in the United States during 1975-2007.[4] They show that the calibrated model can match the observed 30-year increase of housing price dispersion. Their framework relies on two main mechanisms, the labor mobility in response to local wage shocks and inelastic housing supply due to regulatory constraints. The housing price in Van Nieuwerburgh and Weill (2010) is equal to the expected present value of rents net of depreciation. And their model predicts that the magnitude of the increase in the dispersion of rental prices is similar to that of the increase in the dispersion of housing prices. However, this prediction is not supported by data.

[4] Eeckhout, Pinheiro, and Schmidheiny (2014) use a spatial equilibrium model to investigate the skill distribution in cities.




(a) The CV of housing prices and that of rental prices (balanced panel)

(b) The CV of housing prices and that of rental prices (unbalanced panel)

Figure 2: Housing Price Dispersion vs. Rental Price Dispersion

Notes: The left panel is based on data from a balanced panel of 81 MSAs. The right panel is based on data from an unbalanced panel of 330 MSAs. Both housing prices and rental prices are deflated by the consumer price index (excluding shelters). The original data source is Van Nieuwerburgh and Weill (2010). See Appendix A for details.

In panel (a) of Figure 2, we plot the CV of housing prices and that of rental prices in the United States. Our balanced panel of housing prices contains 81 major MSAs during 1975-2007. But data of rental prices are only available during 1984-2007. In panel (b) of Figure 2, we also present the CV of housing prices and that of rental prices for an unbalanced panel of 330 MSAs. Housing prices have much higher values of the CV than rental prices. This implies that distributions of housing prices are more dispersed than those of rental prices. Moreover, the growth of the CV of housing prices is more rapid than that of the CV of rental prices. The CV of housing prices in 1984 is 1.94 times that of rental prices, while this number reaches 2.75 in 2007.[5]

In this paper, we generate a bubble component of the house price-to-rent ratio. Our model is able to match the rapid rising dispersion of housing prices given a moderate increase in the CV of rental prices. While labor mobility plays a crucial role in Van Nieuwerburgh and Weill (2010), it is completely prohibited in our island economy. In our model rational bubbles caused by speculations within each island are the main driver of the increase in the

[5]The CV computed in Figure 2 is unweighted. We also compute the weighted CV using housing units in each MSA as the weight. The pattern does not change.

dispersion of housing prices.

Gyourko, Mayer, and Sinai (2013) shows that the widening dispersion of housing prices can be explained by inelastic land supply in some individual-preferred locations combined with an increasing number of high-income households nationally. Their empirical results show that the growing number of rich families in the United States during 1970-2000, can capture more than 80% of the excess growth of housing price in superstar cities in that period. However, due to the lack of data on local income distributions, it is difficult to empirically investigate the transmission mechanism from the change of the national income distribution to the rapid increase in the dispersion of housing prices.

Froot and Obstfeld (1991) generates intrinsic bubble, which only depends on dividends, in an asset-pricing model by removing the transversality condition. Intrinsic bubble can cause asset prices to overreact to changes in fundamentals, and thus can help to explain excess volatility of stock prices.[6] Froot and Obstfeld (1991) assumes that the growth rate of dividends is independent and identically distributed over time. Lansing (2010) permits autocorrelation of the growth rate of dividends and generalizes the intrinsic rational bubble. He shows that the rational bubble component of the price-dividend ratio can evolve as a geometric random walk without drift, such that the mean of the bubble growth rate is zero. Granziera and Kozicki (2015) investigate whether expectations that are not fully rational can explain the evolution of the housing price index and the price-to-rent ratio in the United States during 1987-2011.

In terms of the agent's decision problem, our paper share the same framework as Lansing (2010) and Granziera and Kozicki (2015). Similar to these two papers, ours also makes use of the solution of the asset price other than the fundamental price itself. However, applications are different in these three papers. Lansing (2010) uses the model to study the price-dividend ratio of the stock market in the United States during 1871-2008. Granziera and Kozicki (2015) investigates the housing price index in the United States. Our paper exams the cross-sectional dispersion of house prices. In Granziera and Kozicki (2015) the near rational bubble solution, which produces a stationary price-to-rent ratio, replicates the

[6] See LeRoy (2004) for a survey about rational bubbles. Glaeser and Nathanson (2015) review recent literatures on housing bubbles.

moments of the price-to-rent ratio well.[7] We find that the rational bubble solution, which produces a non-stationary price-to-rent ratio, generates the rapid rising dispersion of house prices given a moderate increase of rental price CV.

The rest of this paper is organized as follows. Two plausible explanations for rising dispersion of housing prices are examined in Section 2. Empirical findings of housing prices and price-to-rent ratios are reported in Section 3. An asset-pricing island model is presented in Section 4. Section 5 contains calibration and simulation results. We extend the benchmark model to allow bubble burst in Section 6. Section 7 concludes the paper.

## 2 Alternative mechanisms

In this section we examine three possible explanations for the pattern displayed in Figure 1. Investigating of these explanations we find that it is difficult to fully explain the rapid rise in the dispersion of housing prices during 1975-2007.

*Explanation 1. The increasing population concentration leads to the rise in the dispersion of housing prices.*

When population distribution becomes more concentrated, demands for houses are more uneven across MSAs. They could in turn lead to the rise in the dispersion of housing prices. However, data do not support this hypothesis. Using the decennial census during 1970-2010 in the United States, we construct data of populations and housing units at the MSA level. When we use the CV to measure the population concentration, we indeed observe a declining trend in the dispersion of populations. We also find that the CV of housing units decreases over the period. The results are robust both in a balanced panel of 81 MSAs and in an unbalanced panel of 330 MSAs. We report these results in Table 1. The CV of populations steadily decreases from 1.15 (1.94) in 1970 to 0.90 (1.64) in 2010 in our sample of 81 (330) MSAs. Similarly, the CV of housing units steadily decreases from 1.20 (2.01) in 1970 to 0.86 (1.56) in 2010 in our sample of 81 (330) MSAs.

[7] The data in Granziera and Kozicki (2015) include the remarkable downturn of the housing market in the United States after 2007. This could be one of the reasons why a stationary price-to-rent ratio fits their data.

Table 1: The CV of Populations and Housing Units

| | CV of Population | | CV of Housing Units | |
|---|---|---|---|---|
| Year | 81 MSAs | 330 MSAs | 81 MSAs | 330 MSAs |
| 1970 | 1.1461 | 1.9366 | 1.1991 | 2.0110 |
| 1980 | 1.0085 | 1.7452 | 1.0144 | 1.7595 |
| 1990 | 0.9605 | 1.7046 | 0.9226 | 1.6586 |
| 2000 | 0.9310 | 1.6740 | 0.8825 | 1.5970 |
| 2010 | 0.9020 | 1.6357 | 0.8552 | 1.5608 |

Notes: Numbers are from authors' computations. See Appendix A for details of our data.

*Explanation 2. The divergence in income growth rates across MSAs leads to the rise in the dispersion of housing prices.*

When average family incomes in different MSAs increase at different rates, housing prices could potentially have different growth rates. Thus we could observe a rapid increase in the dispersion of housing prices across MSAs. We use data from Gyourko, Mayer, and Sinai (2013) to test this mechanism. We calculate growth rates of the housing price and of the average family income between 1970 and 2000 in each MSA. To investigate the divergence in income growth rates, we define superstar MSAs and non-superstar MSAs as in Gyourko, Mayer, and Sinai (2013).[8] There are 21 superstar MSAs and 296 non-superstar MSAs.[9] In Table 2 we report the mean growth rate of the housing price and the mean growth rate of the average family income within the group of superstar MSAs and those means within the group of non-superstar MSAs. Comparing the column of superstar MSAs with that of non-superstar MSAs, we find that the moderate difference of growth rates of average incomes cannot explain the large difference of growth rates of housing prices. Thus this hypothesis alone cannot fully explain the rapid increase in the dispersion of housing prices.[10]

[8]In Gyourko, Mayer, and Sinai (2013), an MSA is a superstar in a particular year of 1970, 1980, 1990, or 2000, if it has high demands for houses and low elasticity of housing supply over the prior two decades. They calculate the growth rates of housing prices and housing units over four time periods: 1950-1970, 1960-1980, 1970-1990, and 1980-2000. They identify high-demand MSAs by checking whether the sum of the growth rates of housing prices and housing units in one MSA is above the sample median. They define inelastic-supply MSAs as those in which the ratio of housing price growth rate-to-housing unit growth rate is in the top decile of the sample. An MSA has time-invariant superstar status if it is a superstar in any two decades between 1970 and 2000 (See pp.179-180 of Gyourko, Mayer, and Sinai (2013)). In Table 2 we use time-invariant superstar status of Gyourko, Mayer, and Sinai (2013).

[9]The MSA definitions of Gyourko, Mayer, and Sinai (2013) are based on 1990 county boundaries.

[10]The divergence in the income growth rates of top income groups across MSAs could also lead to the rise in the dispersion of housing prices. The fast growth of incomes of the rich people in some particular MSAs could cause the rapid growth of the housing prices in these MSAs. However, it is difficult for us to

Table 2: Growth Rates of Housing Prices and of Average Family Incomes: 1970-2000

| | Superstar MSAs | Non-superstar MSAs |
|---|---|---|
| Housing Prices | 131.80% | 63.52% |
| | (0.805) | (0.325) |
| Average Family Incomes | 42.51% | 32.13% |
| | (0.173) | (0.126) |

Notes: Numbers are from authors' computations. The data source is Gyourko, Mayer, and Sinai (2013). Terms in parentheses are standard deviations.

# 3 Empirical findings

In this section we conduct panel-data tests on housing prices, rental growth rates, and house price-to-rent ratios. For these empirical tests we directly obtain the data from Van Nieuwerburgh and Weill (2010).

We find that housing prices have a unit root. And housing prices grow at different rates across MSAs. These findings are consistent with the rapid rise in the dispersion of housing prices. We also find that rental growth rates are stationary. While different MSAs have different growth rates of housing prices, they have the same average growth rate of rental prices. These observations imply that rental prices may not be able to explain the rapid rise in the dispersion of housing prices. To separate the impact of rental prices on housing prices, we further investigate house price-to-rent ratios. We find that house price-to-rent ratios are also non-stationary. Inspired by this empirical finding we set up a theoretical model in Section 4, where rational bubbles can generate non-stationary price-to-rent ratios.

## 3.1 Housing prices

In this subsection we first use panel data to show that housing prices have a unit root. Then we show that housing prices grow at different rates across MSAs.

Let $p_{i,t}$ denote the housing price of MSA $i$ in year $t$. To have a consistent sample for housing prices and rental prices, we use the balanced panel including 81 MSAs for the period

test this hypothesis due to the limited information on the income distribution within an MSA. The major challenge is top-coding. For example, incomes in the Census data and in the American Community Survey suffer from the top-coding problem. This leads to the difficulty to estimating the top incomes at the MSA level. Gyourko, Mayer, and Sinai (2013) uses the number of families in different income bins as the proxy of the income distribution at the MSA level. But this method cannot help us to estimate the income growth rates of the top income groups in MSAs.

during 1984-2007.[11] Let $I = 81$, then $i = 1, 2, ...I$.

We perform a unit-root test of $\log p_{i,t}$ in our balanced panel. We first run several cross-sectional dependence tests for $\log p_{i,t}$ and find that there exists cross-sectional dependence in the panel data. We report test results in Appendix B.1. Therefore, we cannot adopt conventional panel-data unit-root tests, such as Im-Pesaran-Shin (IPS) test, which rely on the assumption that there does not exist cross-sectional dependence. We then proceed to the cross-sectionally augmented IPS (CIPS) test developed by Pesaran (2007). We conduct the CIPS test against the null hypothesis that the Log of housing prices carries a unit root. The cross-sectionally augmented Dickey-Fuller (CADF) regression takes the following form,

$$\Delta \log p_{i,t} = g_i + h_i \log p_{i,t-1} + m_i \overline{\log p_{t-1}} + \sum_{k=0}^{K} n_{ik} \Delta \overline{\log p_{t-k}} + \sum_{k=1}^{K} l_{ik} \Delta \log p_{i,t-k} + v_{it}, \quad (1)$$

where $\Delta$ is the first-difference operator, $\overline{\log p_t}$ is the cross-sectional average of $\log p_{i,t}$, and $K$ is the lag length. Due to the constraint of the sample length, we set $K = 1$. The CIPS test generates CADF statistics for each individual MSA and computes the CIPS test statistic which is the simple cross-sectional average of individual CADF statistics. We reprot our CIPS test statistic and its critical values at 10%, 5%, and 1% significance levels in Table 3. We find that the CIPS test statistic, $-1.129$, is smaller than the critical value of the 10% level in absolute terms. Therefore, we cannot reject the null hypothesis which assumes that there exists a unit root in the panel data of $\log p_{i,t}$.

Table 3: A Panel Unit-root Test for Housing Prices

| CIPS | Critical Value 10% | Critical Value 5% | Critical Value 1% |
|---|---|---|---|
| $-1.129$ | $-2.080$ | $-2.160$ | $-2.300$ |

Notes: The table presents the CIPS test statistic and its critical values for 10%, 5%, and 1% significance levels. the CIPS test statistic, $-1.129$, is smaller than the critical value of the 10% level in absolute terms. Therefore, we cannot reject the null hypothesis which assumes that there exists a unit root in the panel data of $\log p_{i,t}$.

Next we test whether different MSAs have the same average growth rate of housing prices. We conduct a joint hypothesis test after running a panel regression of the growth rates with

[11] We have housing price data during 1975-2007 while we only have rental price data during 1984-2007.

city fixed effects,

$$\Delta \log p_{i,t} = \chi_i + \vartheta \text{Year}_t + e_{i,t}, \tag{2}$$

where $\text{Year}_t$ is the year dummy and $e_{i,t}$ is the error term. Therefore, $\chi_i$ represents the time-invariant city fixed effect in MSA $i$. A joint hypothesis test against the null hypothesis, $H_0 : \chi_1 = \chi_2 = ... = \chi_I$, can be conducted. Table 4 shows that the chi-squared test statistic 136.00 is larger than its critical value at the 1% significance level, 112.329. And the $p$-value is 0.0001. Thus test results reject the null hypothesis. Housing prices grow at different rates across MSAs. These findings are consistent with the rapid rise in the dispersion of housing prices.

Table 4: Test Statistics for City Fixed Effects in Housing Prices

| Chi-square Statistic | $p$-value |
|---|---|
| 136.00 | 0.0001 |

Notes: The table presents the statistics of the joint hypothesis test against the null hypothesis, $H_0 : \chi_1 = \chi_2 = ... = \chi_I$. The chi-squared test statistic 136.00 is larger than its critical value at the 1% significance level, 112.329. And the $p$-value is 0.0001. Thus test results reject the null hypothesis. Housing prices grow at different rates across MSAs.

## 3.2 Rental growth rates

We also conduct a panel-data unit-root test on rental growth rates. We find that rental growth rates are stationary. More importantly, we show that different MSAs have the same average growth rate of rental prices.

Let $x_{i,t}$ denote the rental growth rate of MSA $i$ in year $t$. In Appendix B.2 we find that there exists cross-sectional dependence in the panel data of rental growth rates. We also conduct the CIPS panel unit-root test for rental growth rates. The CADF regression takes the following form,

$$\Delta x_{i,t} = \widetilde{g}_i + \widetilde{h}_i x_{i,t-1} + \widetilde{m}_i \overline{x}_{t-1} + \sum_{k=0}^{K} \widetilde{n}_{ik} \Delta \overline{x}_{t-k} + \sum_{k=1}^{K} \widetilde{l}_{ik} \Delta x_{i,t-k} + \widetilde{\nu}_{it}, \tag{3}$$

where $\Delta$ is the first-difference operator, $\overline{x}_t$ is the cross-sectional average of $x_{i,t}$, and $K$ is the lag length. Due to the constraint of the sample length, we set $K = 1$. The null hypothesis assumes that rental growth rates have a unit root. We reprot our CIPS test statistic and its critical values at 10%, 5%, and 1% significance levels in Table 5. We find that the CIPS test statistic, −4.318, is larger than these three critical values in absolute terms. Therefore, we reject the null hypothesis that there exists a unit root in the panel data of $x_{i,t}$.

Table 5: A Panel Unit-root Test for Rental Growth Rates

| CIPS | Critical Value 10% | Critical Value 5% | Critical Value 1% |
|---|---|---|---|
| −4.318 | −2.080 | −2.160 | −2.300 |

Notes: The table presents CIPS test statistic and its critical values for 10%, 5%, and 1% significance levels. The CIPS test statistic, −4.318, is larger than these three critical values in absolute terms. Therefore, we reject the null hypothesis that there exists a unit root in the panel data of $x_{i,t}$.

Knowing that rental growth rates are stationary, we propose the following panel regression with city fixed effects for the panel of 81 MSAs during 1985-2007,

$$x_{i,t} = \zeta_i + \eta x_{i,t-1} + \varpi \text{Year}_t + \varepsilon_{i,t}, \tag{4}$$

where $\text{Year}_t$ is the year dummy and $\varepsilon_{i,t}$ is the error term. We also omit year fixed effects in an alternative specification. The estimate of the autocorrelation term $\eta$ is 0.044 in the specification with year fixed effects. And it is 0.122 in the specification without year fixed effects. To test whether different MSAs have the same average growth rate of rental prices, we conduct a joint hypothesis test against the null hypothesis, $H_0 : \zeta_1 = \zeta_2 = ... = \zeta_I$. Table 6 shows that the chi-squared test statistics, 22.06 and 28.78, are smaller than their critical value at the 10% significance level, 96.578. Thus the test results in both specifications cannot reject the null hypothesis. Different MSAs have the same average growth rate of rental prices.

Table 6: Test Statistics for City Fixed Effects in Rental Growth Rates

| | Chi-square Statistic | p-value |
|---|---|---|
| Without Year Fixed Effects | 22.06 | 1.0000 |
| With Year Fixed Effects | 28.78 | 1.0000 |

Notes: The table presents the statistics of the joint hypothesis test against the null hypothesis, $H_0 : \zeta_1 = \zeta_2 = ... = \zeta_I$. The chi-squared test statistics, 22.06 and 28.78, are smaller than their critical value at the 10% significance level, 96.578. Thus the test results in both specifications cannot reject the null hypothesis. Different MSAs have the same average growth rate of rental prices.

## 3.3 House price-to-rent ratios

In this subsection we conduct a panel-data unit-root test on house price-to-rent ratios and show that they are non-stationary. We also find that different MSAs have the same average growth rate of price-to-rent ratios.

Let $d_{i,t}$ be the rental price of MSA $i$ in period $t$. Let $y_{i,t} \equiv p_{i,t}/d_{i,t}$ denote the price-to-rent ratio of MSA $i$ in period $t$. In Appendix B.3 we show that there exists cross-sectional dependence in the panel data of house price-to-rent ratios. Therefore, we conduct the CIPS panel unit-root test for house price-to-rent ratios. The CADF regression takes the following form,

$$\Delta \log y_{i,t} = \tilde{\tilde{g}}_i + \tilde{\tilde{h}}_i \log y_{i,t-1} + \tilde{\tilde{m}}_i \overline{\log y_{t-1}} + \sum_{k=0}^{K} \tilde{\tilde{n}}_{ik} \Delta \overline{\log y_{t-k}} + \sum_{k=1}^{K} \tilde{\tilde{l}}_{ik} \Delta \log y_{i,t-k} + \tilde{\tilde{v}}_{it}, \quad (5)$$

where $\Delta$ is the first-difference operator, $\overline{\log y_t}$ is the cross-sectional average of $\log y_{i,t}$, and $K$ is the lag length. Due to the constraint of the sample length, we set $K = 1$. The null hypothesis assumes that the Log of price-to-rent ratios has a unit root. We reprot our CIPS test statistic and its critical values at 10%, 5%, and 1% significance levels in Table 7. We find that the CIPS test statistic, $-1.577$, is smaller than the critical value of the 10% level in absolute terms. Therefore, we cannot reject the null hypothesis that there exists a unit root in the panel data of $\log y_{i,t}$.

Table 7: A Panel Unit-root Test for House Price-to-rent Ratios

| CIPS | Critical Value 10% | Critical Value 5% | Critical Value 1% |
|---|---|---|---|
| $-1.577$ | $-2.080$ | $-2.160$ | $-2.300$ |

Notes: The table presents the CIPS test statistic and its critical values for 10%, 5%, and 1% significance levels. The CIPS test statistic, $-1.577$, is smaller than the critical value of the 10% level in absolute terms. Therefore, we cannot reject the null hypothesis that there exists a unit root in the panel data of $\log y_{i,t}$.

To test whether different MSAs have the same average growth rate of price-to-rent ratios, we conduct a joint hypothesis test after running a panel regression of the growth rates with city fixed effects,

$$\Delta \log y_{i,t} = \omega_i + \xi \text{Year}_t + \iota_{i,t}, \quad (6)$$

where $\text{Year}_t$ is the year dummy and $\iota_{i,t}$ is the error term. A joint hypothesis test against the null hypothesis, $H_0 : \omega_1 = \omega_2 = ... = \omega_I$, is conducted. Table 8 shows that the chi-squared test statistic 85.11 is smaller than its critical value at the 10% significance level, 96.578. And the $p$-value is 0.3270. Thus test results cannot reject the null hypothesis. Different MSAs have the same average growth rate of price-to-rent ratios.

Table 8: Test Statistics for City Fixed Effects in House Price-to-rent ratios

| Chi-square Statistic | $p$-value |
|---|---|
| 85.11 | 0.3270 |

Notes: The table presents the statistics of the joint hypothesis test against the null hypothesis, $H_0 : \omega_1 = \omega_2 = ... = \omega_I$. The chi-squared test statistic 85.11 is smaller than its critical value at the 10% significance level, 96.578. And the $p$-value is 0.3270. Thus test results cannot reject the null hypothesis. Different MSAs have the same average growth rate of price-to-rent ratios.

## 4 An island model

We present our theoretical framework in this section. Time is discrete and infinite with $t = 0, 1, 2...$ The economy is divided into $I > 1$ segmented islands. Each island is populated with infinitely-lived agents of measure 1. Agents are not allowed to move across islands. Agents on all islands share identical preferences. On each island there is an asset-pricing problem as in Lucas Jr (1978).

Houses are the only asset in the economy. An agent can only purchase houses on the island where she lives. Each unit of house delivers a stream of stochastic rents. Let $c_t$ denote the consumption in period $t$ and $U(c_t)$ be the period utility function. A representative agent on island $i$ chooses a sequence of consumption and a sequence of housing units to maximize the expected present value of her lifetime utility,[12]

$$E_0 \sum_{t=0}^{\infty} \beta^t U(c_t),$$

where $\beta \in (0, 1)$ denotes the time discount factor. The agent's budget constraint in period $t$ can be written as

$$c_t + p_t s_t = (p_t + d_t) s_{t-1}, \text{ with } c_t, s_t > 0,$$

[12]Since no mobility across islands is allowed, we omit the subscript for island $i$ to simplify the notation.

where $s_t$ represents the unit of houses purchased, $p_t$ is the housing price, and $d_t$ is the stochastic housing rent in period $t$. The growth rate of housing rents, $x_t \equiv \log(d_t/d_{t-1})$, is assumed to follow the process,

$$x_t - \bar{x} = \rho(x_{t-1} - \bar{x}) + \varepsilon_t, \ \ \varepsilon_t \sim N\left(0, \sigma_\varepsilon^2\right), \tag{7}$$

where $|\rho| < 1$. We assume that each island in our model has one unit of houses.[13]

From the first order conditions of the agent's problem, we have

$$p_t = \beta E_t \left[\frac{U'(c_{t+1})}{U'(c_t)}(p_{t+1} + d_{t+1})\right]. \tag{8}$$

When there is no technology to store housing rents, and the housing supply is normalized to be 1 in each island, we have $c_t = d_t$ for all $t$. The individual utility function takes a constant relative risk aversion (CRRA) form, i.e. $U(c_t) = \frac{(c_t)^{1-\alpha}-1}{1-\alpha}$, where the parameter $\alpha > 0$ is the coefficient of relative risk aversion. Thus from equation (8) we have

$$p_t = \beta E_t \left[\left(\frac{d_{t+1}}{d_t}\right)^{-\alpha}(p_{t+1} + d_{t+1})\right]. \tag{9}$$

Let us denote by $y_t \equiv p_t/d_t$ the house price-to-rent ratio in period $t$. Together with the definition of $x_t$, equation (9) is equivalent to

$$y_t = \beta E_t \left[(y_{t+1} + 1)\exp\left[(1-\alpha)x_{t+1}\right]\right]. \tag{10}$$

Following Lansing (2010), we let $z_t \equiv \beta(y_t + 1)\exp[(1-\alpha)x_t]$, which is a composite variable that depends on both the growth rate of rents and the price-to-rent ratio. Equation (10) leads to

$$z_t = \beta\left[E_t(z_{t+1}) + 1\right]\exp\left[(1-\alpha)x_t\right]. \tag{11}$$

[13] Favara and Song (2014) show that housing prices not only have different trends in different cities, but also display heterogeneous short-run dynamics. They use a user-cost model to study how dispersed information affects the equilibrium housing price. They also assume that housing supply is inelastic in each MSA and agents are not allowed to move across MSAs. Thus each MSA in their model can also be viewed as a closed island economy.

Equation (10) also implies that $y_t = E_t(z_{t+1})$.

## 4.1 The fundamental solution

In this subsection, we investigate the fundamental solution of price-to-rent ratios and its implications for housing prices.

Under the transversality condition,

$$\lim_{T\to\infty} E_t\left[\beta^T \frac{U'(c_{t+T})}{U'(c_t)} p_{t+T}\right] = 0, \tag{12}$$

the iteration of equation (8) leads to the following housing price equation,

$$p_t = E_t \sum_{j=1}^{\infty}\left[\beta^j \frac{U'(c_{t+j})}{U'(c_t)} d_{t+j}\right],$$

which is the counterpart of equation (10) in Van Nieuwerburgh and Weill (2010). The fundamental housing price can be completely supported by rents.

Let $z_t^f$ be the fundamental solution to equation (11), i.e. the solution satisfying both equation (11) and the transversality condition. We have the following proposition from Lansing (2010) on an approximate analytical expression of the fundamental solution to equation (11):

**Proposition 1** *The fundamental solution to equation (11) can be approximated by*

$$z_t^f = \exp\left[a_0 + a_1\rho\left(x_t - \bar{x}\right)\right], \tag{13}$$

*where $a_1$ solves*

$$a_1 = \frac{1-\alpha}{1-\rho\beta\exp\left[(1-\alpha)\bar{x} + \frac{1}{2}a_1^2\sigma_\varepsilon^2\right]}, \tag{14}$$

*and*

$$a_0 = \log\left[\frac{\beta\exp\left[(1-\alpha)\bar{x}\right]}{1-\beta\exp\left[(1-\alpha)\bar{x} + \frac{1}{2}a_1^2\sigma_\varepsilon^2\right]}\right], \tag{15}$$

*provided $\beta\exp\left[(1-\alpha)\bar{x} + \frac{1}{2}a_1^2\sigma_\varepsilon^2\right] < 1$.*

**Proof** See Proposition 1 of Lansing (2010). ∎

From Proposition 1 we have the fundamental solution of the price-to-rent ratio,

$$y_t^f = E_t\left(z_{t+1}^f\right) = \exp\left(a_0 + a_1\rho\left(x_t - \bar{x}\right) + \frac{1}{2}a_1^2\sigma_\varepsilon^2\right). \tag{16}$$

Formula (16) implies that $y_t^f$ is stationary. In Section 3.3 we show that the house price-to-rent ratio process is non-stationary in data. Thus the fundamental solution of housing prices cannot match data. We will introduce a non-stationary part to the price-to-rent ratio in the model.

## 4.2 Rational bubble solutions

The fundamental solution to equation (11), $z_t^f$, satisfies the transversality condition. There are other solutions to equation (11) which do not satisfy the transversality condition.

We consider a bubble component of $z_t$, $z_t^b$. Let

$$z_t = z_t^f + z_t^b. \tag{17}$$

Substituting equation (17) into equation (11) yields

$$z_t^b = \beta E_t\left(z_{t+1}^b\right)\exp\left[\left(1-\alpha\right)x_t\right]. \tag{18}$$

We define the bubble component of the price-to-rent ratio, $y_t^b$, as $y_t^b \equiv \frac{1}{\beta}z_t^b\exp\left[-\left(1-\alpha\right)x_t\right]$, such that equation (18) implies that $y_t^b = E_t\left(z_{t+1}^b\right)$. Thus we have

$$\underbrace{E_t\left(z_{t+1}\right)}_{y_t} = \underbrace{E_t\left(z_{t+1}^f\right)}_{y_t^f} + \underbrace{E_t\left(z_{t+1}^b\right)}_{y_t^b},$$

which implies that any $y_t$ obtained by $y_t = y_t^f + y_t^b$, satisfies equation (10). The following proposition can be obtained from Lansing (2010).

**Proposition 2** *There exists a continuum of intrinsic rational bubbles of the form*

$$z_t^b = z_{t-1}^b\exp\left[\lambda_0 + \lambda_1\left(x_t - \bar{x}\right) + \lambda_2\left(x_{t-1} - \bar{x}\right)\right], \ \textit{with}\ z_0^b > 0,$$

*where $\lambda_0$, $\lambda_1$, and $\lambda_2$ satisfies the following two conditions,*

$$\lambda_2 = -\left(\rho\lambda_1 + 1 - \alpha\right), \tag{19}$$

*and*

$$\frac{1}{2}\lambda_1^2\sigma_\varepsilon^2 + (1-\alpha)\,\bar{x} + \log(\beta) + \lambda_0 = 0. \tag{20}$$

**Proof** See Proposition 2 of Lansing (2010). ∎

We have multiple equilibria since we have three unknowns $\lambda_0$, $\lambda_1$, and $\lambda_2$ for equations (19) and (20).

The equations above have two roots of $\lambda_1$. The solution with negative $\lambda_0$ will eventually shrink to zero, while the solution with positive drift implies the price-to-rent ratio will grow unboundedly.

We thus have the bubble component of the price-to-rent ratio,

$$y_t^b = y_{t-1}^b \exp\left[\lambda_0 + (\lambda_1 - (1-\alpha))\,(x_t - \bar{x}) + (\lambda_2 + 1 - \alpha)\,(x_{t-1} - \bar{x})\right], \text{ with } \; y_0^b > 0. \tag{21}$$

Formula (21) shows that $y_t^b$ has stochastic growth rates. Thus the price-to-rent ratio,

$$y_t = y_t^f + y_t^b,$$

can easily generate an increasing dispersion. This is the key mechanism, through which our model could match the rapid growth in the dispersion of housing prices. Through calibrating the model, we illustrate this mechanism in Subsection 5.3. Even after we pick values for $\lambda_0$, $\lambda_1$, and $\lambda_2$ in formula (21), we still need finding the initial $y_0^b$. In our quantitative analyses, we calibrate $y_0^b$ for each MSA from data.

# 5 Quantitative analyses

In this section, we calculate housing prices in the island economy. Each island corresponds to an MSA. We intend to investigate whether our model can generate the rapid growth in the dispersion of housing prices. In the following, we first discuss parameterizations and calibration strategies. After we calculate housing prices with only the fundamental solution, we add bubbles to our results. We also conduct some robustness checks.

## 5.1 Calibrations

The model is calibrated to the housing market in the United States during 1975-2007. Each period in the model corresponds to one year in data. The time discount factor, $\beta$, generally lies in the interval $[0.95, 0.99]$ in the literature. We set it to 0.97. We set the coefficient of relative risk aversion, $\alpha$, to 1.2, which is also within the reasonable range in existing studies.[14] We summarize these predetermined parameters in Table 9.

Table 9: Benchmark Parameterization

| Model | Parameter | Value |
|---|---|---|
| Time discount factor | $\beta$ | 0.97 |
| Risk aversion coefficient | $\alpha$ | 1.2 |

Notes: The values of parameters in the table are predetermined.

We pin down the rental growth process by running the following panel regression,

$$x_{i,t} = \kappa + \rho x_{i,t-1} + \varepsilon_{i,t},\ \ \varepsilon_{i,t} \sim N(0, \sigma_\varepsilon^2). \tag{22}$$

The estimation yields $\widehat{\rho} = 0.132$, which is statistically significantly different from zero. Equation (22) also implies that the estimate of the average rental growth rate $\bar{x}$ is equal to $\widehat{\kappa}/(1 - \hat{\rho})$. We report these estimated parameter values in Table 10.

Table 10: Parameter Values of the Rental Growth Process

| Model | Parameter | Value |
|---|---|---|
| Autocorrelation coefficient | $\rho$ | 0.132 |
| Mean of $x$ | $\bar{x}$ | 0.00369 |
| Standard deviation of the error term | $\sigma_\varepsilon$ | 0.0404 |

Notes: The values of parameters in the table are estimated by running a panel regression.

The theoretical framework in Section 4 enables us to calculate the house price-to-rent ratio. We then obtain the housing price by multiplying the price-to-rent ratio with the

[14] The quantitative results are fairly robust when we try different combinations of $\alpha$ and $\beta$.

corresponding rental price in each MSA. Thus we need the rental price in each MSA. However, due to the limited availability of rental data, we estimate rental prices during 1975-1983. In Figure 3, we present a snapshot of the kernel density for rental price data during 1984-2007.

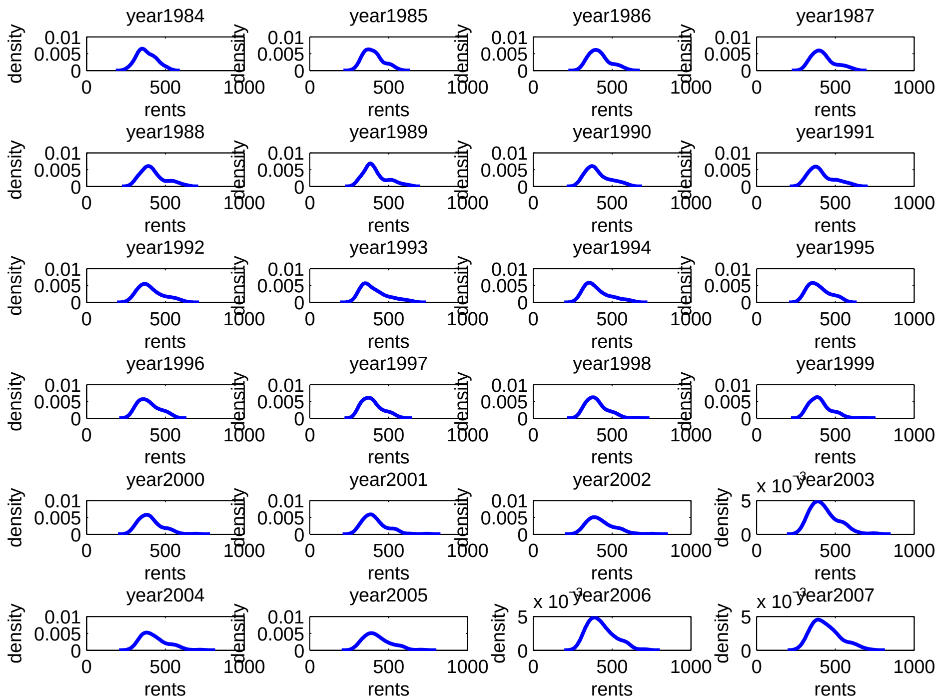



Figure 3: Rental Price Distributions: 1984-2007

Notes: The figure presents a snapshot of the kernel density for rental price data during 1984-2007.

The patterns in Figure 3 suggests that rental prices across MSAs follow a log-normal distribution. But means and variances potentially change over time. Therefore, we assume that rental prices in year 1974 and year 1975 follow a joint log-normal distribution $\ln N(\mu, \Sigma)$, where $\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$ denotes the average rental prices in 1974 and in 1975 respectively. And $\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix}$ is the variance-covariance matrix, whose diagonal elements represent the variances of rental prices in 1974 ($\sigma_1^2$) and in 1975 ($\sigma_2^2$) respectively. The off-diagonal element ($\sigma_{12}$) captures the correlation between rental prices in 1974 and those in 1975. Given this assumed joint distribution of rental prices in 1974 and 1975, we can use the rental growth process (22) to simulate rental prices during 1976-2007. We estimate $\mu_1$, $\mu_2$, $\sigma_1$, $\sigma_2$, and $\sigma_{12}$ using the simulated method of moments (SMM). We jointly estimate their values to minimize the distance between means and variances of simulated rental prices during 1984-2007 and their counterparts in data. We report estimated parameter values in Table 11.

Table 11: SMM Estimation Results

| Parameter | Value |
|---|---|
| $\mu_1$ | 3.81 |
| $\mu_2$ | 5.53 |
| $\sigma_1$ | 0.016 |
| $\sigma_2$ | 0.025 |
| $\sigma_{12}$ | 0.0199 |

Notes: The values of parameters in the table are estimated by using the simulated method of moments (SMM).

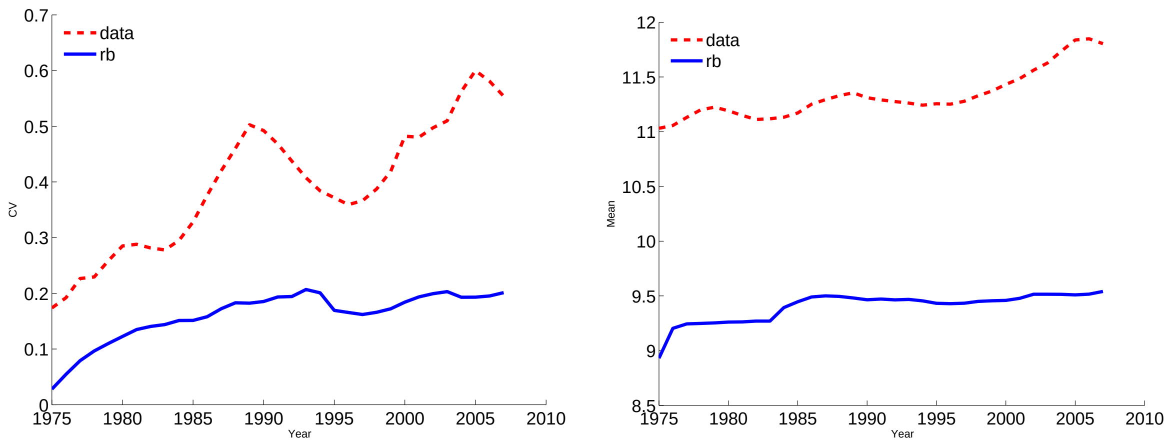
We randomly draw rental prices of 81 MSAs in 1974 and in 1975 from the log-normal distribution specified in Table 11. These enable us to compute the initial rental growth rate for each MSA in 1975. Given the initial distribution $\{x_{i,0}\}$, we can then simulate a panel of $\{x_{i,t}\}$ during 1976-1984 using equation (22). Then we can back out rental prices during 1976-1984. However, our simulated rental prices during 1975-1984 are "anonymous." In order to connect the simulated rental prices to the rental data in each MSA, we label simulated rental prices such that the simulated rentals in 1984 obey the same rank as the 1984 rental data.

## 5.2 The fundamental solution and excessive dispersion

We first calculate fundamental housing prices for each MSA. As in Section 4.1, we can approximate the fundamental solution of the price-to-rent ratio by using equation (16). We determine values of $a_1$ and $a_0$ by formulae (14) and (15). Given our benchmark parameterization, we find $a_1 = -0.23$ and $a_0 = 3.45$.

After we obtain housing prices for 81 MSAs, we can calculate the cross-sectional mean and CV of housing prices for each year during 1975-2007. In Figure 4, we plot the mean and the CV of the simulated fundamental housing prices and their counterparts in data.

As depicted in Figure 4, simulated fundamental housing prices fail to match either the




(a) CV: fundamental housing prices vs. data (b) Mean: fundamental housing prices vs. data

Figure 4: Fundamental Housing Prices

Notes: We approximate the analytical solution for the fundamental price-to-rent ratio according to equation (16). We then obtain the housing price by multiplying the price-to-rent ratio with the corresponding rental price in each MSA. The y-axis of panel (b) is in the logarithm scale.

mean or the CV in data. Fundamental housing prices cannot replicate the rapid increase in the dispersion of housing prices across MSAs.

Panel (a) of Figure 4 also shows that the cross-sectional CV of housing prices is larger than that of housing prices implied by the fundamental solution for each year during 1975-2007. Shiller (1981) and LeRoy and Porter (1981) show that time series of shock prices exhibit excess volatility, i.e. the CV of time series of stock prices is larger than the CV of the expected present value of future real dividends. Bulkley, Snell, and Tonks (1996) study the cross-sectional dispersion of individual company share prices. They find that stock prices of a large sample of firms in the United States are excessively dispersed compared with ex post rational stock prices calculated from subsequent dividend realizations. Through our calculations, we find excessive dispersion for housing prices in the United States.

## 5.3 Results with a rational bubble

In Section 5.2 we find that fundamental housing prices cannot match the rapid rise in the dispersion of housing prices across MSAs. Then we add a bubble component to the house price-to-rent ratio. Thus we have the house price-to-rent ratio, $y_{i,t}$,

$$y_{i,t} = y^f_{i,t} + y^b_{i,t},$$

for $i = 1, 2, \cdots I$.

We have calculated the fundamental component of the price-to-rent ratio $\left\{y^f_{i,t}\right\}$ in Section 5.2. Now we compute the bubble component of the price-to-rent ratio $\left\{y^b_{i,t}\right\}$ by using equation (21).

Given the initial value of the fundamental component of the price-to-rent ratio $y^f_{i,0}$ for each MSA, we can derive the initial value of the bubble component of the price-to-rent ratio $y^b_{i,0}$ by

$$y^b_{i,0} = y_{i,0} - y^f_{i,0},$$

where $y_{i,0}$ is the house price-to-rent ratio in year 1975 in data.

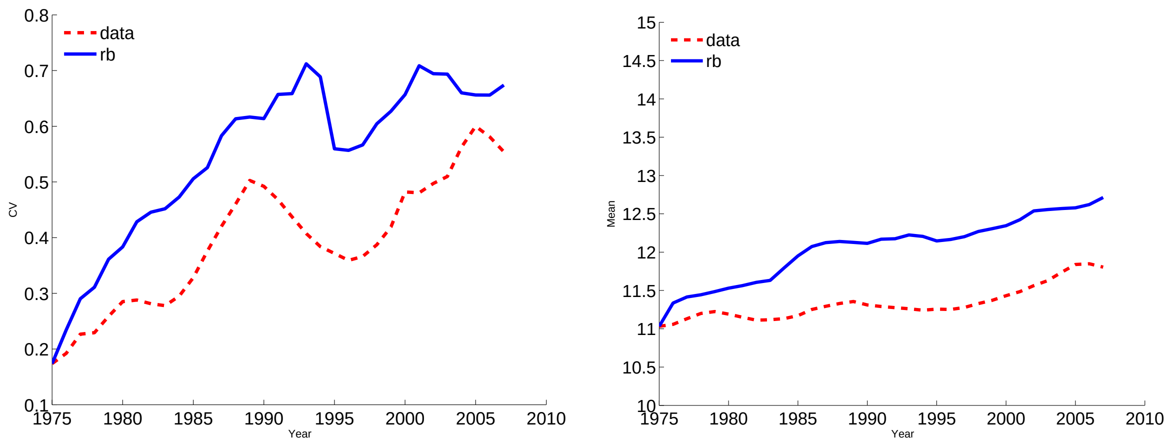
Applying equation (21) we also need to know $\lambda_0$, $\lambda_1$, and $\lambda_2$. We pin down $\lambda_1$ and $\lambda_2$ from equations (19) and (20). We calibrate the drift term $\lambda_0$ to minimize the distance between the average growth rate of housing prices in the model and their counterparts in the panel of 81 MSAs during 1985-2007. $\lambda_0$ solves

$$\min_{\lambda_0} \quad \sum_{t=1985}^{2007} \left[ \frac{\sum_{i=1}^{I} (p^m_{i,t}/p^m_{i,t-1} - 1)}{I} - \frac{\sum_{i=1}^{I} (p^d_{i,t}/p^d_{i,t-1} - 1)}{I} \right]^2,$$

where $p^d_{i,t}$ is the housing price in MSA $i$ at time $t$ in data and $p^m_{i,t}$ is the corresponding housing price in the model. We find $\lambda_0 = 0.030$, $\lambda_1 = 1.305$, and $\lambda_2 = 0.028$.

After we calculate $y_{i,t}$, we then multiply $y_{i,t}$ with its corresponding rental price, to generate the housing price. To calculate the cross-sectional mean and CV of housing prices, we assign an equal weight to each MSA for benchmark results. But the results are fairly robust when we use housing units as weights. In Figure 5, we plot benchmark results together with their counterparts in data.

Evidently, our calculated series of the mean and the CV of housing prices match data well. When we use the CV to measure the dispersion of housing prices, data show a rapid increase in the CV of housing prices during 1975-2007. The CV in data increases from 0.17




(a) CV: model vs. data

(b) Mean: model vs. data

Figure 5: Benchmark Results of Housing Prices: 1975-2007

Notes: Our calculations incorporate a balanced panel of 81 MSAs. Housing prices are plotted during 1975-2007. We apply true rental price data during 1984-2007, and use simulated rental prices during 1975-1983 for each MSA. The y-axis of panel (b) is in the logarithm scale.

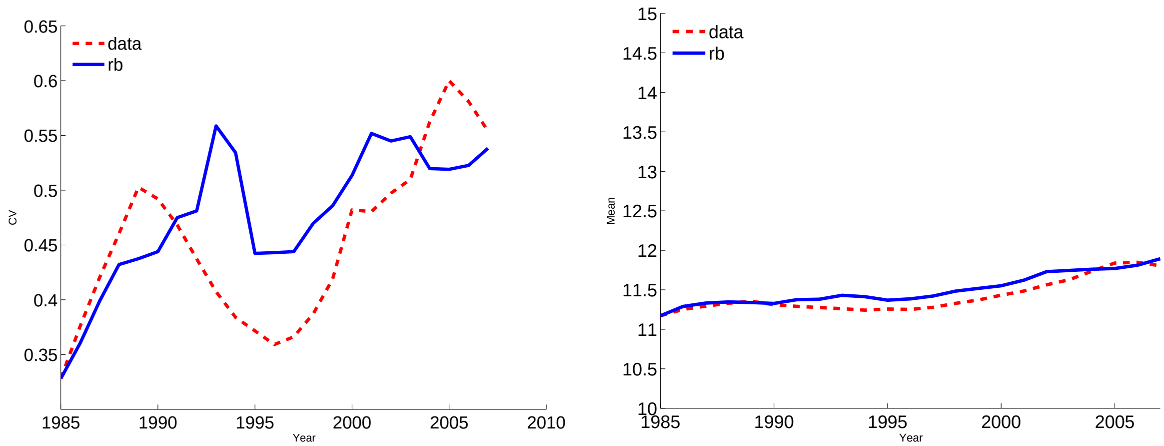
in 1975 to 0.55 in 2007. Our model predicts that the CV rises from 0.17 in 1975 to 0.67 in 2007. In 1975 our model has the same CV as in data since we use housing prices in 1975 in data as initial housing prices in our model. We are successful in predicting the rapid increase in the CV of housing prices during this period.

## 5.4 Robustness

In this section, we perform several robustness exercises. Firstly, we evaluate how well our model can match data after 1985, when we do not need to perform any estimation of rental prices. Secondly, we also simulate housing prices using the entire simulated rental prices, while we use actual rental data after 1985 in the benchmark simulation. Finally, we try to produce housing prices for the unbalanced panel of 330 MSAs. We find that our results are robust for all these alternatives.

**Calculations during 1985-2007** We estimate rental prices during 1975-1983 in the benchmark simulation, because we do not have rental data for that period. To make sure that our results are not driven by the estimation strategy, we now only calculate housing prices using rental prices from data without any estimation. As rental data are available after 1984, we calculate housing prices after 1985, when we can obtain initial rental growth rates,

$\{x_{i0}\}$.[15] We calculate the house price-to-rent ratio $y_{i,t} = y^f_{i,t} + y^b_{i,t}$ for $i = 1, 2, \cdots I$ and $t = 1985, 1986, \cdots 2007$. Then we multiply $\{y_{i,t}\}$ with the corresponding rental price in each MSA to obtain the housing price. Results are presented in Figure 6. The CV in data increases from 0.33 in 1985 to 0.55 in 2007. Our model predicts that the CV rises from 0.33 in 1985 to 0.54 in 2007.[16] We also replicate the rapid rise in the CV of housing prices during this period.




(a) CV: model housing prices vs. data

(b) Mean: model housing prices vs. data

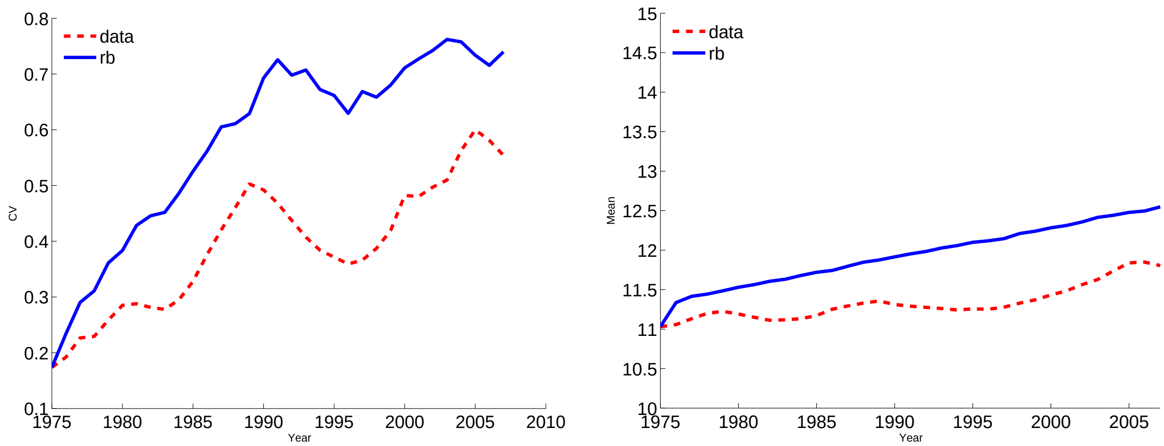
Figure 6: Robustness Results: 1985-2007 without Estimations

Notes: We directly calculate the house price-to-rent ratio, and then obtain the housing price by multiplying it with the rental price of each MSA. Data of rental prices are available since 1984. Thus we do not need to perform estimations after 1985. The y-axis of panel (b) is in the logarithm scale.

**Purely simulated rental prices** In our benchmark simulation, we use actual rental price data after 1985. Since rental prices are important inputs for our simulations, we now inject the entire simulated rental price panel into the economy instead, and evaluate how estimation results may vary accordingly. Results are reported in Figure 7. The CV in data increases from 0.17 in 1975 to 0.55 in 2007. Our model predicts that the CV rises from 0.17 in 1975 to 0.74 in 2007. We also generate the rapid growth in the CV of housing prices during this period.

[15] We use the value of $\lambda_0$ as in our benchmark simulation. There $\lambda_0$ is calibrated to minimize the distance between the average growth rate of housing prices in the model and their counterparts in data during 1985-2007.

[16] In 1985 our model has the same CV as in data since we use housing prices in 1985 in data as initial housing prices in robustness exercise.




(a) CV: model housing prices vs. data

(b) Mean: model housing prices vs. data

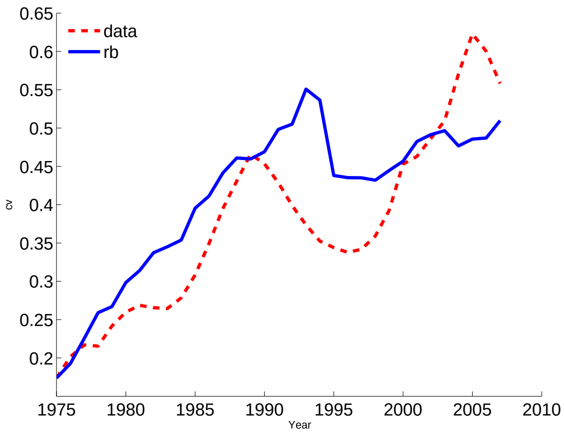
Figure 7: Robustness Results: All Rental Prices are Simulated

Notes: In this robustness check, the entire sequence of rental prices during 1975-2007 is from simulations. The y-axis of panel (b) is in the logarithm scale.

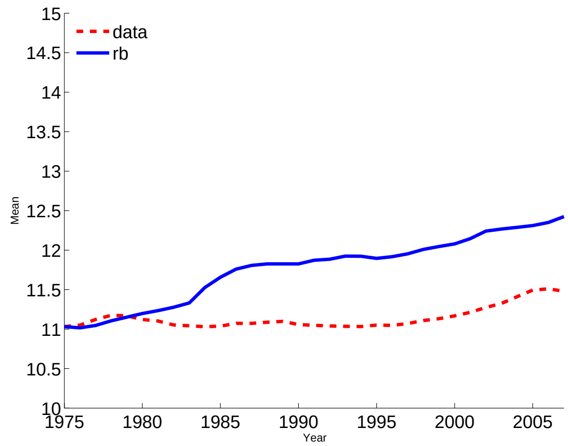
**An unbalanced panel** We also extend our analysis to the unbalanced panel of 330 MSAs. We re-estimate equation (22) for the rental growth rate process, $\{x_{i,t}\}$, in the unbalanced panel. The estimate of the autocorrelation coefficient, $\rho$, now is 0.059. The estimate of the average rental growth rate, $\bar{x}$, now is 0.147%. And the estimate of the standard deviation of the error term, $\sigma_\varepsilon$, now is 0.042. As in Section 5.3, we find values for $\lambda_0$, $\lambda_1$, and $\lambda_2$ for 330 MSAs. Now $\lambda_0 = 0.030$, $\lambda_1 = 0.633$, and $\lambda_2 = 0.163$. We then replicate housing prices in the unbalanced panel of 330 MSAs. Simulation results are reported in Figure 8. The CV in data increases from 0.17 in 1975 to 0.56 in 2007. Our model predicts that the CV rises from 0.17 in 1975 to 0.51 in 2007. We also produce the rapid rise in the CV of housing prices during this period.

# 6 An extension: sunspots and bubble burst

It is widely documented that the incidence of the 2007 global financial crisis was caused by the burst of housing bubbles in the United States. It seems that there are different regimes of housing price dynamics. Since the price-to-ratio has a stochastic growth component in our benchmark model, and thus it eventually leads to explosive dispersion of housing prices.




(a) CV: model housing prices vs. data




(b) Mean: model housing prices vs. data

Figure 8: Robustness Results: an Unbalanced Panel

Notes: We re-estimate the process $\{x_{i,t}\}$ for 330 MSAs, and then replicate housing prices of the unbalanced panel. The y-axis of panel (b) is in the logarithm scale.

In this section, we present an extension of the benchmark model. In the extension, bubbles eventually burst in the long-run. But before bubbles burst, there is a rapid growth in the dispersion of housing prices.

To model regime switching of housing price dynamics, we introduce an extrinsic uncertainty into our benchmark model. Kashiwagi (2014) finds a rational expectations sunspot equilibrium in a model based on search and matching theory. His model can generate a stable path of rental prices and a rapid growth of housing prices. However, he does not investigate the implication of a sunspot equilibrium on the cross-sectional dispersion of housing prices. As in Kashiwagi (2014), the extrinsic uncertainty in our model represents the confidence state, which could be either the high $(H)$ state or the low $(L)$ state. This sunspot event is unrelated to fundamentals of the economy.

Equations (19) and (20) are two key equations generating the price-to-rent ratio in our rational bubble economy. We need these two equations to pin down three parameters $\{\lambda_0, \lambda_1, \lambda_2\}$, and this allows us one degree of freedom. In quantitative analyses of our benchmark model, we calibrate $\lambda_0$. In this section, we instead let the parameter $\lambda_0$ take four possible values, which depend on confidence states of the current and the next period. We denote by $\lambda_0^{ij}$ $(i, j \in \{H, L\})$ these four values. Superscripts of $\lambda_0^{ij}$ mean that the confidence state in the current period is $i$ and the confidence state in the next period is $j$. We also denote by $\psi_t$ the confidence state in period $t$. Let $z_t^b(\psi_t)$ represent $z_t^b$ in the confidence state $\psi_t$. Let $\theta \equiv 1-\alpha$. The following proposition characterizes the sunspot equilibrium.

**Proposition 3** *There exists a continuum of rational bubbles of the form,*

$$z_{t+1}^b(H) = \begin{cases} z_t^b(H)\exp\left(\lambda_0^{HH} + \lambda_1(x_{t+1}-\bar{x}) + \lambda_2(x_t - \bar{x})\right), \text{ if } \psi_t = H \\ z_t^b(L)\exp\left(\lambda_0^{LH} + \lambda_1(x_{t+1}-\bar{x}) + \lambda_2(x_t - \bar{x})\right), \text{ if } \psi_t = L, \end{cases}$$

*and*

$$z_{t+1}^b(L) = \begin{cases} z_t^b(H)\exp\left(\lambda_0^{HL} + \lambda_1(x_{t+1}-\bar{x}) + \lambda_2(x_t - \bar{x})\right), \text{ if } \psi_t = H \\ z_t^b(L)\exp\left(\lambda_0^{LL} + \lambda_1(x_{t+1}-\bar{x}) + \lambda_2(x_t - \bar{x})\right), \text{ if } \psi_t = L. \end{cases}$$

*The confidence state follows a Markov process, and $\pi_{ij}$ represents the transition probability from the state $i$ in the current period to the state $j$ in the next period. Constants $\lambda_0^{HH} > 0$, $\lambda_0^{HL} < 0$, $\lambda_0^{LL} < 0$, $\lambda_0^{LH} >> 0$, $\lambda_1$, $\lambda_2$, $\pi_{HH}$, $\pi_{HL}$, $\pi_{LL}$, and $\pi_{LH}$ satisfy*

$$\lambda_2 = -(\lambda_1\rho + \theta), \tag{23}$$

$$\beta\pi_{LH}\exp\left(\lambda_0^{LH} + \frac{1}{2}(\lambda_1\sigma_\varepsilon)^2 + \theta\bar{x}\right) + \beta\pi_{LL}\exp\left(\lambda_0^{LL} + \frac{1}{2}(\lambda_1\sigma_\varepsilon)^2 + \theta\bar{x}\right) = 1, \tag{24}$$

$$\beta\pi_{HH}\exp\left(\lambda_0^{HH} + \frac{1}{2}(\lambda_1\sigma_\varepsilon)^2 + \theta\bar{x}\right) + \beta\pi_{HL}\exp\left(\lambda_0^{HL} + \frac{1}{2}(\lambda_1\sigma_\varepsilon)^2 + \theta\bar{x}\right) = 1, \tag{25}$$

$$\pi_{HH} + \pi_{HL} = 1, \tag{26}$$

*and*

$$\pi_{LL} + \pi_{LH} = 1. \tag{27}$$

**Proof** See appendix C. ∎

Thus in a sunspot equilibrium we have

$$z_t(H) = z_t^f + z_t^b(H),$$

and

$$z_t(L) = z_t^f + z_t^b(L),$$

where $z_t^f$ is given by equation (13).

Introducing a sunspot event into the benchmark model, we greatly expand our degree of freedom. In total we have ten unknowns and five equations. From equations (24) and (25) we can express $\lambda_0^{LH}$ and $\lambda_0^{HL}$ as

$$\lambda_0^{LH} = \log\left(\frac{1 - \beta\pi_{LL}\exp\left(\lambda_0^{LL} + \frac{1}{2}(\lambda_1\sigma_\varepsilon)^2 + \theta\bar{x}\right)}{\beta\pi_{LH}}\right) - \frac{1}{2}(\lambda_1\sigma_\varepsilon)^2 - \theta\bar{x},$$

and

$$\lambda_0^{HL} = \log\left(\frac{1 - \beta\pi_{HH}\exp\left(\lambda_0^{HH} + \frac{1}{2}(\lambda_1\sigma_\varepsilon)^2 + \theta\bar{x}\right)}{\beta\pi_{HL}}\right) - \frac{1}{2}(\lambda_1\sigma_\varepsilon)^2 - \theta\bar{x}.$$

A sufficient condition to guarantee $\lambda_0^{HL} < 0$ is

$$1 - \beta\pi_{HH}\exp\left(\lambda_0^{HH} + \frac{1}{2}(\lambda_1\sigma_\varepsilon)^2 + \theta\bar{x}\right) < \beta\pi_{HL}.$$

This could help us to generate a scenario of bubble burst.

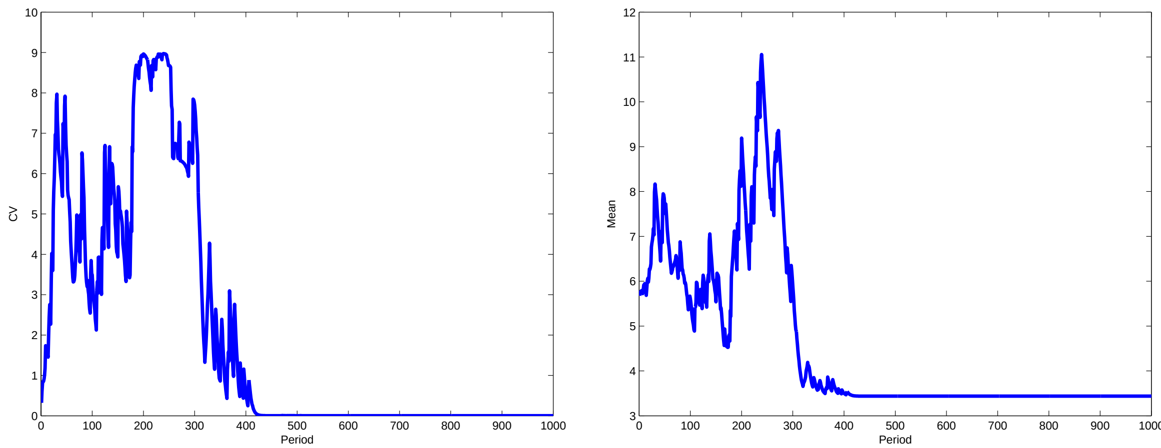
We illustrate a scenario of bubble burst in a simple numerical exercise. We first pick up values for some parameters in Table 12.

Table 12: Parameterization in the Model with Sunspots

| Parameter | Value |
|---|---|
| $\lambda_1$ | 0.5 |
| $\pi_{LH}$ | 0.1 |
| $\pi_{HL}$ | 0.09 |
| $\lambda_0^{LL}$ | $-0.2$ |
| $\lambda_0^{HH}$ | 0.2 |

Notes: The values of parameters in the table are predetermined.

Then we solve $\lambda_0^{LH}$, $\lambda_0^{HL}$ and $\lambda_2$ from equations (23), (24), and (25). We concentrate on the sample of 81 MSAs here. The initial year is still 1975. The process of the rental price growth rate and initial rental price distributions are the same as in quantitative analyses of the benchmark model. Moreover, we set confidence states of all 81 MSAs in 1975 to the state $L$. In each year after 1975, different MSAs may have different realizations of confidence states, even though the confidence state in each MSA follows the same Markov process. We simulate the model economy for $1,000$ periods. Through simulations, we calculate the cross-sectional mean and CV of housing prices. These are reported in Figure 9.




(a) The CV of housing prices

(b) The mean of housing prices

Figure 9: Simulations of a Sunspot Equilibrium

Notes: We simulate sunspot events for 81 MSAs from year 1975. The confidence states of all 81 MSAs in 1975 are set as $L$, but may have different realizations afterwards even if they follow the same Markov process. We simulate the model economy for 1,000 periods. The y-axis of panel (b) is in the logarithm scale.

# 7 Conclusion

We investigate the rapid growth in the dispersion of housing prices across MSAs in the United States during 1975-2007. We first examine several intuitively plausible explanations for this pattern, and find that it is difficult to fully explain it.

We then conduct econometric analyses of panel data. We empirically show that housing prices are non-stationary and housing prices grow at different rates across MSAs. We find that rental growth rates are stationary. While different MSAs have different growth rates of housing prices, they have the same average growth rate of rental prices. Through a panel unit-root test we also find that the Log of price-to-rent ratios follows a random walk process.

To investigate further the rapid growth in the dispersion of housing prices, we set up a parsimonious asset-pricing island model. Each island corresponds to an MSA. We first study the fundamental solution of the asset pricing model. Our calculations show that the cross-sectional CV of housing prices is larger than that of housing prices implied by the fundamental solution for each year during 1975-2007. Housing prices in the United States display excessive dispersion. Also we find that the growth in the dispersion of fundamental housing prices is too slow relative to the pattern in data. Incorporating rational bubble solutions, our calibrated model can simultaneously match four stylized facts in the United States during 1975-2007, the rapid growth in the dispersion of housing prices, the moderate increase in the dispersion of rental prices, the rising mean of housing prices, and the rising mean of rental prices.

One of the important mechanisms that we do not take into account explicitly in this paper is the credit channel of houses. Kiyotaki and Moore (1997) show that collateral channel of assets could amplify the volatility of their prices. Brumm, Grill, Kubler, and Schmedders (2015) find that borrowing against collateral substantially increases the return volatility of long-lived assets. Gelain, Lansing, and Natvik (2015) investigate different impacts of shifting lending standards and movements in the mortgage interest rate, on the boom-bust cycle of the housing market in the United States during 1995-2012. Favilukis, Ludvigson, and Van Nieuwerburgh (2016) find that a relaxation of financing constraints leads to a large boom in house prices. We could follow these researches to investigate the implication of the credit channel on the cross-sectional dispersion of housing prices. We leave this for our future research.


# Acknowledgements

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. We thank Sungbae An, Wen-Tai Hsu, Ji Huang, Nicolas Jacquet, Kevin Lansing, Haoming Liu, Lin Ma, Thomas Sargent, Yifan Shen, Aloysius Siow, Jinli Zeng, and Xin Zheng for helpful discussions and comments.

# References


Brumm, J., M. Grill, F. Kubler, and K. Schmedders (2015): "Collateral requirements and asset prices," *International Economic Review*, 56(1), 1–25.

Bulkley, G., A. Snell, and I. Tonks (1996): "Excessive stock price dispersion: A regression test of cross-sectional volatility," *LSE Financial Markets Group Working Paper 246*.

De Hoyos, R. E., V. Sarafidis, et al. (2006): "Testing for cross-sectional dependence in panel-data models," *Stata Journal*, 6(4), 482.

Eeckhout, J., R. Pinheiro, and K. Schmidheiny (2014): "Spatial sorting," *Journal of Political Economy*, 122(3), 554–620.

Favara, G., and Z. Song (2014): "House price dynamics with dispersed information," *Journal of Economic Theory*, 149, 350–382.

Favilukis, J., S. Ludvigson, and S. Van Nieuwerburgh (2016): "Macroeconomic Implications of Housing Wealth, Housing Finance and Limited Risk Sharing in Equilibrium," *Journal of Political Economy, Forthcoming*.

Froot, K. A., and M. Obstfeld (1991): "Intrinsic bubbles: The case of stock prices," *The American Economic Review*, pp. 1189–1214.

Gelain, P., K. J. Lansing, and G. J. Natvik (2015): "Explaining the boom-bust cycle in the US housing market: A reverse-engineering approach," *Federal Reserve Bank of San Francisco Working Paper 2015-02*.

Glaeser, E. L., and C. G. Nathanson (2015): "Housing Bubbles," in *Handbook of Regional and Urban Economics*, ed. by G. Duranton, J. V. Henderson, and W. Strange, vol. 5B, chap. 11, pp. 702–751. North Holland, Amsterdam.

Granziera, E., and S. Kozicki (2015): "House price dynamics: Fundamentals and expectations," *Journal of Economic Dynamics and Control*, 60, 152–165.

Gyourko, J., C. Mayer, and T. Sinai (2013): "Superstar cities," *American Economic Journal: Economic Policy*, 5(4), 167–99.

Kashiwagi, M. (2014): "Sunspots and self-fulfilling beliefs in the US housing market," *Review of Economic Dynamics*, 17(4), 654–676.

Kiyotaki, N., and J. Moore (1997): "Credit cycles," *Journal of Political Economy*, 105(2), 211–248.

Lansing, K. J. (2010): "Rational and near-rational bubbles without drift," *The Economic Journal*, 120(549), 1149–1174.

LeRoy, S. F. (2004): "Rational exuberance," *Journal of Economic literature*, 42(3), 783–804.

LeRoy, S. F., and R. D. Porter (1981): "The present-value relation: Tests based on implied variance bounds," *Econometrica*, 49(3), 555–574.

Lucas Jr, R. E. (1978): "Asset prices in an exchange economy," *Econometrica: Journal of the Econometric Society*, pp. 1429–1445.

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.

Shiller, R. (1981): "Do stock prices move too much to be justified by subsequent Changes in dividends?," *The American Economic Review*, 71(3), 421–436.

Van Nieuwerburgh, S., and P.-O. Weill (2010): "Why has house price dispersion gone up?," *The Review of Economic Studies*, 77(4), 1567–1606.

Wolff, E. N. (2006): "Changes in household wealth in the 1980s and 1990s in the United States," in *International Perspectives on Household Wealth*, ed. by E. N. Wolff, chap. 4, pp. 107–150. Edward Elgar, Cheltenham.

# Appendices

## A Data sources and variable definitions

**The metropolitan statistical area (MSA):** The metropolitan statistical area (MSA) in the paper is based on 2006 MSA definitions. To account for size differences among MSAs, we follow Van Nieuwerburgh and Weill (2010)) to replace 11 largest MSAs by their constituent metropolitan divisions (MSAD). Our primary sample consists of 81 MSAs during 1975-2007. We include 330 MSAs with unbalanced housing prices in an extended sample for quantitative analyses. The 81 MSAs in the balanced panel are listed in Table A.1.[17] We present those MSAs in our balanced and unbalanced panels in Figure A.1.

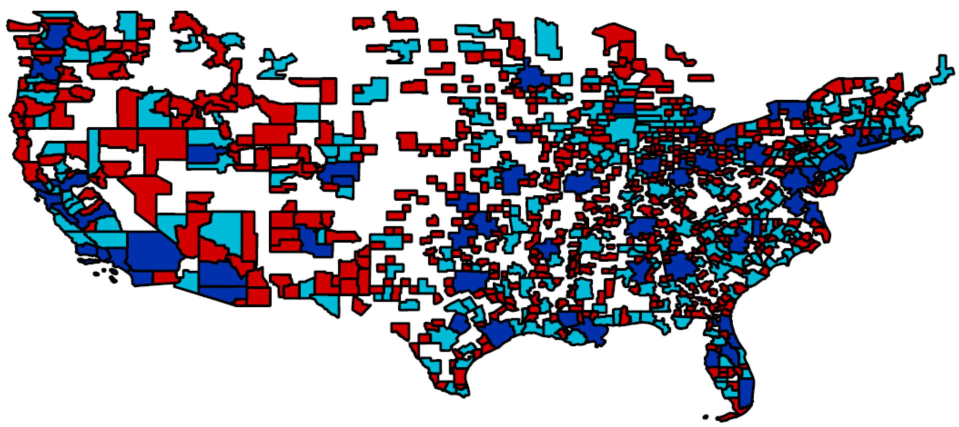


Figure A.1: MSAs in our balanced and unbalanced panel

Notes: Missing: red; Unbalanced panel: light blue; Balanced panel: dark blue

**Housing prices:** We obtain nominal housing price data directly from Van Nieuwerburgh and Weill (2010)). The nominal home value of each MSA is constructed through combining the median single-family home value from the 2000 Census with the Freddie Mac Conventional Mortgage Home Price Index (CMHPI), a repeated-sale housing price index. In 1975 housing prices are only available in 81 MSAs. Data of more and more MSAs have become

[17]We list 330 MSAs in the unbalanced panel in a separate spreadsheet.

available over time. By 1996 all 330 MSAs have housing price data.[18] We deflate nominal home values into 1983 dollars using regional non-housing price indices, which are also taken from Van Nieuwerburgh and Weill (2010)).

**Rental prices:** We take nominal rental data directly from Van Nieuwerburgh and Weill (2010)). We deflate nominal rentals into 1983 dollars using regional non-housing price indices, which are also taken from Van Nieuwerburgh and Weill (2010)). We use rental data in all 330 MSAs during 1984-2007.

**Population:** The original data are from Table 24 of *Population and Housing Unit Counts, United States Summary: 2010*. The United States Census Bureau releases five waves of population data from 1970 to 2010 at the county level. We collect data and aggregate them to the MSA level according to 2006 MSA definitions.

**Housing units:** The original data are from Table 24 of *Population and Housing Unit Counts, United States Summary: 2010*. The United States Census Bureau releases five waves of housing unit data from 1970 to 2010 at the county level. We collect data and aggregate them to the MSA level according to 2006 MSA definitions.

[18] See Appendix D.2 of Van Nieuwerburgh and Weill (2010)) for more details.

Table A.1: The List of 81 MSAs in the Balanced Panel

| MSA Code | Metropolian Statistical Area | MSA Code | Metropolian Statistical Area |
|---|---|---|---|
| 10420 | Akron, OH | 33124 | Miami-Fort Lauderdale-Pompano Beach, FL |
| 10740 | Albuquerque, NM | 33340 | Milwaukee-Waukesha-West Allis, WI |
| 12060 | Atlanta-Sandy Springs-Marietta, GA | 33460 | Minneapolis-St. Paul-Bloomington, MN-WI |
| 12420 | Austin-Round Rock, TX | 34980 | Nashville-Davidson-Murfreesboro-Franklin, TN |
| 12580 | Baltimore-Towson, MD | 35004 | New York-Northern New Jersey-Long Island, NY-NJ-PA |
| 12940 | Baton Rouge, LA | 35084 | New York-Northern New Jersey-Long Island, NY-NJ-PA |
| 13644 | Washington-Arlington-Alexandria, DC-VA-MD-WV | 35300 | New Haven-Milford, CT |
| 13820 | Birmingham-Hoover, AL | 35380 | New Orleans-Metairie-Kenner, LA |
| 14484 | Boston-Cambridge-Quincy, MA-NH | 35644 | New York-Northern New Jersey-Long Island, NY-NJ-PA |
| 14860 | Bridgeport-Stamford-Norwalk, CT | 36084 | San Francisco-Oakland-Fremont, CA |
| 15380 | Buffalo-Niagara Falls, NY | 36420 | Oklahoma City, OK |
| 15764 | Boston-Cambridge-Quincy, MA-NH | 36740 | Orlando-Kissimmee, FL |
| 15804 | Philadelphia-Camden-Wilmington, PA-NJ-DE-MD | 37100 | Oxnard-Thousand Oaks-Ventura, CA |
| 15940 | Canton-Massillon, OH | 37964 | Philadelphia-Camden-Wilmington, PA-NJ-DE-MD |
| 16740 | Charlotte-Gastonia-Concord, NC-SC | 38060 | Phoenix-Mesa-Scottsdale, AZ |
| 16974 | Chicago-Naperville-Joliet, IL-IN-WI | 38300 | Pittsburgh, PA |
| 17140 | Cincinnati-Middletown, OH-KY-IN | 38900 | Portland-Vancouver-Beaverton, OR-WA |
| 17460 | Cleveland-Elyria-Mentor, OH | 39300 | Providence-New Bedford-Fall River, RI-MA |
| 17820 | Colorado Springs, CO | 40060 | Richmond, VA |
| 18140 | Columbus, OH | 40140 | Riverside-San Bernardino-Ontario, CA |
| 19124 | Dallas-Fort Worth-Arlington, TX | 40380 | Rochester, NY |
| 19380 | Dayton, OH | 40900 | Sacramento-Arden-Arcade-Roseville, CA |
| 19660 | Deltona-Daytona Beach-Ormond Beach, FL | 41180 | St. Louis, MO-IL |
| 19740 | Denver-Aurora, CO | 41620 | Salt Lake City, UT |
| 19780 | Des Moines-West Des Moines, IA | 41740 | San Diego-Carlsbad-San Marcos, CA |
| 19804 | Detroit-Warren-Livonia, MI | 41884 | San Francisco-Oakland-Fremont, CA |
| 20764 | New York-Northern New Jersey-Long Island, NY-NJ-PA | 41940 | San Jose-Sunnyvale-Santa Clara, CA |
| 22420 | Flint, MI | 42044 | Los Angeles-Long Beach-Santa Ana, CA |
| 22744 | Miami-Fort Lauderdale-Pompano Beach, FL | 42060 | Santa Barbara-Santa Maria-Goleta, CA |
| 23104 | Dallas-Fort Worth-Arlington, TX | 42220 | Santa Rosa-Petaluma, CA |
| 23420 | Fresno, CA | 42644 | Seattle-Tacoma-Bellevue, WA |
| 24660 | Greensboro-High Point, NC | 45300 | Tampa-St. Petersburg-Clearwater, FL |
| 25540 | Hartford-West Hartford-East Hartford, CT | 45780 | Toledo, OH |
| 26180 | Honolulu, HI | 46060 | Tucson, AZ |
| 26420 | Houston-Sugar Land-Baytown, TX | 46140 | Tulsa, OK |
| 26900 | Indianapolis-Carmel, IN | 47260 | Virginia Beach-Norfolk-Newport News, VA-NC |
| 27260 | Jacksonville, FL | 47644 | Detroit-Warren-Livonia, MI |
| 28140 | Kansas City, MO-KS | 47894 | Washington-Arlington-Alexandria, DC-VA-MD-WV |
| 29404 | Chicago-Naperville-Joliet, IL-IN-WI | 48424 | Miami-Fort Lauderdale-Pompano Beach, FL |
| 30780 | Little Rock-North Little Rock-Conway, AR | 48620 | Wichita, KS |
| 31084 | Los Angeles-Long Beach-Santa Ana, CA | | |

Notes: The metropolitan statistical area (MSA) is based on 2006 MSA definitions.

# B Cross-sectional dependence tests

To test the cross-sectional dependence in panel data of $y_{it}$, we run a panel regression,

$$y_{it} = \alpha_i + \beta' \mathbf{x}_{it} + u_{it},\ \ i = 1, ..., I \text{ and } t = 1, ..., T,$$

where $\mathbf{x}_{it}$ is a vector of regressors, $\beta$ is a vector of coefficients, and $\alpha_i$ represents time-invariant fixed effects. The error term $u_{it}$ is assumed to be independent and identically distributed ($i.i.d.$) along time.

The null hypothesis assumes that error terms are independent across $i$. Thus we have

$$H_0 : \rho_{ij} = \rho_{ji} = cor\left(u_{it},\ u_{jt}\right) = 0 \text{ for } i \neq j,$$

versus

$$H_1 : \rho_{ij} = \rho_{ji} \neq 0 \text{ for some } i \neq j,$$

where $\rho_{ij}$ is the product-moment correlation coefficient of $u_{it}$ and $u_{jt}$,

$$\rho_{ij} = \rho_{ji} = \frac{\sum_{t=1}^{T} u_{it} u_{jt}}{\left(\sum_{t=1}^{T} u_{it}^2\right)^{1/2} \left(\sum_{t=1}^{T} u_{jt}^2\right)^{1/2}}.$$

Our panel data have the property of short $T$ and large $I$. Thus we conduct Pesaran's test and Friedman's test for the cross-sectional dependence.[19]

## B.1 Tests for the Log of housing prices

We conduct the cross-sectional dependence tests for the Log of housing prices after running the following panel regression with city fixed effects,

$$\log p_{i,t} = \chi_i + \phi \log p_{i,t-1} + e_{i,t}, \tag{B.1}$$

[19] See De Hoyos, Sarafidis, et al. (2006)) for more details.

where $e_{i,t}$ is the error term. We report test results in Table B.1. Small $p$-values suggest that we should reject the null hypothesis that there does not exist cross-sectional dependence in the panel data of $\log p_{i,t}$.

Table B.1: Cross-sectional Dependence Tests for Housing Prices

| | Pesaran's Test | Friedman's Test |
|---|---|---|
| Test statistic | 92.170 | 646.861 |
| $p$-value | 0.0000 | 0.0000 |

Notes: The table presents the statistics of the cross-sectional dependence test. Small $p$-values suggest that we should reject the null hypothesis that there does not exist cross-sectional dependence in the panel data of $\log p_{i,t}$.

## B.2 Tests for rental growth rates

We conduct cross-sectional dependence tests for rental growth rates after running the following panel regression with city fixed effects,

$$x_{i,t} = \zeta_i + \eta x_{i,t-1} + \varepsilon_{i,t}, \tag{B.2}$$

where $\varepsilon_{i,t}$ is the error term. We report test results in Table B.2. Small $p$-values suggest that we should reject the null hypothesis that there does not exist cross-sectional dependence in the panel data of $x_{i,t}$.

Table B.2: Cross-sectional Dependence Tests for Rental Growth Rates

| | Pesaran's Test | Friedman's Test |
|---|---|---|
| Test statistic | 39.707 | 476.832 |
| $p$-value | 0.0000 | 0.0000 |

Notes: The table presents the statistics of the cross-sectional dependence test. Small $p$-values suggest that we should reject the null hypothesis that there does not exist cross-sectional dependence in the panel data of $x_{i,t}$.

### B.3 Tests for the Log of price-to-rent ratios

We conduct cross-sectional dependence tests for the Log of price-to-rent ratios after running the following panel regression with city fixed effects,

$$\log y_{i,t} = \omega_i + \delta \log y_{i,t-1} + \iota_{i,t}, \tag{B.3}$$

where $\iota_{i,t}$ is the error term. We report test results in Table B.3. Small $p$-values suggest that we should reject the null hypothesis that there does not exist cross-sectional dependence in the panel data of $\log y_{i,t}$.

Table B.3: Cross-sectional Dependence Tests for Price-to-rent Ratios

| | Pesaran's Test | Friedman's Test |
|---|---|---|
| Test statistic | 66.794 | 466.055 |
| $p$-value | 0.0000 | 0.0000 |

Notes: The table presents the statistics of the cross-sectional dependence test. Small $p$-values suggest that we should reject the null hypothesis that there does not exist cross-sectional dependence in the panel data of $\log y_{i,t}$.

## C Proof of Proposition 3

Proof: We guess the law of motion of $\{z_t^b(\psi_t)\}$ as follows,

$$z_{t+1}^b(H) = \begin{cases} z_t^b(H)\exp\left(\lambda_0^{HH} + \lambda_1(x_{t+1}-\bar{x}) + \lambda_2(x_t-\bar{x})\right), \text{ if } \psi_t = H \\ z_t^b(L)\exp\left(\lambda_0^{LH} + \lambda_1(x_{t+1}-\bar{x}) + \lambda_2(x_t-\bar{x})\right), \text{ if } \psi_t = L, \end{cases}$$

and

$$z_{t+1}^b(L) = \begin{cases} z_t^b(H)\exp\left(\lambda_0^{HL} + \lambda_1(x_{t+1}-\bar{x}) + \lambda_2(x_t-\bar{x})\right), \text{ if } \psi_t = H \\ z_t^b(L)\exp\left(\lambda_0^{LL} + \lambda_1(x_{t+1}-\bar{x}) + \lambda_2(x_t-\bar{x})\right), \text{ if } \psi_t = L, \end{cases}$$

which depends on the state of the economy in the previous period.

Thus we have

$$E_t z_{t+1}^b(H) = z_t^b(L)\exp\left(\lambda_0^{LH} + (\lambda_2+\lambda_1\rho)(x_t-\bar{x}) + \frac{1}{2}(\lambda_1\sigma_\varepsilon)^2\right),$$

and

$$E_t z_{t+1}^b(L) = z_t^b(L)\exp\left(\lambda_0^{LL} + (\lambda_2 + \lambda_1\rho)(x_t - \bar{x}) + \frac{1}{2}(\lambda_1\sigma_\varepsilon)^2\right),$$

for $\psi_t = L$. And

$$E_t z_{t+1}^b(H) = z_t^b(H)\exp\left(\lambda_0^{HH} + (\lambda_2 + \lambda_1\rho)(x_t - \bar{x}) + \frac{1}{2}(\lambda_1\sigma_\varepsilon)^2\right),$$

and

$$E_t z_{t+1}^b(L) = z_t^b(H)\exp\left(\lambda_0^{HL} + (\lambda_2 + \lambda_1\rho)(x_t - \bar{x}) + \frac{1}{2}(\lambda_1\sigma_\varepsilon)^2\right),$$

for $\psi_t = H$.

Thus

$$z_t^b(L) = \beta e^{\theta x_t}\left(\pi_{LH}E_t z_{t+1}^b(H) + \pi_{LL}E_t z_{t+1}^b(L)\right),$$

and

$$z_t^b(H) = \beta e^{\theta x_t}\left(\pi_{HH}E_t z_{t+1}^b(H) + \pi_{HL}E_t z_{t+1}^b(L)\right),$$

imply that

$$\lambda_2 = -(\lambda_1\rho + \theta),$$

$$\beta\pi_{LH}\exp\left(\lambda_0^{LH} + \frac{1}{2}(\lambda_1\sigma_\varepsilon)^2 + \theta\bar{x}\right) + \beta\pi_{LL}\exp\left(\lambda_0^{LL} + \frac{1}{2}(\lambda_1\sigma_\varepsilon)^2 + \theta\bar{x}\right) = 1,$$

and

$$\beta\pi_{HH}\exp\left(\lambda_0^{HH} + \frac{1}{2}(\lambda_1\sigma_\varepsilon)^2 + \theta\bar{x}\right) + \beta\pi_{HL}\exp\left(\lambda_0^{HL} + \frac{1}{2}(\lambda_1\sigma_\varepsilon)^2 + \theta\bar{x}\right) = 1.$$

These establish the proof. ■