The Role of Contagion in the Last American Housing Cycle

This Draft: September 23, 2013

Anthony DeFusco, Wenjie Ding, Fernando Ferreira** and Joseph Gyourko**
The Wharton School, University of Pennsylvania and NBER**

Abstract

Using proprietary micro data on the complete set of housing transactions between 1993 and 2009 in 99 metropolitan areas, we investigate whether contagion was an important factor in the last housing cycle. We define contagion as the price correlation between two different housing markets following a shock to one market that is above and beyond that which can be justified by common aggregate trends. Our estimates deal with the following empirical challenges: (a) defining the timing of local housing booms in a non-*ad hoc* way; (b) addressing specification search bias that arises when only one aggregate series is used to estimate both the timing of the housing boom and the magnitude of price volatility during that period; and (c) controlling for common variation in economic conditions. We find strong evidence of contagion during the housing boom, but not during the bust. These effects appear to arise mostly from the closest neighboring metropolitan area, with the price elasticity ranging from 0.10 to 0.27. This is large enough to account for up to 30% of the jump in prices at the beginning of local booms, on average. Estimated elasticities are greater when transmitted from a larger to a smaller market, and also more important for the most elastically-supplied markets. Finally, local fundamentals and expectations of future fundamentals have very limited ability to account for our estimated effect, suggesting a potential role for non-rational forces.

The authors thank the Research Sponsors Program of the Zell/Lurie Real Estate Center at Wharton for financial support. We also appreciate the comments of Amit Seru as well as participants in presentations at the NBER Conference on Housing and Financial Crisis, the University of Miami, and the University of California-Berkeley.

I. *Introduction*

One of the striking features of the recent U.S. housing cycle is its heterogeneity across markets. Both the magnitudes and timing of price swings varied greatly across metropolitan areas (Sinai (2012)). Figure 1 plots the geography and timing of the start of housing booms at the metropolitan area level from 1993 to 2009 based on estimates reported in Ferreira and Gyourko (2011).[1] The top left panel marks the 15 primarily rust belt and interior markets that never boomed. The other panels show that the remaining markets boomed at very different times over a nearly decade-long period from 1997-2005 and that the timing of these booms was non-random. The housing boom spread from what were initially highly concentrated areas on the two coasts, with the earliest booms beginning between 1997-1999 in California and the mid-New England region. On the west coast, housing booms eventually spread inland towards central California and to neighboring states to the east and north. On the east coast, housing booms spread to other markets in New England and then to neighboring regions, eventually reaching the majority of Florida markets by 2004 and 2005. These patterns are suggestive of spillover effects that disseminate positive housing price shocks from one market to another.

In this paper we investigate whether such spillovers were an important element of the last American housing cycle and also directly test which mechanisms may have contributed to them. In the financial economics literature, a spillover of this type is often referred to as contagion when it is found following a negative shock to one or more countries or markets.[2] While we focus on spillovers from a positive shock in much of the paper, we use the terms spillover and contagion interchangeably in order to emphasize the close intellectual linkage of our work with the analysis of contagion in financial economics. We define contagion as the price correlation across space between two different housing markets following a shock to one market that is above and beyond that which can be justified by common aggregate trends.[3]

---

[1] They define the beginning of a metropolitan area's housing boom by the quarter in which there is a structural break (a discrete positive jump in this case) in the market's house price appreciation rate. This methodology and the rationale behind it are discussed more fully below in Section II.

[2] See Forbes (2012) for an excellent recent review of that literature, and Dungey, et. al. (2005) for a more technical analysis of the challenges involved in convincingly estimating contagion or spillover effects. Previous work on financial market contagion includes studies of the 1987 U.S. stock market crash (King and Wadhwani (1990); Lee and Kim (1993)), the 1994 Mexican peso crisis (Calvart and Reinhart (1996)), and the Hong Kong stock market and Asian currency crisis of 1997 (Corsetti, et. al. (2005)).

[3] There is no single, agreed-upon definition of contagion, but our definition is similar in spirit to many used in the financial economics literature. See, for example, Forbes (2012), which emphasizes the distinction between

The nature of the housing market and richness of our data allow us to address several empirical concerns that plague previous contagion-related research. One example involves the determination of the relevant period(s) in which to study spillovers. For example, a non-*ad hoc* procedure for identifying the timing of a shock is preferred to an arbitrary choice of time period 'after the fact'. This typically is not feasible in most studies of stock market or currency crises because there is little or no variation across countries in the onset of those events. Fortunately, this was not the case during the most recent housing cycle which saw substantial variation across markets (Ferreira and Gyourko (2011); Sinai (2012)). In addition, we are able to appeal to existing theory as a guide in helping to determine how to time the beginning of a local boom. As is described more fully in the next section, implications of Glaeser et. al.'s (2012) dynamic version of the classic model of spatial equilibrium in urban economics lead us to date the beginning of a given market's boom by whether and when there was a structural break in that area's price appreciation rate. These estimates were shown in Figure 1 and provide us with substantial variation in the timing of the start of local market booms that can be used in testing for the existence of contagion.

A second advantage is provided by the use of a voluminous micro-level data on U.S. housing transactions.[4] We have over 23 million observations on individual home sales in 99 metropolitan areas dating back to the early 1990s in most cases. This data enables us to address specification search bias of the type identified by Leamer (1978), which arises when the same data is used to identify both the timing of a shock and the magnitude of the volatility during that period. Our strategy uses randomly split samples to separately identify the timing of booms and the magnitude of price volatility in those periods.[5] Most contagion studies in financial economics are not able to deal with this issue because they use a single aggregate stock index in each country. Doing so increases the likelihood of falsely concluding that there are more and bigger booms (or crisis periods in the financial economics literature) than truly exist.

Third, the richness of our data and the variation in the timing of booms across markets also helps us deal with omitted variable biases in several ways. We use the time line of a

---

contagion and interdependence, with the latter term reflecting when events in one country affect others in all states of the world, not just after severe negative events.

[4] The property transaction data is collected by Dataquick or by intermediaries from county assessor's offices captures the universe of sales.

[5] This is the same strategy followed by Card, Mas and Rothstein (2008) in their study of tipping points in residential segregation models.

neighbor's boom as our source of variation in the data to identify contagion effects. Our baseline specification involves regressing a focal market's price changes on a series of indicators reflecting whether the relevant neighboring market is booming and how proximate a given period is in time to the start of that market's boom.[6] The added degrees of freedom afforded by the multiple, non-contemporaneous booms we observe also allow us to control for omitted factors that might reflect common economic shocks. We do so through the inclusion of time by census division dummies, lagged price changes, as well as a host of local fundamentals. We also are able to address the Forbes and Rigobon (2002) critique regarding heteroskedasticity, whereby increased volatility in the 'crisis period' (when the boom starts, in our context) generates upward bias in correlation coefficients across markets. We adjust for the volatility in prices being higher than normal when the boom starts by directly controlling for the time line of the focal market's boom. Finally, in an alternative specification, we also address the potential for reverse causality with an instrumental variable approach using further lags of close neighbor's price changes.

Our main conclusion is that contagion played a statistically and economically significant role in the development of the most recent housing boom. The elasticity of prices in a typical focal market with respect to those in the closest metropolitan area in the year following the beginning of the boom in that neighboring area ranges from 0.10 to 0.27.[7] Empirically, the upper end of our elasticity range implies that from one-fourth to one-third of the average jump in price growth at the start of a typical local boom was due to contagion effects. This average impact is driven entirely by the physically closest neighbor and is only detected if the nearest neighbor had a statistically significant housing boom. There is no evidence of spillovers on prices arising from more geographically distant markets. In addition, this impact does not vary materially with the number of miles between the focal market and its nearest neighbor. As a robustness check, we investigated whether there is evidence of contagion on the extensive margin. Hazard models show that the probability of a boom beginning this quarter is indeed influenced by close metros that boomed in the previous quarter.

---

[6] As expected, this is very important empirically. Naively regressing 'price on price' yields contagion estimates that are 3-5 times larger than the results we report below from our preferred specification.

[7] We also report evidence of on-going contagion effects as the boom builds, but those results are potentially confounded by feedback effects. This is not a concern at the start of local booms because the data show no pre-trends, with markets appearing to be on their equilibrium paths before the initial jump in price growth when the booms commenced.

We also investigated heterogeneity in the price contagion along a number of other non-distance related dimensions. For example, one might expect contagion to be stronger when the transmission is from larger neighbor to a smaller focal market. That is precisely what we find. The magnitude of the contagion effect also varies with the elasticity of supply of the focal metropolitan area. Specifically, there is no evidence of contagion in inelastically supplied markets, but the estimated impact in the most elastically supplied markets is double the average effect discussed above.

What mechanisms could explain the observed price contagion effect? Fundamental factors including local income, migration flows between focal and neighboring markets, lending behavior in both markets, and speculative behavior in the focal market, do not show much variation after the beginning of the boom of the nearest neighbor. We also include a simple form of expectations of these local market fundamentals in our main econometric model to see if this materially affects the magnitude of the estimated contagion effect. It does not. This indicates that at least some of the contagion we estimate may be due to forces not related to the fundamentals analyzed in this paper. This has potentially important implications for policy makers. To the extent that the spread of a housing boom is even partially due to non-fundamental forces (e.g., some type of irrational exuberance or otherwise mistaken perceptions of the influence of a neighboring market), it may be worth rethinking the advisability of policy makers not responding to a boom.

Despite finding important spillovers during the run up of the U.S. housing boom, we report mixed evidence that contagion played a role in the bust. Estimated elasticities are zero at the beginning of the bust, but they increase to about 0.15 by the third year of the housing bust of the nearest neighbor. We also cannot detect any impact of contagion on the extensive margin during the bust. This is perhaps not all that surprising since the timing of the bust across MSAs is heavily concentrated in an 18 month period during 2006 and 2007, while the buildup of the housing boom took almost a decade. This highlights difficulties in detecting spillovers during economic or financial crashes/booms that quickly spread across countries, regions, or firms.[8]

[8] In addition, we do not have the added advantage of relying on a prediction from external theory to date the beginning of the bust. In this part of the analysis, we follow the tradition in the contagion literature and date the bust's beginning in an *ad hoc* manner based on the assumption that it begins in the quarter after nominal price levels peaked in a given metropolitan area.

While our research is motivated by prior research in financial economics noted above, it is also part of a growing body of work on the most recent housing cycle. One important strand of that research tries to understand whether the most recent cycle was a bubble.[9] Another more voluminous body of papers analyzes the bust and its consequences.[10] There are also some prior studies of price spillovers in the housing market.[11] In this paper we focus on one particular facet of the cycle, the role of contagion, but contribute to the literature by: a) looking at contagion over the full time span of the cycle, b) improving empirical identification in many respects, c) estimating heterogeneity in the contagion elasticity, and d) investigating the potential causes of contagion across housing markets.

The plan of the paper is as follows. The next section motivates our use of an urban economic approach to analyzing contagion in housing markets and discusses our method for dating the beginning of the boom. Section III then describes the various data sources employed and the variables created. Section IV discusses the different types of specifications estimated, reports results, and explores potential mechanisms that might explain the contagion effects. In Section V we estimate alternative specifications, look at the extensive margin of contagion, and whether we can find contagion at the housing bust. There is a brief conclusion.

II. An Urban Economic Motivation and Definition of Contagion

---

[9] Shiller (Chapter 2, 2005) provides perhaps the most famous characterization of the boom as a non-rational event. Others recently have estimated rational expectations general equilibrium models to try to explain the national aggregate price data (e.g., Favilukis, Ludvigson, and Van Nieuwerburgh (2011)) or the serial correlation and volatility of prices and quantities within and across metropolitan areas (e.g., Glaeser et al. (2012)). Related work includes Arce and Lopez-Salido (2011), Burnside, Eichenbaum and Rebelo (2011), Lai and Van Order (2010), and Wheaton and Nechayev (2008).

[10] Much of this research focuses on the subprime sector (e.g., Bajari, Chu and Park (2008), Danis and Pennington-Cross (2008), Demyanyk and Van Hemert (2011), Gerardi, Shapiro and Willen (2007), Goetzmann, Peng and Yen (2009). Mayer and Pence (2008), Haughwout, et. al. (2011)), mortgage securitization (e.g., Bubb and Kaufman (2009), Keys et. al. (2010)), the default/foreclosure crisis (e.g., Adelino, Gerardi and Willen (2009), Campbell, Giglio and Pathak (2011), Foote, Gerardi and Willen (2008), Gerardi et. al. (2008), Mayer, Pence and Sherlund (2009), Mian and Sufi (2009), Mian, Sufi and Trebbi (2010), Piskorski, Seru and Vig (2009)) or the role of government regulation (e.g., Avery and Brevoort (2010), Bhutta (2009), Ho and Pennington-Cross (2008)).

[11] Some early examples include Clapp, et al. (1995) and Dolde and Tirtiroglu (1997), who use data on towns in Connecticut and San Francisco (respectively) to test for the existence of cross-market linkages in price movements. More recently, Holly, et al. (2010, 2011) use data on U.S. states and U.K. regions to study the spatial and temporal diffusion of changes in house prices during the most recent cycle. Fuss, et. al. (2011) and Cotter, et. al. (2011) use publicly available, metropolitan area-level house price series to test for the existence of some form of contagion during the most recent housing cycle. However, all studies use of the same aggregate market-level price data to determine both the timing of the crisis period and to measure the magnitude of volatility changes during that period makes their estimates susceptible to specification search bias. In addition, the timing of the shock is usually defined in an ad-hoc way and there are questions about how that research deals with omitted factors.

II.a. Timing of local housing booms

Any analysis of possible contagion effects in the spreading of the recent housing boom first requires knowledge of the timing of the beginning of the boom in different markets. Our data on the timeline of local booms come from estimates discussed above and reported in Ferreira and Gyourko (2011). In that work, the start of local booms is determined by when there was a structural break in each area's price appreciation rate series. The justification for that strategy is based on implications of the dynamic spatial equilibrium model developed in Glaeser, et al. (2012). In particular, that model implies that in steady state each local market will exhibit constant and continuous growth paths for house prices, new construction and population.[12]

Empirically, this means that we should see house prices in a given market growing at a constant rate unless there is a shock to local productivity, amenities, or expectations, in which case we would then observe a discrete jump in the appreciation rate for that market. The data are generally consistent with this predicted pattern. As a particularly stark example, Figure 2 plots annualized house price appreciation rates over time for the Las Vegas market. This graph shows that fast-growing market to be appreciating at a high, but roughly constant rate for many years before house price growth escalates sharply at the beginning of its boom. Informally, our approach defines the beginning of the housing boom in a local market as the point at which house price growth rates exhibit this type of discrete jump.

To formalize this idea, we start with the following reduced form model of house price growth in MSA i at time t:

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

Glaeser, et. al. (2012) implies that $d_{i,t} = d_{i,0}$ for all t if the market is on its steady-state growth path. However, if there is a shock to local productivity, amenities, or expectations at time t then the price growth rate will exhibit a discrete jump in that period. The beginning of a local housing

[12] Glaeser, et. al. (2012) introduce dynamics into Rosen's (1979) and Roback's (1982) classic static model of sptial equilibrium. In this compensating differential framework, house prices (P_i) are the entry fee paid to access the wages (W_i, which reflect productivity) and amenities (A_i) of labor market area *i*. Their model is closed with an assumption that there is some elastically supplied reference market area which is always open to another household. The utility level available in the reference market is given by U*, and establishes the lower bound on utility provided in any market. In the long run, perfect mobility ensures that U* is achieved in all markets, so that in equilibrium, no one has an incentive to move to another place which offers higher utility. A very simple, linear version of this framework would imply that $U^* = W_i + A_i - P_i$, so that $dP_i = dW_i + dA_i$ in equilibrium. The steady state rate of price appreciation need not be zero. Secular trends in house prices can come from an underlying trend in housing demand as long as the market is not in perfectly elastic supply. It can also arise from trends in physical construction costs under certain conditions.

boom can thus be identified by testing for the existence of one or more structural breaks in the parameter $d_{i,t}$. To carry out this test we follow well-established methods in the time series literature for estimating structural breaks.

Borrowing heavily from Estrella's (2003) notation, the null hypothesis is that $d_{i,t}$ is constant for the entire sample period

$$H_0: d_{i,t} = d_{i,0}, \qquad t = 1, \ldots T.$$

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*. Specifically the alternative hypothesis is

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

For any given π_i, it is straightforward to carry out this hypothesis test. However, things are slightly more complicated when π_i is unknown and the determination of its value is the primary interest, as the case here.

To see how we estimate the value of π_i and assess its statistical significance, 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 break point is the value t^* from the set S_i that maximizes the likelihood ratio statistic from a test of H_1 against H_0.[13] That is, for every $t \in S_i$ we construct the likelihood ratio statistic corresponding to a test of H_1 against H_0 for that value of t, and we take the t that produces the largest test-statistic as our estimated break point for MSA i.

Assessing the statistical significance of this breakpoint estimate requires knowing the distribution of the supremum of the likelihood ratio statistic as calculated from among the values in S_i. Let $\xi_i = sup_{S_i} LR$ denote this supremum. Andrews (1993) shows that this distribution can be written as

$$(2)\ P\ (\xi_i > c) = P\left(sup_{\pi_i \in \Pi_{\mathrm{i}}} Q_1(\pi_i) > c\right) = P\left(sup_{1<s<\lambda_i} \frac{\|B_1(s)\|}{s^{1/2}} > c^{1/2}\right),$$

where $\|B_1(s)\|$ is the Bessel process of order 1, $\lambda_i = \pi_{i,2}(1-\pi_{i,1})/\pi_{i,1}(1-\pi_{i,2})$, and

[13] We use the terms supremum and maximum interchangeably in this exposition. Technically, all of the results are in terms of the supremum of the likelihood ratio statistic.

$$Q_1(\pi_i) = \frac{(B_1(\pi_i) - \pi_i B_1(1))'(B_1(\pi_i) - \pi_i B_1(1))}{\pi_i(1-\pi_i)}.$$

Direct calculation of the probability in (2) is non-trivial and prior research has relied on approximations that are typically 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 this method to calculate p-values for the estimated break point, π_i, for each MSA in the sample.

Note that this method does not provide an unbiased estimate of the magnitude of the change in price growth rates at the breakpoint, $d_{i,2}$. Under the null hypothesis that there is no break point, the estimate of $d_{i,2}$ has a nonstandard distribution and OLS estimates of its magnitude will be upwardly biased in absolute value. This problem can lead to an increased chance of falsely concluding that $d_{i,2} \neq 0$ and is a form of specification search bias arising from the fact that the same data is being used to estimate both the timing and the magnitude of the structural break.

Several approaches for adjusting the estimate of the magnitude of structural break have been suggested and are typically based on simulations of the distribution of $d_{i,2}$ under the null hypothesis of no break point (Andrews (1993), Hansen (2000)). Our approach to correcting the estimates of $d_{i,t}$ follows the recently suggested method used by Card, Mas, and Rothstein (2008) of randomly splitting the underlying sample of housing transactions into two and using one sample to estimate the timing of the boom and the other to estimate the magnitude of price changes around that time. The idea is that if the two subsamples are independent, then estimates of $d_{i,2}$ from the second sample, which was not used to estimate the location of the break point, will have a standard distribution even under the null hypothesis of no structural break in the first sample. In practice, we randomly split our sample of unique houses into two and create separate price growth series for each sample of houses. The first price series is used to estimate the timing of the boom following the method just discussed, while the second is used to analyze the magnitude of price changes following housing booms in neighboring markets.[14]

[14] Not accounting for this issue can result in large biases. Card, Mas and Rothstein (2008) noted that their estimates from the full sample were somewhat larger than the estimates from the split sample approach. In our case, if we use the full sample of transactions to estimate both the timing of the beginning of the boom and the magnitude of the jump in price growth at that time, the estimated jumps are almost X% larger, on average, than those arising out of

The procedure above will generate a breakpoint estimate regardless of whether the structural break represents a positive or negative change in the price growth rate. In the cases where the estimated break point is either insignificant or implies a negative change in growth rates, we conclude that the market did not have a boom. That is the case for the 15 interior markets shown in the first panel of Figure 1. For those locations where we do find evidence of a statistically significant and positive break point, we also test for the existence of two breaks against the null hypothesis of only one. To do so, we closely follow Bai (1999) and Bai and Perron (1998) and we refer the reader to Appendix A1 for the details of this procedure. About half of the MSAs were found to have experienced more than one structural break. However, for many of those cases, the secondary breaks were either small economically or not significantly different from zero. The estimation of a secondary break generally does not displace the location of the main structural break either. Moreover, comparison of histograms of timing of local booms based on one-break or two-break methods lead to similar distributions of local booms over time. Given these facts, we simplify our analysis by only using the one-break method in the empirical study below. When necessary, we also report robustness tests based on the multiple break method.

II.b. Contagion in local markets

A potential role for neighbors to influence house price growth in a focal market arises naturally within the dynamic urban model of spatial equilibrium referenced above. While users of that framework typically presume that shocks originate from own market fundamentals, they could arise from neighboring markets as well. In our case, we are interested in whether neighbors that just had housing booms influence housing market outcomes in the focal metropolitan area, all else constant. Note that any such contagion effects could arise from fundamental or behavioral factors.

An example of a fundamental factor generating spatial spillovers is a positive industry or income shock that triggers a housing boom in a local labor market. For example, if there is such a shock in the Silicon Valley, house prices in the San Jose-Sunnyvale-Santa Clara MSA will increase in the short run given supply constraints, and perhaps start a housing boom.

---

the split sample estimation. Part of this difference, of course, could also be attributed to additional random variation that arises when reducing sample sizes by half. This may be an issue for our smaller cities that do not have as many transactions per quarter as the big cities, and are more subject to the influence of potential outliers.

Neighboring areas, such as the San Francisco-Oakland-Fremont MSA (or smaller, but more distant, metropolitan areas in the central valley of California), may eventually benefit from that positive income shock, as some of the Silicon Valley jobs could migrate to nearby areas. Even though such fundamental spillovers may occur with lags, house prices in the neighboring markets should immediately capitalize the expectation of future economic growth.

Housing market shocks could also be disseminated through the credit market channel. In the example above, Silicon Valley lenders may achieve extra profits since foreclosures and delinquencies tend to decline during an economic boom. If those lenders decide to reinvest profits and expand market shares in a nearby MSA – possibly because it is less costly to expand business to nearby communities – then the San Francisco metro may observe a shift in the availability of credit, which will boost its housing market both in the short and long runs.[15] A similar mechanism could exist for land owners and housing investors in the Silicon Valley. Their wealth increases after the beginning of the Silicon Valley housing boom, which could trigger an expansion of investments into neighboring MSAs.[16]

Spillovers could arise even in the absence of expected future fundamental changes in San Francisco or the central valley. Residents in those neighboring markets may be right to think that some type of positive income spillover will occur from the Silicon Valley boom, but they may incorrectly predict its magnitude. Those biased expectations can lead to short-run increases in their housing prices.[17] In addition to this type of behavior, irrational factors may lead residents in the focal market to have not only incorrect, but also non-fundamentally based expectations about future price growth in their market following a shock to a nearby neighboring area. Finally, the housing boom in the Silicon Valley may generate another type of behavioral spillover effect—namely, an increase in Silicon Valley housing prices may lead its own local residents to pay more attention to what is happening in neighboring markets, such as the city of

[15] Ortalo-Magne and Rady (2006) show how increases in the availability of credit to marginal buyers can lead to overreaction in house prices.

[16] Chinco and Mayer (2012) report that out of town speculators played a significant role in the housing boom and that their presence may have exacerbated price increases in some markets. Bayer et al. (2011) similarly document the rise of speculative activity during the housing boom.

[17] A similar mechanism underlies the analysis in King and Wadhwani (1990) who document that contagion in financial markets can arise as a result of attempts by rational agents to infer information from price changes in other markets. Similarly, Clapp et al. (1995) document characteristics of house price dynamics that could be consistent with rational learning. Burnside et al. (2011) and Favara and Song (2011) show how the presence of optimistic agents in the housing market can lead to increases in house price levels and volatility.

San Francisco. Therefore, a shock that makes the focal housing market and its interactions with neighbors more salient to investors may itself lead to stronger contagion.

Regardless of its source, contagion would manifest itself in the form of abnormally large increases in focal market price growth rates immediately following a boom in a neighboring market. Our empirical approach makes use of the estimates of the timing of local booms discussed above in order to test for this type of spillover and its potential mechanisms. However, before discussing the econometric specifications we estimate, we first give a brief description of the data underlying our analysis.

III. Data

Our house price data come from DataQuick, a private data vendor which collects the universe of housing transactions from county recorder's offices in markets across the country. The sample used is for 99 metropolitan areas, with information on over 23 million individual observations ranging from the first quarter of 1993 (1993(1)) through the third quarter of 2009 (2009(3)). We randomly split the sample into two and in each subsample we create a constant quality quarterly price index for each MSA. [18] From these indices we create the annualized growth series used in estimating the timing of the boom and in assessing how neighboring booms affect price growth in focal markets. The mean, standard deviation, and interquartile range for the price index we use to measure magnitudes are reported in the first row of Table 1.

We also create a number of variables to measure fundamentals that are potentially correlated with house price growth and the timing of the beginning of local housing booms. These are reported in subsequent rows of Table 1. We consider three types of fundamentals: (1) demand shifters, such as the average income of mortgage applicants, MSA-level unemployment rates, and net migration flows; (2) buyer characteristics and property traits, including the percentage of speculators, the percentage of minority buyers and the average square footage of transacted housing units; and (3) credit market conditions, measured by the average loan-to-

[18] We create a MSA-level constant quality house price series by quarter using hedonic regressions. Price, in logarithmic form, is modeled as a function of the square footage of the home entered in quadratic form, the number of bedrooms, the number of bathrooms, and the age of the home. We also created a version of the Case and Shiller (1987) repeat sales price index for 14 Case-Shiller markets that overlap with the DataQuick files, and found that the simple correlation of appreciation rates on the two different indexes based on DataQuick is usually higher than 0.9. We employ hedonic price indexes because their data requirements are much less onerous.

value ratio of home purchases, the percentage of mortgages originated by subprime lenders and those insured by the FHA.

To construct many of the demographic measures of home buyers, we merge the DataQuick files with Home Mortgage Disclosure Act (HMDA) data, which provide information on the income and race of all mortgage applicants. In each time period, we calculate the average income of all local loan applicants as reported in HMDA. Similarly, the "Percent Minority" variable reflects the fraction of African-American and Hispanic loan applicants as coded in the HMDA files. Because these measures reflect the characteristics of all mortgage applicants, and not only the set who end up purchasing a home, we take them to be an accurate description of the race and income of potential homebuyers in each market.

MSA-level unemployment rates come from the Bureau of Labor Statistics' *Local Area Unemployment Statistics* series, and net migration flows are calculated using data on county-to-county migration patterns provided on an annual basis by the Internal Revenue Service.

"Percent speculators" refers to the fraction of transactions involving a speculator on either the buyer or the seller side of the transaction. We leverage the fact that we observe the names of both the buyer and seller for each transaction in order to define speculators in a similar way as Bayer, et al. (2011). Specifically we define a person as a speculator if he or she is observed to have 'flipped' at least two homes in the same metropolitan area during the entire course of the sample where a flip is defined as a purchase and sale of the same home within a two-year period. We then consider all transactions in which that person is involved as either the buyer or seller as being "speculative".

Credit market variables include the average loan-to-value ratio (LTV) among homebuyers in DataQuick, the fraction of FHA-insured loans, and the fraction of subprime loans. We use information on the names of the underlying mortgage lenders from the DataQuick files to calculate the share of subprime loans. More specifically, we obtained lists of the top twenty subprime lenders from 1990-onward in a publication now called *Inside Mortgage Finance*.[19] "Percent subprime lenders" is then defined as the share of mortgages issued by these top twenty lenders.

[19] This publication claims to capture up to 85% of all subprime originations in most years. Previously, it was named *B&C Mortgage Finance*. See Chomsisengphet and Pennington-Cross (2006) for more details on these lenders and lists. Other papers such as Mian and Sufi (2009) and Keys, et al. (2010) have access to micro-level FICO scores and use that to define subprime borrowers.

IV. *Econometric model and estimates*

A. Econometric Model

As discussed in Section II, the implications of the underlying dynamic model of urban economics are readily extended to include shocks from neighboring housing markets. We estimate the following reduced form model to gauge the impact of housing market booms of close neighbors on the prices of the focal MSA *m* in Census Division *d* and quarter *t:*

$$(3)\ \log(P_{m,d,t}) = \sum_{r=-4,r\neq 0}^{4} \theta_r^1 * \psi_{m,r}^1 + \sum_{r=-4,r\neq 0}^{4} \theta_r^2 * \psi_{m,r}^2 + \sum_{r=-4,r\neq 0}^{4} \theta_r * \psi_{m,r} + \sum_{\tau=1}^{4} \theta_\tau * \log(P_{m,d,t-\tau}) + \gamma_{d,t} + \epsilon_{m,d,t}$$

The first term on the right-hand side of equation (3) contains the primary variables of interest, the $\psi_{m,r}^1$'s, which are indicators for the years relative to the beginning of the housing boom (r) of the closest neighboring market (neighbor number 1). The coefficients, θ_r^1, on these indicator variables describe how prices in the focal market evolve over the course of its nearest neighbor's housing boom. We define Relative Year 0 to be the 12-month period prior to the beginning of the neighboring market's boom.[20] Relative year 1 then includes the quarter in which the boom starts as well as the subsequent three quarters. Relative years from -2 to +3 are entered individually, with all relative years greater (less) than those numbers binned together.[21] This allows us to see whether there are any important pre-trends and to track the build-up of the boom after it starts. The second term on the right-hand side of (3) includes an analogous set of controls, denoted $\psi_{m,r}^2$ for the second closest neighbor. In a set of robustness tests presented below, we also control for log prices of other near MSAs, to make sure that our elasticity estimates from the nearest neighbor are not confounded by other neighbors, or other neighboring shocks.

---

[20] We work with 12 month periods because there is noise in the quarterly data that is not due solely to error in the estimation of the break point. For example, it is common for there to be at least a one quarter difference between the time that a transactions price is agreed upon and when the actual closing occurs. In addition, we know that prices in housing markets do not follow a random walk, but move slowly and are strongly positively correlated over short horizons (Case and Shiller (1987, 1989)). Locations for which we do not estimate a statistically significant boom are still assigned a relative year according to their estimated break points.

[21] We did estimate all our models with lengthier spans of individual relative years controlled for, but they did not yield any new insights beyond those reported below.

Physical proximity is the most natural measure of distance, and the specifications reported below assign the nearest neighbor based on the number of miles between the centroids of the relevant metropolitan areas.[22] We also experimented with a measure of proximity based on migration flows between pairs of markets. Those results were qualitatively and quantitatively similar, so we do not report them for space reasons.[23]

This specification also controls for the time line of the focal market's boom via the third term on the right-hand side of equation (3). By including the relative year effects of the focal MSA (denoted $\psi_{m,r}$, where the absence of a superscript indicates the variable refers to the focal market itself), we control for the average increase in prices of the focal market over the course of its own boom. This vector serves a similar role to the adjustments made in related financial economics research to deal with upward bias in contagion estimates arising from volatility being higher in all markets during 'crisis' periods (e.g., Forbes and Rigobon (2002)).

The remaining terms of equation (3) include four quarterly lags of (log) focal market prices to control for potentially unobserved time-varying characteristics of MSA m[24], as well as Census Division-by-quarter fixed effects to deal with common regional shocks that might influence close neighbors simultaneously. Finally, note that we do not control for contemporaneous focal market fundamentals in this baseline specification. This is because they could represent intermediate outcomes through which the contagion effect may be operating. In the mechanisms section below, we will directly estimate the impact of the neighbor's housing boom on those intermediate outcomes, and also test whether their inclusion in equation (3) mitigates the contagion effect.

B. Main estimates

[22] Distances are calculated using the full set of MSAs according to the 2000 Census. Because we only have price data for 99 of these MSAs, some data for nearest neighbors remain empty in 24 cases. In our regressions, we create an indicator for whether we have price data for the nearest neighbor, and interact this indicator with the relevant relative years. The results are qualitatively similar when we drop MSAs with missing neighbors' price data and also when we calculate distances using only the 75 MSAs for which we have data.

[23] This economic measure of distance based on migration flows between metropolitan areas has been used in other research (Sinai and Souleles (2009)). It is strongly positively correlated with geographic distance between markets. For example, the probability of the physically closest neighbor also being the closest economic neighbor is 57%, and the probability it is one of the two closest economic neighbors is 76%. Hence, it is not surprising these two measures of distance yield similar results.

[24] We also considered specifications that dispense with the lags of the dependent variable in favor of MSA fixed effects. Results from these specifications are qualitatively similar. However, we believe that the lagged dependent variable specification is more appropriate given that omitted time-varying common factors are more likely to confound the contagion effect than unobservable but fixed MSA-specific characteristics.

Column 1 of Table 2 reports baseline estimates of equation (3) for the metropolitan areas whose nearest neighbors had statistically significant booms. The reported coefficients show that prices were relatively stable in the three years prior to the beginning of the boom in the nearest neighboring MSA, so that there is no evidence of a pre-trend. In the year that the neighboring MSA begins its boom (Relative Year [1]), focal market prices then jump 0.87 percentage points and remain almost 1% higher for another couple of years. Column 2 reports the analogous coefficients on the relative year dummies for the set of neighboring MSAs for which we do not estimate a statistically significant or positive break point. These results confirm that a positive effect is only detected when the neighboring MSA actually had a housing boom. Hence, we find spillovers on focal market prices only if the nearest neighbor actually experienced a significantly positive shock. Columns 3 and 4 show analogous estimates for the second closest neighbors. The coefficients on the timeline of the boom for 2nd nearest neighbor generally are not statistically different from zero regardless of whether this particular neighbor had a housing boom. Hence spillovers mostly arise from the closest neighbor, and we will focus on those results in the remaining of the paper.[25]

In order to determine the elasticity of focal market housing price growth with respect to near neighbors' price growth, we need to gauge the magnitude of the housing boom for the nearest neighbors. The starting point of this exercise is to estimate a version of equation (3) that uses the log price of the nearest neighbor as the dependent variable. Table 3 reports those results. Pre-boom prices are trending down a bit in these results, and then jump 3% in the first year of the housing boom. By the third year of the boom, prices are 8% higher than the pre-boom period, and are more than 11% higher in subsequent years.[26]

An upper bound on the implied elasticity can be computed by using only the estimates of price changes in the first year of the boom under the assumption that agents in these markets are myopic. Combining these figures with the estimates from Table 2 yields an elasticity of 0.27.[27]

---

[25] If we do not control for the time line of the focal market's boom, the point estimates of the nearest neighbor's contagion effects are about 20% higher than those reported just below for our baseline specification in Table 2. However, the results are not statistically different once the standard errors are taken into account. While this is consistent with price volatility being artificially high when the focal market itself is booming, we note that including these indicators could also be controlling for intermediate outcomes.

[26] These magnitudes are similar to those in Ferreira and Gyourko (2011) who conduct a similar exercise using the full set of MSAs, not just those who serve as closest neighbors to some other MSA in the sample.

[27] This is the result of dividing the 0.008688 spillover estimate from column 1 of Table 2 by 0.03224 from column 1 of Table 3.

Smaller elasticities result if we consider cumulative price changes by the third year of the shock. In that case, the elasticity falls to about 0.10.

While differences in data and methodologies make it difficult to directly compare our estimates of contagion to others in the literature, our magnitudes appear to be smaller and we find that spillovers only arise from the closest neighbors.[28] We suspect this is partially due to our empirical strategy that minimizes specification search bias and includes various controls to deal with omitted factors related to common shocks.

A rough gauge of the relative importance of contagion in fomenting local booms can be made as follows. Ferreira and Gyourko (2011) concluded that jumps in one fundamental, local income growth, could account for one-half of the magnitude of the jump in price growth at the beginning of local booms (on average). Our estimates indicate that contagion did not play as important a role, but it still was economically meaningful, as it can account for over one-quarter of the jump at the start of the boom.[29] Using different and more conservative figures and assumptions reduce the share, but nothing reasonable can drive it below 10%.[30]

Finally, the economic interpretation of the contagion estimates for the years after the beginning of the boom may be more complicated due to potential feedback effects. Feedback effects are less of a concern in the first year of the boom, as we showed that prices, roughly speaking, are in equilibrium right before that moment. It may not play a major role in subsequent years either, especially if only 10% or less of the main price effect propagates across close neighbors. Nonetheless, contagion estimates for relative years two and three, for example, are better thought of as reduced form estimates that include the impact of recent contagion, but

[28] For example, Cotter et al. (2011) regress housing price appreciation in eleven MSAs near San Francisco on the contemporaneous and 3 quarterly lags of San Francisco's housing appreciation rate and report coefficients ranging from 0.05-0.67 on lagged housing price appreciation of San Francisco. Fuss et al. (2011) model the volatility spillover intensity and suggest that a 1 percentage point shock in housing return in Las Vegas could result in an eventual 0.19 percentage point increase in housing return in San Diego.

[29] To see this more clearly, start with the 0.27 elasticity just discussed. Given the 3.2% average jump in Relative Year [1] reported in Table 3, that aforementioned elasticity implies that about 0.9 points, or about 27 percent of the jump at the start of the boom, can be explained by contagion.

[30] It is not clear what figure to use for the average level of price growth in the denominator of this ratio, which is why we focus more on the elasticity. The 3.2% number used here based on Table 3's results is close to the log price changes reported in Ferreira and Gyourko (2011). However, one also could use the 6.5% average jump in price growth rate (not the change in the log prices) also reported by those authors. That number arises from an estimate that does not control for any year or metropolitan area fixed effects. With that denominator, the 0.9 points amounts to about 14% of the jump at the start of the boom. As noted in the text, there are no reasonable assumptions one could make that drive the share below 10%.

that also embed a share of contagion from the complete path of price appreciation since the beginning of the boom.

C. Heterogeneity in the Contagion Effect

We test for heterogeneity in the average contagion effect along a number of dimensions. The first is distance. We already found that only the nearest neighbor matters. A natural extension is to ask whether the strength of the contagion impact associated with the closest neighbor increases with its proximity to its focal market. The first two columns of Table 4 show that the answer is no. Those figures are the output from a regression like that in equation (3) which further interacts the neighbor's relative year dummies with an indicator for whether that neighbor is more or less than the median distance of about 40 miles away from the focal market. Although the estimates tend to be imprecise, the point estimates show little difference between relatively close and farther away nearest neighbors, especially around the time when the neighbor's boom begins. Thus, contagion effects arise only from the nearest neighbor, but they do not vary materially based on how close that nearest neighbor is.[31]

It also seems natural to ask whether contagion impacts depend upon the relative sizes of the focal and neighbor markets. To investigate this, we classified focal MSAs whose population sizes are within 50% of the population of the closest neighbor as being of similar size. They are considered larger if they have at least 50% more population, and smaller if they have less than 50% of the population of the closest neighbor. Appendix Figure 2 shows the distribution of relative population size, and the thresholds used to determine groups of MSAs for both geographic and economic neighbors. The estimates reported in the third and fourth columns of Table 4 indicate that the contagion effect on a focal market is larger if the nearest neighbor is substantially bigger than the focal area. Prices in the focal market are 1.3% higher immediately when the large near neighbor booms, but are little changed when the focal market is bigger (row 3 for Relative Year [1]). Prices stay higher in subsequent years for focal markets being influenced by large neighbors. In sum, size does matter and in an intuitive way in the sense that

[31] The interquartile range of distances between neighboring markets runs from 30-48 miles, so there is not much variation for much of the sample. The mean is larger at 74 miles, but that reflects the influence of Honolulu, whose nearest neighbor is over 2,000 miles away. The next biggest distance is 111 miles. We also experimented with alternative groupings such as dividing markets into whether their nearest neighbor was less than 30 miles away, from 36-60 miles away, and greater than 60 miles away. The results were no different from those reported here.

contagion effects are much larger (and consistently statistically significant) if the nearest neighbor is large relative to the focal market.

The final dimension along which we investigated whether there was any heterogeneity was by the degree of the focal market's elasticity of housing supply. For this test, we split the focal MSAs into three groups according to the supply elasticities provided by Saiz (2010).[32] Results are reported in the final two columns of Table 4 for the bottom third (supply inelastic) and top third (supply elastic) groups. Note that prices do not jump when the closest neighbor of an inelastically-supplied market begins to boom. However, for the most elastically-supplied metros, prices in the focal market are about 2% higher if its nearest neighbor begins to boom. The gap increases to about 3% by the third year of the boom. Performing calculations analogous to those discussed above for the economic importance of the average contagion effect show that spillovers could account for nearly two-thirds of the jump in prices at the beginning of booms in the most elastically supplied markets. Basic economics suggests that any contagion effects would be more likely to be capitalized in the inelastically supplied markets, *ceteris paribus*, so this outcome may seem counterintuitive at first glance. However, all else is not constant in this case. It turns out that a disproportionately large share of these markets are large coastal metropolitan areas with relatively small neighbors.[33] So, at least some of the variation we document by degree of supply elasticity could have been driven by the size results just discussed.[34]

D. Mechanisms

[32] Saiz's supply elasticity estimates are available for only 76 of our metropolitan areas, so we start with a smaller sample for this particular analysis.

[33] Included in this most inelastic tercile are the metropolitan areas of Barnstable Town, MA, Boston-Cambridge-Quincy, MA-NH, Bridgeport-Stamford-Norwalk, CT, Cleveland-Elyria-Mentor, OH, Deltona-Daytona Beach-Ormond Beach, FL, Eugene-Springfield, OR, Jacksonville, FL, Los Angeles-Long Beach-Santa Ana, CA, New Haven-Milford, CT, New York-Northern New Jersey-Long Island, NY-NJ-PA, Oxnard-Thousand Oaks-Ventura, CA, Palm Bay-Melboune-Titusville, FL, Portland-Vancouver-Beaverton, OR-WA, Port St. Lucie-Fort Pierce, FL, Riverside-San Bernardino-Ontario, CA, San Diego-Carlsbad-San Marcos, CA, San Francisco-Oakland-Fremont, CA, Santa Rosa-Petaluma, CA, Sarasota-Bradenton-Venice, FL, Seattle-Tacoma-Bellevue, WA, Tampa-St. Petersburg-Clearwater, FL, and Vallejo-Fairfield, CA. Each of these areas also has one of the 99 markets in our sample as its closest market.

[34] We also investigated whether contagion effects varied by the relative timing of booms in order to see if contagion impacts occurred primarily when the focal market itself was booming. Appendix Figure 1 shows a histogram of the difference between the timing of the boom of the focal MSA and its closest neighbor. There is wide variation in timing of the booms of these pairs of markets. We find that contagion effect seems concentrated in focal MSAs that were already booming when the closest neighbor started to boom. This result suggests salience being a feature of the contagion effect. But as with the heterogeneity by supply elasticity, relative timing is correlated with other factors, such as market size.

In this subsection, we investigate some of the potential mechanisms that could account for the influence of a near neighbor's boom on house prices in the focal market. We are particularly interested in whether our contagion effects are fundamentally-based in the sense discussed in Section II. If not, the relevance of our results for policy makers is increased, as they well may want to reevaluate their past practice of not intervening in response to asset booms in housing markets if their spread is based on some type of irrational exuberance or otherwise mistaken expectations.

We begin by asking whether there are visible economic changes in the focal market that may be driven by the neighboring market boom. We then alter our baseline specification to account for these potential fundamental drivers. Four fundamentals are investigated, with each having received prominent mention in previous academic research or by policy makers and the popular press. They are focal market income, mortgage market activity, net migration flows into the focal market and the share of house 'flips' in overall market sales. Table 5 reports results using these local market traits as the dependent variable in a specification similar to that in equation (3), with one difference being that here we include MSA fixed effects rather than own price lags.

Column 1 reports estimates for focal market income, where income is defined as the average income for all mortgage applicants in that market and quarter. If what is driving our contagion result is a real spillover such as focal market income going up because of boom in a nearby market, then we should see it changing along the timeline of the neighbor's boom. Table 5's results show that there is a jump from zero to 1% in Relative Year [1], but this impact is not precisely estimated (the t-statistic is 1.3). The results are similar for the second and third relative years. It is only at least four years after the nearest neighbor booms that we see focal market income higher by a statistically significant amount. One would not want to interpret these coefficients as proving that contagion does not operate via spillovers onto focal market income, but they also provide no robust evidence to the contrary.

The next two columns investigate whether the contagion effect might operate through credit markets in some fashion. We approach this question by examining two aspects of credit lender activities. First, we investigate whether the mortgage lender bases become more similar during and after the nearest neighbor booms. The intuition is that if lenders observe a boom in that neighbor, they might increase their activity in the focal market for reasons just discussed in

Section II. We use a proportional index to measure lender similarity.[35] Second, we investigate whether lenders speed up mortgage lending during and after the closest neighbor booms by calculating the rate at which each lender increases mortgage issuance in the focal market. The regression results in Column 2 and 3 indicate that on average, both lender similarity and lending amount largely are unaffected by the housing boom of the nearest neighbor.[36] Therefore, based on our two measures of lender activities, we did not find robust and material role of lenders in causing contagion.[37]

The fourth and fifth columns report analogous specifications using net migration between two MSAs (which uses IRS data on annual tax records) and focal market flippers (based on the fraction of transactions conducted by speculators) as the dependent variable. Both sets of results show no discernible effect before or after the beginning of the nearest neighbor's housing boom.[38]

The fact that only focal market income shows any correlation with the timeline of the nearest neighbor's boom, especially its beginning, indicates that these fundamental factors are unlikely to be able to account for our estimated contagion elasticity. Table 6 provides additional support for this conclusion with alternative specifications similar to our baseline equation (3) that use (log) focal market price as the dependent variable. The first column reports results from a

---

[35] This lender similary index is calculated as $1 - \frac{1}{2}\sum_{l=1}^{L}|m_l - n_l|$, where *m* and *n* are the market shares of lender *l* in the focal and the nearest neighbor markets, respectively. A value of zero implies no similarity, while a value of one means that each lender has the same shares in both markets.

[36] We also found that lending amount by subprime lenders also does not respond to the neighbor's boom. If we restrict the analysis to the top 5 lenders with the highest lending amount for each market, we observe a jump in lender similarity index in Relative Year [1]. However, this result is not robust when we go from top 5 to top 10 or top 3 lenders. Including only top lenders also leads to a jump in lending growth rate in Relative Year [2] (the coefficient is around 0.5), but not in Relative Year [1]. However, lending growth also is higher before neighbor's boom, with coefficients on Relative Year [-1] and Relative Year [-2] being in the range of 0.2-0.4. Given this noticeable pre-trend, it is difficult to come to any conclusion that major credit lenders are the channel in disseminating positive housing market shocks.

[37] This average result, however, does not exclude the possibility of major credit lenders responding to neighbor's booms in other dimensions. The increase in similarity index among top lenders and the dip in their lending amount right at neighbor's boom may suggest some bank-level spillovers or substitution effects going on across markets. To fully understand the role of credit markets in this contagion context would require a closer examination of a full spectrum of lender behaviors, including those at the corporate level (e.g., mergers and acquisitions, and shocks to other business sectors of the lenders), which is beyond the scope of this paper.

[38] We also investigated the shares of experienced and inexperienced speculators, respectively. Speculators are defined as experienced if having flipped at least four homes during the sample period, or inexperienced if otherwise. We did not find significant jumps in either experienced or inexperienced flippers when the closest neighbor begins its boom. However, the market share of experienced flippers declined by up to 0.5 percentage point (which is statistically significant) since two years after the neighbor booms, which is in line with those types of flippers being more sophisticated in timing the housing cycle and maximizing their return (Bayer et al., 2011).

model that includes focal market fundamental controls and the log average price of near neighbors on the right-hand side, in addition to the standard controls from the baseline specification in Table 2. Our fundamental controls are local incomes, migration, subprime and FHA lending market shares, percentage of speculative buyers, percent minority, average LTV, average square footage, and the local unemployment rate. Note that these new estimates are very similar to the baseline results presented in Table 2. Controlling for these fundamentals does not change the magnitudes or time pattern of estimated contagion effects very much, indicating that the spatial spillovers are not being transmitted via the fundamentals that we consider here.

Thus far, we have abstracted from expectations of future fundamental factors, effectively treating actors as myopic. The second column of each panel in Table 6 begins to address this issue by adding four leads of own market income to the previous specification. Effectively, this presumes that local residents can fully predict the path of local incomes over the next four quarters. The inclusion of such stylized expectations does not change the estimated contagion effect. The third and final column of each panel reports results from adding four quarterly leads of all fundamentals, not just income. Once again, the magnitudes of the point estimates as well as the time pattern are relatively unchanged.

That we find no evidence that these spatial spillovers work through fundamental factors has potentially important implications for how policy makers should view intervention during housing booms. To the extent the booms are spread by non-fundamental forces (e.g., some type of Keynesian/Shillerian irrational exuberance), they might want to try to stop them from growing in scale and scope. Of course, fundamentals could encompass more than just income growth (or speculators, migration, etc.), and we would like to control for expectations as generally as possible. One extreme way that does eliminate the impact of contagion is to presume that the future path of price growth in the focal market is known with certainty in every period. In this context, the effect of a nearby housing boom on future prices in the focal MSA is immediately known and expectations are simply equal to the future value of price growth in the focal area. Adding four quarterly price growth leads completely wipes out the contagion impact. This indicates that the spillover could be operating primarily via expectations. Unless those expectations are based solely on fundamentals, the implications just discussed still hold. The likelihood that contagion operates via effects on expectations makes that issue an essential component of future research, but that clearly is beyond the scope of the current paper.

V. Additional Analysis: Instrumental Variable, Extensive Margin, and Bust

A. Alternative Model and Instrumental Variable

In this subsection, we relate price changes in the focal MSA with price changes from all neighbors, not just the nearest two. This direct estimation of the contagion elasticity has the benefit of allowing us to use an instrumental variable strategy to deal with omitted factors. Also, by interacting the price changes of the closest neighbor with a set of relative year dummies, we are able to explore how this effect varies over the course of a neighbor market's housing boom. The downside of this approach is that the specification does not allow us to fully observe the dynamic pattern of contagion, as we restrict the effect of neighboring price changes to operate through only one quarterly lag.

More specifically, we group neighboring locations into bins based on their distance from the focal market and estimate the following equation:

$$(4)\ \Delta P_{m,d,t} = \sum_{r=-4,r\neq 0}^{4} \rho_r^1 * \psi_{m,r}^1 * \Delta\bar{P}_{m,t-1}^1 + \sum_{k=2}^{K} \rho^k * \Delta\bar{P}_{m,t-1}^k + \sum_{\tau=1}^{4} \rho_\tau * \Delta P_{m,d,t-\tau} + \sum_{r=-4,r\neq 0}^{4} \rho_r * \psi_{m,r} + \gamma_{d,t} + \xi_{m,d,t}$$

where $\Delta\bar{P}_{m,t-1}^k$ is the lagged average price growth among neighboring MSAs falling into bin k for focal MSA *m*. In theory, we could allow each neighbor to be in its own bin based on how close it was to the focal area. However, that turns out not to be practical due to data limitations, so we bin neighbors based on distance rankings 1,2, 3-5, 6-10, 11-50, and 51+.[39] This makes the coefficients, ρ^k, the elasticity of focal area current price growth with respect to the *average* of lagged price growth among neighbors in bin *k*. For the closest neighbor, we further interact the lagged price growth variable with relative years to that neighbor's boom -- resulting in a coefficient, ρ_r^1, for each of the neighbor's relative years, as shown in the first term on the right-hand side of (4) Relative year 0 again is the omitted category in all specifications. Thus, the

[39] As in our main regressions, we include an indicator for whether we have price data for a given bin and interact this indicator with the relevant lagged average price variable. Results are qualitatively similar when we drop MSAs with missing neighbors' price data and also when we calculate distances using only the 99 MSAs for which we have data.

coefficients on the lagged average price growth of the closest neighbor are interpreted relative to the effect in the 12-month period prior to that neighbor's boom.

One concern is that, even after including lags of the dependent variable (the third term on the right-hand side of (4)) and area-by-time fixed effects (the $\gamma_{d,t}$ vector), there still could be some common omitted factors helping drive the observed correlations. Ideally, we would like an instrumental variable that shifts the lagged average price growth of the focal MSA's closest neighbor, but does not directly affect the contemporaneous appreciation rate in the focal market itself. If taken literally, our estimating equation implies that further lags of the neighbor's price growth could potentially serve as an instrument because those variables would only affect the focal market's contemporaneous price growth through their impact on the lagged neighbor's price growth. This leads us to instrument for the lagged average price growth in each group of neighbors using one further lag of the average price growth among the relevant neighboring areas.[40]

Table 7 reports the results of estimating equation (4). Once again, there is a clear pattern that shows a shift in the importance of contagion right after the first year that close neighbors boom. Estimates for the first year of each of the neighbor's boom are 0.148, or approximately half the elasticity derived from our baseline estimates when considering the first year of the housing boom. Estimates for subsequent relative years fade relatively slowly, reaching a 0.1 elasticity that is similar to the baseline results.

B. A Hazard Model of Housing Boom Contagion

Our work above focuses entirely on the magnitude of contagion. In this section, we give empirical content to the extensive margin on the timing of booms that was suggested by Figure 1. We estimate simple hazard models to see if the probability of a focal market booming is influenced by the fact that neighbors boomed previously, after controlling for a host of covariates that also might account for the beginning of a boom.

Recent work on technology diffusion provides an intuitive way to generate contagion that is particularly appropriate for the specifications estimated in this subsection. In some of that

[40] It turns out that the second quarterly lag of neighbor's price growth has a small and statistically insignificant impact on the focal MSA price appreciation, after controlling for all other covariates. In the specifications that interact the lagged average price growth of the neighbors with the focal MSA relative year indicators, we also interact lagged neighbors income growth and the second lag of neighbor's price growth with the focal area's relative year and include the full set of these interactions as instruments.

research, contagion refers to a process in which people adopt a new technology when they physically meet with others who have already adopted it (Young (2009); Comin *et al*. (2012)). In our context of housing booms, this suggests that the probability of one metropolitan area entering a housing boom increases with its connection to other such areas that already have boomed. A standard assumption is that the intensity of the connection decays in distance between MSA pairs at a constant rate. Given that assumption, it is straightforward to generate the following two conclusions from a simple model: (a) if MSA q enters a housing boom at time t, that increases the hazard of MSA $m \neq q$ having a housing boom at time $t+1$; and (b) this contagious effect is larger the closer is MSA m to MSA q.[41]

To investigate these implications, we begin by estimating whether the beginning of a boom in neighboring MSAs affects the hazard of focal MSAs entering a boom. We consider the following proportional hazard model relating the hazard of each focal MSA entering a boom in quarter $t, h_m(t)$, to a series of factors as noted in the following equation.

$$(5)\ h_m(t) = h_0(t)\exp\left\{\sum_{k=1}^{K}\gamma_k \sum_{j\in N_m^k} Boom_{j,t-1} + \Delta X_{m,t}\beta + \sum_{\tau=1}^{4}\theta_\tau\,\Delta P_{m,t-\tau} + \eta PctBoomed_{t-1} + \delta_R\right\}$$

The primary coefficients of interest are γ_k 's on the variable $\sum_{j\in N_i^k} Boom_{j,t-1}$ which reflects the number of MSAs among neighbors in bin k that began their boom in the previous quarter. We allow for multiple groups of neighbors (indexed by *k*) based on how close they are to the focal MSA. As before, we rank each focal MSA's neighbors and group them into K mutually exclusive bins based on these ranks, with N_m^k denoting the set of neighbors in bin k for focal MSA m. A positive coefficient γ_k would suggest a positive contagion effect. If the contagion effect decays with distance, we should also expect the γ_k's of closer neighbors to be larger than those of farther away neighbors. As with the estimation of magnitudes, we control for many potential correlates of a housing boom through the vector $\Delta X_{m,t}$.

We use the same bins as in Table 7. Table 8 then lists summary statistics on the number of booms started across different bins of neighbors. The very small means for the bins

[41] See Appendix A2 for a more formal presentation of how that approach generates specifications of the type we estimate in this section.

containing five or fewer markets document how unlikely it is that even a single boom began in any given quarter among the few markets in those bins. Thus, the lagged value of this variable ($\sum_{j \in N_m^k} Boom_{j,t-1}$) will contain many zeroes, making it difficult to estimate precise contagion effects about the timing of the beginning of local market booms. Nevertheless, we use these bins initially given the differential importance of the two closest neighbors in the contagion magnitude results and discuss findings later that use larger bin sizes.

Our baseline results are from a common parametric specification, the exponential model, which assumes a flat baseline hazard $h_0(t) = \exp(\alpha_0)$. Implied changes in the probability of the focal area experiencing a boom this quarter (i.e., the hazard ratio) from one more neighbor having experienced a boom last quarter are reported in Table 9. The unconditional hazard, which is reported in column 1, indicates that the probability of the focal MSA booming this quarter more than triples if one more of its two physically closest neighbors boomed last quarter. The coefficients on the next three bins (Neighbors[3-5], Neighbors[6-10], and Neighbors[11-50]) are much smaller and none is statistically significant at conventional levels. That the lack of significance might be due to the nature of these variables having so little natural variation is indicated by the statistically significant impact of the group of furthest away markets (Neighbors[51+]). Still, it is lagged booms among the two closest neighbors that have the strongest correlation with a contemporary boom in the focal market.

Controlling for a standard set of covariates lowers the estimated effect substantially, as reported in column two. This time, the coefficients on the bins for all but the two closest neighbors are close to or below 1, indicating they have no positive impact on the probability of the focal market booming this period. And, the coefficient on the bin for the two closest neighbors falls by more than half. Even the 49% increase in probability implied by the coefficients on Neighbors[1-2] is not statistically significant, although that could be due to the nature of our data as discussed above. This conclusion is supported by the findings reported in column 3, which uses a Weibull hazard.[42] It reports a statistically significant correlation between lagged booms in very near neighbors and contemporaneous booms in the focal markets. These results show that if one of the two closest neighbors boomed last quarter, the probability of the focal market booming this quarter is about 70%-80% higher. The standardized marginal effect is

[42] A Weibull hazard model presumes a monotonic baseline hazard $h_0(t) = pt^{p-1}\exp(\alpha_0)$. We experimented with different functional forms to see if the pattern of results was materially affected.

smaller, of course, given the 0.18 standard deviation on Neighbors[1-2]. Using the middle of the range estimate from the Weibull hazard model, it is about 30% (i.e., 1.69*0.18).

These results only allow for a contagion effect from a single quarterly lag of neighbors' booms. We also estimated models that allowed for an increase in the number of booming neighbors over the past 6 to 12 months. There is a modest increase in the hazard (to 32%) if we allow for booms in any of the previous two quarters, but the result does not increase further if we allow for booms in any of the previous four quarters. Thus, the spatial effect estimated here appears to happen fairly quickly. This is also consistent with the elasticity estimates reported above.

The magnitude of the economic impact is hard to gauge on its own. We can gain some useful perspective by comparing it to the impacts on the hazard ratio of standard deviation changes in other variables. Among the underlying controls, the focal market's current income growth and previous quarter's price growth also were highly statistically significant predictors of a higher probability of booming. A one standard deviation higher own income growth rate is associated with about a 28% higher probability of the focal market booming. A one standard deviation higher rate of lagged house price appreciation is much more influential, as it is associated with an 87% higher hazard ratio. Thus, the standardized marginal effect of a boom in a very close neighbor appears to be quite influential and on a par with a standard deviation increase in its own income growth rate.[43] And, that we still find a meaningful influence of lagged near neighbors after controlling for everything else shows that the implications of the 'eyeball econometrics' from Figure 1 are not entirely due to many other potentially important factors.[44]

C. Was There Contagion During the Spread of the Housing Bust?

---

[43] The growth rate in the percentage of buyers with mortgages insured by the FHA also is a very powerful control. As expected, it is associated with a lower probability of booming. Specifically, a one standard deviation increase in that share is associated with an 86% fall in the hazard. Other statistically significant controls include the growth rate of the metropolitan area unemployment rate, as well as the second and fourth quarterly lags of focal market house price appreciation.

[44] Conditional hazard model estimates using more aggregate bins (Neighbors[1-10], Neighbors[11-50] and Neighbors[51+]) do not show any impact on the timing of the beginning of the boom in focal markets. Averaging across the ten closest neighbors masks the distinct impact of the two closest neighbors. As in the analysis of the magnitude of contagion, only the closest neighbors appear to matter.

Our results indicate that contagion played a statistically and economically meaningful role in the timing and magnitude of the spread of housing booms across metropolitan areas. For completeness, we also explored the extent to which the same is true for the bust. In many respects, analysis of the bust is more challenging. One of the challenges is in deciding how it is defined and determined. We choose it to be the quarter in which nominal house prices peaked in the relevant MSA. While that may be intuitive, it also is much more *ad hoc* than our definition of the beginning of the boom, which is based on an external prediction of an economic model. Another challenge is that the bust was much more temporally correlated across markets than was the boom. This can be seen in Figure 3 which plots histograms of the quarters in which local market booms and busts began. This plot includes all markets that had a statistically significant boom. Note that the temporal concentration of busts is much greater, with every market experiencing a peak in prices within a 2.5-3 year period from mid-2005 through early 2008.

Figure 4 then provides more geographical detail with its plots of market busts over time. The first panel in this figure is identical to the first panel in Figure 1 and plots the MSAs for which we never estimate a statistically significant boom. The remaining panels show the timing of the bust among those MSAs that experienced booms. Unlike Figure 1's plots of the time line of metropolitan area booms, we see markets in all parts of the country, not just in coastal California and upper New England, with early price peaks between 2003-2005. The largest fraction of those peaks happening in last two quarters of 2005. Thus, the beginning of the bust is more national in scope than was the beginning of the boom. The subsequent plots in this figure do show a spreading out of the 'busts' to nearby markets. In the west, prices tended to peak earlier in interior markets and then spread to the coast. In Florida, the first price peaks were in markets on both coasts of that state. Peaking then occurred in a few other coastal markets before spreading to interior markets.

While this has the flavor of contagion seen in the start of the boom, more detailed analysis shows this not to be the case. For example, Table 10 reports hazard model estimates akin to those in Table 9, except here the dependent variable is the start of the bust, not the start of the boom. Unconditionally, lagged busts of neighbors are positively correlated with contemporary busts in focal markets, and near neighbors matter the most (column 1). However, column 2 of this table shows this conclusion of 'eyeball econometrics' from Figure 4 does not survive the inclusion of covariates.[45]

In Table 11 we estimated price specifications for the bust akin to those in Table 2. First, using geographic distance, we do not see significant jumps in the magnitude of the contagion effect in relative year one. But focal prices decline by 1.2% and 3.8% in relative years two and three respectively. Results using economic distance follow a similar pattern. When compared to the magnitude of the price decline in the neighboring MSA, we find an elasticity of around 0.15. Those results are surprisingly similar to the ones observed during the spread of the housing boom.

VI. Conclusions

We provide estimates of the role of contagion in the most recent American housing boom and bust. We find a statistically and economically meaningful role for contagion during the beginning and spread of the housing boom, and mixed evidence of contagion for the bust. Our key results are as follows. First, contagion impacts arise only from the very closest neighbors. There is no evidence of spillovers associated with more distant neighbors. The elasticity of focal market prices with respect to changes in its nearest neighbor's prices is in the range of 0.10-0.27. Back-of-the-envelope calculations suggest this is large enough to account for up to 30% of the jump in prices at the beginning of local booms, on average.

Finally, we found that local fundamentals and expectations of future fundamentals have very limited ability to account for our estimated contagion effect. That contagion transmission is not associated with local fundamentals suggests a potential role for non-rational forces. That is an issue in urgent need of new research because, if contagion does reflect some type of irrational exuberance, policy makers may want to rethink their past policy of not intervening to stop the spread of asset booms in housing.

[45] All markets are used in this particular estimation, including those that did not boom. The results are virtually identical if we restrict the sample to those that did boom.

References

Adelino, Manuel, Kristopher Gerardi and Paul Willen. "Why Don't Lenders Renegotiate More Home Mortgages? Redefaults, Self-Cures and Securitization", FRB Atlanta Working Paper 2009-17.

Andrews, Donald. "Tests for Parameter Instability and Structural Change with Unknown Change Point", *Econometrica*, Vol. 61, No.1 (July 1993): 821-856.

Arce, Oscar and David Lopez-Salido. "Housing Bubbles", *American Economic Journal: Macroeconomics*, Vol. 3 (January 2011): 212-241.

Avery, Robert and Kenneth Brevoort. "The Subprime Crisis: How Much Did Lender Regulation Matter". Division of Research and Statistics. Board of Governors of the Federal Reserve System. August 2010.

Bai, Jushan and Pierre Perron. "Estimating and Testing Linear Models with Multiple Structural Changes", *Econometrica*, Vol. 66, No.1 (January 1998): 47-78.

Bai, Jushan. "Likelihood Ratio Tests for Multiple Structural Changes", *Journal of Econometrics*, Vol. 91, No. 2 (August 1999): 299-323.

Bajari, Patrick, Chenghuan Sean Chu and Minjung Park. "An Empirical Model of Subprime Mortgage Default from 2000-2007", NBER Working Paper no. 14625, December 2008.

Banerjee, Abhijit V., "A Simple Model of Herd Behavior," *The Quarterly Journal of Economics*, Vol. 107, No. 3 (1992): 797–817.

Bayer, Patrick, Christopher Geissler and James Roberts. "Speculators and Middlemen: The Role of Flippers in the Housing Market", NBER Working Paper No. 16784, February 2011.

Bhutta, Neil. "Regression Discontinuity Estimates of the Effects of the GSE Act of 1992", *Finance and Economics Discussion Series*, *Divisions of Research & Statistics and Monetary Affairs, Federal Reserve Board, Washington, D.C.*, March 2009.

Bikhchandani, David, Hirshleifer Sushil and Ivo Welch, "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades," *The Journal of Political Economy*, Vol. 100, No. 5 (October 1992): 992–1026.

Bubb, Ryan and Alex Kaufman. "Securitization and Moral Hazard: Evidence from a Lender Cut-Off Rule", working paper, 2009 (SSRN).

Burnside, Craig, Martin Eichenbaum and Sergio Rebelo. "Understanding Booms and Busts in Housing Markets", NBER Working Paper No. 16734, January 2011.

Calvo, Sarah, and Carmen M. Reinhart, “Capital Flows to Latin America: Is There Evidence of Contagion Effects?” in Guillermo A. Calvo, Morris Goldstein, and Eduard Hochreiter, eds.: Private Capital Flows to Emerging Markets After the Mexican Crisis **(**Institute for International Economics, Washington, DC**)**.

Campbell, John, Stefano Giglio and Parag Pathak. “Forced Sales and House Prices”, *American Economic Review*, Vol. 101, No. 5 (2011): 2108-2131.

Case, Karl and Robert Shiller. “Prices of Single Family Homes Since 1970: New Indexes for Four Cities”, *New England Economic Review* (1987): 46-56.

________________________, “The Efficiency of the Market for Single Family Homes”, *American Economic Review*, Vol 70, no. 1 (1989): 125-137.

________________________, “Is There a Housing Bubble?” *Brookings Papers on Economic Activity*, 2003.

Case, Karl, Robert Shiller and Anne Thompson, “What Have They Been Thinking? Home Buyer Behavior in Hot and Cold Markets”, NBER Working Paper No. 18400, September 2012.

Cipriani, Marco and Antonio Guarino, “Estimating a Structural Model of Herd Behavior in Financial Markets,” Working Paper (2010).

Clapp, Walter, John M. Dolde and Dogan Tirtiroglu, “Imperfect Information and Investor Inferences From Housing Price Dynamics,” *Real Estate Economics*, Vol. 23, No. 3 (1995): 239–263.

Comin, Diego, Mikhail Dmitriev and Esteban Rossi-Hansberg. “The Spatial Diffusion of Technology”, Working Paper, May 2012.

Corsetti, Giancarlo, Marcello Pericoli, and Massimo Sbracia. “Some Contagion, Some Interdependence: More Pitfalls in Tests of Financial Contagion”, *Journal of International Money and Finance*, Vol. 24 (2005): 1177-1199.

Cotter, John, Stuart Gabriel and Richard Roll, “Integration and Contagion in US Housing Markets”. Working Paper, October 2011.

Danis, Michelle A. and Pennington-Cross, Anthony. "The delinquency of subprime mortgages," *Journal of Economics and Business*, vol. 60 (2008): 67-90.

Demyanyk, Yuliya and Otto Van Hemert. “Understanding the Subprime Mortgage Crisis”, *Review of Financial Studies*, Vol.24, No. 6 (2011): 1848-1880.

Dolde, Walter and Dogan Tirtiroglu. “Temporal and Spatial Information Diffusion in Real Estate Price Changes and Variances,” *Real Estate Economics*, 1997, Vol. 25, No. 4 (1997): 539–565.

Dungey, Mardi, Renee Fry, Brenda Gonzalez-Hermosillo, and Vance Martin. "Empirical Modeling of Contagion: A Review of Methodologies", *Quantitative Finance*, Vol. 5, No. 1 (February 2005): 1-16.

Estrella, Arturo. "Critical Values and *P* Values of Bessel Process Distributions: Computation and Application to Structural Break Tests", *Econometric Theory*, Vol. 19, No. 6 (2003): 1128-1143.

Favara, Giovanni and Zheng Song. "House Price Dynamics with Dispersed Information", Working Paper, January 2010

Favilukis, Jack, Sydney Ludvigson and Stijn Van Nieuwerburgh. "Macroeconomic Implications of Housing Wealth, Housing Finance and Limited Risk Sharing in Equilibrium", NYU Stern Working Paper, January 2011.

Ferreira, Fernando and Joseph Gyourko. "Anatomy of the Beginning of the Housing Boom: U.S. Neighborhoods and Metropolitan Areas, 1993-2009", NBER Working Paper No. 17374, August 2011 (latest version is May 22, 2012).

Forbes, Kristin. "The "Big C": Identifying Contagion", NBER Working Paper No. 18465, October 2012.

Forbes, Kristin and Roberto Rigobon. "No Contagion, Only Interdependence: Measuring Stock Market Comovements", *Journal of Finance*, Vol. LVII, No. 5 (October 2002): 2223-2261.

Foote, Chris, Kris Gerardi and Paul Willen. "Negative Equity and the Foreclosure: Theory and Evidence", *Journal of Urban Economics*, Vol. 14, no. 2 (2008): 234-245.

Fuss, Nico B., Roland Rottke and Bing Zhu, "Correlated Risk and Volatility Spillovers Across U.S. Regional Housing Markets," Working Paper, January 2011.

Gerardi, Kristopher, Andreas Lehnert, Shane Sherlund and Paul Willen. "Making Sense of the Subprime Crisis", *Brookings Papers on Economic Activity*, Fall 2008: 69-145.

Gerardi, Kristopher, Adam Shapiro and Paul Willen. "Subprime Outcomes, Risky Mortgages, Homeownership Experiences and Foreclosures", Federal Reserve Bank of Boston Working Paper, 2007.

Glaeser, Edward, Joseph Gyourko, Eduardo Morales and Charles Nathanson. "Housing Dynamics", June 2011 (update of NBER Working Paper 13704).

Goetzmann, William, Liang Peng and Jacqueline Yen. "The Subprime Crisis and House Price Appreciation", NBER Working Paper no. 15334, September 2009.

Hansen, Bruce E., "Approximate Asymptotic P Values for Structural-Change Tests", *Journal of Business and Economic Statistics*, Vol. 15 (1997): 60-67.

Haughwout, Andrew, Ebiere Okah and Joseph Tracy. "Second Chance: Subprime Mortgage Modification and Re-Default", *Federal Reserve Bank of New York Staff Reports*, report no. 417, Revised August 2010.

Ho, Giang and Anthony Pennington-Cross. "Predatory Lending Laws and the Cost of Credit", *Real Estate Economics*, Vol 36, no. 2 (2008): 175-211.

Holly, Sean, M. Hashem Pesaran, and Takashi Yamagata, "A Spatio-Temporal Model of House Prices in the USA," *Journal of Econometrics*, Vol. 158 (2010): 160–173.

– , – , and , – "The Spatial and Temporal Diffusion of House Prices in the UK," *Journal of Urban Economics*, Vol. 69 (2011): 2–23.

Keys, Benjamin, Tanmoy Mukherjee, Amit Seru and Vikrant Vig. "Did Securitization Lead to Lax Screening? Evidence from Subprime Loans", *Quarterly Journal of Economics*, Vol. 125, no. 1 (February 2010): 307-362.

King, Mervyn A. and Sushil Wadhwani, "Transmission of Volatility between Stock Markets," *The Review of Financial Studies*, Vol. 3, No. 1 (1990): 5–33.

Lai, Rose and Robert Van Order. "Momentum and House Price Growth in the United States: Anatomy of a Bubble", *Real Estate Economics*, Vol. 38, no. 4 (2010): 753-773.

Lee, Sang B., and Kwang Jung Kim, "Does the October 1987 Crash Strengthen the Comovements Among National Stock Markets?", *Review of Financial Economics*, Vol. 3 (1993): 89–102.

Mayer, Christopher and Karen Pence. "Subprime Mortgages: What, Where and to Whom?", Finance and Economics Discussion Series, Divisions of Research & Statistics and Monetary Affairs, Federal Reserve Board, Washington, D.C., No. 2008-29 (2008).

Mayer, Christopher, Karen Pence and Shane Sherlund. "The Rise of Mortgage Defaults", *Journal of Economic Perspectives*, Vol. 23, no. 1 (Winter 2009): 27-50.

Mian, Atif and Amir Sufi. "The Consequences of Mortgage Expansion: Evidence from the U.S. Mortgage Default Crisis", *Quarterly Journal of Economics*, Vol 124, no. 4 (November 2009): 1449-96.

Mian, Atif, Amir Sufi and Francesco Trebbi. "The Political Economy of the U.S. Mortgage Default Crisis", *American Economic Review*, Vol.100, no. 6 (December 2010): 1967-1998.

Ortalo-Magne, Francois and Sven Rady. "Housing Market Dynamics: On the Contribution of Income Shocks and Credit Constraints", *The Review of Economic Studies*, Vol. 73 no. 2 (2006): 459-85.

Piskorski, Tomacz, Amit Seru and Vikrant Vig. "Securitization and Distressed Loan Renegotiation: Evidence from the Subprime Mortgage Crisis", Chicago Booth School of Business Research Paper No. 09-02, September 2009.

Roback, Jennifer. "Wages, Rents, and the Quality of Life", *Journal of Political Economy*, Vol. 90, no. 4 (December 1982): 1257-78.

Rosen, Sherwin. "Wage-Based Indexes of Urban Quality of Life". In Current Issues in Urban Economics, edited by Peter Mieszkowski and Mahlon Straszheim. Baltimore: Johns Hopkins Univerity Press, 1979.

Shiller, Robert. *Irrational Exuberance*. Princeton, NJ: Princeton University Press, 2005 (2nd edition).

Sinai, Todd and Nicholas Souleles. "Can Owning a Home Hedge the Risk of Moving?", NBER Working Papers 15462, National Bureau of Economic Research, Inc., 2009.

Sinai, Todd. "House Price Moments in Boom-Bust Cycles", NBER Working Paper No. 18059, May 2012.

Wheaton, William and Gleb Nechayev. "The 1998-2005 Housing 'Bubble' and the Current 'Correction': What's Different This Time?", *Journal of Real Estate Research*, Vol. 30, no. 1 (2008): 1-26.

Young, H. Peyton. "Innovation Diffusion in Heterogeneous Populations: Contagion, Social Influence, and Social Learning", *American Economic Review*, Vol. 99, No. 5 (December 2009): 1899-1924.

Figure 1: Timing of Housing Boom by MSA

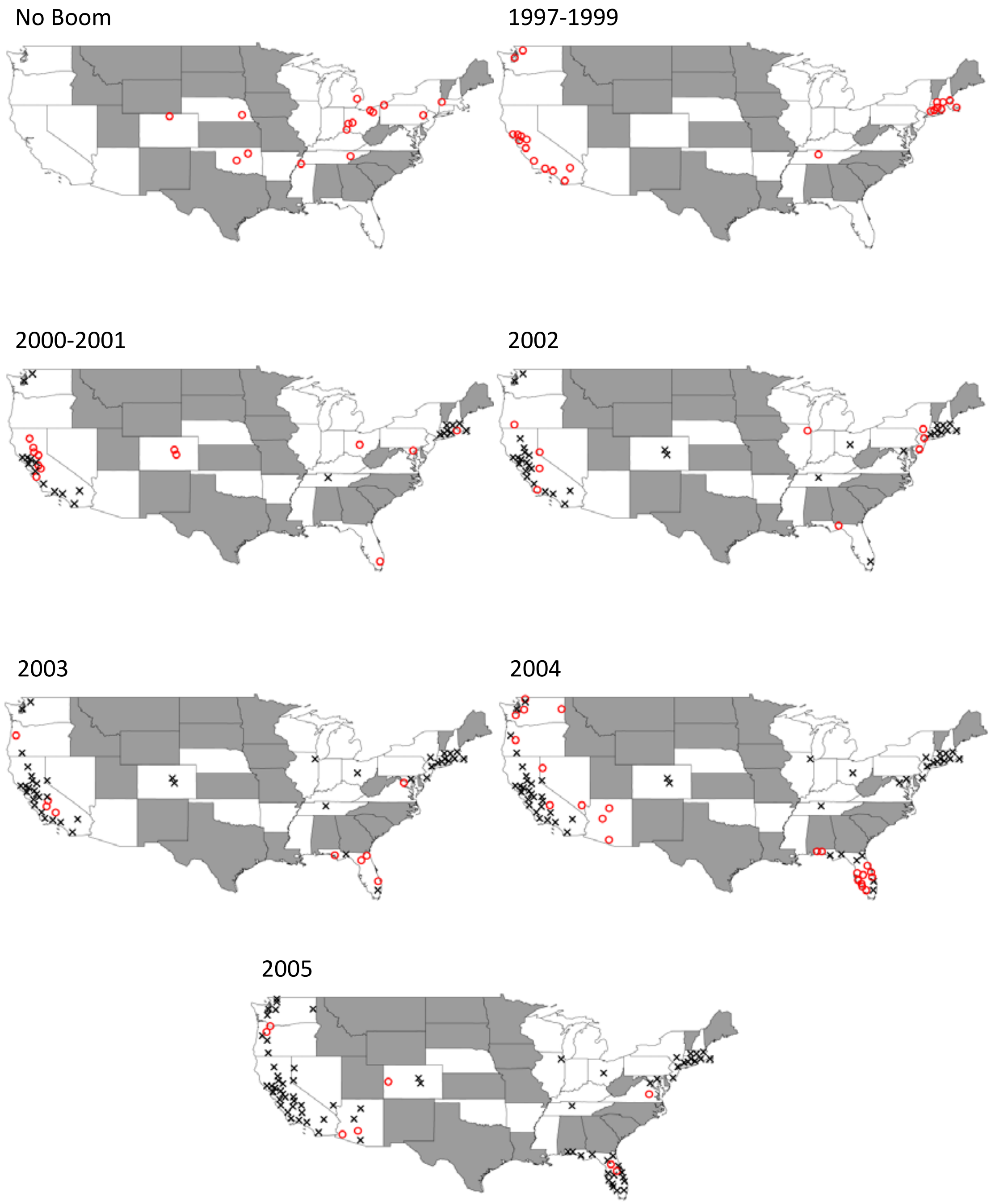
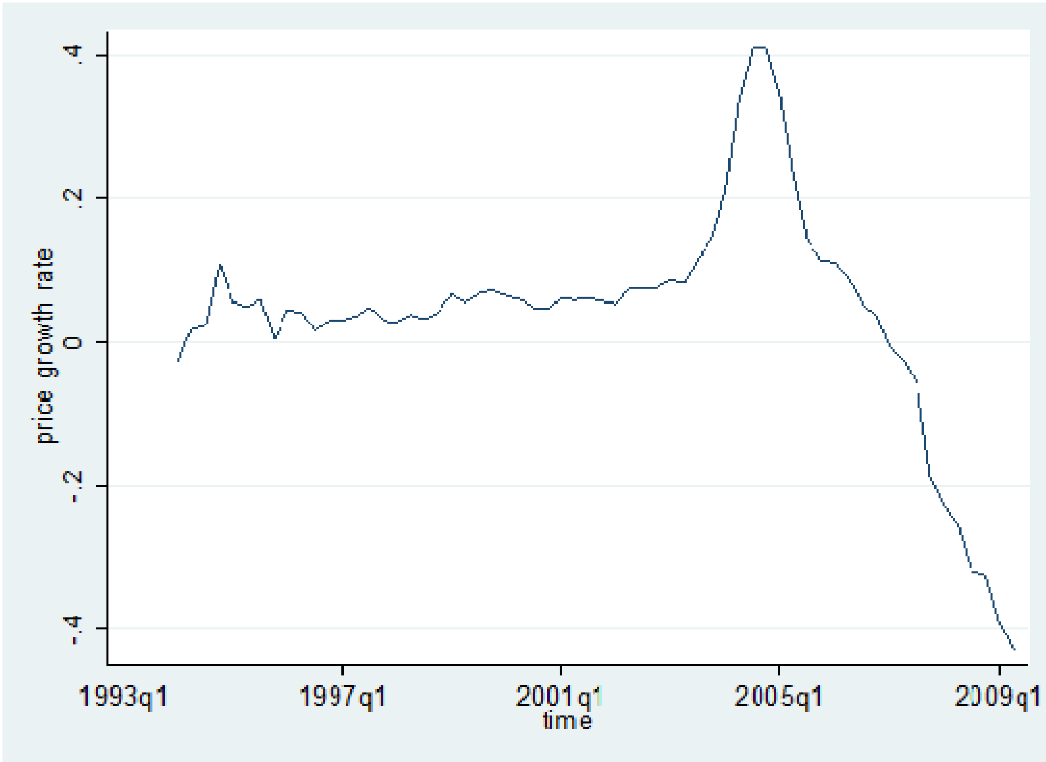


Figure 2: Las Vegas' Constant Growth Rate Before Booming



Source: Ferreira and Gyourko (2011, Figure 2)

Figure 3: Histograms of the Beginning of Booms and Busts, MSAs

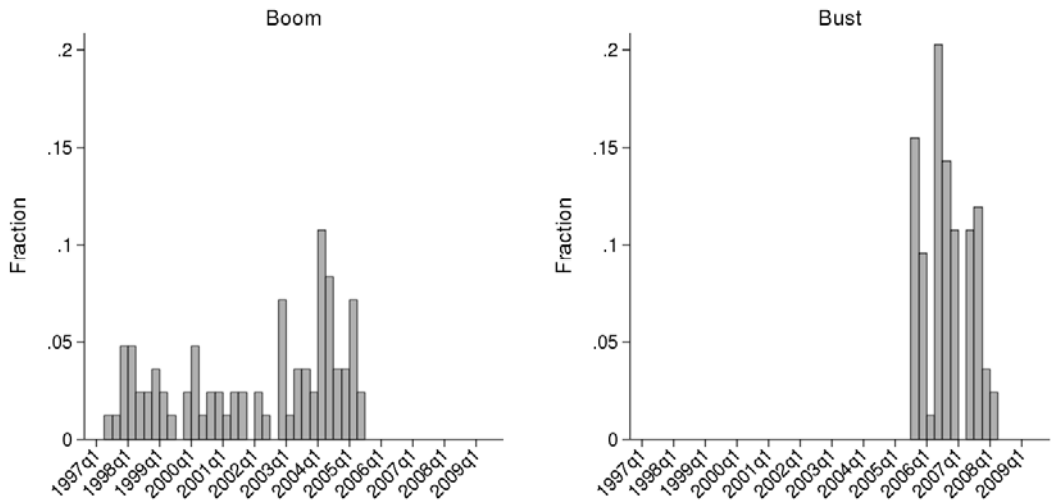


Figure 4: Timing of Housing Busts by MSA

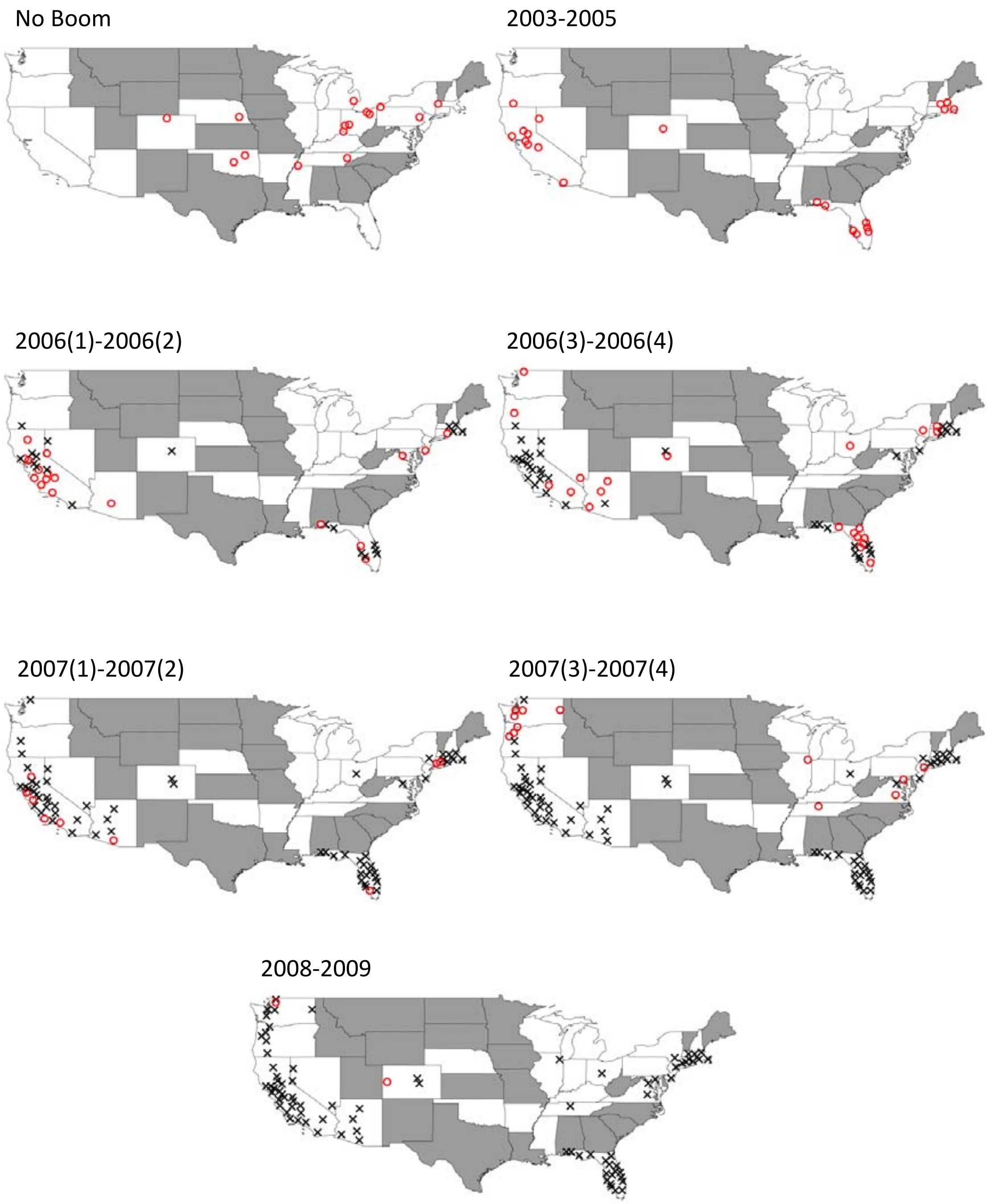


Table 1: Summary Statistics

	Mean	Std. Dev.	25th Percentile	75th Percentile
Price Index	132	52	96	156
Average Income ($1000's)	75	26	56	86
Percent Minority	.18	.13	.075	.25
Percent Speculators	.054	.032	.031	.071
Percent FHA Insured	.096	.097	.014	.14
Percent Subprime Lenders	.14	.075	.087	.19
Average LTV	.73	.1	.69	.79
Average Square Footage	1672	161	1584	1764
Unemployment Rate	5.9	3.3	4	6.6
Net Migration	266	6878	-668	1408
N.	5043			

Columns present descriptive statistics for all MSA-quarter observations in our sample. Observation counts in regressions will vary depending on the specification and control variables used.

Table 2: The impact of nearest neighbor housing boom on log focal market price

Dep. Var: Log Focal Market Price	Neighbor1		Neighbor 2	
Nearest Neighbors' Relative Years	Neighbor Boom Significant	Neighbor Boom Insignificant	Neighbor Boom Significant	Neighbor Boom Insignificant
Relative Year [-2]	-.003029	-.003357	-.00013	.0114
	(.004662)	(.01543)	(.005425)	(.008886)
Relative Year [-1]	-.000076	-.03365**	.002323	-.003854
	(.003184)	(.01526)	(.005658)	(.006118)
Relative Year [1]	.008688**	.00297	.006096	.000575
	(.004191)	(.00563)	(.006421)	(.00634)
Relative Year [2]	.00977**	-.02102*	.01487***	.002127
	(.004236)	(.01091)	(.005638)	(.01006)
Relative Year [3]	.008542*	-.01698***	.01029	.001398
	(.004985)	(.005502)	(.006735)	(.005476)
Relative Year [>=4]	.005247	-.006412	.004154	.005613
	(.004774)	(.004954)	(.00551)	(.005984)
Quarter-by-Census Division FE	Y			
Focal Market Relative Year FE	Y			
Four Lags of Focal Log Price	Y			
Neighbor-2 Relative Year FE	Y			
N	4584			

Notes: Cells represent the coefficient on the dummy variable for indicated relative years of the closest or second closest geographic neighbor. Relative year 0 indicates the 12 month period preceding the boom of the neighboring MSA and is the omitted category. Specification also includes dummy variable(s) indicating whether the closest neighbor(s) are in the sample. Standard errors are clustered at the Census division by year level and are reported in parentheses. Significance levels 10%, 5%, and 1% are denoted by *, **, and

Table 3: The impact of nearest neighbor housing boom on nearest neighbor log price

Dep. Var: Log Nearest Neighbor's Price

Nearest Neighbor Relative Years	
Relative Year [-2]	.02815**
	(.01333)
Relative Year [-1]	.01751
	(.0153)
Relative Year [1]	.03224
	(.02066)
Relative Year [2]	.06701***
	(.02199)
Relative Year [3]	.08197***
	(.02188)
Relative Year [>=4]	.1179***
	(.01976)
Quarter-by-Census Division FE	Y
Focal Market Relative Year FE	Y
Four Lags of Focal Log Price	Y
Neighbor-2 Relative Year FE	Y
N	3583

Notes: Cells represent the coefficient on the dummy variable for indicated relative years of the closest geographic neighbor. Relative year 0 indicates the 12 month period preceding the boom of the neighboring MSA and is the omitted category. Relative year effects are interacted with a dummy for whether the neighboring market had a statistically significant and positive break point and only the coefficients for the MSAs that actually had a boom are reported here. Specification also includes dummy variables indicating whether the closest neighbors are in the sample. Standard errors are clustered at the Census Division by year level and are reported in parentheses. Significance levels 10%, 5%, and 1% are denoted by *, **, and ***, respectively.

Table 4: Heterogeneity in the impact of nearest neighbor housing boom on log focal market price

Dep. Var: Log Focal Market Price	Distance		Relative Size		Focal Market Elasticity	
	(1)	(2)	(3)	(4)	(5)	(6)
Nearest Neighbor Relative Years	<= 40 Miles	> 40 Miles	Focal Larger	Nbr. Larger	Inelastic	Elastic
Relative Year [-2]	-.004239	-.001194	-.001048	-.005334	-.001111	.001165
	(.004747)	(.005436)	(.004448)	(.005874)	(.006739)	(.007681)
Relative Year [-1]	-.001823	.001363	-.001269	.000373	-.003162	.003788
	(.003802)	(.004727)	(.003668)	(.003878)	(.004464)	(.006904)
Relative Year [1]	.008405	.008564	.00461	.01288**	.00769	.01877*
	(.005087)	(.006187)	(.00417)	(.005703)	(.007545)	(.0103)
Relative Year [2]	.008409	.009215	.01116***	.009466*	.008632	.02683**
	(.005425)	(.006344)	(.004208)	(.005423)	(.005739)	(.01069)
Relative Year [3]	.008389	.005968	.007858	.0104**	.00978	.02897***
	(.007093)	(.005006)	(.005809)	(.004913)	(.007775)	(.0102)
Relative Year [>=4]	-5.5e-06	.0118**	.005298	.008329	.01061*	.01055
	(.005004)	(.005328)	(.004551)	(.005509)	(.006347)	(.01232)
Quarter-by-Census Division FE	Y		Y		Y	
Focal Market Relative Year FE	Y		Y		Y	
Four Lags of Own Log Price	Y		Y		Y	
Neighbor-2 Relative Year FE	Y		Y		Y	
N	4647		4584		3586	

Notes: Cells represent the coefficient on the dummy variable for indicated relative years of the closest geographic neighbor interacted with dummies for whether the focal market is in the category indicated in the column header for various measures of heterogeneity. Relative year 0 indicates the 12 month period preceding the boom of the neighboring MSA and is the omitted category. All specifications also include dummy variables indicating whether the closest neighbors are in the sample. Standard errors are clustered at the Census Division by year level and are reported in parentheses. Significance levels 10%, 5%, and 1% are denoted by *, **, and ***, respectively.

Table 5: The impact of nearest neighbor housing boom on focal market fundamentals using geographic distance

	Dependent Variable				
	(1)	(2)	(3)	(4)	(5)
Nearest Neighbor Relative Years	Focal Market Income	Lender Similarity	Lending Amount	Net Migration	Focal Market Flippers
Relative Year [-2]	-0.013	.0127*	0.0205	84.43	.0078***
	(.009)	(.0065)	(0.0609)	(132.6)	(0.0029)
Relative Year [-1]	-0.006	0.005	0.0572	91.49	0.0039
	(.008)	(.0051)	(0.0562)	(132.9)	(0.0026)
Relative Year [1]	0.011	0.0025	0.0239	-43.23	-0.0004
	(.008)	(.0039)	(0.0505)	(85.75)	(0.0032)
Relative Year [2]	0.016	-0.0015	0.0491	-93.54	-0.0012
	(.0115)	(.0047)	(0.0416)	(78.01)	(0.003)
Relative Year [3]	0.0135	0.0004	-0.0230	-165.4	-0.0038
	(.0132)	(.0071)	(0.0448)	(114)	(0.0031)
Relative Year [>=4]	.0229*	0.0048	0.0194	-251.6**	-0.0037
	(.013)	(.0073)	(0.0618)	(101.4)	(0.0032)
Quarter-by-Census Division FE	Y	Y	N	N	Y
Year-by-Census Division FE	N	N	Y	Y	N
MSA FE	Y	Y	Y	Y	Y
Focal Market Relative Year FE	Y	Y	Y	Y	Y
Neighbor-2 Relative Year FE	Y	Y	Y	Y	Y
Focal Market Price and LTV Growth	N	N	Y	N	N
N	4877	4219	996941	1215	4829

Notes: Cells represent the coefficient on the dummy variable for indicated relative years of the closest geographic neighbor. Relative year 0 indicates the 12 month period preceding the boom of the neighboring MSA and is the omitted category. All specifications also include dummy variables indicating whether the closest neighbors are in the sample. Standard errors are clustered at the Census Division by year level and are reported in parentheses. Significance levels 10%, 5%, and 1% are denoted by *, **, and ***, respectively.

Table 6: The impact of nearest neighbor housing boom on log focal market price, controlling for local fundamentals and expectations

Dep. Var: Log Focal Market Price			
Nearest Neighbor Relative Years			
Relative Year [-2]	-.003202	-.003899	-.002159
	(.004133)	(.003979)	(.004018)
Relative Year [-1]	-.000332	-.000577	.000934
	(.003053)	(.003016)	(.00288)
Relative Year [1]	.007821**	.008673**	.007878**
	(.003796)	(.003624)	(.003235)
Relative Year [2]	.008964**	.01033**	.008651**
	(.004278)	(.004157)	(.003704)
Relative Year [3]	.00804	.008275	.005047
	(.005421)	(.006464)	(.005994)
Relative Year [>=4]	.003512	.006228	.005626
	(.004924)	(.004178)	(.004507)
Quarter-by-Census Division FE	Y	Y	Y
Focal Market Relative Year FE	Y	Y	Y
Four Lags of Own Log Price	Y	Y	Y
Neighbor-2 Relative Year FE	Y	Y	Y
Focal Market Fundamental Controls	Y	Y	Y
Log Average Price of Neighbor Groups	Y	Y	Y
Four Leads of Own Income	N	Y	Y
Four Leads of All Fundamental Controls	N	N	Y
N	4538	4142	4142

Notes: Cells represent the coefficient on the dummy variable for indicated relative years of the closest geographic neighbor. Relative year 0 indicates the 12 month period preceding the boom of the neighboring MSA and is the omitted category. All specifications also include dummy variables indicating whether the closest neighbors are in the sample. Standard errors are clustered at the Census Division by year level and are reported in parentheses. Significance levels 10%, 5%, and 1% are denoted by *, **, and ***, respectively.

Table 7: The impact of nearest neighbor price changes on focal market price changes—IV Results

Dep. Var: Focal Market Price Changes

Nearest Neighbor Price Changes	OLS	IV
Relative Year [-2]	0.139*	-0.045
	(0.072)	(0.228)
Relative Year [-1]	0.094**	0.041
	(0.042)	(0.131)
Relative Year [1]	0.135***	0.148***
	(0.038)	(0.037)
Relative Year [2]	0.102**	0.125***
	(0.040)	(0.035)
Relative Year [3]	0.090**	0.090**
	(0.039)	(0.037)
Relative Year [>=4]	0.089***	0.088***
	(0.027)	(0.028)
Quarter-by-Census Division FE	Y	Y
Focal Market Relative Year FE	Y	Y
Four Lags of Focal Log Price	Y	Y
Neighbor-2 Relative Year FE	Y	Y
N	4156	4131

Notes: Table reports the results from a regression of annualized focal market price growth rates on the lag of the nearest neighbor's annualized price growth interacted with dummies for the indicated relative years of the neighboring market. In the second column, lagged neighbor's price growth is instrumented using one further lag.Standard errors are clustered at the Census Division by year level and are reported in parentheses. Significance levels 10%, 5%, and 1% are denoted by *, **, and ***, respectively.

Table 8: Summary statistics on Lagged Number of Booms for the Hazard Estimation

	Lagged Number of New Booms				
	Mean	Sd	Min	Max	% Zero Boom
Neighbors [1-2]	0.0297	0.1777	0	2	97.20%
Neighbors [3-5]	0.0393	0.2058	0	3	96.20%
Neighbors [6-10]	0.0626	0.2692	0	3	94.30%
Neighbors [1-10]	0.1317	0.4246	0	5	89.00%
Neighbors [11-50]	0.3051	0.6079	0	7	76.20%
Neighbors [51+]	1.3055	1.6134	0	6	39.50%

Note: Neighbors are ranked with respect to their geographic distance from the focal MSA.

Table 9: Hazard Model Estimates of MSA Neighbors on the Probability of Booming by Geographic Distance

Independent Variables	Unconditional Hazard (1)	Baseline Results (Proportional Hazard) (2)	Weibull Hazard (3)
Lagged Number of New Booms			
Neighbors[1-2]	3.49***	1.49	1.69*
Neighbors[3-5]	1.31	0.79	0.84
Neighbors[6-10]	1.41	1.26	1.25
Neighbors[11-50]	1.14	0.86	0.8
Neighbors[51+]	1.23***	1.02	1.05
Other Controls			
Lagged % MSAs that Already Boomed	N	Y	Y
Focal Market Fundamental Controls	N	Y	Y
Four Lags of Focal Market Price Growth	N	Y	Y
Census Region FE	N	Y	Y
N	2114	2114	2114
Log-Likelihood	-29.93	29.43	34.14

Note: Implied Hazard Ratios are reported along with indicators of statistical significance of the underlying regression coefficients. Significance levels 10%, 5% and 1% are denoted by *, ** and ***, respectively.

Table 10: Hazard Model Estimates of MSA Neighbors on the Probability of Busting

Independent Variables	Unconditional Hazard (1)	Proportional Hazard with Controls (2)
Lagged Number of New Busts		
Neighbors[1-2]	1.66*	0.96
Neighbors[3-5]	1.14	0.74
Neighbors[6-10]	1.55**	1.08
Neighbors[11-50]	1.12*	0.96
Neighbors[51+]	1.18***	1
Other Controls		
Lagged % MSA that Already Busted	N	Y
Focal Market Fundamental Controls	N	Y
Four Lags of Focal Market Price Growth	N	Y
Census Region FE	N	Y
N	3,637	3,637
Log-Likelihood	2.181	100.1

Note: We define the time of housing bust as the quarter when price peaks in our sample period. Implied hazard ratios are reported along with indicators of statistical significance of the underlying regression coefficients. Significance levels 10%, 5% and 1% are denoted by *, ** and ***, respectively.

Table 11: The impact of nearest neighbor housing bust on log focal market price

Dep. Var: Log Focal Market Price

Nearest Neighbor Relative Years	
Relative Year [-2]	-.01169**
	(.004509)
Relative Year [-1]	-.003846
	(.003101)
Relative Year [1]	-.006359
	(.005119)
Relative Year [2]	-.01228**
	(.006151)
Relative Year [3]	-.03804***
	(.007463)
Quarter-by-Census Division FE	Y
Focal Market Relative Year FE	Y
Four Lags of Focal Log Price	Y
Neighbor-2 Relative Year FE	Y
N	5178

Notes: Cells represent the coefficient on the dummy variable for indicated relative years of the closest geographic neighbor. Relative year 0 indicates the 12 month period preceding the bust of the neighboring MSA and is the omitted category. All specifications also include dummy variables indicating whether the closest neighbors are in the sample. Standard errors are clustered at the Census Division by year level and are reported in parentheses. Significance levels 10%, 5%, and 1% are denoted by *, **, and ***, respectively.

Appendix A1: Estimating Multiple Breakpoints

In estimating the break points, we allow for the possibility that a given market might experience more than one housing boom during the course of our sample period. Our method is recursive in that we first test for the existence of one break point against the null hypothesis of zero. Given the existence of at least one break point, we can then test the hypothesis of $m + 1$ break points against the null of m using the results from Bai (1999). Bai and Perron (1998) show that the test for one break is consistent in the presence of multiple breaks, which is what allows for this sequential estimation procedure.

More specifically, let $0 < \varphi_{i,1} < \cdots < \varphi_{i,m} < 1$ mark the proportions of the sample generated by the m break points estimated under the null hypothesis for MSA i. For technical reasons, we require that $\varphi_{i,j} - \varphi_{i,j-1} > \pi_{i,0}$ for some small $\pi_{i,0}$ where we define $\varphi_{i,0} = 0, \varphi_{i,m+1} = 1$. Further, let $\eta_{i,j} = \frac{\pi_{i,0}}{\varphi_{i,j}-\varphi_{i,j-1}}, j = 1, \ldots, m + 1$. The likelihood ratio test compares the maximum of the likelihood ratio obtained when allowing for $m + 1$ breaks to that from only allowing for m. The distribution of this likelihood ratio statistic is given by

$$(A1.1)\ \ P(LR > c) = 1 - \prod_{i=1}^{m+1} \left(1 - P\left(sup_{\pi_i \in [\eta_{i,j}, 1-\eta_{i,j}]} Q_1(\pi_i) > c\right)\right),$$

which we calculate by recursive application of the method provided in Estrella (2003).

We apply this procedure to test for the existence of two break points against the null of one as well as three against the null of only two among those MSAs for which we find at least two statistically significant break points. There are some noteworthy practical issues involved with carrying out this procedure. We have not until this point said where the sample proportions $\pi_{i,0}, \pi_{i,1}, \pi_{i,2}$ come from. In practice, we restrict the full sample period for each MSA to lie between the first quarter in the data and the peak of price growth. We then do not allow any break points to lie in either the first or last two quarters of this sample for each MSA. This determines the fractions $\pi_{i,1}$ and $\pi_{i,2}$ which, because different MSAs have a different number of quarters, will vary across areas.

When estimating multiple break points, we further require that any two break points be at least four quarters apart. This determines the fraction $\pi_{i,0}$ which, again, will vary across areas due to differing sample sizes. Because of these restrictions, we are not able calculate p-values for many MSAs in the case of multiple breaks. The reason for this can be seen from the expression in (A1.1). Because this expression requires that $\eta_j < 0.5$, we must require that $\frac{\pi_{i,0}}{\varphi_{i,j}-\varphi_{i,j-1}} > 0.5$ for all j. This implies that we will not be able to calculate p-values for the two-break case in MSAs (neighborhoods) where the first break is less than $\pi_{i,0}/0.5$ from the beginning of the sample period. Naturally, this restriction is more burdensome when trying to calculate p-values in the three break case.

Appendix A2: A Hazard Model of Housing Boom Contagion

Consider an economy with N metro areas. The probability of MSA m entering a housing boom increases if it "meets" other MSAs which already have entered a housing boom. Assume α is the frequency of such a meeting and that the connection between MSA pairs decays in distance at rate δ. Following Comin, et. al. (2012), we can write MSA m's probability of not entering a housing boom at time $t + h$ conditional on not having a housing boom in time t as

$$(A2.1)\quad P\left(0, m, t + \hbar\right) = P(0, m, t)\left[\frac{\sum_{q\neq m} P(0, q, t)e^{-\delta r_{mq}}}{\sum_{q\neq m} e^{-\delta r_{mq}}}\right]^{\alpha\hbar}$$

where r_{mq} denotes the distance between MSA m and MSA q. Taking $h \to 0$, we have

$$(A2.2)\quad \frac{\partial\, ln\, P(0, m, t)}{\partial t} = \alpha\, ln\left(\sum_{q\neq m} P(0, q, t)e^{-\delta r_{mq}}\right) - \alpha\, ln\left(\sum_{q\neq m} e^{-\delta r_{mq}}\right)$$

By assumption, $P(0, q, t) = 0, \forall t \leq \tau$ if MSA q enters a housing boom at time τ. As long as some MSA enters a boom, equation (A2.2) implies that $\frac{\partial \ln P(0,m,t)}{\partial t} < 0$, so that the hazard of entering a housing boom increases over time. To consider the contagious effect of housing booms, suppose MSA q booms at time t. That increases the hazard of MSA m having a boom because

$$(A2.3)\quad \frac{\partial \ln P(0, m, t)}{\partial t \partial P(0, q, t)} = \frac{\alpha e^{-\delta r_{mq}}}{\sum_{q\neq m} P(0, q, t)e^{-\delta r_{mq}}} > 0$$

In addition, the contagious effect of a housing boom in MSA q decreases over geographical distance, because

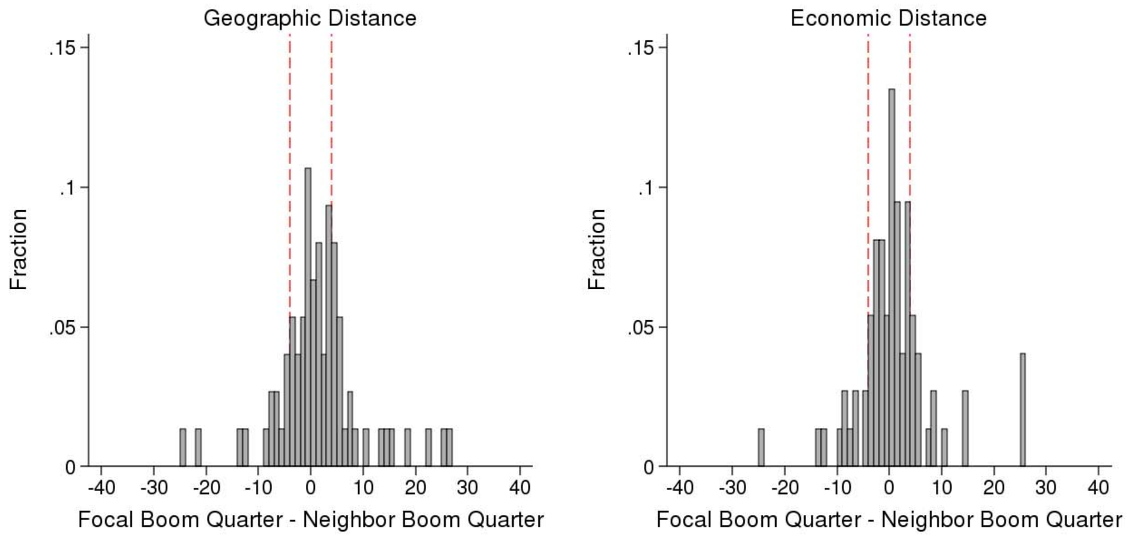
$$(A2.4)\quad \left.\frac{\partial^2\, ln\, P(0, m, t)}{\partial t \partial r_{mq}}\right|_{P(0,q,t)=0} = \frac{\alpha\delta e^{-\delta r_{mq}}}{\sum_{q\neq m} e^{-\delta r_{mq}}} > 0$$

In sum, this framework provides two relevant implications for our purposes:

Implication 1: The fact that MSA q enters a housing boom at time t increases the hazard of MSA m $(m \neq q)$ having a housing boom at time $t + 1$.

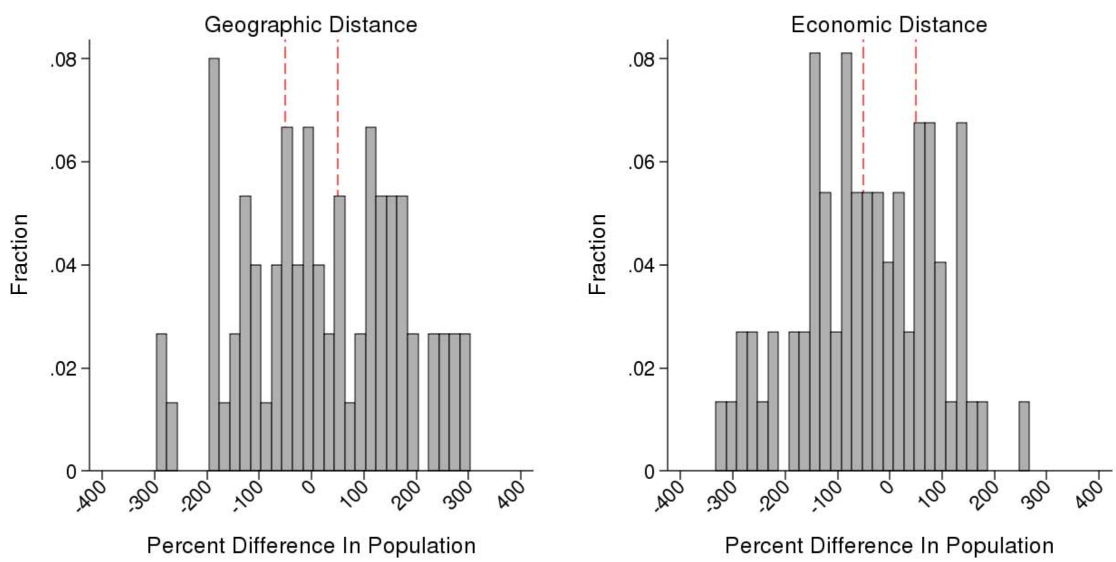
Implication 2: Contagion effect in ***Implication 1*** becomes larger when MSA q is closer to MSA m.

Appendix Figure 1: Histogram of number of quarters between timing of the booms of focal markets and nearest neighbors



Note: The level of observation is the MSA-quarter.

Appendix Figure 2: Histogram of percentage difference in population of focal markets and nearest neighbors



Note: The level of observation is the MSA-quarter.