Spatial Effects and House Price Dynamics in the U.S.A.

Jeffrey P. Cohen[1]

Yannis M. Ioannides[2]

and

Win (Wirathip) Thanapisitikul[3]

October 30, 2015

Abstract

While an understanding of spatial spillovers and feedbacks in housing markets could provide valuable information for location decisions, little known research has examined this issue for the U.S. Metropolitan Statistical Areas (MSAs). Also, it is unknown whether there can be differences in the spatial effects before and after a major housing "bust". In this paper we examine spatial effects in house price dynamics. Using panel data from 363 US MSAs for 1996 to 2013, we find that there are significant spatial diffusion patterns in the growth rates of urban house prices. Lagged price changes of neighboring areas show greater effects after the 2007-08 housing crash than over the entire sample period of 1996-2013. In general, the findings are robust to controlling for potential endogeneity, and for various spatial weights specifications (including contiguity weights and migration flows). These results underscore the importance of considering spatial spillovers in MSA-level studies of housing price growth.

[1] Center for Real Estate, University of Connecticut. Jeffrey.Cohen@business.uconn.edu
[2] Department of Economics, Tufts University. Yannis.Ioannides@tufts.edu
[3] Associated with Tufts University, when the research started, now with Lazada.co.th

1. Introduction

Spatial effects in economic phenomena have recently attracted particular attention by economists. For instance, during periods of housing booms, house prices in metropolitan areas in the Northeast of the U.S. show distinct patterns as they appreciated at similarly high rates, while house prices in metropolitan areas in the Midwest did not experience comparably high appreciation rates during those same periods. Thus, house price dynamics might have spatial features. Spatial features of the housing bust associated with the Great Recession of 2007-2009 are currently receiving particular attention. This paper looks empirically at such spatial aspects in house price dynamics in more detail and by utilizing data that include several housing market booms and busts including the Great Recession.

A contribution of this paper is our focus on spatial effects in assessing the impacts of lagged Metropolitan Statistical Area (MSA for short) house price growth. Another contribution is our use of MSA to MSA migration panel data to assess the importance of other MSAs' house price growth on a particular MSA's contemporaneous house price growth. We also demonstrate that in the periods following the Great Recession of 2007-09, the spatial effects appear to be stronger than before the crisis. These spatial spillover results can have important implications for both residential and business location decisions.

While our focus here is on MSA level house price growth, economists have studied the role of space in house price dynamics at various levels of aggregation. Some studies look at spatial effects at the level of submarkets within a particular MSA. The fact that different urban neighborhoods are often developed at the same time, dwellings in them may share similar structural characteristics and neighborhoods may offer similar amenities suggest that house prices may react similarly to exogenous shocks. Basu and Thibodeau (1998) confirm this intuition by using data from submarkets within metropolitan Dallas. Also working at a similar local level, Clapp and Tirtiroglu (1994) estimate a spatial diffusion process in house price for towns in the Hartford, CT, MSA. They find that lagged house price changes in a submarket affect the current house price of a contiguous submarket positively and more strongly than the lagged house price changes for noncontiguous submarkets. However, such evidence of spatial effects at the submarket level of aggregation may be MSA-specific, and the conclusions for a single MSA may not necessarily be valid for other MSAs. Further, there could be spatial effects between submarkets that are geographically contiguous and belong to the same economic area and housing market, but not in the same MSA.

A second level of aggregation at which economists have examined spatial effects is at the US Census division. Pollakowski and Ray (1997) find that although house prices in one division are affected by lagged house prices in other divisions, there is no clear spatial diffusion pattern that describes such processes. That is, lagged house price changes of adjacent Census divisions do not provide greater explanatory power in explaining current house prices than those in non-adjacent divisions. However, when those authors look at house prices within the greater New York's primary metropolitan statistical areas (PMSAs), they find evidence of spatial pattern consistent with that by Clapp and Tirtiroglu (1994).

Given mixed results from previous research on spatial diffusion patterns in house prices at different aggregation levels, this paper has the following aims; one, to examine house price interactions at the MSA level of aggregation across the entire continental U.S; two, to bring geographic and economic distances into the analysis in addition to adjacency as measures of proximity.

The paper employs the consolidated house price index, which has been published (since March 1996) by the Office of Federal Housing Enterprise Oversight (OFHEO). The data cover almost 400 MSAs across the entire US, with data for all MSA's going back to 1995 (Calhoun, 1996). We use these data along with appropriate geographic information to study house price dynamics across the entire U.S. at the MSA level of aggregation and to compare the results with other levels of aggregation.

The organization of this paper is as follows: section 2 reviews the existing literature; section 3 introduces the econometric models used in this study; section 4 describes the data and methodology; section 5 presents the results and impulse response diagnostics, and section 6 concludes.

2. Literature Review

The modern view of housing emphasizes its role as an asset in household portfolios (Henderson and Ioannides, 1983). Economists predict that asset prices in informationally efficient markets react rapidly to new information. Over the last twenty years, however, empirical studies have established that this might not be the case for housing markets (Rayburn *et al*., 1987; Guntermann and Norrbin, 1991). Economists have come to believe that households may be backward-looking in housing markets. Therefore, past house price changes can be used to explain future prices changes. Case and Shiller (1989) use their own house price indices for four major cities and find that one-year house price lag in a city is statistically and economically significant in forecasting that city's current house price.

Informational inefficiency in housing markets, as exhibited by temporal and spatial persistence in house prices, is not surprising when one considers the potential frictions affecting real estate markets. For instance, real estate markets do not clear immediately after a shock to the economy. The process of matching buyers with sellers of existing houses takes time. It also takes time for developers to bring new houses to the market, after an increase in demand, and to liquidate inventories when demand weakens. Speculative inventory holding is very costly. Transaction costs in housing markets are also higher than other asset markets. Case *et al*. (2005) find that selling costs, such as five to six percent brokerage fees nominally charged in the US, combine with the physical and the psychological costs of moving (e.g. moving across neighborhoods and changing schools) to generate substantial transaction costs. Such costs limit arbitrage opportunities for rational investors, and thus lead to pricing inefficiencies.

Brady (2014) reviews some recent studies on spatial aspects of housing prices that have incorporated Vector Autoregressions (VAR) as an approach to model simultaneity.[4] In earlier work, Anselin (2001) and Pace *et al*. (1998) introduce general methodologies for incorporating the time dimension in spatial models.[5] These papers followed Clapp and Tirtiroglu (1994), which was one of the pioneering works on spatial effects in house price dynamics that we referred to earlier, using house price data from Hartford, CT.

[4] These include Pesaran and Chudik (2010), Holly *et al*. (2011), Beenstock and Felsenstein (2007), and Kuethe and Pede (2011).

[5] One advantage of our approach over the approaches of Pace et al. (1998), Anselin (2001), and others who utilize a VAR approach is that ours is well-suited for a study where contemporaneous spatial lags are not present. Our approach is also parsimonious, which is helpful in estimating versions of the model with time lags that contain several types of spatial lags.

These authors regress excess returns, defined as the difference between the return of a submarket within an MSA and the return at the MSA level, on the lagged excess returns of a group of neighboring towns and on a "control group" of non-neighboring towns. The authors find that estimated coefficients were significant for excess returns in neighboring towns, but insignificant for non-neighboring towns. Their results suggest that house price diffusion patterns exist within an MSA and are consistent with one form of a positive feedback hypothesis, where individuals would be expected to place more weight on past price changes in their own and neighboring submarkets and less weight on those further away.

Pollakowski and Ray (1997) also examine house price spatial diffusion patterns, but at a much higher aggregation level. Using vector autoregressive models with quarterly log house price changes from the nine U.S. Census divisions, these authors are unable to identify a clear spatial diffusion pattern. Past growth rates in neighboring Census divisions were correlated with a particular division's current growth rates for some divisions. However, upon examining neighboring PMSAs within the greater New York CMSA, like Clapp and Tirtiroglu (1994), the authors find evidence in support of a positive feedback hypothesis. Shocks in housing prices in one metropolitan area are likely to Granger-cause subsequent shocks in housing prices in other metropolitan areas.

More recently, Brady (2008), using spatial impulse response function and VAR models, finds that spatial autocorrelations in house prices across counties in California are highly persistent over time. The average housing price in a Californian county is positively affected by bordering counties for up to 30 months. Brady (2011) examines how fast and how long a change in housing prices in one region affect its neighbors.

Using an impulse response function with a panel of California counties, he finds that the diffusion of regional housing prices across space lasts up to two and half years. Brady (2014) estimates the spatial diffusion of housing prices across U.S. states over 1975-2011 using a single equation spatial autoregressive model. He shows that for the 1975 to 2011 period spatial diffusion of housing prices is statistically significant and persistent across US states and so is in the four US Census regions for the United States. He shows that the persistence of spatial diffusion may be more pronounced after 1999 than before.

Holly *et al*. (2010) also find spatial correlations in both housing prices across the eight Bureau of Economic Analysis (BEA) regions and housing prices across the contiguous states within the U.S. They find evidence of house price departures from the long run growth rates for markets in California, Massachusetts, New York, and Washington, even after accounting for spatial effects between states. Holly *et al*. (2011) work with data for the UK, but also allow for spillovers from the US economy (to express financial markets interdependence), in order to explain the spatial and temporal diffusion of shocks in real house prices within the UK economy at the level of regions. They model shocks that involve a region specific and a spatial effect. By focusing on London they allow for lagged effects to echo back to London, which in turn is influenced by international economic conditions via its link to New York and other financial centers. They show that New York house prices have a direct effect on London house prices. Their use of generalized spatio-temporal impulse responses allows them to highlight the diffusion of shocks both over time and over space.

An interesting recent development in this literature is the utilization of data on individual transactions. DeFusco *et al.* (2013) use micro data on the complete set of housing transactions between 1993 and 2009 in 99 US metropolitan areas to investigate contagion in the last housing cycle. By defining 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, their estimations allow them to determine the timing of local housing booms in a non-*ad hoc* way. The evidence for contagion is strong during the boom but not the bust phase of the cycle. They show that these effects are due to interactions between closest neighboring metropolitan areas, with the price elasticity ranging from 0.10 to 0.27. The impact of larger markets on smaller markets imply greater elasticities. They show local fundamentals and expectations of future fundamentals are limited in accounting for their estimated effects, suggesting a potential role for non-rational forces.

Bayer *et al*. (2014) utilize data from a detailed register of housing transactions in the greater Los Angeles metropolitan Area from 1988 to 2012. Properties in their data set are fully geocoded, which allows those authors to readily merge with 2011 county tax assessor data and obtain information on property attributes. For some of the data, additional information may be obtained from the Home Mortgage Disclosure Act (HMDA) forms on purchaser/borrower income and race. Bayer *et al*. find evidence of strong spillovers within neighborhoods: homeowners were much more likely to speculate both after a neighbor had successfully “flipped” a house, and when a house had been successfully flipped in their neighborhood. Social contagion appears to be at work and to involve amateur investors, whose share of the market reached a record high at the same time as the market reached its peak, with equity losses following in the ensuing crash.

3. Econometric Models

Following the literature, we postulate that housing prices (in our case, defined at the MSA level) in a given period depend on lagged housing prices and MSA specific effects. By differencing the housing prices, we are left with the annual house price growth rate of MSA_i at year *t*[6] as a dependent variable. We utilize this measure to calculate one-year through four-year lagged house prices growth of MSA_i in year *t*. Similarly, we calculate the one-year, two-year, three-year, and four-year time lags of the spatial lags of the dependent variable. As we explain below, these spatial lags are obtained using the contiguity neighbor weights, the inverse distance weights for several distance intervals, and the migration weights. The results presented in the tables are for the growth-rate regressions, the growth-rate regressions with the spatial lags for the contiguous weights, and the growth rate regressions with the MSA to MSA migration weights. While we also estimate models for the growth rates with inverse distance weights, due to potential multicollinearity concerns between the inverse distance weighted house price lags for various intervals that lead to many insignificant t-statistics, and these results are available from the authors upon request.

The growth rate of house prices in MSA_i is defined in the standard fashion as the difference in logs, $G_{i,t} = \log(HPI_{i,t}) - \log(HPI_{i,t-1})$. We compute $G_{i,t}$ by using values

[6] The MSA fixed effects would then drop out, although for completeness we present two sets of results – one with and one without MSA fixed effects.

for each year's first quarter.[7] Such differencing eliminates unobserved additive heterogeneity at the log house price levels in the form of MSA specific fixed effects.

Baseline Model: Own-Lag Effects

We start with an autoregressive model up to order 4, AR(4), for the house price annual growth rates as a baseline for comparison against spatial effects:

$$G_{i,t} = \beta_0 + \beta_1 G_{i,t-1} + \beta_2 G_{i,t-2} + \beta_3 G_{i,t-3} + \beta_4 G_{i,t-4} + \varepsilon_{i,t} \quad , \tag{1}$$

where $\varepsilon_{i,t}$, the error term, may include MSA-specific time-invariant fixed effects. Our choice of 4-year lags is based on the literature discussed above, where other researchers have found significant correlation between lagged and current growth-rates. In our estimation of model (1), all four lags are statistically significant. Our results based on the Akaike Information Criterion (AIC) imply the four lag specification is preferred over models with additional lags.[8]

Incorporating Geography into the Model by Means of Cross-Lag Effects

We control for spatial effects in several ways. First, we add as a regressor in the r.h.s. of (1) the average annual growth rate in the HPI of all MSAs that border MSA_i at

[7] We think this choice rather than the fourth quarter is more appropriate in order to be close to individuals' circumstances, given the fact that the IRS-based data that we use for the migration weights originate in income tax returns.

[8] One could argue that the lagged dependent variables can give rise to time series autocorrelation. It is for this reason that we have utilized the Heteroskedasticity Autocorrelation Consistent (HAC) estimator.

time period t, $A_{i,t}$, and up to four of its lags.[9] That is, we estimate dynamic fixed effect spatial autoregressive models up to order 4, SAR(4):[10]

$$G_{i,t} = \eta + \beta \vec{\phi}_{i,t} + \lambda \vec{\delta}_{i,t} + \varepsilon_{i,t}, \tag{2}$$

where η is a constant intercept, $\vec{\phi}_{i,t}$ is a vector of four own-lag growth rates, $\vec{\delta}_{i,t}$ is a vector of the four lags of the mean annual growth rate of all MSAs that border MSA_i, and β and λ are the respective vectors of parameters.

Second, we introduce physical distances between MSAs in order to examine whether spatial dependence in MSA house prices is related to distance among them. For each MSA_i, we group the remaining 362 MSAs into 5 groups. Group 1 includes all MSAs that are less than 100 km away from MSA_i; Group 2 includes all MSAs that are between 100 km and under 200 km from MSA_i; and Groups 3, 4, and 5 include MSAs that lie within 200 km to 350 km, 350 km to 500 km, and 500 km to 1000 km from MSA_i, respectively. We calculate the inverse distance weighted (IDW) growth rates for each of the 5 intervals for each MSA_i at time period *t*. The use of inverse distance weights imposes a particular spatial attenuation of interactions. The lagged IDW growth rates for each of the 6 intervals are then incorporated into (1), such that we have,

$$G_{i,t} = \eta + \beta \vec{\phi}_{i,t} + \lambda \vec{\delta}_{i,t} + \mu \vec{\omega}_{i,t} + \varepsilon_{i,t}, \tag{3}$$

where $\vec{\omega}_{i,t}$ is a vector of the four lags for the annual inverse distance weighted growth rate of all MSAs in each of the 5 distance intervals relative to MSA *i*.

[9] Using adjacency as a measure of geographical proximity has also been used in previous literature. Dobkins and Ioannides (2001) use adjacency for comparing MSA population growth rates. Pollakowski and Ray (1997) and Holly *et al.* (2011) use it at the regional and the state level, respectively.
[10] Anselin (2001) and Pace *et al.* (1998) provide overviews of space-time models.

Finally, to proxy for economic interaction we add an additional set of spatial weights based on migration data. Specifically, we denote the *i*th element at time *t* of $\vec{\alpha}_{i,t}$ as $\Sigma_j \psi_{ij,t} G_{j,t}$, where $\psi_{jj,t}$ =0, and *i*,*j*=1,2,…,363. $\psi_{ij,t}$ is the year t total migration – the sum of migration inflows and outflows – between MSA_j and MSA_i. We then modify equation (3) to incorporate the time lags of the migration weighted housing price growth rates:

$$G_{i,t} = \eta + \beta \vec{\phi}_{i,t} + \lambda \vec{\delta}_{i,t} + \mu \vec{\omega}_{i,t} + \gamma \vec{\alpha}_{i,t} + \varepsilon_{i,t} , \tag{4}$$

where γ is a vector of parameters. We estimate this specification for the time period 1996-2011, because those are the only years covered by the migration data set.

Given the potential for multicolinearity among the inverse distance spatial lags (since one might speculate greater migration flows occur between larger cities that are far away), we estimate a variation of equation (4) by dropping the term $\mu \vec{\omega}_{i,t}$. The resulting equation includes only spatial lags for contiguous neighbors and migration spatial lags:

$$G_{i,t} = \eta + \beta \vec{\phi}_{i,t} + \lambda \vec{\delta}_{i,t} + \gamma \vec{\alpha}_{i,t} + \varepsilon_{i,t} . \tag{5}$$

MSA-specific effects

The housing literature suggests that it may be important to consider MSA-specific effects in house prices dynamics. We recall the maxim “location-location-location” in connection with real estate and apply it at the MSA level. Indeed, intuitively, growth rates in house prices in MSA’s in a relatively warm state such as California could have different fundamental characteristics than houses in colder MSA’s in New England. In addition, Gyourko *et al*. (2013) identify “superstar cities,” defined as cities with extraordinarily high growth in real income and fixed housing supply, whose housing prices exhibit greater volatility and faster appreciation rates. Thus, there may well be unobservable idiosyncratic differences among different MSAs. Further, results reported in Thanapisitikul (2008) suggest widely varying patterns in boom and bust periods of the housing price cycle across MSAs. In addition to estimating these models without fixed effects, we also adopt fixed effects in the stochastic structure of Equations (2), (3), and (4), because such effects could capture MSA-specific patterns that may be correlated with other regressors.

Since the own-lags of the dependent variable are included as regressors, our model is susceptible to endogeneity bias. This is in principle quite important, especially because we use annual growth. Such bias may be due to a common contemporaneous component, which may arise from an exogenous shock in the previous period. Following Arellano and Bond (1991) and Glaeser and Gyourko (2006), we use the Arellano-Bond estimator. Briefly, this procedure utilizes the generalized method of moments (GMM) to instrument with the dependent variable's own lag, $G_{i,t-1}$, with $G_{i,t-2}$. This process is repeated for the dependent variable's own second, third, and fourth lags. That is, $G_{i,t-2}$ is instrumented by $G_{i,t-3}$, $G_{i,t-3}$ by $G_{i,t-4}$, and $G_{i,t-4}$ by $G_{i,t-5}$. The same procedure is repeated for the possibly endogenous spatial regressors.

It is important to note that the Arellano-Bond estimator[11], like all instrumental estimation methods, hinges on two assumptions. First, $G_{i,t-h}$ must be correlated with

[11] As described by Arellano and Bover (1995) and Blundell and Bond (1998), lagged levels are not good instruments for first differenced variables. While they suggested using lagged differences *and* lagged levels, we chose to use the lagged differences of the housing price indexes and lagged differences of all the lagged spatial variables. Given the large number of instruments with lags of all the lagged differences, and the documented poor performance of the levels as instruments, we chose this approach rather than also including lagged levels.

$G_{i,t-h-1}$, where h denotes a lag. Second, the instrument, $G_{i,t-h-1}$, must be *uncorrelated* with the model's error term.[12] One potential problem that may arise is when the instruments of the spatial regressors are correlated with the time lag of the dependent variable, which could result in multicollinearity. As discussed in Brady (2008), this issue is less of a concern when there is high variability in the dependent variable. Since we use panel data from 363 MSAs, in which the cross-sectional variability in house price growth rates is high, this helps mitigate the problem of multicollinearity in the instruments and in the time lags of the dependent variable.

4. Data Description[13]

Our rich data set, which results from merging the house price data with contextual information from a number of different sources including three different measures of spatial proximity, is a distinct contribution of the present paper.

4.1 House Price Data

House price data are taken from the Office of Federal Housing Enterprise Oversight (OFHEO). We define a panel dataset that is comprised of annual (based on 1st quarter) house price indices from 363 MSAs within the continental U.S. The panel data run from the first quarter of 1975 until the first quarter of 2013 and they were first published in March of 1996. However, because OFHEO requires that an MSA must have at least 1,000 total transactions before the MSA's Housing Price Index (HPI) may be published, the panel is unbalanced, with only roughly half of the 363 MSAs have HPIs

[12] We examined the correlation between the instruments and the error term for the AB estimation of equation (1) below. The correlations range from 5.79x10^-19 to 1.07x10^-16. These correlations appear sufficiently small that we are confident the errors are uncorrelated with the instruments.

[13] See Table 1 for the complete list of MSAs included in this study, which are based on the U.S. Census Bureau's 2009 MSA definitions; and the summary statistics.

that begin in 1975. We are able to obtain a balanced panel beginning in 1995 through 2013.[14] We use the first quarter data from each year during this period to construct the balanced panel. A map of the U.S. and all of the 363 MSA's is in Figure 1, and a list of the 363 MSA's is in Table 1.[15]

4.2 Spatial Data

We use three measures of spatial proximity. The first is a contiguity or adjacency matrix, a 363x363 matrix $\boldsymbol{W}$; the second is a 363x363 matrix of physical distances between metro areas, $\boldsymbol{D}$.[16] The third, also a 363x363 time-varying matrix, $\boldsymbol{\psi}_{t}$, is a set of migration weights defined by using migration data. This migration data was originally collected by the Internal Revenue Service (IRS). We use a compiled version of the IRS data that we obtained from Telestrian.com.

Regarding the contiguity weights, any pair of MSAs that border one another, the value "1" is entered into $\boldsymbol{W}$, otherwise the default value is "0". $\boldsymbol{W}$ is normalized such that the sum of each row equals one.

For the distance matrix, the U.S. Census Bureau provides centroids (reference points at the center of each polygon) for all of the 363 MSAs. The physical distances, measured in kilometers, between any pair of centroids define the entries in $\boldsymbol{D}$.

---

[14] As we detail further below, this conforms to the availability of the migration data, which is based on the US Internal Revenue Service information.

[15] Note that since we focus on the continental U.S., Figure 1 includes some MSA's that we did not include in our sample, such as those in Alaska, Hawaii, and Puerto Rico. Also, Figure 1 also shows Micropolitan Statistical Areas, but we restrict our attention to the MSAs due to data consistency and availability over our entire sample period of 1996-2013.

[16] All spatial distance data are calculated based on 2009 Tiger Line files from the U.S. Census Bureau, which include latitude and longitude for the MSA centroids. We use the Haversine distance formula to calculate the distances between each pair of MSAs. For the contiguity matrix, an ArcGIS script is used to identify whether any polygons (MSA boundary) pairs border one another.

The migration data are annual, covering the period 1996-2011 (16 years). The migration weights that MSA j have on MSA i in a given year are based on the sum of migration inflows and outflows between i and j in that year. Since our annual migration data cover the period 1996-2011, we construct a separate migration weights matrix for each of these years, $\boldsymbol{\psi}_t$. We then row-normalize this migration weights matrix, and place them into a larger, block diagonal matrix that is dimension (16 times 363) by (16 times 363).

4.3 Consumer Price Index Data

The Consumer Price Index (CPI) data are taken from the Bureau of Labor Statistics (BLS). The data are the Urban Consumer CPI for All Items from 1995 to 2013, for each of 4 regions in the U.S. Each MSA is classified into one of these 4 regions, then the appropriate regional CPI is used to deflate its HPI. Using these regional CPI deflators avoids some limitations with the MSA-level CPI data that are noteworthy. First, the BLS only publishes CPI data for 27 metropolitan areas.[17] There are only 39 MSAs that fall inside (completely and partially) the boundaries of the 27 metropolitan areas. Therefore only 39 MSAs have a reported CPI, while many of the 39 MSAs also share common CPI. Second, the frequency of the published CPI varies (monthly, bimonthly, and semiannually) for different metropolitan areas. These variations do not necessary coincide with the timing of OFHEO's HPI, which is reported on a quarterly basis.

Descriptive statistics for the housing price growth and spatial lags of housing price growth are presented in Table 2. The typical MSA experienced year-over-year mean price growth of approximately 0.4% and a twice as large median value. Its

[17] See www.bls.gov for the list of the 27 metropolitan areas.

neighbors' house price growth ranged from an average (mean and median) of between 2% and 3%, depending on the definitions of neighbors. The largest year-over-year increase in an MSA over the years 1996-2013 was 28%, while the largest drop was 45%. This range is somewhat smaller for the neighboring MSA's maximum and minimum year-over-year price growth.

5. Empirical Results

Table 2 reports summary statistics for the data. The average annual real growth rate of housing prices in MSAs is approximately 0.43%, while the average of the contiguity "neighbor" MSA price growth was 0.37% and the migration weighted MSA price growth rate average was approximately 0.72%.

Tables 3 and 4 report the estimation results for the OLS and Arellano-Bond regressions, respectively. Columns 1, 4, 7, and 10 present estimates that only include time lags of the house price growth rates ("lagged growth") as explanatory variables. Columns 2, 5, 8, and 11 of Tables 3 and 4 include time lags of the contiguity-based spatial lags ("spatial lag growth"), in addition to the own lagged growth variables. Columns 3, 6, 9, and 12 of Tables 3 and 4 present results for time lags of the migration-based spatial lags (spatial migration growth), in addition to the "lagged growth" and "spatial lagged growth" variables. The first two rows of each of Tables 3 and 4 indicates whether or not the results in each corresponding column of the table were estimated with MSA-level fixed effects ("Yes" if the model included fixed effects, "No" if not); and whether the model is estimated based on the post-2007 sample (indicated by "Yes") or for the entire sample period (indicated by "No"). The entire sample period covers the years 1996-2013 (except for the regressions containing spatial lags for migration flows in columns 3, 6, 9, and 12 of Tables 3 and 4, which covers the period 1996-2011, due to data unavailability for 2013).

5.1 Own-Lag Effects: Baseline Results

For our entire sample of 1996-2013, we first report the results of the own-lags regression. We first focus on the results in Table 3, which reports the baseline results from estimating equation (1) with only own-lags effects and MSA fixed effects, for the CPI-deflated housing price growth regression. Column 1 of Table 3 reports the coefficient estimates of the own-lagged model with fixed effects. The coefficients on the first 2 lags are positive, less than 1, and statistically significant. The coefficients on the three-year and four-year lags are negative, less than 1 in absolute value, and statistically significant. The sign of the one-year lag coefficient is consistent with previous literature, suggesting that the first own-lag of house prices has explanatory power in forecasting the next period's house prices. Since the third, and fourth own-lags are all also highly statistically significant but negative, though smaller in absolute value than the first and second own-lags, once again this may suggest some degree of mean reversion.

We introduce four time lags of the spatially lagged growth rates of all MSAs contiguous to MSA_i as additional regressors as in equation (2), with fixed effects. The results are reported in column 2 of Table 3. The first three of these lagged spatial lag coefficients are positive, while the second and fourth time lags are statistically significant at all significance levels. We perform a likelihood ratio test between the restricted model (OLS estimates of own-lags without the presence of adjacent spatial regressors) and the unrestricted models (equation (2)). The LR test statistic is approximately equal to 100, while the χ^2 statistic (critical value) with 4 degrees of freedom is 9.49. Hence, the LR test strongly rejects the null hypothesis at all levels of statistical significance that spatial dependence is *not* present in the residuals.

In column 3 of Table 3, we add time lags of spatial lags using migration weights, for the OLS model as in equation (5). Since our migration data covers the years 1996-2011, we estimate equation (5) for this time period. In the OLS case for the time lags of the spatially lagged migration weights in column 3 of Table 3, the second time lag is positive and significant, while the fourth time lag is negative and significant; the other two time lags are insignificant. Column 4 of Table 3 presents the growth rates with fixed effects regression results, for the estimated coefficients of own-time lags.[18] These results are similar to the results without fixed effects in column 1, in terms of their signs and significance. Column 5 of Table 3 presents the fixed effects results for the model with own-time lags and time lags of the contiguity spatial lags. In contrast to the OLS results in column 2, these results are similar in terms of the signs and significance of the coefficients. Finally, in column 6 of Table 3 we add time lags of spatial lags using migration weights with fixed effects, as in equation (5). The signs and significance of these coefficients are similar to the OLS version of this model in column 3 of Table 3.

A number of remarks are in order. First, spatial effects from growth rates of housing prices in neighboring MSAs are clearly present. Second, in general, lagged house

[18] We also estimate models where we add time lags of inverse distance weights spatial lags, as in equation (3). Once again, all four of the own-price lags are statistically significant. But in this version of the model, all of the contiguity neighbor spatial lags, and all of the inverse distance spatial lags with the exception of the second and fourth year lags of the 500 to 1000 km spatial lags, are statistically insignificant (which we attribute to multicollinearity). We do not present these inverse distance results in the paper, however the detailed results are available from the authors upon request.

price growth rates for adjacent MSAs have a comparable degree of explanatory power relative to the own-lags in explaining MSA_i's current growth rates. Third, including fixed effects does not substantially affect the estimates.[19] And fourth, the own-lag effects and cross-lag effects could be co-determined. Accordingly, we re-estimate these models with the Arellano-Bond estimator, which accounts for potential endogeneity of the regressors.

Column 1 of Table 4 reports the results using the Arellano-Bond correction, for OLS. The signs for all 4 lag coefficients remain unchanged. [20] The 4-period lagged growth rate coefficient is now positive.

Column 2 of Table 4 reports the results including both the own-lags and the time lags of the spatial variables with the Arellano-Bond estimator. Here the first and third year time lags of the own lags are significant, but none of the spatial lags is significant.[21] In contrast, the Arellano-Bond estimations with fixed effects have similar signs and significance, with the exception of the fourth spatial lag, which is now significant.

The results for equation (3) with the Arellano-Bond estimator are not as good as those obtained with OLS. Specifically, none of the spatial parameter estimates are significant. We examine the correlations between the various inverse distance and contiguity spatial lag variables, and find that there is likely a high degree of multicollinearity that is leading to insignificant parameter estimates. Given the lack of

[19] For completeness, we include the fixed effects estimates in Table 3. Obviously, if the original model is a fixed effects specification with the house price index in levels, these fixed effects would drop out when obtaining the first differences.

[20] Although the magnitude of the coefficient on the one-year lag is greater than 1.0, it is not statistically significantly greater than 1.0.

[21] The one-period lag on both the own-lag and the spatial lag are greater than 1.0 in magnitude, but not statistically significantly greater than 1.0.

significance of most of the parameters when we include inverse distance weighted spatial lags, we omit those results from Table 4.[22]

Column 3 of Table 4 presents the results for equation (5) estimated by the Arellano-Bond estimator, without fixed effects. All parameter estimates are highly insignificant. While one might anticipate that there can be persistence over time in MSA-to-MSA migration that leads to time series autocorrelation and contributing to the high standard errors (and in turn, low t-statistics), all of these estimates are based on heteroskedasticity and autocorrelation consistent (HAC) standard errors. Thus, it is likely in the full sample that when we include time lags of both types of spatial lags and control for potential endogeneity, there is little evidence of spatial spillovers. One might conjecture, however, that this result does not hold during the period following a “bust”, so we turn our attention to the post-2007 period.

Column 4 of Table 4 estimates the own-lag model with fixed effects for the Arellano-Bond estimator. While the signs are similar to the OLS estimates in column 1, the magnitudes are somewhat larger in column 1 than those in column 4. Column 5 presents the results for the Arellano-Bond fixed-effects estimation where we include own-lags and time lags of spatial contiguity neighbors. In column 5, there is one additional own-lag and one additional spatial lag that is significant, compared with the OLS model in equation 2. Finally, column 6 of Table 4 presents the Arellano-Bond estimation results including fixed effects, and once again, all parameter estimates are highly insignificant.

[22] These results are available from the authors upon request.

5.2. Post-2007 Results

Since one might expect the results to differ after the onset of the Great Recession in 2007, we re-estimated the models described above, for the period 2008-2013 (and for equation (5), for the period 2008-2011, because of the unavailability of migration flows data after 2011).

The post-2007 sample OLS estimates of the model with own-time lags, shown in Table 3 column 7, are similar in terms of signs and significance of the parameter estimates for the entire sample OLS estimates in column 1 of Table 3. In the post-2007 sample with own- and contiguity neighbor spatial lags in Table 3, column 8, there are 2 more significant spatial lags, but one fewer significant own-lag, than for the results in the full sample in column 2. In column 9 of Table 3 for the post-2007 sample, there is one additional significant migration spatial lag, and two additional significant contiguity neighbor spatial lag, compared with the results for the full sample in column 3 of Table 3. When we incorporate fixed effects for the post-2007 sample, the results for the model with only own-time lags in column 10 is similar to the results without fixed effects for the entire sample in column 1; the fixed effects post-2007 model in column 11 of Table 3 with contiguity spatial lags has one more significant contiguity spatial lag than the corresponding model for the full sample in column 2; and there are two more significant contiguity spatial lags and two more significant migration spatial lags in column 12 of Table 3 compared with column 6.

In the Arellano-Bond estimation for the post-2007 sample, the results for the model with only own-time lags in column 7 of Table 4 are similar to the results for the full sample in column 1 of Table 4. In the model with own-time lags and contiguity neighbor spatial lags for the post-2007 sample, there are two additional significant own-time lags and one additional significant spatial time lag in column 8 of Table 4, compared with column 2 of Table 4. For the time lags of the migration weighted spatial lags estimated with the Arellano-Bond approach in equation (5), shown in column 9 of Table 4, two of the own-price lags are significant, as are the two-, three- and four-period lagged migration weighted price changes (whereas all of the variables are insignificant in the full sample Arellano-Bond estimations in column 3 of Table 4). These larger magnitudes in the post-2007 sample imply contagion may have been more important after the crisis of 2007-08. More importantly, spatial contagion in this post-2007 model is present in the three-period lag of the spatial lag, which is significant and not in the one-, two-, and four-period lags of the spatial lag in column 8 of Table 4, which are insignificant.[23]

In column 10 of Table 4 for the post-2007 sample in the Arellano-Bond model with fixed effects, there is one additional significant own-time lag compared with the full sample in column 1 of Table 4. In column 11 of Table 4, there are two additional own-time lags and one additional contiguity spatial lag that is significant in this post-2007 Arellano-Bond fixed effects model compared with the full sample Arellano-Bond results without fixed effects in column 3 of Table 4. Virtually all own- and spatial time lags in the post-2007 Arellano-Bond fixed effects model are highly insignificant in column 12 of Table 4, while two of the own-price growth lags are significant, two of the spatial migration growth lags are significant, and all 4 contiguity price growth spatial lags are insignificant. In contrast, for the OLS post-2007 model with fixed effects in column 12 of

[23] None of the four spatial lags are statistically significant in the full sample, while one of the spatial lags are statistically significant in the post-2007 sample.

Table 3, three out of the 4 own-price lags are significant, while all 4 migration spatial lags are significant, and 2 out of 4 contiguity neighbor spatial lags are significant.[24] These post-2007 OLS fixed effects results in Table 3 are very similar to the post-2007 Arellano-Bond fixed effect results in Table 4.

5.2.1 Structural Breaks after 2007

As an additional diagnostic, we examine models with structural breaks for the post-2007 period while including the entire sample from 1996-2013 in the data set.[25] This allows us to determine whether or not such a model is preferred over our approach in the previous section with the post-2007 data set only. As a robustness check, we re-estimate each of the 3 basic models including a set of structural break interaction terms. These three models are OLS models (with MSA fixed effects and HAC standard errors): one, with own-time lags only; two, with spatial lags of neighbors with own- and neighbor-time lags; and three, spatial lag of neighbors and migration spatial lag, together with own- and neighbor- and migration- time lags). For the most part, the signs and statistical significance of the parameter estimates in each model are not substantially different when we estimate each model with structural breaks compared with its counterpart in the estimations with the post-2007 sample only. A major difference, however, is that in all three of the basic models, the R-squared (as well as the adjusted R-squared) is substantially higher in the post-2007 sample estimation than in the structural breaks

---

[24] With the inverse distance weights model in equation (3) for post-2007, only 3 of the time-lagged inverse distance spatial lags are significant, and all of the own time lags are statistically significant. This is a slight improvement over the results for the entire sample, but still somewhat disappointing, likely due to the multicollinearity between the inverse distance-weighted spatial lags. In the version of the model with the Arellano-Bond estimator for post-2007, the one-year and two-year own time lags are the only two statistically significant variables in the model. These detailed coefficient estimates are available from the authors upon request.

[25] While we do not report these results in the Tables, we summarize them in this section. Detailed results for these models are available upon request from the authors.

estimation with the entire sample period. For this reason, we have chosen not to present detailed tables of the structural breaks estimation results (these results are available from the authors upon request). We do include some discussion of these results below.

In the version of the model with the own-MSA time lags only, the break dummy is statistically significant, and the first two lagged coefficients are not substantially different than they are for the model with no structural breaks. After factoring the interaction terms for the structural break, the last two lags (years 3 and 4) are significantly less than they are in the model without the breaks. But the R-squared is lower in the model with structural breaks (0.578) opposed to the model with the post-2007 data only (0.689).

In the model with neighbor spatial lags and own- and neighbor- time lags, there are 3 own-time lags interaction terms that are negative and significant, and there are 2 spatial own-time lags interaction terms that are negative and significant. The pre-2008 coefficients are smaller than they were in the full sample model with no structural break terms. Once again, the model with the sample post-2007 has a higher R-squared than the model with the post-2007 interaction terms (0.712 opposed to 0.594).

Finally, we include the structural break interaction terms for neighbor spatial lags, migration spatial lags, and own-, neighbor-, and migration- time lags. In this variation of the model, two migration spatial weighted time lagged break variables are positive and significant, while only one is negative and significant, implying a large effect of migration on house price growth in the post-2007 period. Once again, the R-squared value for the sample that includes only post-2007 observations (R-squared=0.7399) is substantially higher than the R-squared=0.6051 for the model estimated with structural break interaction terms.

5.2.2 Moran I test for Spatial Dependence

We conduct and report here a Moran I test before and after the Great Recession of 2007-2009, in order to examine whether spatial effects are more pronounced following the crisis. We repeat two cross-sectional regressions for equation (1), one of which is for all 363 MSA's in 2007 and the other is for the same MSA's in 2012. The Moran I test cannot reject the null hypothesis of no spatial autocorrelation (P-value=0.2800) for the year 2007. But we find the opposite result for the 2012 regression, with a P-value of 0.0012. These Moran I test results confirm our claim that spatial effects are particularly pronounced following the Great Recession of 2007-2009.

5.3 Impulse Response Functions

Here we report the simulations that utilize the estimates discussed in the previous section (based on column 4 of Table 3 for the own-effects, and column 5 of Table 3 for the feedback effects) in order to help us assess the "ripple" effect of a one standard deviation shock to house price growth rates in one area propagate across space, ceteris paribus.[26] The first simulation examines how house price growth rates in one area respond to a shock in that same area, or the own-effect. The results from the first simulation will serve as a benchmark against spatial effects, which is the focus of the second simulation. The second simulation looks at how the own-effects interact with the spatial effects. For both simulations, we assume that there are only two

[26] A referee pointed out that impulse response functions are more commonly used with long time-series data, while here we have a greater number of cross-sectional than longitudinal observations. While we are aware of the fact that our approach stretches the standard practice a bit, we still think it is an interesting diagnostic avenue to pursue.

MSAs, X and Y, and that they are mutually and exclusively adjacent neighbors to one another. The first simulation focuses on the effect that a shock in X has on prices in X. The second simulation focuses on the effect that a shock in X has on Y, and the feedback effects that Y has on X. Finally, we present the total effects – inclusive of the feedbacks and the direct effects – of a shock on X.

5.3.1. Simulation 1: Own-Effect

We compute the impulse response function of house prices in X to an exogenous one standard deviation shock that occurred in X, and follow the effect the shock has on house price growth rates in X over time,

$$G_{x,t} = B_0 + B_1 G_{x,t-1} + B_2 G_{x,t-2} + B_3 G_{x,t-3} + B_4 G_{x,t-4} + \varepsilon_{x,t}\,, \qquad (6)$$

where $G_{x,t}$ is the annual house price growth rate of X at period time t, B_0 to B_4 are the estimated coefficients of equation (1) taken from column 4 of Table 3, and $\varepsilon_{x,t}$ is the model's error term.

Consider an exogenous one standard deviation shock on X at time t = 0, such that $G_{x,0} = 0.0615$, the standard deviation of all house price growth rates from Table 2. We then follow this unit shock and see how it attenuates. The results are reported in Figure 2.

5.3.2 Simulation 2: Effects on Neighboring MSA Y of a One Standard Deviation Shock to MSA X House Prices (with Feedback Effects)

We report the impulse response function of house price growth rates in Y to an exogenous one standard deviation shock on house price growth rates in X, and follow the effect of this shock over time. We also follow the effects of this shock on X over time, that is, we allow for feedback effects between X and Y,

$$G_{y,t} = \eta_0 + \beta_1 G_{y,t-1} + \beta_2 G_{y,t-2} + \beta_3 G_{y,t-3} + \beta_4 G_{y,t-4} + \lambda_1 G_{x,t-1} + \lambda_2 G_{x,t-2} + \lambda_3 G_{x,t-3} + \lambda_4 G_{x,t-4} + \varepsilon_{x,t} \text{ , (7)}$$

where $G_{x,t}$ and $G_{y,t}$ are the annual house price growth rate of X and Y at period time t,respectively, η_0 ,β_1 to β_4, and λ_1 to λ_4 are the constant, the estimated coefficients of the own-lags, and the estimated coefficient of the adjacent spatial lags in equation (3), based on the results in column 5 of Table 3, respectively.

The second simulation begins with an exogenous one standard deviation shock on house prices in X such that $G_{y,0}$ = 0.0615. We follow this unit shock and examine the period t = 1 effects, that linger to affect house prices in X in the subsequent years. This simulation also allows for the spillover effect on house prices in X to feedback and affects house prices in Y. For example, the impulse response function will estimate how an exogenous unit shock on house prices in Boston in 2000 has spillover effects on house prices in Providence in 2001, and how the spillover effect lingers and affects house prices in Providence in 2002 and the subsequent years. Moreover, the spillover effect from Boston to Providence in 2001 will have a feedback effect on house prices in Boston in 2002. This feedback effect also lingers and continues to affect house prices in Boston in 2003, 2004, and so on. In addition, the lingering feedback effect from Providence to Boston in 2003 will spillover and affect house prices in Providence in 2004. Both the feedback and spillover effects will continue to interact in the same manner as the previous year, but the effects will greatly attenuate since the coefficients β_1 through β_4 and λ_1 to λ_4 are all smaller than 1, and will ultimately disappear. The results of this experiment are reported in Figure 2.

Figure 2 pictures the changes in the annual house price growth rate of X, $G_{x,t,}$, across time for both simulations. It also pictures the spillover effects on Y of the shock to X. The first simulation reveals that a one standard deviation (approximately 6.15%) exogenous shock in X has positive but attenuated impacts on the house price growth rate in X for the first three years and the growth rate becomes negative in the fourth year. The house price growth rate in X hits a trough in the sixth year at about -2% before it begins to recover to around 1% and finally levels out after approximately 9 years.

The second simulation focuses only on spatial effects, that is, how an exogenous, one standard deviation shock on house prices in area X affects house prices in Y. To have a more complete understanding of how house prices interact across space, this simulation allows for feedback effects – in both directions – between house prices in X and Y. It is noteworthy that the positive spatial effect in time period 1 is lower than the own-effect of the same period, which is calculated in the first simulation. The impact of the unit shock also attenuates slower than those in the first simulation; the growth rate of house prices in X becomes negative between the fifth and the sixth year. The growth rate hits the trough at -0.75 percent during the seventh year, before it recovers and moves towards the growth rate of the MSA X house prices. We also simulate the effects (including feedbacks) of this shock on MSA X in Figure 2. Finally, we graph the total effect of a shock in X. That is, it reflects the sum of the direct effect in

MSA X and the effect in MSA Y, where we allow for feedback effects between the two adjacent MSAs.

It is important to note that we have simplified these simulations by treating the reference MSAs, X and Y, as mutually and exclusively adjacent neighbors. In reality, there are many MSAs that have more than one neighbor, and thus the results from our simulations are merely indicative. Nonetheless, the simulations help us conceptualize how spatial effects have significant and lasting effects over time in house prices.

6. Conclusions

We argue that studying house price dynamics at the MSA level can be more informative than those at higher levels of aggregation, such as the national and the Census region levels. This is, in part, because at the higher levels of aggregation the regional own-lag effects obscure the own-lag effects and spatial effects of MSAs within the respective US Census division.

Using panel data from 363 MSAs across the U.S. from 1996 to 2013, we establish that there is a notable spatial diffusion pattern in inter-MSAs house prices. Specifically, information on lagged price changes in neighboring (i.e., contiguous) MSAs, in addition to an MSA’s own time-lagged price changes, helps explain current house price changes. Consideration of migration flows can also be a driver for spatial interdependence in housing prices across MSAs. Such spatial attenuation is intuitively appealing, but has not been documented in earlier research. We use our estimation results to obtain a variety of diagnostics including impulse response functions for the house price dynamics.

Overall, we find that spatial effects play a significant role in explaining house prices even after controlling for own-lag effects. These spatial effects are particularly pronounced when considering the post-2007 sample period. In other words, our results imply greater spatial contagion following the Great Recession of 2007-09.

Our finding of very significant spatial effects underscores the need to incorporate the spatial dimension into existing house price dynamic equilibrium models, such as Glaeser and Gyourko (2006). The spatial threshold effects in spatial house price interactions that we identify suggest that policy makers and economic agents must take into account spatial attenuation and interactions when making economic decisions. For instance, it may be important to consider the spatial diffusion pattern in inter-MSA house prices when formulating business development policies. The impact of such a policy in some areas may spillover into the housing markets of neighboring areas, an intuitively appealing notion.

Our results on the richness of the dynamics in house prices are relevant in assessing arguments about housing price bubbles. Robert Shiller writing in the *New York Times* on "How a Bubble Stayed under the Radar" (Shiller, 2008) argues that the fundamental problem in verifying the existence of a housing market bubble is that

> "[the] information obtained by any individual — even one as well-placed as the chairman of the Federal Reserve — is bound to be incomplete. If people could somehow hold a national town meeting and share their independent information, they would have the opportunity to see the full weight of the evidence. Any individual errors would be averaged out, and the participants would collectively reach the correct decision."

Shiller goes on to state that

> "Of course, such a national town meeting is impossible. Each person makes decisions individually, sequentially, and reveals his decisions through actions — in this case, by entering the housing market and bidding up home prices" [*ibid*.]

Our paper provides some evidence on Shiller's argument. Specifically, information from neighboring areas is very important. Naturally, overreactions are smaller when people can share information, even if the combined information is still incomplete.

References

Anselin, Luc. 2001. Spatial Econometrics. *A Companion to Theoretical Econometrics, 310330*.

Arellano, Manuel, and Stephen Bond. 1991. "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations." *Review of Economic Studies*. 58(2):277-97.

Arellano, M. and Bover, O. 1995. "Another look at the instrumental variable estimation of error-components models." *Journal of Econometrics*, *68*(1):29-51.

Bayer, Patrick, Kyle Mangum, and James Roberts. 2014. "Speculative Fever: Micro Evidence for Investor Contagion in the Housing Bubble." Duke working paper, preliminary. June.

Basu, Sabyasachi, and Thomas G. Thibodeau. 1998. "Analysis of Spatial Autocorrelation in House Prices," *Journal of Real Estate Finance and Economics*. 17: 61-85.

Beenstock, Michael, and Daniel Felsenstein. 2007. "Spatial Vector Autoregressions." *Spatial Economic Analysis.* 2(2): 167-196.

Blundell, R., and Bond, S. 1998. "Initial conditions and moment restrictions in dynamic panel data models." *Journal of Econometrics*, *87*(1):115-143.

Brady, Ryan B. 2008. "Measuring the Persistence of Spatial Autocorrelation: How Long does the Spatial Connection between Housing Markets Last?" United States Naval Academy Department of Economics Working Papers 2008-19.

Brady, Ryan B. 2011. "Measuring the Diffusion of Housing Prices across Space and over Time." *Journal of Applied Econometrics*, 26: 213-231.

Brady, Ryan B. 2014. "The Spatial Diffusion of Regional Housing Prices across U.S. States." *Regional Science and Urban Economics*. 46:150-166.

Calhoun, Charles A. 1996. "OFHEO House Price Indexes: HPI Technical Description." Office of Federal Housing Enterprise Oversight.

Case, Karl E., and Robert J. Shiller. 1989. "The Efficiency of the Market for Single-Family Homes." *American Economic Review*. 79(1):125-37.

Case, Karl E., John M. Quigley, and Robert J. Shiller. 2005. "Comparing Wealth Effects: The Stock Market vs. the Housing Market.” *Advances in Macroeconomics*. 5(1). Article 1.

Clapp, John M., and Dogan Tirtiroglu. 1994. “Positive Feedback Trading and Diffusion of Asset Price Changes: Evidence from Housing Transactions.” *Journal of Economic Behavior and Organization*. 24(3):337-55.

Cutler, David M., James M. Poterba, and Lawrence H. Summers. 1991. “Speculative Dynamics and the Role of Feedback Traders,” *American Economic Review*. 80(2): 63-68.

DeFusco, Anthony, Wenjie Ding, Fernando Ferreira, and Joseph Gyourko. 2013. “The Role of Contagion in the Last American Housing Cycle.” Wharton School working paper.

Dobkins, Linda H., and Yannis M. Ioannides. 2001. “Spatial Interactions among US Cities: 1900–1990” *Regional Science and Urban Economics*. *31*(6): 701-731.

Glaeser, Edward L., and Joseph Gyourko. 2006. “Housing Dynamics,” National Bureau of Economic Research, Inc, NBER Working Paper No. W12787.

Gyourko, Joseph, Christopher Mayer, and Todd Sinai. 2013. "Superstar Cities." *American Economic Journal: Economic Policy*. 5(4): 167-99.

Guntermann, Karl L., and Stefan C. Norrbin. 1991. Empirical Tests of Real Estate Market Efficiency,” *Journal of Real Estate Finance and Economics*, 4(3): 297-313.

Henderson, J. Vernon, and Yannis M. Ioannides. 1983. “A Model of Housing Tenure Choice,” *American Economic Review*. 73(1): 98-113.

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

Holly, Sean, M. Hashem Pesaran, and Takashi Yamagata. 2011. “Spatial and Temporal Diffusion of House Prices in the UK.” *Journal of Urban Economics*. 69(1):2-23.

Kuethe, Todd H., and Valerien O. Pede. 2011. “Regional housing price cycles: a Spatio-temporal Analysis using US State-level Data.” *Regional Studies*. *45*(5):563-574.

Pace, R. Kelley, Ronald Barry, John M. Clapp, and Mauricio Rodriquez. 1998. "Spatial Temporal Autoregressive Models of Neighborhood Effects." *Journal of Real Estate Finance and Economics*, 17(1): 15-33.

Pesaran, M. Hashem, and Alexander Chudik. 2010. "Econometric Analysis of High Dimensional VARs Featuring a Dominant Unit." ECB Working Paper No. 1194.

Pollakowski, Henry O., and Traci S. Ray. 1997. "Housing Price Diffusion Patterns at Different Aggregation Levels: An Examination of Housing Market Efficiency." *Journal of Housing Research*. 8(1):107-24.

Rayburn, William, Michael Devaney, and Richard Evans. 1987. "A Test of Weak-Form Efficiency in Residential Real Estate Returns," *American Real Estate and Urban Economics Association Journal*. 15(3):220-33.

Shiller, Robert J. 2008. "How a Bubble Stayed Under the Radar," *The New York Times*, p. B1, March 2.

Thanapisitikul, Wirathip. 2008. "U.S. House Prices: Dynamics and Spatial Interactions, 1975-2007," B.S. Thesis, Tufts University.

Figure 1 – Metropolitan Statistical Areas of the U.S., 2009 definitions (Source: U.S. Census Bureau)

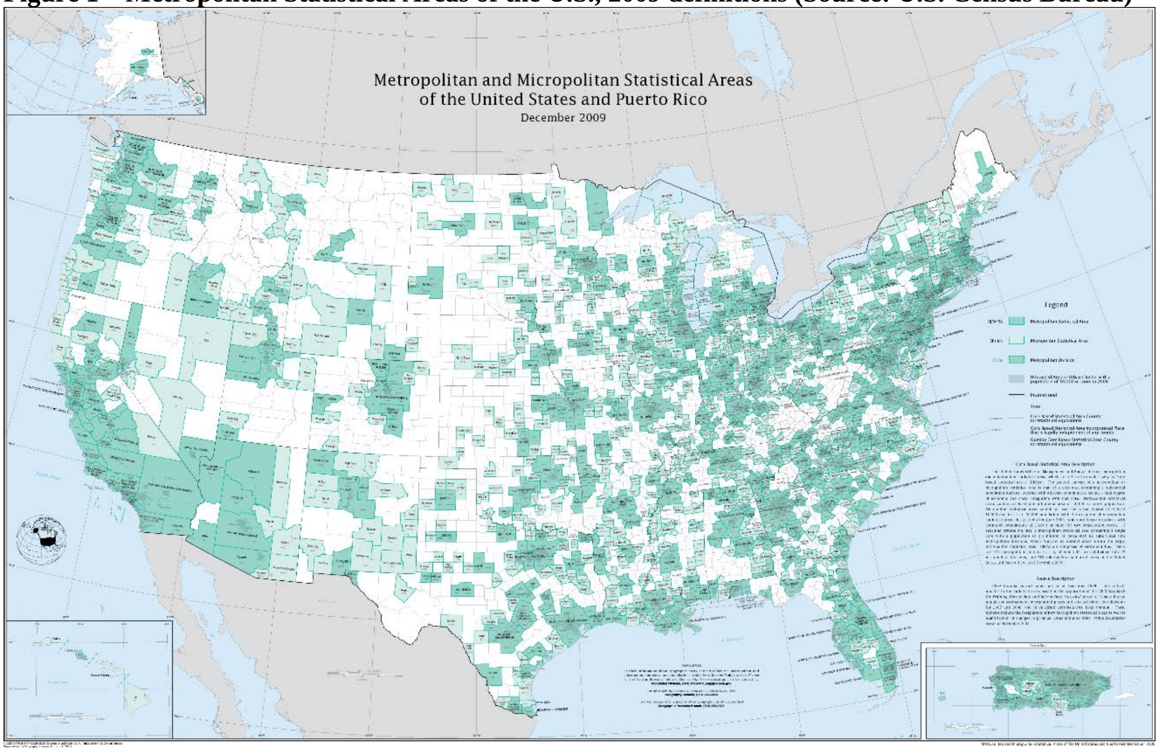
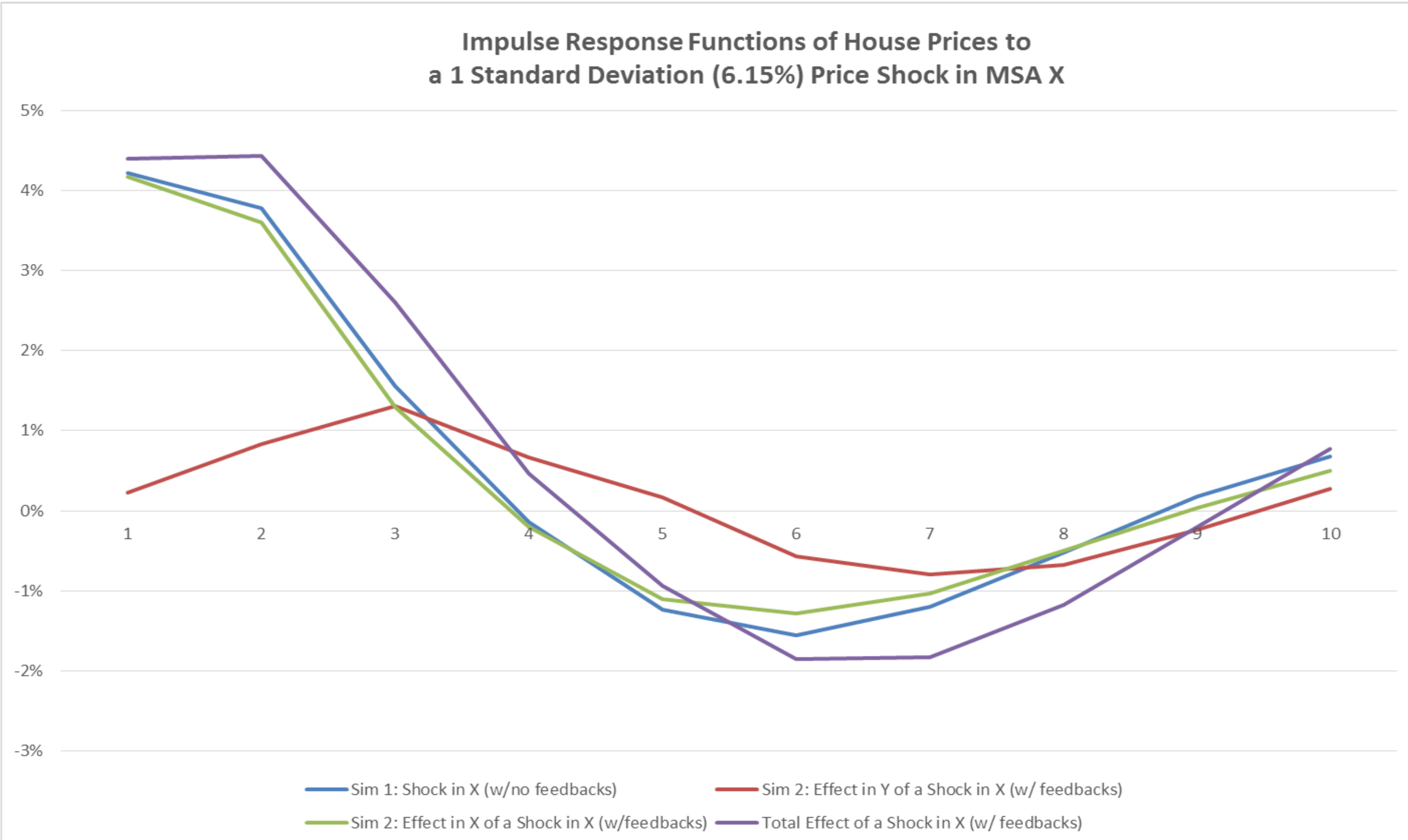


Figure 2



Source: Authors' Own Calculation

Table 1: 363 MSA's in the continental U.S. (2009 U.S. Census Delineations)

Abilene, TX
Akron, OH
Albany, GA
Albany-Schenectady-Troy, NY
Albuquerque, NM
Alexandria, LA
Allentown-Bethlehem-Easton, PA-NJ
Altoona, PA
Amarillo, TX
Ames, IA
Anderson, IN
Anderson, SC
Ann Arbor, MI
Anniston-Oxford, AL
Appleton, WI
Asheville, NC
Athens-Clarke County, GA
Atlanta-Sandy Springs-Marietta, GA
Atlantic City-Hammonton, NJ
Auburn-Opelika, AL
Augusta-Richmond County, GA-SC
Austin-Round Rock, TX
Bakersfield, CA
Baltimore-Towson, MD
Bangor, ME
Barnstable Town, MA
Baton Rouge, LA
Battle Creek, MI
Bay City, MI
Beaumont-Port Arthur, TX
Bellingham, WA
Bend, OR
Billings, MT
Binghamton, NY
Birmingham-Hoover, AL
Bismarck, ND
Blacksburg-Christiansburg-Radford, VA
Bloomington, IN
Bloomington-Normal, IL
Boise City-Nampa, ID
Boston-Cambridge-Quincy, MA-NH
Boulder, CO
Bowling Green, KY
Bradenton-Sarasota-Venice, FL
Bremerton-Silverdale, WA
Bridgeport-Stamford-Norwalk, CT
Brownsville-Harlingen, TX
Brunswick, GA
Buffalo-Niagara Falls, NY
Burlington, NC
Burlington-South Burlington, VT
Canton-Massillon, OH
Cape Coral-Fort Myers, FL
Cape Girardeau-Jackson, MO-IL
Carson City, NV
Casper, WY
Cedar Rapids, IA
Champaign-Urbana, IL
Charleston, WV
Charleston-North Charleston-Summerville, SC
Charlotte-Gastonia-Concord, NC-SC
Charlottesville, VA
Chattanooga, TN-GA
Cheyenne, WY
Chicago-Naperville-Joliet, IL-IN-WI
Chico, CA
Cincinnati-Middletown, OH-KY-IN
Clarksville, TN-KY
Cleveland, TN
Cleveland-Elyria-Mentor, OH
Coeur d'Alene, ID
College Station-Bryan, TX
Colorado Springs, CO
Columbia, MO
Columbia, SC
Columbus, GA-AL
Columbus, IN
Columbus, OH
Corpus Christi, TX
Corvallis, OR
Cumberland, MD-WV
Dallas-Fort Worth-Arlington, TX
Dalton, GA
Danville, IL
Danville, VA
Davenport-Moline-Rock Island, IA-IL
Dayton, OH
Decatur, AL
Decatur, IL
Deltona-Daytona Beach-Ormond Beach, FL
Denver-Aurora-Broomfield, CO
Des Moines-West Des Moines, IA
Detroit-Warren-Livonia, MI
Dothan, AL
Dover, DE
Dubuque, IA
Duluth, MN-WI
Durham-Chapel Hill, NC
Eau Claire, WI
El Centro, CA
El Paso, TX
Elizabethtown, KY
Elkhart-Goshen, IN
Elmira, NY
Erie, PA
Eugene-Springfield, OR
Evansville, IN-KY
Fargo, ND-MN
Farmington, NM
Fayetteville, NC
Fayetteville-Springdale-Rogers, AR-MO
Flagstaff, AZ
Flint, MI
Florence, SC
Florence-Muscle Shoals, AL
Fond du Lac, WI
Fort Collins-Loveland, CO
Fort Smith, AR-OK
Fort Walton Beach-Crestview-Destin, FL
Fort Wayne, IN
Fresno, CA
Gadsden, AL
Gainesville, FL
Gainesville, GA
Glens Falls, NY
Goldsboro, NC
Grand Forks, ND-MN
Grand Junction, CO
Grand Rapids-Wyoming, MI
Great Falls, MT
Greeley, CO
Green Bay, WI
Greensboro-High Point, NC
Greenville, NC
Greenville-Mauldin-Easley, SC
Gulfport-Biloxi, MS
Hagerstown-Martinsburg, MD-WV
Hanford-Corcoran, CA
Harrisburg-Carlisle, PA
Harrisonburg, VA
Hartford-West Hartford-East Hartford, CT
Hattiesburg, MS
Hickory-Lenoir-Morganton, NC
Hinesville-Fort Stewart, GA
Holland-Grand Haven, MI
Hot Springs, AR
Houma-Bayou Cane-Thibodaux, LA
Houston-Sugar Land-Baytown, TX
Huntington-Ashland, WV-KY-OH
Huntsville, AL
Idaho Falls, ID
Indianapolis-Carmel, IN
Iowa City, IA
Ithaca, NY
Jackson, MI
Jackson, MS
Jackson, TN
Jacksonville, FL
Jacksonville, NC
Janesville, WI
Jefferson City, MO
Johnson City, TN
Johnstown, PA
Jonesboro, AR
Joplin, MO
Kalamazoo-Portage, MI
Kankakee-Bradley, IL
Kansas City, MO-KS
Kennewick-Pasco-Richland, WA
Killeen-Temple-Fort Hood, TX
Kingsport-Bristol-Bristol, TN-VA
Kingston, NY
Knoxville, TN
Kokomo, IN
La Crosse, WI-MN
Lafayette, IN
Lafayette, LA
Lake Charles, LA
Lake Havasu City-Kingman, AZ
Lakeland-Winter Haven, FL
Lancaster, PA
Lansing-East Lansing, MI
Laredo, TX
Las Cruces, NM
Las Vegas-Paradise, NV
Lawrence, KS
Lawton, OK
Lebanon, PA
Lewiston, ID-WA
Lewiston-Auburn, ME
Lexington-Fayette, KY
Lima, OH
Lincoln, NE
Little Rock-North Little Rock-Conway, AR
Logan, UT-ID
Longview, TX
Longview, WA
Los Angeles-Long Beach-Santa Ana, CA
Louisville-Jefferson County, KY-IN
Lubbock, TX
Lynchburg, VA
Macon, GA
Madera-Chowchilla, CA
Madison, WI
Manchester-Nashua, NH
Manhattan, KS
Mankato-North Mankato, MN
Mansfield, OH
McAllen-Edinburg-Mission, TX
Medford, OR
Memphis, TN-MS-AR
Merced, CA
Miami-Fort Lauderdale-Pompano Beach, FL
Michigan City-La Porte, IN
Midland, TX
Milwaukee-Waukesha-West Allis, WI
Minneapolis-St. Paul-Bloomington, MN-WI
Missoula, MT
Mobile, AL
Modesto, CA
Monroe, LA
Monroe, MI
Montgomery, AL
Morgantown, WV
Morristown, TN
Mount Vernon-Anacortes, WA
Muncie, IN
Muskegon-Norton Shores, MI
Myrtle Beach-North Myrtle Beach-Conway, SC
Napa, CA
Naples-Marco Island, FL
Nashville-Davidson--Murfreesboro--Franklin, TN
New Haven-Milford, CT
New Orleans-Metairie-Kenner, LA
New York-Northern New Jersey-Long Island, NY-NJ-PA
Niles-Benton Harbor, MI
Norwich-New London, CT
Ocala, FL
Ocean City, NJ
Odessa, TX
Ogden-Clearfield, UT
Oklahoma City, OK
Olympia, WA
Omaha-Council Bluffs, NE-IA
Orlando-Kissimmee, FL
Oshkosh-Neenah, WI
Owensboro, KY
Oxnard-Thousand Oaks-Ventura, CA
Palm Bay-Melbourne-Titusville, FL
Palm Coast, FL
Panama City-Lynn Haven-Panama City Beach, FL
Parkersburg-Marietta-Vienna, WV-OH
Pascagoula, MS
Pensacola-Ferry Pass-Brent, FL
Peoria, IL
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD
Phoenix-Mesa-Scottsdale, AZ
Pine Bluff, AR
Pittsburgh, PA
Pittsfield, MA
Pocatello, ID
Port St. Lucie, FL
Portland-South Portland-Biddeford, ME
Portland-Vancouver-Beaverton, OR-WA
Poughkeepsie-Newburgh-Middletown, NY
Prescott, AZ
Providence-New Bedford-Fall River, RI-MA
Provo-Orem, UT
Pueblo, CO
Punta Gorda, FL
Racine, WI
Rapid City, SD
Reading, PA
Redding, CA
Reno-Sparks, NV
Richmond, VA
Riverside-San Bernardino-Ontario, CA
Roanoke, VA
Rochester, MN
Rochester, NY
Rockford, IL
Rocky Mount, NC
Rome, GA
Sacramento--Arden-Arcade--Roseville, CA
Saginaw-Saginaw Township North, MI
Salem, OR
Salinas, CA
Salisbury, MD
Salt Lake City, UT
San Angelo, TX
San Antonio, TX
San Diego-Carlsbad-San Marcos, CA
San Francisco-Oakland-Fremont, CA
San Jose-Sunnyvale-Santa Clara, CA
San Luis Obispo-Paso Robles, CA
Sandusky, OH
Santa Barbara-Santa Maria-Goleta, CA
Santa Cruz-Watsonville, CA
Santa Fe, NM
Santa Rosa-Petaluma, CA
Savannah, GA
Scranton--Wilkes-Barre, PA
Seattle-Tacoma-Bellevue, WA
Sebastian-Vero Beach, FL
Sheboygan, WI
Sherman-Denison, TX
Shreveport-Bossier City, LA
Sioux City, IA-NE-SD
Sioux Falls, SD
South Bend-Mishawaka, IN-MI
Spartanburg, SC
Spokane, WA
Springfield, IL
Springfield, MA
Springfield, MO
Springfield, OH
St. Cloud, MN
St. George, UT
St. Joseph, MO-KS
St. Louis, MO-IL
State College, PA
Stockton, CA
Sumter, SC
Syracuse, NY
Tallahassee, FL
Tampa-St. Petersburg-Clearwater, FL
Terre Haute, IN
Texarkana, TX-Texarkana, AR
Toledo, OH
Topeka, KS
Trenton-Ewing, NJ
Tucson, AZ
Tulsa, OK
Tuscaloosa, AL
Tyler, TX
Utica-Rome, NY
Valdosta, GA
Vallejo-Fairfield, CA
Victoria, TX
Vineland-Millville-Bridgeton, NJ
Virginia Beach-Norfolk-Newport News, VA-NC
Visalia-Porterville, CA
Waco, TX
Warner Robins, GA
Washington-Arlington-Alexandria, DC-VA-MD-WV
Waterloo-Cedar Falls, IA
Wausau, WI
Weirton-Steubenville, WV-OH
Wenatchee-East Wenatchee, WA
Wheeling, WV-OH
Wichita Falls, TX
Wichita, KS
Williamsport, PA
Wilmington, NC
Winchester, VA-WV
Winston-Salem, NC
Worcester, MA
Yakima, WA
York-Hanover, PA

Youngstown-Warren-Boardman, OH-PA
Yuba City, CA

Yuma, AZ

Table 2 - Descriptive Statistics, 363 U.S. Metropolitan Statistical Areas, Annual Growth Rates

	Real House Price Growth	Neighbor (contiguity weights) Real House Price Growth	Neighbor (Migration Weights) Real House Price Growth
Mean	0.0043	0.0037	0.0072
Median	0.0087	0.0029	0.0179
Maximum	0.2847	0.2781	0.2079
Minimum	-0.4586	-0.4586	-0.1739
Std. Dev.	0.0615	0.0497	0.0443
Skewness	-0.5448	-0.4651	-0.6685
Kurtosis	8.1719	8.0560	3.7618
Jarque-Bera	7605.4184	7195.1214	572.9798
Probability	0.0000	0.0000	0.0000
Sum	28.1163	24.0774	42.0466
Sum Sq. Dev.	24.7439	16.1550	11.3989
Observations	6534	6534	5808
Years	1996-2013	1996-2013	1996-2011

Table 3 - 1996-2013, OLS estimates (regressions using migration data cover the years 1996-2011)

Dependent Variable: House Price Index Growth between year t and t-1, for MSA i

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
MSA-Level Fixed Effects?	**No**	**No**	**No**	**Yes**	**Yes**	**Yes**	**No**	**No**	**No**	**Yes**	**Yes**	**Yes**
Post-2007 sample?	**No**	**No**	**No**	**No**	**No**	**No**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
constant	0.0038	0.0046	0.0061	0.0013	0.0014	0.0023	-0.0219	-0.0239	-0.0212	-0.0992	-0.1027	-0.1109
	3.4879	***4.3707***	***6.0110***	***0.0679***	***0.1411***	***0.2254***	***-11.6479***	***-13.4886***	***-12.9309***	***-6.6274***	***-7.0855***	***-8.0289***
Lagged Growth (t-1)	0.6725	0.6646	0.6557	0.6645	0.6542	0.6436	0.4787	0.5409	0.6030	0.0845	0.0980	0.0404
	18.7082	***17.6752***	***16.1218***	***18.0750***	***48.4598***	***44.9028***	***10.8214***	***11.8537***	***11.5192***	***3.8790***	***4.2363***	***1.6346***
Lagged Growth (t-2)	0.1363	0.1174	0.0890	0.1372	0.1179	0.0883	0.1222	0.0342	-0.0513	0.1344	0.0450	-0.0509
	3.5566	***3.0593***	***2.1310***	***3.4508***	***7.4371***	***5.1947***	***2.7476***	***0.7173***	***-0.9276***	***7.4445***	***2.3037***	***-2.4031***
Lagged Growth (t-3)	-0.2708	-0.2784	-0.2776	-0.2712	-0.2799	-0.2798	-0.3070	-0.2313	-0.1602	-0.2240	-0.1695	-0.0999
	-10.2602	***-9.8116***	***-8.9641***	***-9.9521***	***-17.8217***	***-16.5397***	***-10.6354***	***-6.9375***	***-4.2012***	***-12.1182***	***-8.6546***	***-4.8241***
Lagged Growth (t-4)	-0.1003	0.0768	-0.0449	-0.1108	-0.0879	-0.0570	-0.1336	-0.1460	-0.1587	-0.2892	-0.3282	-0.3916
	-4.0124	***3.0905***	***-1.7460***	***-4.2813***	***-6.2859***	***-3.9069***	***-4.7688***	***-5.5903***	***-6.0219***	***-16.7853***	***-18.5013***	***-21.5490***
Spatial Lag Growth (t-1)		0.0082	-0.0049		0.0138	-0.0018		-0.2425	-0.1211		-0.0921	-0.1983
		0.3770	***-0.1961***		***0.8030***	***-0.0810***		***-7.7079***	***-3.9729***		***-3.2241***	***-5.7566***
Spatial Lag Growth (t-2)		0.0770	-0.0155		0.0777	-0.0184		0.2592	0.0626		0.2621	0.0396
		2.7036	***-0.4862***		***3.7260***	***-0.6399***		***7.0226***	***1.6401***		***10.9748***	***1.2349***
Spatial Lag Growth (t-3)		0.0345	0.0235		0.0365	0.0235		-0.1320	0.0353		-0.1037	0.0394
		1.2952	***0.8429***		***1.7232***	***0.7882***		***-3.7028***	***0.9791***		***-4.2561***	***1.2286***
Spatial Lag Growth (t-4)		-0.1341	0.0096		-0.1318	0.0118		-0.0493	-0.0586		0.0156	-0.0747
		-5.4314	***0.3651***		***-6.7540***	***0.4464***		***-1.8794***	***-2.1875***		***0.6868***	***-2.6416***
Spatial Migration Growth (t-1)			0.0164			0.0221			-0.2243			0.1587
			0.4303			***0.7894***			***-3.9291***			***4.2938***
Spatial Migration Growth (t-2)			0.1809			0.1873			0.3869			0.4047
			3.3937			***5.1580***			***5.8120***			***9.3184***
Spatial Migration Growth (t-3)			0.0254			0.0290			-0.3275			-0.2051
			0.5648			***0.7840***			***-4.9808***			***-4.7835***
Spatial Migration Growth (t-4)			-0.3069			-0.3037			-0.0220			0.0904
			-6.1895			***-8.7183***			***-0.4374***			***2.2960***
R-squared	**0.4975**	**0.5044**	**0.5148**	**0.5027**	**0.5099**	**0.5205**	**0.5160**	**0.5426**	**0.5586**	**0.6893**	**0.7115**	**0.7399**

Numbers in bold italics are t-statistics; all standard errors are HAC consistent

Table 4 - 1996-2013, Arellano-Bond Estimation (regressions using migration data cover the years 1996-2011)

Dependent Variable: House Price Index Growth between year t and t-1, for MSA i

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
MSA-Level Fixed Effects?	**No**	**No**	**No**	**Yes**	**Yes**	**Yes**	**No**	**No**	**No**	**Yes**	**Yes**	**Yes**
Post-2007 sample?	**No**	**No**	**No**	**No**	**No**	**No**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Constant	0.0002	0.0030	0.0290	-0.0015	0.0049	0.0215	0.0087	0.0004	-0.0029	0.0991	0.0678	0.0417
	0.0706	***0.7645***	***0.3694***	***-0.0362***	***0.0705***	***0.1530***	***0.9541***	***0.0644***	***-0.4297***	***1.2442***	***0.9246***	***0.6974***
Lagged Growth (t-1)	2.5980	4.7649	16.9758	2.5766	3.9537	7.9512	1.8243	-1.5373	1.4168	2.3640	1.9901	1.7368
	3.1737	***1.5845***	***0.3229***	***3.5235***	***2.2607***	***0.8757***	***5.0538***	***-5.6363***	***5.2540***	***3.3191***	***2.9915***	***3.0643***
Lagged Growth (t-2)	-1.2409	-2.7375	-11.1459	-1.2219	-2.1540	-4.8760	-0.6470	0.5790	-0.6114	-0.7392	-0.6839	-0.6978
	-2.0603	***-1.3034***	***-0.3076***	***-2.2688***	***-1.7709***	***-0.7869***	***-2.8910***	***3.1559***	***-3.2741***	***-2.5095***	***-2.6135***	***-3.1427***
Lagged Growth (t-3)	-0.4775	-0.4118	-0.6716	-0.4741	-0.3955	-0.4431	-0.3676	-0.2098	-0.0534	-0.4633	-0.2299	-0.0516
	-4.4231	***-2.7177***	***-0.5000***	***-4.7545***	***-3.4086***	***-1.4269***	***-6.6315***	***-3.8213***	***-0.8676***	***-5.6964***	***-4.0185***	***-0.8133***
Lagged Growth (t-4)	0.4784	0.9798	3.9734	0.4761	0.8036	1.8163	0.3116	0.1496	0.0369	0.5062	0.3042	0.1411
	1.8554	***1.2640***	***0.3074***	***2.0069***	***1.6840***	***0.7812***	***2.8343***	***1.9570***	***0.5461***	***2.1299***	***1.4375***	***0.8056***
Spatial Lag Growth (t-1)		-2.9138	-0.2449		-2.1175	-0.3542		-0.1604	-0.1159		-0.4883	-0.2989
		-1.2376	***-0.0617***		***-1.6699***	***-0.2753***		***-0.6053***	***-0.4156***		***-0.7780***	***-0.3896***
Spatial Lag Growth (t-2)		2.0804	0.1022		1.4961	0.2133		0.3570	0.0484		0.4458	0.0886
		1.2836	***0.0344***		***1.7556***	***0.2292***		***1.9821***	***0.2514***		***1.7943***	***0.3040***
Spatial Lag Growth (t-3)		-0.1196	0.4683		-0.0997	0.2038		-0.3668	0.0573		-0.3884	0.0765
		-0.7921	***0.2837***		***-0.8953***	***0.5945***		***-5.1805***	***1.0260***		***-4.3886***	***1.1907***
Spatial Lag Growth (t-4)		-0.7466	-0.5192		-0.6039	-0.2626		0.0997	-0.0712		-0.0051	-0.1339
		-1.4773	***-0.3355***		***-2.0787***	***-0.6742***		***1.0291***	***-0.7795***		***-0.0251***	***-0.5407***
Spatial Migration Growth (t-1)			-9.3060			-3.7966			-0.2181			-0.2723
			-0.2900			***-0.7021***			***-1.2430***			***-0.7370***
Spatial Migration Growth (t-2)			6.2083			2.5456			0.6116			0.6385
			0.2971			***0.7470***			***2.9660***			***2.9254***
Spatial Migration Growth (t-3)			-2.4819			-1.0708			-0.7133			-0.7857
			-0.2926			***-0.6995***			***-4.2043***			***-3.0033***
Spatial Migration Growth (t-4)			-1.4366			-0.7917			0.1918			0.1909
			-0.3507			***-0.9822***			***2.3847***			***1.8521***
J-statistic	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**

Numbers in bold italics are t-statistics; all standard errors are HAC consistent