Anatomy of the Beginning of the Housing Boom Across U.S. Metropolitan Areas*

Fernando Ferreira and Joseph Gyourko

The Wharton School, University of Pennsylvania

& NBER

This Version: November 11, 2021

Abstract

We provide novel estimates of the location, timing, magnitude, and determinants of the start of the last U.S. housing boom. The last housing cycle cannot be interpreted as a single, national event, as different markets began to boom across a decade-long period, some of them multiple times. A fundamental factor, income of prospective buyers, can account for half of the initial jump in price growth, while expansion of purchases by underrepresented minorities cannot. The start of the boom also was financed conventionally, not by subprime mortgages. The latter's share did rise sharply over time, but only after a multi-year lag.

JEL CODES: R2, R3, E3, and G2.

---

* Contact information: Fernando Ferreira: email: fferreir@wharton.upenn.edu, address: The Wharton School, University of Pennsylvania, 3620 Locust Walk, Philadelphia, PA 19104. Joseph Gyourko: email: gyourko@wharton.upenn.edu, The Wharton School, University of Pennsylvania, 3620 Locust Walk, Philadelphia, PA 19104. We thank the Research Sponsors Program of the Zell/Lurie Real Estate Center at Wharton for financial support. Qize Chen, Matthew Davis, Anthony DeFusco, Wenjie Ding, Anna Gao, Hye Jin Lee, Gordon MacDonald, Xuequan Peng, Cindy Soo, Boyuan Zhang, and Fan Xia provided outstanding research assistance. This version of the paper is much improved for the efforts of the Editor (Bryan Graham) and two anonymous referees. We also thank Matthew Gentzkow, Ed Glaeser, Erik Hurst, Alvin Murphy, Emily Oster, Paul Willen, and the participants in seminars at the Board of Governors of the Federal Reserve System, the NBER Summer Institute, the University of Chicago, UC-Berkeley, and UCLA for comments on previous versions of the paper. A previous version of this work appeared as the NBER Working Paper #17374 titled "Anatomy of the Beginning of the Housing Boom: U.S. Neighborhoods and Metropolitan Areas, 1993-2009".

I. Introduction

The last U.S. housing cycle was among the defining economic events of the past half century. While there is much research on the bust and the factors that helped build the boom near its peak,[1] there is virtually no work on when the last housing boom began. In this paper we provide a framework for determining when housing booms begin and estimate it for a large sample of U.S. metropolitan areas. Empirical estimation relies on finding structural breaks in the house price appreciation time series for individual housing markets. Since a housing market can have more than one boom (and bust) during the sample period, we allow for multiple breaks.

We use a large proprietary micro data base from DataQuick (now CoreLogic) with over 23 million housing transactions across 94 metropolitan areas from 1993-2011. Access to rich micro data is critical for our empirical design, as it permits the creation of two random subsamples to separately estimate the timing and magnitude of the beginning of the boom, thereby avoiding the specification search bias that arises when the same data are used to identify both the timing and size of a breakpoint jump (Leamer (1983)).

Our results illuminate key stylized facts about the great housing boom, the first of which is that the boom itself was not a single national event that built smoothly over time. Rather, it was a series of individual markets' price growth jumping at specific times across a decade-long period. New tests for a common break across groups of metropolitan areas (MSAs) based on recent work by Kim et al. (2020) strongly suggest there was no common national break. Another important stylized fact is that the initial price spurts are economically meaningful, averaging

[1] The massive amount of work on the bust is too voluminous to be cited here. See Ferreira and Gyourko (2015) for a summary on research related to the housing bust itself, and Brunnermeier (2009). Many factors were proposed to explain the large increase in prices, including irrational exuberance (Shiller, 2005), subprime mortgages (Mian and Sufi, 2009), expansion in middle class credit (Adelino, Schoar and Severino (2016), Foote and Willen (2016), and Albanesi, De Giorgi, and Nosal (2017)), expansion of credit via bank deregulation (Favara and Imbs (2015), and reductions in mortgage rates (Bhutta and Keys, 2016).

about 11% unconditionally, and 5% after controlling for time and market fixed effects. Elevated price growth above pre-boom levels persists for many years.

We exploit the variation in timing and magnitude of local booms to show that the beginning of the last boom was fundamentally based to a significant extent. Growth in incomes of the pool of potential homebuyers can account for about 50% of the initial jump in house values. This positive change in fundamentals is short lived though, as the income growth of potential buyers reverts back to its pre-trend in year three of the boom. Changes in buying by underrepresented minorities are not found to be empirically important. The starts of local market booms also were conventionally financed with prime loans. Subprime shares increase only later. If the most recent housing boom is typical of the nature of housing cycles, increasing mortgage and financial risk is a characteristic of a mature boom, not a new one.

Section II presents the strategy for identifying the beginning of a housing boom. Section III describes the data, results are presented in section IV, and section V concludes.

II. Theory and empirical methods

Our empirical strategy involves using structural breaks in the time series of local housing price appreciation rate series to pin down the start of a housing boom. The motivation for this approach arises from the core model of spatial equilibrium in urban economics. That compensating differential framework models the location decision as a function of utility based on the wages (W_i) one earns by working in labor market area *i* (wages are a function of the area's productivity), the amenities (A_i) consumed in the same market, and the house price (H_i) one has to pay to live there. Equalizing utility implies that housing prices are the entry fee a household must pay to access the productivity and amenities of a local labor market.

A dynamic version of this Rosen-Roback model was introduced in Glaeser, et. al. (2014). Housing demand is determined in equilibrium by equating the net flow costs of living in the reservation market that always provides the minimum utility level to the net benefits of living there over the same time period. It also includes a transversality condition preventing bubbles. These assumptions generate steady state values for house prices, new construction, and city size for each market. If there are no shocks to any local traits (i.e., $dW_{i,t}$=0 and $dA_{i,t}$=0 for all *t*), then those steady state values fully describe location *i,* with each market having constant growth path values for house prices, new construction and population. House price growth should then exhibit a discrete jump if there is a positive shock to local productivity or amenities. That jump marks the beginning of the boom.

Empirically, we search for structural breaks in the price appreciation series as markers for when local booms begin. Consider the following specification of house price growth (PG) in MSA i at time t:

$$(1)\quad PG_{i,t} = d_{i,t} + \epsilon_{i,t}, t = 1, \dots, T.$$

The discussion above implies that $d_{i,t} = d_{i,0}$ for all t if the market is on its steady-state growth path. The null hypothesis is: $H_0\colon d_{i,t} = d_{i,0}, t = 1, \dots T$. The alternative is that $d_{i,t}$ changes at some proportion, $0 < \pi_i < 1$, of the sample which marks the beginning of a housing boom in market i. The alternative hypothesis is

$$H_1\colon\ d_{i,t} = \begin{cases} d_{1,i}(\pi_i), t = 1, \dots, \pi_i T \\ d_{2,i}(\pi_i),\ \ t = \pi_i T + 1, \dots T. \end{cases}$$

Let $\Pi_\mathrm{i} = [\pi_{i,1}, \pi_{i,2}]$ be a closed interval in (0,1) and let S_i be the set of all observations from $t = int(\pi_{i,1}T)$ to $t = int(\pi_{i,2}T)$, where $int(\cdot)$ denotes rounding to the nearest integer. The estimated breakpoint is the value t^* from the set S_i that maximizes the likelihood ratio statistic from a test of H_1 against H_0. Direct calculation of the probabilities in the likelihood ratio is non-trivial and prior research has relied on approximations that are based on simulation or curve-fitting methods (Andrews (1993), Hansen (1997)). However, Estrella (2003) provides a numerical procedure for calculating exact p-values that does not rely on these types of approximations. We use his method to calculate p-values for the estimated break point, π_i, for each MSA in the sample.

Our procedure will generate a breakpoint estimate regardless of whether the structural break represents a positive or negative change in the price growth rate. When an estimated breakpoint is insignificant, we conclude that the market did not have a boom. When we find a statistically significant breakpoint, we also test for the existence of two breaks against the null hypothesis of only one, closely following Bai (1999) and Bai and Perron (1998). If we can reject the null hypothesis of one break against the alternative of two, we discard the one-break model estimate and keep the results from the two-break model. Similarly, for those MSAs with two breaks, we also estimate and test for the significance of three breaks relative to two. Allowing for multiple breaks helps address the possibility that certain MSAs had more than one shock or cycle during the sample period.[2]

Under the null hypothesis that there is no breakpoint, the estimate of $d_{i,t}$ has a nonstandard distribution and OLS estimates of its magnitude will be upwardly biased in absolute value. This can lead to an increased chance of falsely concluding a break exists, and is a form of specification search bias arising from the same data being used to estimate both the timing and the magnitude of the structural break (Leamer (1983)). Our approach to correcting this follows

[2] We also tested a Markov switching model that estimates probabilities of being in different regimes (Hamilton (2016)). Those probabilities largely line up with our estimated breakpoints and are available upon request. We prefer using the breakpoint methodology because it provides sharp estimates that allow us to see what other forces might have changed along a well-defined time line of the boom.

Card, Mas, and Rothstein (2008) in randomly splitting the underlying sample of houses transactions into two, one for estimating breakpoints and the other to estimate the magnitude of price changes (and also the magnitude of changes in other variables, such as income, race, and mortgage characteristics).[3] The holdout panel includes all markets (i) and quarters (t) available.

Magnitudes of the changes in price growth rates (PG) are estimated for all pre- and post-boom years, based on the following panel equation:

$$(2) \quad PG_{i,t} = \sum_{\rho=-6}^{6} \sum_{b \in B} \theta_{\rho}^{b} \, \mathbf{1}\{t - t^{*,b} = \rho\} + \, q_t + \, a_i + \epsilon_{i,t},$$

where $\mathbf{1}\{t - t^{*,b} = \rho\}$ is a set of years relative to the quarter of the beginning of the housing boom ($t^{*,b}$). For example, relative year 1 includes the quarter of the breakpoint plus three subsequent quarters. Relative year 2 dummy includes the next four quarters (#4-#7 since the boom started), and so on. We show all estimates up to relative year five, as after that there is a large drop in the number of MSAs with relevant data (all models also include dummies for relative years 6 or above). To distinguish between markets that had a housing boom and those that did not, we estimate separate coefficients for breakpoints that are positive and statistically significant ($b = boom$), and those that are either negative or insignificant. The coefficients, θ_{ρ}^{boom}, then describe how prices in a market that had a boom evolve over the course of time. These estimates are relative to the 12-month period prior to the beginning of the boom (the dummy for that period is excluded from the estimation). Additionally, q_t are year-quarter fixed effects, and a_i are MSA fixed effects.

Our approach estimates the beginning of the boom for each MSA, but it does not rule out that some MSAs may share a common break. Few attempts have been made to implement a

[3]If the two subsamples are independent, then estimates of $d_{i,t}$ from the second sample, which was not used to estimate the location of the breakpoint, will have a standard distribution even under the null hypothesis of no structural break in the first sample.

global hypothesis test for a large number of time series because of the heavy computational burden and unknown asymptotic properties. We are able to build on recent work by Kim, et. al. (2020) that tests for whether two sequences share one commonly timed break. We extend their approach to three and four sequences of MSAs.[4] Below we report results indicating that it is highly unlikely a large number of housing markets experienced a common shock to start the last great housing boom. Before getting to results, we briefly discuss the data used in the estimation.

III. Data

Our primary source of housing market data comes from DataQuick (now CoreLogic), which provides micro observations on home purchases collected from deeds records. The final sample used in this analysis contains more than 23 million arms-length, single-family and condominium housing transactions in 94 metropolitan areas spread across 29 states, from 1993(Q1) to 2011(Q4) – see the Online Appendix for more details. Online Appendix Table 1 reports the representativeness of our final sample. Using these data we create a MSA-level constant quality house price series by quarter using a hedonic model.[5] Online Appendix Figure 1's plot of MSA price indexes confirms that dispersion in prices changed dramatically over the cycle.

We merge DataQuick with the Home Mortgage Disclosure Act (HMDA) files in order to capture information on the income and race of all loan applicants, including applicants whose

[4] We do not explore a higher number of MSAs due to the exponentially increasing computational burden. To test m time series having a common break, we have to run OLS regression and examine the variance-covariance matrix under each possible date for each m series. The total number of regressions required to obtain the break date by maximizing the log-likelihood function is $C_n^m \times T^m$, where T is the length of the time series and C_n^m is the combinatorial number of different choices of choosing m MSAs out of n.

[5] We employ hedonic price indexes because their data requirements are less onerous than repeat sales indexes, but we have investigated both methodologies and find they yield very similar results. See the Online Appendix for details.

mortgages were denied and homeowners that refinance existing mortgages.[6] We also look at labor market conditions and demographics for all residents in the metropolitan area, including all renters.[7] CoreLogic also provides information on the amounts of up to three loans used to finance the purchase of each home, as well as the names of the buyers, sellers, and lender(s). Our subprime measure is the share of all homeowners (not just recent homebuyers) with current mortgages issued by subprime lenders. We separately identify borrowers whose loans were guaranteed by FHA or VA (regardless of lender identity).[8] All remaining owners with debt are Prime borrowers by definition. Homes purchased without a mortgage are classified as cash-only. Finally, we measure two other proxies for riskier loans: i) the average loan-to-value ratio of all transactions in a quarter, which is based on sum of the mortgage balance of up to three initial loans per property divided by the transaction price, and ii) the share of home purchase mortgages in a quarter that have some type of variable interest rate structure.[9] Online Appendix Table 2 reports descriptive statistics on each of these variables.

IV. Results

Figure 1 plots the histogram of the timing of breaks for all statistically significant positive structural breaks. This shows the last boom to be a series of local booms that ran from the mid-

---

[6] Both data sets were merged using a sequential matching process. In total, 93.2% of all sales transactions in DataQuick were matched at some point in the procedure, with approximately 60% considered "high quality" matches, as there was a unique match based on tract ID, year of transaction, precise house value, and lender name. More details are available in the Online Appendix.

[7] The local unemployment rate is measured monthly by the Bureau of Labor Statistics (BLS). From the Bureau of Economic Analysis (BEA), we collect yearly data on overall personal income by MSA. Finally, from the Common Core of Data (CCD), we collect the yearly demographic composition of all students in each metropolitan area. See the online appendix for more details and summary statistics.

[8] Subprime borrowers are identified based on annual lists of lenders published in a prominent trade magazine. See the online appendix for more detail

[9] Interest rates themselves are not used in our empirical analysis, as our empirical strategy is best suited to examining factors that vary across space and time; mortgage rates are roughly constant across local markets.

1990s to the mid-2000s, rather than a single national event. The modal quarter is 2004(Q2), but it contains less than 10% of all breaks. Most quarters see at least one market start to boom, and many have from two to four percent of the total number of breaks.[10]

Table 1 summarizes the results of estimating equation (2). Unconditional price growth is 10.7% higher in relative year 1. Column 2's results include time dummies, so they represent the variation remaining after sweeping out the impacts from nationwide changes that are common across all locations. The estimated jump is now 4.5%, so over 50% of the variation in unconditional price changes can be explained by factors that are not exclusively local. Even though the timing of the start of local booms does not seem to follow a national pattern, yearly price changes do. Column 3's results are from our preferred specification that also includes area fixed effects. Price growth rates jump by 5.1% at the beginning of the boom relative to the pre-boom year, then goes to 6.8%, 5.1%, 3.4%, and 2.2% in subsequent years.[11] Columns 4 and 5 also report estimates for significantly negative breaks (i.e., busts) and for statistically insignificant breaks. Prices were mostly flat in markets without a break, and negative for markets with a bust. In the remainder of the paper we focus solely on the timeline of the booms.

---

[10] More detail on breakpoints in all markets is reported in Online Appendix Table 3. There were not any statistically significant jumps in the rate of price growth in six markets, with most being in the Midwest or upstate New York. Negative and statistically significant jumps were found in various places such as Cleveland and Detroit. Eighty-two markets had at least one statistically significant boom, with starting dates ranging widely in time from the last quarter of 1995 in Boston, MA, to the third quarter of 2005 in Salem, OR, and Mt. Vernon-Anacortes, WA. Fifty-eight markets had at least two jointly statistically meaningful jumps in price growth, and in 21 markets we estimated three jointly significant structural breaks. They are usually coastal markets that had an early boom prior to 2000 and a second boom in the mid-2000s.

[11] These results show a modest pre-trend, with a small, but statistically significant, negative growth rate in relative year -1. Further analysis suggests this likely is an artifact of measurement error in the estimation of breakpoints, as the breaks have a non-zero probability of being one quarter earlier or later than the dates shown in Online Appendix Table 3. To investigate this more fully, we estimated another specification that added a dummy for relative quarters -1 and -2. In that specification, the baseline period becomes relative quarters -3 to -6, which are less likely to be affected by measurement error. The results, which are reported in the first column of Online Appendix Table 4, show the complete absence of a pre-trend.

This still begs the question of whether one can reject the null of a common break across all markets. As noted above in footnote 4, is it not feasible computationally to perform that test. We did run tests to determine how often we could reject the null of a common break across all pairs of metro areas, all triplets and all quadruplets. Even though we know from Figure 1 that it is common to observe two markets booming at the same time, we still reject the null from 50%-58% of the time depending upon the type of test: Lagrange Ratio (LR), Wald GLS and Wald OLS. For triplets, the rejection rates increase to 73%-83%. For quadruplets, the rejection rate never is lower than 86% (LR) and reaches 93% (Wald GLS). Moreover, for the quadruplets that we cannot conclusively reject, the vast majority have negative, or a mix of positive and negative breaks, while the remaining positive breaks have small price changes (<2%). Given how quickly rejection rates reach very high levels, it seems highly unlikely that our full sample (or any larger, double-digit number of markets) has an economically meaningful common break.

IV.a. Incomes, Demographics, and Risky Mortgages

Table 2 reports results regarding how other factors changed across the timelines of local housing booms (using only the holdout samples for all variables). Starting with income, there is no evidence of a pre-trend in the reported earnings of those applying for a mortgage. However, their incomes do jump by a statistically significant 3.3% at the beginning of the local area's housing boom and stay elevated throughout the following year. Using estimates from the housing literature suggesting an income elasticity of demand ranging from 0.75-1.0, income changes can account for around half of the initial jump in prices at the beginning of the boom. While the beginning of a boom is at least partially fundamentally grounded, income growth reverts back to pre-boom levels from relative year 3-onward, so it cannot account for the lengthy boom.

Columns 2-3 of Table 2 report results on the variation in the racial composition of buyers across the housing cycle timeline. There is an economically small, but statistically significant 0.8 percentage point change in White buyer share (on a sample of 60.9%) in the year a boom begins, but this does not persist in a significant fashion for another year, and the White share has reverted back to pre-boom levels entirely by relative year 3. Column 3 reports a drop in minority share as the boom begins, but it is small and insignificant.[12]

Columns 4 to 9 of Table 2 document how the shares of homeowners using different types of mortgages vary over the timeline of housing booms. Note that there is no evidence of meaningful pre-trends in the years prior to the start of a local market boom for both prime and subprime mortgages. Only prime lending starts to ramp up when housing booms start, with the bulk of the increase being taken from the FHA-insured sector. This is consistent with the concomitant rise in the income of current homebuyers, as households with rising incomes reasonably could be viewed as less risky. Subprime shares increase only in relative year 2, and do not mean revert back to pre-boom levels throughout our timeline. Finally, the cash-only buyer share is flat in the first and second years of the typical boom, but it does start to drop in a statistically significant fashion three years into the boom. This increase in riskier lending with a lag from the start of a boom is corroborated by other proxies for risky mortgages reported in the final two columns of Table 2.[13]

---

[12] We also estimated magnitudes for incomes and race using different measures based on all households in a market, including renters. Those results are reported in Online Appendix Table 5. These estimates do not show a significantly positive jump in relative year 1 or 2. This is likely explained by measurement error biasing the coefficients towards zero. They are annual measures (not quarterly), and they reflect a much larger population, including a large fraction of renters not necessarily interested in becoming an owner-occupier in any given year. We also show how overall unemployment rates varied over the cycle. That rate declines by a statistically significant 0.3 percentage points in the first year of the boom (compared to the pre-boom year), and by 0.5 percentage points in relative years 2 and 3, before slowly starting to return to pre-boom levels in subsequent years.

[13] In Online Appendix Table 6 we also show magnitudes for the volume of transactions and refinances. We find a statistically significant jump in transactions volume when the boom first begins, and that high volume is maintained for the next couple of years. Refinancings are higher than baseline in almost all years, and become especially large

V. Conclusion

Given the difficulties in conceptualizing and estimating the beginning of a boom, it is not surprising that the recent literature has focused more on the later part of the boom and its bust. Guided by the intuition of a dynamic model of spatial equilibrium across housing markets, we provide the first timeline of the last boom based on breakpoints in local market price growth rate series. The methods developed here further our understanding by providing a framework that exploits the underlying heterogeneity in housing booms – which markets are booming at any point in time, their timelines, and price changes - and that more precisely links the variation from each location with fundamental or credit market factors. Moreover, the timeline of local housing booms may also be useful in understanding the interaction of housing cycles with other aspects of the economy, such as the educational achievement of individuals (Charles, Hurst and Notowidigdo, 2016) or the determinants of local public finances (Davis and Ferreira, 2017).

References

Adelino, M., A. Schoar, and F. Severino. 2016. "Loan Originations and Defaults in the Mortgage Crisis: The Role of the Middle Class." *The Review of Financial Studies* 29(7): 1635-70.

Albanesi, S., G. De Giorgi, and J. Nosal. 2017. "Credit Growth and the Financial Crisis: A New Narrative." NBER Working Paper 23740.

Andrews, D. 1993. "Tests for Parameter Instability and Structural Change with Unknown Change Point." *Econometrica* 61(4): 821-56.

Bai, J.. 1999. "Likelihood ratio tests for multiple structural changes." *Journal of Econometrics* 91(2): 299-323.

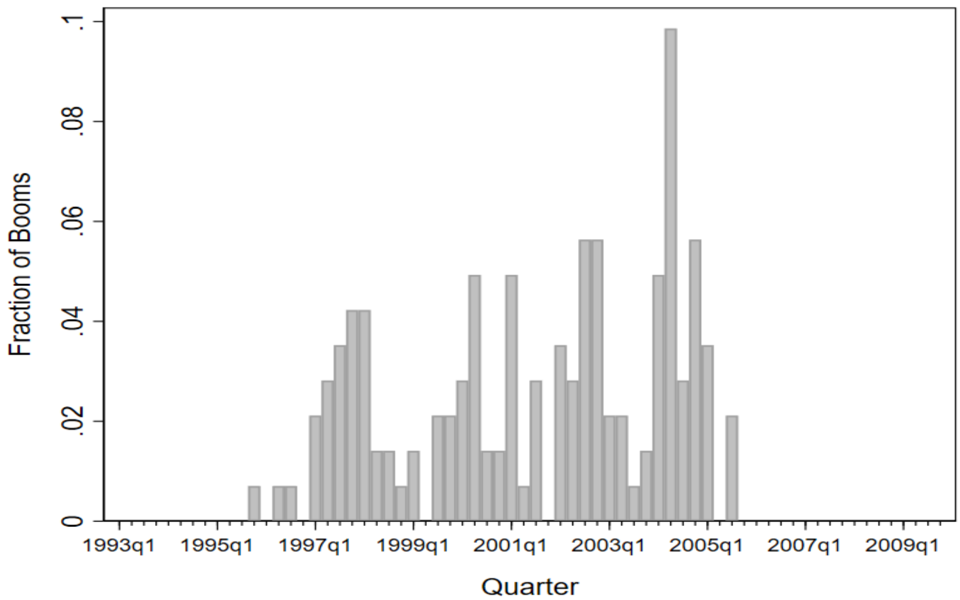
Bai, J., and P. Perron. 1998. "Estimating and Testing Linear Models with Multiple Structural Changes." *Econometrica* 66(1): 47-78.

---

in relative years 2 and 3, coinciding with the peak of price growth rates. Pre-trend estimates are noisier for both volume variables, but somewhat consistent with DeFusco, Nathanson and Zwick (2017) who document a lead-lag relationship between changes in trading volume and price changes in U.S. housing markets. The Online Appendix presents other sensitivity analysis of our main results, such as estimating similar models for variables in growth rates, testing for the impact of the boom in other variables, such as the share of speculators, the share of mortgages with LTV greater than 95%, and several migration flows.

Bhutta, N., and B. Keys. 2016. "Interest Rates and Equity Extraction during the Housing Boom." *American Economic Review* 106(7): 1742-74.
Brunnermeier, M. 2009. “Deciphering the Liquidity and Credit Crunch 2007-2008.” *Journal of Economic Perspectives* 23(1): 77-100.
Card, D., A. Mas and J. Rothstein. 2008. “Tipping and the Dynamics of Segregation.” *Quarterly Journal of Economics* 123(1): 177-218.
Charles, K., E.Hurst, and M. Notowidigdo. 2016. "The Masking of the Decline in Manufacturing Employment by the Housing Bubble." *Journal of Economic Perspectives* 30(2): 179-200
Davis, M. and F. Ferreira. 2017. “Housing Disease and Public School Finance.” NBER Working Paper 24140.
DeFusco, A., C. Nathanson, and E. Zwick. 2017. “Speculative Dynamics of Prices and Volume.” NBER Working Paper 23449.
Estrella, A. 2003. "Critical Values and P Values of Bessel Process Distributions: Computation and Application to Structural Break Tests." *Econometric Theory* 19(6): 1128-43.
Favara, G., and J. Imbs. 2015. "Credit Supply and the Price of Housing." *American Economic Review* 105(3): 958-92.
Ferreira, F. and J. Gyourko. 2015. “A New Look at the U.S. Foreclosure Crisis: Panel Data Evidence of Prime and Subprime Borrowers from 1997 to 2012”. NBER WP 21261.
Foote, C., and P. Willen. 2016 (Revised January 2020). “Cross-Sectional Patterns of Mortgage Debt during the Housing Boom: Evidence and Implications.” Federal Reserve Bank of Boston Research Department Working Papers 16-12.
Glaeser, E., J. Gyourko, E. Morales and C. Nathanson. 2014. “Housing Dynamics: An Urban Approach.” *Journal of Urban Economics* 81: 45-56.
Hamilton, J.. 2016. "Macroeconomic Regimes and Regime Shifts." In *Handbook of Macroeconomics*, Vol 2, edited by John B. Taylor and Harald Uhlig, 163-201. Amsterdam: North Holland.
Hansen, B. 1997. "Approximate Asymptotic P Values for Structural-Change Tests." *Journal of Business & Economic Statistics* 15(1): 60-67.
Kim, D., T. Oka, F. Estrada, and P. Perron. 2020. “Inference Related to Common Breaks in a Multivariate System with Joined Segmented Trends with Applications to Global and Hemispheric Temperatures.” *Journal of Econometrics*, 214(1): 130-152.
Leamer, E. 1983. “Let’s Take the Con Out of Econometrics.” *American Economic Review* 73(1): 31-43.
Mian, A. and A. Sufi. 2009. “The Consequences of Mortgage Credit Expansion: Evidence from the U.S. Mortgage Default Crisis.” *Quarterly Journal of Economics* 124(4): 1449-96.
Shiller, R.. 2005. *Irrational Exuberance*. 2ed. Princeton: Princeton University.

Figure 1. Beginning of local booms histogram



Notes: The histogram plots the fraction of booms that start in a given quarter.

Table 1: Magnitude of house price appreciation changes

Dependent Variable	Price Growth Rate				
	Positive Breaks	Positive Breaks	Positive Breaks	Negative Breaks	Non-Significant Breaks
	(1)	(2)	(3)	(4)	(5)
Relative Year = -2	-0.009	-0.022	-0.006	-0.021	0.002
	(0.005)	(0.004)	(0.005)	(0.011)	(0.018)
Relative Year = -1	-0.001	-0.023	-0.012	-0.004	0.009
	(0.005)	(0.003)	(0.004)	(0.010)	(0.015)
Relative Year = 0	--	--	--	--	--
Relative Year = 1	0.107	0.045	0.051	-0.084	-0.000
	(0.007)	(0.004)	(0.005)	(0.016)	(0.030)
Relative Year = 2	0.131	0.061	0.068	-0.076	0.042
	(0.007)	(0.005)	(0.006)	(0.015)	(0.045)
Relative Year = 3	0.096	0.045	0.051	-0.098	0.004
	(0.007)	(0.006)	(0.008)	(0.017)	(0.036)
Relative Year = 4	0.046	0.030	0.034	-0.089	-0.013
	(0.010)	(0.007)	(0.008)	(0.019)	(0.026)
Relative Year = 5	-0.011	0.018	0.022	-0.049	0.016
	(0.011)	(0.006)	(0.008)	(0.019)	(0.028)
R-squared	0.36	0.73	0.75	0.75	0.75
Number of observations	5,846	5,846	5,846	5,846	5,846
Dependent variable mean	0.082	0.082	0.082	0.082	0.082
Time FEs		X	X	X	X
MSA FEs			X	X	X

Notes: The table shows points estimates for price growth rates around the timeline of the housing booms and busts. All dummies for relative year zero are omitted in the estimation. Columns 1 and 2 report point estimates for different models based on equation (2). Columns 3, 4, and 5 are estimates from the same model that follow equation (2) and include time and MSA fixed effects.

Table 2: Income, race, and financial factors around the timeline of housing booms

Dependent Variables	HMDA Income Growth Rate	HMDA White Share	HMDA Minority Share	Prime Share	Subprime Share	FHA Share	Cash-only Share	Average Initial LTV	Variable-rate Share
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Relative Year = -2	0.009	0.002	0.006	0.007	-0.005	0.020	-0.013	0.020	0.013
	(0.009)	(0.005)	(0.005)	(0.008)	(0.004)	(0.007)	(0.006)	(0.006)	(0.007)
Relative Year = -1	0.007	0.008	-0.000	0.001	-0.001	0.014	-0.009	0.015	0.003
	(0.009)	(0.004)	(0.004)	(0.006)	(0.003)	(0.005)	(0.005)	(0.005)	(0.005)
Relative Year = 0	--	--	--	--	--	--	--	--	--
Relative Year = 1	0.033	0.008	-0.003	0.012	0.004	-0.007	-0.002	0.002	0.007
	(0.010)	(0.004)	(0.004)	(0.005)	(0.003)	(0.005)	(0.005)	(0.005)	(0.007)
Relative Year = 2	0.025	0.005	-0.001	0.019	0.011	-0.019	-0.005	0.005	0.015
	(0.008)	(0.005)	(0.005)	(0.006)	(0.004)	(0.005)	(0.006)	(0.006)	(0.008)
Relative Year = 3	0.007	0.000	0.003	0.019	0.018	-0.019	-0.013	0.013	0.027
	(0.008)	(0.005)	(0.006)	(0.006)	(0.004)	(0.005)	(0.006)	(0.006)	(0.010)
Relative Year = 4	0.012	-0.004	0.006	0.012	0.024	-0.013	-0.018	0.015	0.029
	(0.009)	(0.005)	(0.006)	(0.006)	(0.004)	(0.006)	(0.005)	(0.005)	(0.011)
Relative Year = 5	0.005	-0.006	0.005	-0.003	0.026	-0.004	-0.017	0.015	0.038
	(0.008)	(0.005)	(0.005)	(0.006)	(0.004)	(0.007)	(0.005)	(0.005)	(0.012)
R-squared	0.310	0.910	0.861	0.837	0.792	0.823	0.818	0.817	0.819
Number of observations	5,846	6,222	6,222	6,222	6,222	6,222	6,222	6,222	6,222
Dependent variable mean	0.062	0.782	0.162	0.531	0.124	0.133	0.176	0.696	0.177
Time FEs	X	X	X	X	X	X	X	X	X
MSA FEs	X	X	X	X	X	X	X	X	X

Notes: The table shows points estimates for income, race, and financial variables around the timeline of the housing booms. Each column reports estimates of a separate model based on equation (2). Dummies for relative year zero are omitted in the estimation.