Local House Price Diffusion

Jeffrey P. Cohen[a]
University of Connecticut

and

Jeffrey Zabel[b,*]
Tufts University
Latest Version: July 2017

Key Words: House Price Diffusion; Spillover Effects

JEL Codes: R31

Abstract

An understanding of the temporal and spatial diffusion processes of local house prices, at different levels of aggregation, can help policymakers and homeowners anticipate the impacts of shocks. To this end, this paper investigates house price diffusion over time and space at the CBSA-level (nationwide), town-level and census tract-level (in the Greater Boston Area). One contribution of our analysis is our focus on diffusion effects at these different levels of aggregation. We estimate fixed effects models of growth rates on lagged growth rates (persistence effects), lagged growth rates from nearby jurisdictions (spillover effects) and growth rates in fundamentals. The estimated persistence and spillover effects are positive and significant at the CBSA-level. We find large ripple (contagion) effects that may have contributed to the recent housing downturn that reached the national-level. When estimating town-level diffusion, we find little evidence of persistence or spillover effects. Hence price diffusion appears to be stronger across than within housing markets. Both fundamentals and the price expectations appear to be the drivers of price diffusion which leaves room for housing bubbles based on households' over-optimism about future house prices.

a Department of Finance, and Center for Real Estate and Urban Economic Studies, School of Business, University of Connecticut, Storrs Connecticut, 06269,
e-mail: Jeffrey.Cohen@business.uconn.edu
b Department of Economics, Tufts University, Medford MA 02155, e-mail: jeff.zabel@tufts.edu
* corresponding author

The authors would like to thank Carolyn Talmadge and Aishwarya Venkat for their help with GIS, the MIT Center for Real Estate, the Warren Group, and CORELOGIC for providing the housing data. We acknowledge participants at the 2016 North American Regional Science Conference and the May 2015 Meeting of the Weimer School of Advanced Studies in Real Estate and Land Economics for their useful comments.

1. Introduction

An understanding of the temporal and spatial diffusion processes of local house prices, at different levels of aggregation, can help policymakers and homeowners anticipate the impacts of shocks. To this end, this paper investigates the movement of house prices over time and space. We look at house price diffusion at three different levels of aggregation: Core Based Statistical Area (CBSA), town, and census tract. This involves modeling house price growth rates as functions of lagged growth rates in both the own and nearby jurisdictions. The latter capture spillover effects. We also include fundamentals that are seen to drive house prices including population, income, and employment on the demand side and new housing on the supply side. We then estimate impulse response functions that result from shocks to the system to estimate short- and long-run persistence and spillover effects.

Most studies that look at price diffusion do so at a more aggregate level such as the MSA, region or county (see references below). Our analysis is different since we examine price shocks occurring at different levels of aggregation; CBSA, town, and census tract, and then we compare the price diffusion processes at these different levels.

At the CBSA level, this study can help to explain the recent major downturn in the housing market that precipitated the Great Recession. We show that there are large ripple or contagion effects such that house price growth rates are affected by price changes in nearby CBSAs. Large and persistent positive or negative growth rates in say, the sand states, can ripple across the CBSA landscape and potentially result in a downturn at a national level as witnessed in the Great Recession. Policymakers might pay closer attention to large changes in house prices in local areas if there is some chance that they can spread across jurisdictions.

This diffusion process is important for understanding the full impact on prices of housing cycle fluctuations and for understanding the impact of local shocks on prices in nearby jurisdictions. It should be of use in evaluating the full impact of policies that target specific areas such as enterprise zones or areas for redevelopment or policies that create affordable housing in certain locations. Another example is the cleanup of hazardous waste sites where it is interesting to look at how this affects prices in nearby areas. Another topic where price diffusion is important is gentrification where rising prices can “price out” longer-term, lower income residents.

It is important to understand the mechanisms by which price changes can propagate over time and space. House price growth can be influenced by changes in fundamentals such as area income, population, and employment. The employment effects could be important due to agglomeration economies (Duranton and Puga, 2004) that drive the growth of urban areas. An example of agglomeration economies is improved labor market matching, which can enhance productivity and lower costs for businesses. This can lead to additional growth and in turn, impact housing prices in and nearby the areas where the agglomeration is occurring. Prices can be expected to rise in the hub town(s), and fall in more distant towns as the higher productivity residents choose to migrate to the bigger towns in search of better opportunities. Shocks can also occur as demand for housing changes over time and certain types of units or certain areas become popular. On the supply side, overbuilding can actually result in a reversal of positive price shocks as witnessed in the recent Great Recession.

Changes in fundamentals can also lead to spillovers due to regional coordination. Zabel (2012) shows that labor demand shocks can lead to relatively small house price spillover effects in nearby MSAs that depend on the house price elasticity of supply. At the local level, Zabel and Dalton (2011) show that changes in minimum lot size restrictions not only affect prices in the town in which they occur but also spill over to nearby towns. Growth rate shocks can also stimulate supply-side effects in neighboring jurisdictions as profits can motivate developers to build new units in these areas.

At the local level, local price shocks can spread through information channels. These include the process by which appraisers value houses. That is, they use transactions from nearby units or "comps" as a means for valuing units. Real estate agents (and real estate websites such as Zillow.com ®) typically follow a similar procedure when valuing units and/or when advising clients about list prices. Hence even a price shock that affects a single unit can spread to nearby units through this process. Real estate agents, in particular, can be the basis of information dissemination as buyers are made aware of "hot" areas. This might cause them to buy in the hot areas and this can lead to persistence in the price shock in this area. They also might buy in nearby areas in the belief that these areas will be the next hot spots which will be self-fulfilling. This can result in a "ripple" effect as the initial price change leads to a series of price changes farther and farther away. For example, if one town increases its school quality this can affect prices in that town and possibly spillover to nearby towns.

At the CBSA level, Defusco et al. (2015) find that changes in fundamentals cannot explain house price changes in the home or neighboring MSAs. They also don't find that the mortgage lending can explain contagion effects. Short of these channels, other non-economic channels include information channels and non-rational behavior that can give rise to bubbles. Glaeser et al. (2008) develop a model of bubbles and endogenous housing supply. They model "irrational" bubbles based on households' over-optimism about future house prices. Their model predicts that bubbles will be more prevalent in areas with inelastic housing supply.

We have few priors about how this price diffusion will propagate across time and space, so we will impose as few assumptions as possible on its form. We include economic fundamentals to test for direct and spillover effects at both the CBSA and town level. We also include measures for school quality and crime at the town level to test if changes in local amenities can explain direct and spillover price effects. We then simulate the impact of price shocks that will result in visual graphs of the price propagation.

We make the following contributions to the literature on house price diffusion. First, we estimate diffusion models at the local level (both town and census tract) and the CBSA level – few, if any, studies estimate diffusion models at the local level much less at multiple levels of geography. Second, we estimate the ripple effect due to a house price shock in a single jurisdiction. Third, we identify persistence and spillover effects at the local level using the split sample IV estimator proposed by Case and Schiller (1989). Fourth, we use Locally Weighted Regressions (LWR) to estimate price indices at the town and census tract level where transaction data can be sparse. We also expand the kernel to smooth over time as well as space.

In addition, we answer the following questions that relate to the underlying diffusion process. Does the house price diffusion process depend on:

- The aggregation level of the data?
- The frequency of the data?
- Positive versus negative price shocks (asymmetric response)?
- Supply-side factors (ease of building)?
- Economic and demographic fundamentals?
- Location: is the spillover effect different across borders of jurisdictions such as census tract or towns?
- At the local level, does the spillover over effect arise from nearby towns as determined by distance or by similarity of local public goods?

The remainder of this paper is structured as follows. Section 2 includes a review of the relevant literature. Section 3 provides a detailed discussion of the data. The price diffusion model and the estimation procedure are developed in Section 4. Results are given in Section 5 and conclusions in Section 6.

2. **Literature Review**

Previous studies on house price diffusion can be categorized into one of three categories: studies that attempt to explain the source of diffusion; research into whether or not diffusion exists; and a combination of testing for diffusion and explaining its source(s).

Early work attempting to explain spatial house price transmission attributes the diffusion to positive attitudes about assets that lead to feedback effects, as well as a form of "rational learning". A linkage between migration and diffusion has also been found.

Clapp and Tirtiroglu (1994) shed light on the potential causes of house price diffusion, finding that positive "news" about asset prices (such as housing) can lead to positive attitudes. The positive attitudes further enhance the direct effects of the news on the asset prices. This is one of the earliest known attempts to explain how the diffusion process takes place. In a subsequent paper Clapp et al (1995) examine several Connecticut and California jurisdictions, and find evidence of spillovers across town boundaries that may be consistent with "rational learning".

Another early study examines the linkage between housing price transmission and migration across regions. Gabriel et al (1992) find that differences in house prices across regions are important explanatory variables in the determinants of inter-regional migration. Substantial variation in migration patterns may also impact housing prices.

Meen (1999) focuses on housing markets in the UK and finds that intrinsic differences across housing markets is a key contributing factor for "ripple effects". Through simulations, the author demonstrates how different patterns of growth across different markets are not as important as the "structural" differences in regional housing markets.

Contrary to the studies described above, Pollakowski and Ray (1997) focus on whether diffusion exists as opposed to attempting to explain the sources of the diffusion. They find evidence of spillover effects in housing prices across MSA's that share a common border. Similarly, Dolde and Tirtiroglu (2003) focus on testing for feedback effects. Their findings are distinct from others in that they find negative feedback effects in the short-run.

Holly et al (2011) develop a vector autoregression (VAR) model with regional and spatial shocks to subsequently simulate impulse response functions. Their goal is to assess for the presence of diffusion between London, UK, and other parts of the UK as well as between other parts of the world. An important finding is that London is a "dominant region" in terms of housing market prices.

The findings of these previous studies are important building blocks in the present analysis for at least two reasons. First, there is clear evidence of diffusion in various parts of the U.S. and the UK, which provides a basis of comparison and a set of hypotheses for us to test. Second, these previous studies generate several potential reasons why we should expect diffusion, which is helpful in motivating our research. Our focus is on the Greater Boston Area (GBA), while most of the previous studies focused on other parts of the U.S. and the rest of the world. Finally, our implementation of Locally Weighted Regresssions (LWR) to "smooth" our price estimates appears to be novel in the literature. One important benefit from using LWR is that we can generate separate smoothing estimates for each jurisdiction, which implies the potential for distinct diffusion effects across each jurisdiction. These various jurisdictions in our analysis are described in the data section.

DeFusco et al (2015) investigate the role of spillover (contagion) effects in the recent housing boom. They look at spillovers that follow a positive shock to a nearby housing market that is identified by estimating structural breaks in house price growth rates that signified the beginning of a boom period. One key to identification is that there was a lot of heterogeneity in the start of housing booms across MSAs in the U.S. To avoid bias that arises when using the same data to measure the beginning of housing booms and the resulting price changes, the authors use a split sample estimator. They find strong evidence that contagion played an important role in the recent housing boom with estimates of spillover effects elasticities in the range of 0.15-0.33. The authors find that the contagion effect is larger in smaller markets when the price shock begins in a larger market. They also find that the contagion effect is limited to markets with larger price elasticities of housing supply.

Defusco et al (2015) offer a number of mechanisms for price spillovers. First, a demand shock in one MSA could affect prices in nearby MSAs due to regional connectedness. The resulting increase in jobs and income might lead to a migration of jobs and or population that could put upward pressure on prices in nearby MSAs. Second, lenders who profit from price growth in one MSA might expand their business in nearby MSAs. This increase in credit availability could lead to an increase in house prices. Third, the same mechanism could work for housing investors. Fourth, price increases in an MSA might lead residents in nearby MSAs to think that prices will rise in their MSAs. Fifth, price growth in an MSA might lead to greater awareness of neighboring housing markets and hence enhance spillover effects.

They test for the viability of these mechanisms by regressing the following five variables on measures of when the nearest neighbor enters a housing boom; home MSA income, the percent of sales due to speculators, net migration flows into the home MSA, the fraction of new mortgage originations by subprime lenders, and the fraction of new mortgage originations insured by the FHA or VA. They find no evidence that any of these variables are significantly affected by the neighbor's boom and hence cannot explain the price spillovers. Defusco et al (2015) also include the fundamentals directly in the spillover model and find that the results do not change.

Agglomeration can be another source of house price variation across geographic areas. For instance, Van Nieuwerburgh and Weill (2010) demonstrate that productivity dispersion – which can be caused by differences in agglomeration across MSA's – can lead to dispersion of house prices. Also, Duranton and Storper (2006) indicate that the "non-market interactions in agglomeration and residential behavior" is one area of interest to both economists and geographers. However, there is a lack of empirical evidence on how concentrations of employment may be a diffusion mechanism that drives house prices nearby. It may be the case that towns with concentrations of employment (or, lower unemployment) lure potential workers towards these towns, which can lead to lower demand for housing (and lower house prices) in the other towns with lower employment densities.

Zabel (2012) shows that generally labor demand shocks can lead to relatively small house price spillover effects in nearby MSAs. But the spillover effects depend on the price elasticity of housing supply in the home MSA. In MSAs where the supply elasticity is high, out-migration is positive. This could arise as residents move out to fill the new jobs in the MSA that received the demand shock. The loss of population results in lower house prices and lower new housing supply. On the other hand, in-migration is positive in MSAs where the supply elasticity is low. This could result from residents moving in as a response to the rise in house prices due to the demand shock in the neighboring MSA. The increase in population results in higher house prices, new housing supply, and employment.

Glaeser et al (2014) find the following stylized facts in their analysis of housing dynamics: 1) positive persistence in the short-run, 2) mean reversion in the long-run, and 3) most variation in house prices is local not national. We examine the extent to which our results are consistent with these stylized facts.

3. **Data**

We look at price diffusion at three different levels; CBSA, town, and census tract. The CBSA-level data come from the quarterly Federal Housing Finance Agency (FHFA) single-family house price index for 100 CBSAs between the first quarter in 1991 and fourth quarter of 2014. This index is based on transaction only data. The index is set to 100 for all 100 CBSAs in the first quarter of 1991. It is put in real terms using the national CPI. Then we calculate the real growth rate in house prices. We calculate the distance to nearby CBSAs and use this as a basis for identifying neighbors and their impacts on prices.

We include a number of economic and demographic variables at the CBSA level over this time period. These include total employment, population, and per capita (real) income (from the BEA). On the housing side, we include the number of building permits (from the Census Bureau) and a measure of the price elasticity of housing supply from Saiz (2009). The latter is time invariant whereas the remaining variables are annual.

We also use transactions of single-family homes in the Greater Boston Area (GBA) for 1987-2012. The data are from the Warren Group for 1987-1994 and CoreLogic for 1995-2012 and cover towns in Bristol, Essex, Middlesex, Norfolk, Plymouth, and Suffolk Counties. These data include the exact date of sale (or at least the month and year) and the exact location (latitude and longitude). This will allow us to accurately account for the price diffusion process over time and space. Given this long time period, we do not have to temporally limit the price propagation. This is important since price impacts can take a long time to fully manifest themselves in the housing market. Another benefit of the data is that they cover multiple housing cycles and we can investigate if price propagation varies with the housing cycle.

Sales that were not standard market transactions such as foreclosures, bankruptcies, land court sales, and intra-family sales are excluded. Further, for each year, observations with the bottom and top 1% sales prices are excluded to further guard against non-arms-length sales and transcription errors. The data include typical house characteristics: age, living space, lot size, the number of bathrooms, bedrooms, and total rooms. The sample is limited to units with at least one bedroom and bathroom, 3 total rooms and 500 square feet of living space and no more than 10 bedrooms and 10 bathrooms, 25 total rooms, 8000 square feet of living space, or 10 acres.

The second transaction is excluded for properties that sold twice within 6 months (similar to Case/Shiller) and for properties with two sales in the same calendar year with the same transaction price (likely duplicate records). Properties for which consecutive transactions occurred in the same year or in consecutive years and where the transaction price changed (in absolute value) by more than 100% are excluded. Similarly, properties where consecutive transactions were in year t and t+j and where the transaction price changed (in absolute value) by more than j00% were excluded for j=2,…,12.

32 towns with less than 100 total observations are dropped and 36 census tracts with less than 10 observations are excluded leaving a total of 145 towns and 833 census tracts for a total of 639,859 observations. Summary statistics are given in Table A1.

When carrying out the analysis at the census tract level, we need to deal with the problem that census tracts change over time; they can split or merge. Using GIS, we have determined the largest origin tract as the consistent tract. For example, if tract A splits into B and C in 2000 we use A as the consistent tract and aggregate sales in B and C starting in 2000. This results in 641 consistent census tracts for this analysis.

We include a number of economic and demographic variables at the town level for 1990-2012. These include total employment, population, and the unemployment rate. On the housing side, we include the number of building permits (from the Census Bureau). We also include variables to capture local public goods; school quality and safety. The measures we use are test scores and crime rates.

Data from the Massachusetts Department of Education (MADOE) are available for the Massachusetts Educational Assessment Program (MEAP) (every other year from 1988 until 1996) and the Massachusetts Comprehensive Assessment System (MCAS) (every year starting n 1998). School quality is measured as the sum of district-level 4th and 8th grade math and reading/ELA exams. The average of the two surrounding years for 1989, 1991, 1993, 1995, and 1997 is used since no state-wide standardized exams were given in these years. Since scores are not comparable across years, the school quality variable is then standardized on an annual basis.

Data on property and violent crimes are obtained from the FBI's Uniform Crime Reporting Statistics. Property crimes include burglary, larceny-theft, and motor vehicle theft. Violent crimes include murder and non-negligent manslaughter, forcible rape, robbery, and aggravated assault. The first principle component of the two crime variables is used as the two measures of crime are highly correlated (correlation = 0.68). This variable is then standardized over the whole sample since units are not meaningful. One of the drawbacks of this crime measure is that 36 of the 145 towns did not report crime statistics and many of the remaining towns do not report data for all years. This is dealt with in the model by including a variable that indicates which towns do not report crime information (and by setting the crime measure to zero for the missing values).

4. **Model and Estimation Procedure**

To capture price diffusion, we specify a model that includes own lags of house price growth, lags of neighbors' house price growth, and growth rates in market fundamentals. We follow the previous literature by using excess returns as the dependent variable so this results in the following model of price diffusion:

$$G\hat{R}_{ij}^{E} = \beta_0 + \sum_{k=1}^{M} G\hat{R}_{j,t-k}^{E}\alpha_k + \sum_{k=1}^{N} f\left(G\hat{R}_{n_j,t-k}^{E}\beta_k\right) + \sum_{k=1}^{P} G\hat{R}_{j,t-k}^{F}\delta_k + \sum_{k=1}^{Q} f\left(G\hat{R}_{n_j,t-k}^{F}\gamma_k\right) + u_j + e_{jt} \quad (1)$$

where $G\hat{R}_{jt}^{E} = G\hat{R}_{jt} - G\hat{R}_{t}$ is the estimated excess return in the growth rate at time t in jurisdiction j, $G\hat{R}_{t}$ is the estimated growth rate at the aggregate level (i.e. if j is the town then this is the growth rate for the Boston market at a whole), $G\hat{R}_{n_j,t}^{E}$ is the estimated excess return in growth rate(s) in time t in jurisdictions that are neighbors to j, $f(\cdot)$ is the spillover process, and $G\hat{R}_{jt}^{F}$ is a vector of growth rates of market fundamentals. At the CBSA level, these include total employment, per capita income, population, and building permits. At the town level, these include total employment, population, the unemployment rate, building permits, test scores, and crime rates.

For the CBSA-level model, f(·) is specified as the distance weighted average of the growth rates in the five nearest CBSAs. In the case of town-level data, we consider two different means for calculating the spillover effect. One is based on the average of growth rates in adjacent towns and the other is based on substitute towns in terms of similar levels of local public goods. Note that we include both price and fundamentals spillovers. This allows for different spillover channels and we can test for different mechanisms to explain potential spillovers.

For the CBSA-level data, we use the quarterly FHFA house price index for 100 CBSAs between the first quarter in 1991 and fourth quarter of 2014. We can then take log-differences across years to obtain growth rates.

To obtain growth rates at the town and census tract level, we first regress the natural log of real house prices on house characteristics and time-by-jurisdiction fixed effects

$$\ln\left(P^{r}_{ijt}\right) = \beta_0 + X_{it}\beta_1 + u_{jt} + e_{ijt} \tag{2}$$

where P_{ijt} is the price for house i, in jurisdiction j, in year t, X_{it} is a vector of house characteristics and u_{jt} is a time by jurisdiction fixed effect. We use the results from equation (2) to obtain the estimates of these fixed effects, $\hat{u}_{jt}$. These can be viewed as jurisdiction-level prices since they are averages of all sales in a given jurisdiction and time period that control for structural characteristics. We can use these estimates to obtain growth rates.

A potential problem with this approach is that there can be very few sales in a given jurisdiction and time period and hence the prices can be measured with considerable error. We address this problem by smoothing over nearby prices using Locally Weighted Regressions, LWR (McMillen and Redfern 2010). This is also commonly referred to as Geographically Weighted Regressions, GWR (Fotheringham et al 2002). While the typical use of LWR only smooths across space, we use a two-dimensional kernel that takes the geographic distance and the time of "nearby" prices into account. The smoothing kernel is based on a generalization of the Gaussian kernel:

$$w_{kj} = \sqrt{\exp\left(-\left(\frac{\delta_{kj}}{b}\right)^2\right)} \cdot \sqrt{\exp\left(-(t_k - t_j)^2\right)} \tag{3}$$

where $w_{kj} = 1$ if $k = j$, δ_{ij} is distance between jurisdictions k and j in geographic space, and b is the bandwidth. While one approach to bandwidth selection is the Silverman (1986) rule of thumb, we choose b optimally using the approach within the routine specified in the GWR.ado file in Stata.

Given that the second exponential expression in equation (3) approaches zero rapidly when (t_k– tj) is greater than 3 (or less than -3), we restrict our attention in the time dimension smoothing to time periods that are within 3 years of the target transaction.

The use of the time dimension in the kernel is motivated by two insights. First, it adds more nearby observations that should increase the accuracy of the smooth. Second, it is likely that observations from the same jurisdiction in different time periods are closer approximations to the true price than are some of the observations in other jurisdictions. This would be particularly true when the other jurisdictions are in other towns (versus census tracts in the same town).

In the context of the diffusion model that we estimate (equation 1), the downside of using the time dimension in the kernel is that it can potentially contaminate the dynamics in the model as previous years' prices are used to estimate current prices and also appear as explanatory variables. We will estimate models with and without the time dimension in the kernel to see how this affects the results.

The nonparametric model we estimate is as follows:

$$Y_{jt} = g(Z_i) + \phi_{jt} \qquad j = 1,\ldots,N_j;\ t = 1987,\ldots,2012, \tag{4}$$

where Y_{jt} and Z_i are estimated fixed effects from equation (2) and the subscript i above spans all jurisdictions and all years in the data set. We weight each observation by the (square root of the) number of transactions in each jurisdiction in each year.

Similar to the approach of McMillen and Redfearn (2010), our marginal effects estimates, denoted as $\hat{d}_{jt}$, are obtained by the following form of weighted least squares:

$$\hat{d}_{jt} = \left(\sum_k w_{kj} Z_i Z_i^{'}\right)^{-1}\left(\sum_k w_{kj} Z_i Y_i\right) \qquad j,k = 1,\ldots,N_j;\ t= 1987,\ldots,2012, \tag{5}$$

This marginal effect $\hat{d}_{jt}$ gives an estimate of the log of the jurisdiction price in each year, which we henceforth denote as $\ln\hat{P}_{jt}$.

Next we take the difference in smoothed log prices to generate the jurisdiction price growth rate

$$G\hat{R}_{jt} = 100 \cdot \left(\ln\hat{P}_{jt} - \ln\hat{P}_{j,t-1}\right) \tag{6}$$

Case and Shiller (1989) point out that a bias arises when including the estimated lagged growth rate in equations (1). First, the growth rate that we use is estimated based on the difference in estimated prices

$$\begin{aligned}
G\hat{R}_{jt} &= 100 \cdot \left(\hat{\ln} P_{jt} - \hat{\ln} P_{j,t-1}\right) \\
&= 100 \cdot \left(\left(\ln P_{jt} - \hat{u}_{jt}\right) - \left(\ln P_{j,t-1} - \hat{u}_{j,t-1}\right)\right) \\
&= 100 \cdot \left(\left(\ln P_{jt} - \ln P_{j,t-1}\right) - \left(\hat{u}_{jt} - \hat{u}_{j,t-1}\right)\right) \\
&= GR_{jt} - 100 \cdot \left(\hat{u}_{jt} - \hat{u}_{j,t-1}\right)
\end{aligned}$$

Second

$$\begin{aligned}\mathrm{Cov}\left(\hat{GR}_{jt}, \hat{GR}_{j,t-1}\right) &= \mathrm{Cov}\left(GR_{jt} - 100\cdot\left(\hat{u}_{jt} - \hat{u}_{j,t-1}\right), GR_{j,t-1} - 100\cdot\left(\hat{u}_{j,t-1} - \hat{u}_{j,t-2}\right)\right) \\ &= \mathrm{Cov}\left(GR_{jt}, GR_{j,t-1}\right) - 100^2\cdot\mathrm{Var}\left(\hat{u}_{j,t-1}\right)\end{aligned}$$

So even if there is no correlation in growth rates over time, there is an induced negative correlation in the estimated growth rates due to the common residual.

Case and Shiller (1989) recommend a split sample IV procedure to solve this problem. This amounts to dividing the sample in two and estimating separate growth rates for each sample. Call these $\hat{GR}^1_{jt}$ and $\hat{GR}^2_{jt}$. We can the alternate $\hat{GR}^1_{jt}$ and $\hat{GR}^2_{jt}$ as lags in the price diffusion equation (1).

Note that we can use the split sample IV procedure when estimating the town-level and tract-level diffusion models but not for the CBSA-level model since we have the transaction-level data to estimate the former models but not the latter. Still, we can use the results from the town- and tract-level models to provide information about the bias that might plague the estimates for the CBSA-level diffusion model.

5. **Results**

We present results for the basic diffusion model (1) at three different levels of aggregation; CBSA, town, and census tract. We also estimate the CBSA-level model using quarterly and annual data.

5.1 CBSA Level

We have quarterly data from the FHFA for price indices for 100 CBSAs for 1991q1 to 2014q4. Hence, we do not go through the above process for generating price indices. But this also means we cannot employ the split sample IV estimator. Figure 1 displays the real FHFA national house price index for this period where prices are discounted using the national CPI. Real prices were constant until 1998 and then increased by 50% between 1998 and 2005. An almost similar decline followed between 2007 and 2011 but prices have rebounded to be about 20% above the level in 1991. Of course, this national index masks great heterogeneity in prices for individual CBSAs. This is apparent in Figure 1 which includes real FHFA price indices for Boston, Cleveland, Phoenix, and San Francisco.

Recall that we use excess returns in the growth rate as the dependent variable. We use growth rates at the nine census divisions as the base against which excess returns at the CBSA level are measured as the FHFA provides house price indices at this level (as is the case for the fundamentals that are included in the next section). This will factor out common trends at the census division level so that spillover effects are not confounded with these regional trends. And as Figure 2 shows there is significant variation in the census division real annual house price growth rates and using the census division house price growth rates as the bases versus the national house price growth rate can factor out these differential trends.

Since the price diffusion model includes jurisdiction-level fixed effects, u_j, and lags of the dependent variable, the fixed effects estimator is inconsistent. One alternative is to use the Arellano-Bond estimator. But, for three reasons, we use the fixed effects estimator. First, since we are not interested in the individual coefficient estimates but only in forecasting impacts on the growth rate in prices from price shocks, the inconsistency of individual coefficient estimators is less important. Second, the impacts when estimated using fixed effects and the Arellano-Bond estimators are generally very similar. Third, the Sargan over-ID test is rejected. This indicates that the instruments used for the Arellano-Bond estimator are not valid and hence these results are also not consistent.

5.1.1 Quarterly Frequency

We first estimate the price diffusion model using quarterly data. We don't include the market fundamentals since they are only observed annually. Later, we will estimate this model at the annual frequency and then include the market fundamentals. We estimate equation (1) with 20 lags of both the own growth rate and the neighbors' growth rate. We standardize all the growth rates so that the coefficient estimates are comparable.[1]

We use the results to simulate the effects from a 1 standard deviation increase in the real national house price growth rate. This means that the growth rates in all CBSAs receive the same shock. We use the coefficient estimates to calculate the response function (similar to the impulse response function). We refer to the impacts from the own lags as the persistence effects, those from the neighbors' lags as the spillover effects, and their sum as the total effects. The results are given in Figure 3. The persistence effects are mostly positive for the first 10 quarters and then mostly negative for the remaining 10 quarters; though the positive responses tend to be larger than the negative ones. The spillover effects are positive for the first five quarters and then are close to zero after that. The total effect shows a positive initial impact on growth rates, followed by a period of smaller negative impacts. The long-run persistence effect (the sum of the lags) is 0.465; the long-run spillover impact is 0.488, and the total long-run impact is 0.953. We bootstrapped the standard errors and the p-value for the long-run persistence effect is 0.027 and is less than 0.001 for the other two statistics. Hence all three long-term effects are positive, large, and significant. The result for the total effect is generally consistent with the stylized facts in Glaeser et al (2014); the impact is positive for the first 12 quarters then negative after that and the long-run effect is positive.

It is possible that there is an asymmetric effect. That is, a positive shock can induce building whereas a negative shock cannot result in a symmetric decline in the housing stock. Furthermore, transaction volume leads price in response to a negative shock as households are reluctant to drop their asking price for various psychological reasons (Leamer (2007) and Genesove and Mayer (2001)). So we allow for separate coefficients for positive and negative growth rates. We use a spline specification since we expect that there is a continuous response at

[1] Defusco et al (2015) also use 5 years of lags. Generally, 6 year and higher lags (more than 20 quarterly lags) are not significant.

a zero growth rate versus a discontinuous jump as there is no real change in regime when the zero growth rate threshold is crossed.

We then simulate the effects from a 1 standard deviation increase and a 1 standard deviation decrease in the real national house price growth rate. In the case of a 1 standard deviation increase, the long-run persistence effect is 0.153, the long-run spillover impact is 0.536, and the total long-run impact is 0.689. In the case of a 1 standard deviation decrease, the long-run persistence effect is -0.716; the long-run spillover impact is -0.492, and the total long-run impact is -1.208. Despite the difference in long-run persistence effects, the period-by-period patterns are fairly similar; though opposite in sign (see Figures A1 and A2). The reason that the persistence effect is small in magnitude in the case of a 1 standard deviation increase in the real national house price growth rate versus a 1 standard deviation decrease is the existence of larger negative impacts in the last few periods in the former case whereas there are not commensurate positive impacts is the latter case. Furthermore, the long-run persistence effects are not that different in a statistical sense given the p-values are 0.54 and 0.24 for the positive and negative shocks, respectively.

As in Zabel (2012) and DeFusco et al (2015), we allow for the impact of a 1 standard deviation increase in the real national house price growth rate to vary with the price elasticity of housing supply. We then evaluate this impact at the 10th and 90th percentile values of this price elasticity distribution; 0.867 and 3.466, respectively. When the value when the price elasticity is low, the long-run persistence effect is 0.483 (p-value = 0.025), the long-run spillover effect is 0.531 (p-value < 0.001) and the total long-run impact is 1.014. When the value when the price elasticity is high, the long-run persistence effect is -0.149 and not significant, the long-run spillover effect is large and significant; 0.849, and the total long-run impact is 0.700. As expected, the price effect is larger when housing is inelastically supplied though the spillover effect is larger when the price elasticity of housing supply is large.

We next consider the ripple effect. To set this up, we consider a 1 standard deviation increase in the growth rate of a single CBSA and then look at how this affects the growth rate in the nearby CBSA (ring 1), the nearby CBSA to the nearby CBSA (ring 2) and so on. The coefficient estimate for the first lag of the growth rate of the nearby CBSA is positive, large and significant (0.183), so the ripple effect is significant. Figure 4 shows the ring effects for 3 waves. Wave 1 increases prices by 0.300 standard deviation in ring 1 (year 1), 0.033 in ring 2 (year 2), 0.006 in ring 3 (year 3) and 0.001 in ring 4 (year 4). Wave 2 starts with a negative impact of 0.120 standard deviation on prices in the center CBSA in year 1, and impacts of 0.080, -0.004, and 0.008 standard deviations in ring 1 (year 2), ring 2 (year 3), and ring 3 (year 4), respectively. Finally Wave 3 starts with an impact of 0.223 standard deviation on prices in the center CBSA in year 2, a positive impact of 0.037 in ring 1 (year 3), and a positive impact of 0.022 in ring 2 (year 4). Hence the total impact on prices in the nearest CBSA is a 0.300 standard deviation increase in prices and in the next CBSA of 0.052. If there are persistent large positive shocks, the ripple effect can be substantial. This significant contagion effect can help explain the recent housing downturn that reached the national level.

5.1.2 Annual Frequency

We next estimate the price diffusion model at the CBSA level using annual data. This allows us to investigate two issues. First, we can see the impact of aggregating from the quarterly to the annual level. Second, we can add fundamentals (that are only available annually) to determine their impact on price growth.

We include 5 lags of all explanatory variables (which is the same as 20 quarters). To see the impact of aggregation on the diffusion process, we first estimate equation (1) without the market fundamental variables. The long-run persistence effect is 0.263, the spillover effect is -0.076, and the total effect is 0.1875 and only the latter effect is significant at the 5% level. All three estimates are smaller than those obtained when using the quarterly data. Thus it appears that aggregating to the annual level results in an under-estimate of the price effects compared to using quarterly data.

We then add in the economic, demographic, and housing supply variables. As we did with the growth rates in house prices, we use growth rates at the nine census divisions as the base against which excess returns for the fundamental variables at the CBSA level are calculated. Using these results, we calculate the long run price responses to a 1 standard deviation shock in the growth rate in each variable in the host CBSA and in the neighboring CBSAs which is the source of the spillover effect.

First consider the changes in fundamentals in the host CBSA. The total impact is the sum of Fundamental, Persistence and Price Spillover effects. The Fundamental effect is the direct effect on prices from the change in the growth rate in the fundamental variable, $G\hat{R}^{F}_{jt}$. The Persistence Effect arises from lagged price effects via $G\hat{R}^{E}_{j,t-k}$ and the Price Spillover Effect is due to the impact of the lagged neighbors' price variable, $G\hat{R}^{E}_{n_j,t-k}$. The latter effect arises as price changes in the host CBSA lead to changes in prices in the neighboring CBSAs which feedback to the host CBSA.

The results are given in Panel A of Table 1. Generally, the impacts are small and not significant. In fact, the only result that is significant at the 5% level is direct effect from a one standard deviation increase in the growth rate in employment; 0.195 (p-value = 0.027)**.** A one standard deviation increase in new building permits for single family units results in a long-run persistence effect of -0.158 but this is only significant at the 10% level (p-value = 0.054).

Long-run spillover effects from changes in growth rates in the fundamentals in nearby CBSAs, $G\hat{R}^{F}_{n_j,t}$, are included in the Panel B of Table 1. These Fundamental Spillover Effects are made up of the same components as the Direct Spillover Effects. The total long-run effects from a change in the growth rate in employment and population are positive and relatively large; 0.297 and 0.274, respectively. These are both dominated by the direct impacts of the changes in the fundamentals in the neighboring CBSAs. This can arise as in-migration from the neighboring CBSAs put upward pressure on prices in the host CBSA. Finally, a one standard deviation increase in new permits in the neighboring CBSAs results in a large 0.315 standard deviation decrease in the long-run growth rate in prices in the host CBSA (p-value = 0.030). This could be due to out-migration from the host CBSA as prices fall in the neighboring CBSAs as a result of the increase in building permits in those areas. We also interacted the fundamentals with the price elasticity of housing supply but found no significant differences in the impacts when evaluated at the 10^{th} and 90^{th} percentile values of this price elasticity distribution.

Thus, of the fundamentals, increases in employment growth and housing supply in the host CBSA has the largest effect on the growth rate in house prices at the CBSA level. Whereas increases in employment growth, population, and housing supply in the neighboring CBSA have even larger and significant impacts on the growth rate in house prices in the host CBSA. Furthermore, these fundamental impacts, both direct and spillover effects are comparable to the price effects when estimated using the data at the annual frequency. So is appears that both fundamentals and other mechanisms, such as price expectations, are important drivers of house price growth at this level.

5.2 Town-Level

We next estimate the price diffusion model using the town-level data from the Greater Boston Area for 1987-2012. Figure 2 includes the real price index for the Boston CBSA. It follows a similar trend as the national index; little growth between 1991 and 1997 and then a period of rapid growth between 1998 and 2005 followed by a drop in prices through 2011 and then a mild recovery where real prices are 50% higher in 2015 than they were in 1991.

First we estimate the hedonic model given in equation (2). The dependent variable is the natural log of house price. House characteristics include the natural logs of lot size and living space and their squares, indicator variables for age less than or equal to 10 years, 10 to 30 years and 30 to 50 years (greater than 50 years is the excluded category), indicators for 2, 3, 4, and 5 or more bedrooms (1 bedroom is the excluded category) , 2 or 3 or more bathrooms (1 bathroom is the excluded category), the number of half baths, and indicators for 5-9 rooms, 10-14 rooms, and 15 or more rooms (3 or 4 rooms is the excluded category).

We recover the town-by-year fixed effects from this regression, TOWN_FE. We then use LWR, weighted by the number of transactions in each town-year, to smooth these fixed effects estimates. As is typically done, we use LWR to smooth across space, LWR1. As developed in Section IV, we also use a version of LWR that smooths across time and space, LWR2. This adds more nearby observations that should increase the accuracy of the smooth. This will be particularly useful when we use the census tract level data where there are fewer sales in a given period. Also, sales in different time periods but in the same jurisdiction are likely to be closer approximations to the true price than are observations in other jurisdictions in the same time period. Smoothing across time might be problematic given that we are estimating dynamic models and hence the dynamics of the smooth might contaminate the estimates of the dynamics of the diffusion model. This is why we also use the kernel in only the spatial dimension (LWR1) as a comparison.

We first use a spillover process that is based on the average of house price growth rates in adjacent towns.[2] We exclude 3 towns from the analysis (Dunstable, Peperell, and Townsend) because, while they are contiguous to each other, they are not connected to the bulk of the towns in the dataset. This leaves 142 towns for 26 years for a total of 3,692 observations. Calculating growth rates leaves a total of 3,550 observations.

We estimate the following version of the diffusion model (1) that includes fundamentals

$$G\hat{R}^{E}_{ij} = \beta_0 + \sum_{k=1}^{M} G\hat{R}^{E}_{j,t-k}\alpha_k + \sum_{k=1}^{N} f\left(G\hat{R}^{A,E}_{j,t-k}\beta_k\right) + \sum_{k=1}^{P} G\hat{R}^{F}_{j,t-k}\delta_k + \sum_{k=1}^{Q} f\left(G\hat{R}^{F,A}_{n_j,t-k}\gamma_k\right) + u_j + e_{jt} \quad (7)$$

where $G\hat{R}^{E}_{jt}$ is either the growth rate based on LWR1 or LWR2 minus the growth rate for the Boston CBSA, $G\hat{R}^{A,E}_{j,t}$ is the average estimated excess return for adjacent towns, and $G\hat{R}^{F,A}_{j,t}$ is the average growth rate of the fundamentals for adjacent towns. Because excess returns are based on the growth rate for the Boston CBSA, the first year of data used to estimate equation (7) is 1992.

We also consider two other sources for growth rates in town house prices. First, we use the non-smoothed, town-by-year fixed effects. This allows us to evaluate the impact of using LWR. Second, we use the town-level house price index generated by the Boston Fed. It is a repeat sales index and is constructed in a similar manner as the Case/Shiller house price index. It is an annual index for 1987-2012. Smaller towns are merged to provide enough transactions to obtain a reliable index. Hence, we can compare results for two different house price indices. Though, note that we cannot apply the split sample IV estimator when using the Boston Fed index.

Table 2 lists the correlations between the different estimates of the town prices in the top panel and the correlations between the respective growth rates in the bottom panel. The town prices using the sample data (LWR1, LWR2, and TOWN_FE) are very highly correlated (≥ 0.93). This is an indication that smoothing does not make much of a difference given that there are generally a lot of transactions in each town-year. These three prices are also quite highly correlated with the Boston Fed index (all around 0.65). The correlations between the four growth rates are all still quite high (Table 2).

We estimate equation (7) with 5 own lags, 5 lags of adjacent neighbors' average house price growth rates, 5 lags of the growth rates of market fundamentals that include total employment, population, the unemployment rate, building permits, test scores, and crime rates and 5 lags of adjacent neighbors' market fundamentals. We standardize the growth rates so that the coefficient estimates are comparable. We estimate 10 models using the growth rates for LWR1, LWR2, TOWN_FE, and the two split sample IV versions of LWR1 (LWR1-1, LWR1-2), LWR2 (LWR2-1, LWR2-2), and TOWN_FE (TOWN_FE-1, TOWN_FE-2) and finally for the Boston Fed index.[3]

[2] We also weighted by the length of the border though this doesn't improve the fit of the regression

[3] The regression results for the four estimators using the full sample are available upon request.

We use these results to simulate the impacts of a 1 standard deviation increase in the GBA market growth rate. The long-run impacts from this simulation exercise are given in Table 3. Standard errors for the four estimators using the full sample are bootstrapped. Since the sample is split randomly for the two split sample IV estimators, we carry out the randomization 100 times and take the average of the resulting point estimates. Standard errors are based on these 100 point estimates.

The long-run persistence impacts for LWR1, LWR2, and TOWN_FE are clearly negatively biased and this is corrected using the split sample IV estimator. Note that there are two versions using $G\hat{R}^{1}_{jt}$ and $G\hat{R}^{2}_{jt}$ as the dependent variables in equation (7). Reassuringly, the results using these two estimators are very similar. What is clear is that once this bias is corrected using split sample IV, all of the estimates (persistence, spillover, and total) are small in magnitude and none are significantly different from zero at the 5% level. So in this case, smoothing, either over space (LWR1) or over time and space (LWR2), does not produce different results than just using the town fixed effects (TOWN_FE).

Finally, the results based on the Boston Fed index appear to be biased though not to the extent of the other three estimators that use the full sample. We cannot mitigate this bias using the split sample IV estimator since we do not have access to the individual transaction data used to estimate the Boston Fed index.

The yearly persistence, spillover, market, and total impacts are displayed in Figures 5a and 5b for the two split-sample estimators LWR1-1 and LWR2-1. The two figures are quite similar. There is little heterogeneity across years in either figure. This shows that both the yearly and long-run impacts are small.

One reason that the estimated town spillover effects are generally small and not significant might be that the towns that are likely to cause spillovers are not adjacent towns but towns that are closer substitutes in terms of the level of local public goods provided. If prices rise in one town, then potential new residents might choose to live in other towns that are similar in terms of school quality and other local public goods rather than the town "next door." To capture this avenue for town spillover effects, we generate a measure of a town's amenities, say town A, and then choose the town with the most similar level of amenities, say town B, as the town with house prices changes that are most likely to affect those in town A. The measure of town amenities we use is the town fixed effect that is estimated from the house price hedonic equation (2) where the fixed effect does not vary by year. That is, we use the full sample but only generate one fixed effect for each town. One can view this fixed effect as the dollar value of the town amenities. Then town B is the one with a fixed effect that is closest to that for Town A. The "distance" between towns A and B is the absolute value of the difference in town fixed effects. Then $G\hat{R}^{A,E}_{j,t}$ in equation (7) is the excess return in town B divided by this distance measure.[4] We re-estimate the price diffusion equation (7) using this town spillover variable. Though the town spillover effects are larger, they are still relatively small and not significant.

[4] Zabel and Dalton (2011) use a similar approach in calculating monopoly zoning power as a means for determining which towns are most able to sustain a price increase that results from stricter land use regulations.

Given that the spillover effects are small, we do not estimate the ripple effect. We also allowed the coefficient estimates to differ for positive and negative growth rates but this has little effect on the results so we do not report them here.

Next, we simulate the impacts of a 1 standard deviation increase in the fundamentals on house price growth rates. These include the impacts due to fundamentals in the same town and the spillover effects from a 1 standard deviation increase in the fundamentals in adjacent towns. The results are listed in Table 4 for the split sample IV versions of LWR1 (LWR1-1, LWR1-2), LWR2 (LWR2-1, LWR2-2), and TOWN_FE (TOWN_FE-1, TOWN_FE-2). Given that the split sample results are similar, we take the average and hence there is only one result presented for each of the three estimators. Again, standard errors are bootstrapped. Similar to the CBSA analysis, the impacts include Fundamental, Persistence, Price Spillover, and Total effects. Generally, the Persistence and Price Spillover (indirect) effects are very small and not significant. The only case where the direct Fundamental effect is significant and of a reasonable magnitude is for the unemployment rate. The direct effect of a 1 standard deviation increase in the change in the unemployment rate in the same town is negative and significant in one of the three cases (-0.23) whereas the direct effect of a 1 standard deviation increase in the change in the unemployment rate in adjacent towns is positive, significant and relatively large (approximately 0.4). The latter could result from an increase in the demand for housing in the hub town as the perceived quality of the adjacent towns diminishes (quality as embodied in neighbors; employed neighbors are of higher quality than are unemployed neighbors; and greater employment opportunities in the hub town draws away productive workers from other towns).

These impacts result from isolated changes in the hub town or the adjacent towns. A change in fundamentals at the MSA level would result in changes in all towns. The resulting impacts would then be the sum of the direct and spillover effects due to changes in the fundamentals. Generally, the direct and spillover effects counteract each other so that the overall impact of a change in fundamentals is even smaller (in magnitude). The only case where the impacts are reinforcing is crime. An increase of one standard deviation in the crime rate change in the MSA will reduce the growth rate in prices in all towns by around 0.15 standard deviations (significant at the 1% level).

5.3 Census Tract-Level

Using census tract-level data allows for spillovers within and across towns. We need to address the problem that census tracts change over time; they can split or merge at each decennial census. Using GIS, we determined the largest origin tract as the consistent tract. For example, if tract A splits into B and C in 2000 we use A as the consistent tract and aggregate sales in B and C starting in 2000. We refer to these as consistent census tracts.

We have data on single family transactions for 535 consistent census tracts in the Greater Boston Area with at least one sale in each year for 1987-2012. But we must have at least two transactions each period for the split-sample IV estimator to work. There are 492 consistent census tracts with at least two sales in each year for 1987-2012.

The version of the diffusion model that we estimate is

$$\hat{GR}^{E}_{cjt} = \beta_0 + \sum_{k=1}^{M} \hat{GR}^{E}_{cj,t-k}\alpha_k + \sum_{k=1}^{N} f\left(\hat{GR}^{A\text{-}TR,E}_{cj,t-k}\beta_k\right) + \sum_{k=1}^{P} f\left(\hat{GR}^{TO,E}_{j,t-k}\gamma_k\right) + u_j + e_{cjt} \qquad (8)$$

where $\hat{GR}^{E}_{cjt}$ is the estimated excess return in census tract c, town j and time t (the growth rate minus the growth rate for the Boston CBSA), $\hat{GR}^{A\text{-}TR,E}_{cjt}$ is the average estimated excess return for other census tracts in the same town, and $\hat{GR}^{TO,E}_{cjt}$ is the estimated excess return for the town with the most similar amenities.[5] Again, because we use the growth rate for the Boston CBSA to calculate excess returns, the first year of data used to estimate equation (8) is 1992.[6]

We estimate equation (8) with 5 own lags and 5 lags of the mean growth rate of the other census tracts in the same town and of the house price growth rate of the nearby town. Growth rates are standardized so that the coefficient estimates are comparable. We provide the estimation results for LWR1, LWR2, and TRACT_FE and for the corresponding two split sample IV estimators.

We use the regression results to simulate the impacts of a 1 standard deviation increase in the market growth rate. The long-run impacts from this simulation exercise are given in Table 4. Standard errors for the three estimators using the full sample are bootstrapped. Since the sample is split randomly for the two split sample IV estimators, we carry out this randomization 50 times and take the average of the resulting point estimates. Standard errors are based on these 50 point estimates.

Interestingly, the biased (full sample) estimators give very different results; the total effect based on LWR1 is -0.051 and not significant, that based on LWR2 is 0.600 and significant, and that based on TOWN_FE is -0.493 and significant. The results using the split sample IV estimators are much more similar. The total effects for the split sample IV estimators based on LWR1 and LWR2 are all approximately 0.35. Despite the similarity in the total effects estimates, differences do exist between LRW1 and LRW2 in terms of the persistence and spillover effects. First, the persistence effect estimates using LWR1 are positive but small and insignificant whereas the ones using LWR2 are negative and significant at the 5% level. The estimates of the tract-level spillover effects are all positive and significant though the ones using LRW2 are larger. The estimates of the town-level spillover effects are very similar; positive and small in magnitude though three of the four are significant at the 10% level or better.

The biggest driver of the total effects are the positive, (relatively) large, and significant tract-level spillover effects; the range is 0.22 – 0.39. Note that we have determined the nearby town to measure spillover effects as the closest substitute in terms of the level of local public goods. Other census tracts in the same town can be viewed as perfect substitutes as they have the same amenities. Hence the significant tract-level spillover effects make sense in this context. This is

[5] In the town-level analysis, we proposed two approaches for determining nearby towns that were used to estimate spillover effects. One was based on adjacent towns and one was based on the town with the most similar amenities. We report the results for the latter as they provide larger estimates of the town-level spillover effects.
[6] We don't include fundamentals in the tract-level model since they are measured at the town level and hence their impacts would be the same as in the town-level model.

also consistent with other literature on "neighborhood effects" such as Rosenthal (2008). Specifically, he describes how similarities in homeowner educational attainment and/or economic status can attract others who are looking for similar "attractive" nearby neighborhoods to live in, which can be a source of spillovers across very local areas.

The split sample IV estimates of the total effects using the town fixed effects are a third the size of those based on the smoothed estimators. This contrasts to the results using the town-level data where the results are similar. This is likely due to the small number of annual sales per census tract that are used to estimate the census tract fixed effects, the annual mean is 39 sales, compared to the larger number used to estimate the town fixed effects; the mean of annual sales is 170. So it appears that the smoothed estimator is particularly valuable when estimating price indices and growth rates for small areas such as census tracts.

Note that, like the town-level results, the persistence and town spillover effects are generally small in magnitude. The big difference from using the tract-level data is the ability to measure the tract-level spillovers. It appears that aggregating the data to the town level misses important dynamics of growth rates that occurs within towns at the census tract level.

One mechanism by which local price shocks can spread is through the process by which appraisers value houses. Similarly, information about nearby house prices can disseminate quickly with the increasing popularity of Zillow.com ®, Trulia.com ®, etc. That is, appraisers, real estate agents, and these websites use transactions from nearby units or "comps" as a means for valuing units. In this case, the comps are probably confined to sales in the same town. This would help explain the significant tract-level spillovers and small town-level spillovers. Furthermore, real estate agents may well follow a similar procedure when valuing units and/or when advising clients about list prices. So, if one part (tract) of a town shows significant price increases, agents probably advise clients to purchase houses in other parts of the town. This would contribute to the tract-level spillovers.

The yearly persistence, spillover, and total impacts are displayed in Figures 6a and 6b for split-sample estimators LWR1-1 and LWR2-1. The long-run impacts, which are fairly similar, mask some heterogeneity in the yearly values. For example, the first-year persistence effect estimated using LWR2-1 is relatively large and positive (0.235) and significant whereas the second-year persistence effect is similar in magnitude but negative (-0.201) and significant. Whereas the persistence effects estimated using LWR1-1 are less than half these impacts (in magnitude). The first-year and second-year tract-level spillover effects estimated using LWR2-1 are positive and significant (0.119 and 0.156) while only the first-year effect is significant when using LWR1-1 (0.164).

Overall, it is not clear that smoothing across space and time (LWR2) is an advantage over just smoothing across space only. While there are some differences in the persistence and tact-level spillover effects, the long-run total effects are almost identical. More examples are needed to determine if adding the time dimension to the smoothing kernel is advantageous.

6. **Conclusion**

This paper investigates the movement of house price growth (shocks) over time and space at three different levels of aggregation: CBSA, town, and census tract. At each level of aggregation we estimate fixed effects models for growth rates on lagged growth rates in prices and fundamentals and lagged growth rates in price and fundamentals for nearby jurisdictions. We use the results to simulate impacts of shocks to growth rates to calculate short-run and long-run persistence and spillover effects.

For the CBSA model using quarterly data, the estimated long-run persistence and spillover effects are positive, large, and significant; 0.465 and 0.488 standard deviations, respectively. The result for the persistence effect is generally consistent with the stylized facts in Glaeser et al (2014); the impact is positive for the first 12 quarters then is mostly negative after that and the long-run effect is positive. We also find that the long-run impact is larger when the price elasticity of housing supply is low (10^{th} percentile = 0.863) versus high (90^{th} percentile = 3.466).

We also look at the ripple effect that results from a 1 standard deviation increase in the growth rate of a single CBSA. The estimated total impact on prices in the nearest CBSA is a 0.300 standard deviation increase in prices and in the next CBSA it is 0.052. If there are persistent large shocks, the ripple effect can be substantial. This significant contagion effect can help explain the recent housing downturn that reached the national level.

We estimate the CBSA model using annual data so that we can add fundamentals (that are only available annually) to determine their impact on price growth. First, we find that the price effects are smaller compared to the estimates using the quarterly data. Second, we find that increases in employment growth and housing supply in the host CBSA have the largest effect on the growth rate in house prices. Whereas increases in employment growth, population, and housing supply in the neighboring CBSA have even larger and significant impacts on the growth rate in house prices in the host CBSA. Furthermore, these fundamental impacts, both direct and spillover effects are comparable to the price effects when estimated using the data at the annual frequency. So is appears that both fundamentals and other mechanisms, such as price expectations, are important drivers of house price growth at this level. It is likely that irrational price expectations fuel price effects across CBSAs (and not as much with CBSAs). This gives credence models of “irrational” bubbles based on households’ over-optimism about future house prices (Glaeser et al. 2008).

To estimate the diffusion process at a more local level, we use annual transaction data for the Greater Boston Area. We estimate the diffusion model at both the town and census tract level to determine the impact of aggregation on the diffusion process. The diffusion model results is biased estimates because there is an induced negative correlation in the estimated growth rate and its lag due to a common residual. The solution proposed by Case and Schiller (1989) is a split sample IV estimator. We use this estimator and do find that the usual fixed effects estimator results in negatively biased estimates of the persistence effects (the coefficient estimates for the lagged growth rates). We find little evidence of persistence or spillover effects at the town level whereas there is evidence of significant tract-level spillovers. Hence, aggregating the data to the town level can miss interesting dynamics that occur with towns.

In terms of fundamentals, only the changes in the town unemployment rate has a statistically and economically significant impact on price growth. These total impacts are positive, implying that house price growth falls as unemployment rates in nearby towns fall. This may support the notion that agglomeration due to improved matching opportunities could be driving nearby house price growth. In other words, agglomeration economies in some of the larger towns may be drawing the more productive residents from other towns to work and live in those larger towns, which may be dampening house prices in the more distant towns. When

When there is a change in fundamentals at the MSA level, the direct and spillover effects counteract each other so that the overall impact of a change in fundamentals is even smaller (in magnitude). The only case where the impacts are reinforcing is crime. The impact a decrease in the growth rate of prices by around 0.15 standard deviations (significant at the 1% level).

In the Introduction, we said that we would answer the following questions that relate to the heterogeneity of the diffusion process. Does the house price diffusion process depend on:

1. The aggregation level of the data?
2. The frequency of the data?
3. Positive versus negative price shocks (asymmetric response)?
4. Supply-side factors (ease of building)?
5. Economic and demographic fundamentals?
6. Location; is the spillover effect different across borders of jurisdictions such as census tract or towns?
7. At the local level, does the spillover effect arise from nearby towns as determined by distance or by similarity of local public goods?

The answers are:

1. We find evidence that the aggregation level of the data can affect the estimates of the persistence and spillover effects; the results at the town level can obscure spillover effects at the census tract level.
2. Our annual estimates of the price effects using the CBSA data are much smaller than when using the quarterly estimates. This is evidence that estimating the diffusion model at different frequencies can produce different results.
3. We find some evidence of asymmetric responses in the long-run persistence effects although the period-by-period patterns are fairly similar until the last few periods; though, of course, opposite in sign. This difference can result from the fact that there can be a positive supply-side response to the positive price shock but not a commensurate negative supply-side response to the negative price shock.
4. We find that long-run price effects are larger in CBSAs with low price elasticities of housing supply compared to CBSAs with high pries elasticities of housing supply.
5. Of the economic, and demographic fundamentals, we find that changes in housing supply and employment and population growth affect house price growth rates at the CBSA level, whereas changes in the unemployment rate and crime affect house price growth rates at the town level. Perhaps agglomeration economies due to labor market matching are underlying these unemployment drivers of diffusion.

6. We find evidence of relatively large and significant within-town spillovers at the census tract-level but little evidence of cross-town spillovers.
7. Using growth rates in nearby towns based on the similarity of local public goods produce somewhat larger but still insignificant town-level spillover effects compared to those based on adjacent towns when using the tract-level data.

The persistence and spillover effects that we estimate at the CBSA level are much larger than those estimated at the local level. Why are these estimates of the diffusion process so different across aggregation levels? One reason is that, because we do not have the transaction-level data that are the basis of the CBSA-level price indices, we cannot apply the split sample IV estimator. But it is noteworthy that the standard fixed effects estimator of the persistence effects will be negatively biased so that, if anything, the true CBSA-level persistence effects are larger than our estimates. The bias in the spillover effects when using the fixed effects estimator is less clear. The results from the town- and tract-level analyses show that there might be a positive bias though it does not appear to be as large in magnitude as the bias in the persistence estimates. It appears that it is not just the bias in the estimators that can explain the differences in the persistence and spillover effects at the CBSA and local levels

The evidence indicates that price diffusion operates across CBSAs and not as much across towns within CBSAs. If we think of CBSAs as housing markets, we find evidence of price diffusion across markets versus within markets. That is, while there is clearly variation in price growth within CBSAs, this doesn't significantly affect growth rates in nearby towns. But there is some evidence of spillovers across tracts within towns. One can view that as information spillovers operating within towns (maybe via real estate agents and/or information available on the internet) or by the way appraisers value houses in their towns.

Given that the Greater Boston Area housing market is not representative of the all housing markets in the U.S., this analysis should be replicated in other housing markets to determine if the within-market diffusion process estimated here is similar in housing markets. Clearly more research is needed to explain the differences in across-market and within-market persistence and spillover effects.

References

Barr, Jason and Jeffrey P. Cohen. 2014. "The floor area ratio gradient: New York City, 1890–2009," **Regional Science and Urban Economics** 48: 110-119.

Case, Karl E. and Robert J. Shiller. 1989. "The Efficiency of the Market for Single-Family Homes," **American Economic Review** 79(1): 125-137

Clapp, John, Walter Dolde, and Dogan Tirtiroglu. 1995. "Imperfect Information and Investor Inferences from Housing Price Dynamics," **Real Estate Economic**s 23(3): 239–269.

Clapp, John and Dogan Tirtiroglu. 1994. "Positive Feedback Trading and Diffusion of Asset Price Changes: Evidence of Housing Transactions," **Journal of Economic Behavior and Organization** 24: 337–355.

DeFusco, Anthony, Wenjie Ding, Fernando Ferreira, and Joseph Gyourko. 2015. "The Role of Contagion in the American Housing Boom" working paper.

Dolde, Walter and Dogan Tirtiroglu. 2003. "Temporal and Spatial Information Diffusion in Real Estate Price Changes and Variances," **Real Estate Economic**s 25(4): 539–565.

Duranton, G. and Storper, M., 2006. Agglomeration and growth: a dialogue between economists and geographers. *Journal of Economic Geography*,*6*(1): 1-7.

Duranton, G. and Puga, D., 2004. Micro-foundations of urban agglomeration economies. *Handbook of regional and urban economics*, *4*: 2063-2117.

Ferreira, Fernando and Joseph Gyourko. 2011. "Anatomy of the Beginning of the Housing Boom: U.S. Neighborhoods and Metropolitan Areas, 1993-2009, unpublished mimeo, The Wharton School, University of Pennsylvania

Ferreira, Fernando and Joseph Gyourko. 2012. "Heterogeneity in Neighborhood-Level Price Growth in the United States, 1993–2009," **American Economic Review** 102(3): 134–140

Fotheringham AS, C Brunsdon, and ME Charlton. 2002. **Geographically Weighted Regression: The Analysis of Spatially Varying Relationships**. John Wiley, Chichester, UK.

Gabriel, Stuart A., Janice Shack-Marquez and William L. Wascher. 1992. "Regional House-Price Dispersion and Interregional Migration," **Journal of Housing Economics** 2: 235-256.

Genesove, David, and Christopher J. Mayer. 2001. "Loss Aversion and Seller Behavior: Evidence from the Housing Market." **Quarterly Journal of Economics** 116(4): 1233–60.

Glaeser, Edward L., Joeseph Gyourko, Eduardo Morales and Charles G. Nathanson. 2014. "Housing Dynamics: An Urban Approach," Journal of Urban Economics 81: 45-56.

Glaeser, Edward L., Joeseph Gyourko, and Albert Saiz. 2008. “Housing supply and housing bubbles,” Journal of Urban Economics 64: 198-217.

Härdle, W. K., Müller, M., Sperlich, S., & Werwatz, A. 2012. **Nonparametric and Semiparametric Models.** Springer Science & Business Media.

Holly, Sean, M. Hashem Pesaran and Takashi Yamagata. 2011. “The spatial and temporal diffusion of house prices in the UK,” Journal of Urban Economics 69: 2-23

Leamer, Edward, 2007, “Housing IS the Business Cycle”, NBER Working Paper #13248.

McMillen, Daniel P. and Christian L. Redfearn. 2010. “Estimation and Hypothesis Testing for Nonparametric Hedonic House Price Functions,” Journal of Regional Science 50(3), 712–733.

Meen, Geoffrey. 1999. “Regional House Prices and the Ripple Effect: A New Interpretation,” **Housing Studies** 14:6: 733-753.

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: 107-124.

Rosenthal, Stuart S. 2008. “Old Homes, Externalities, and Poor Neighborhoods. A Model of Urban Decline and Renewal,” **Journal of Urban Economics** 63: 816-840.

Silverman, B.W. 1986. **Density Estimation for Statistics and Data Analysis**. London: Chapman & Hall/CRC.

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

Zabel, Jeffrey. 2012. “The Role of the Housing Market in the Migration Response to Employment Shocks,” **Journal of Urban Economics** 72: 267–284.

Zabel, Jeffrey and Maurice Dalton. 2011. “The Impact of Minimum Lot Size Regulations on House Prices in Eastern Massachusetts,” **Regional Science and Urban Economics,** 41(2011): 471-483.

Table 1 – Long Run Impacts from a 1 SD Increase in the Fundamentals: CBSA-Level Diffusion Model

Impact	Income	Employment	Population	Permits
Panel A: Direct Effect				
Fundamental	0.091	0.195**	0.049	-0.015
	(0.154)	(0.088)	(0.092)	(0.116)
Persistence	0.014	0.024	-0.003	-0.158*
	(0.067)	(0.038)	(0.037)	(0.082)
Price Spillover	0.020	-0.020	-0.020	0.022
	(0.029)	(0.033)	(0.023)	(0.035)
Total	0.125	0.199	0.026	-0.151
	(0.227)	(0.129)	(0.126)	(0.169)
Panel B: Spillover Effect				
Fundamental	-0.154	0.266***	0.213***	-0.143
	(0.156)	(0.058)	(0.077)	(0.104)
Persistence	-0.016	0.055	0.067**	-0.177**
	(0.067)	(0.042)	(0.032)	(0.080)
Price Spillover	-0.010	-0.024	-0.005	0.005
	(0.025)	(0.021)	(0.017)	(0.042)
Total	-0.180	0.297***	0.274***	-0.315**
	(0.225)	(0.084)	(0.097)	(0.145)

Note: bootstrapped standard errors are in parentheses
*, **, *** - significant at 10%, 5%, 1% significance level

Table 2
Correlations for Town Prices and Growth Rates

Correlations for Town Prices			
	LWR1	LWR2	Town_FE
LWR2	0.96		
Town_FE	0.93	0.98	
Boston_Fed	0.67	0.64	0.63
Correlations for Town Growth Rates			
	DLWR1	DLWR2	DTown_FE
DLWR2	0.80		
DTown_FE	0.70	0.72	
DBoston_Fed	0.65	0.67	0.69

Table 3 – Long Run Impacts from a 1 SD Increase in the Market Growth Rate: Town-Level Diffusion Model

Model	Persistence	Spillover	Total
LWR1	-1.131	0.610	-0.522
	(0.000)	(0.000)	(0.000)
LWR1 – 1	-0.014	0.044	0.030
	(0.166)	(0.112)	(0.125)
LWR1 – 2	0.015	0.040	0.055
	(0.138)	(0.094)	(0.108)
LWR2	-0.749	0.351	-0.398
	(0.000)	(0.000)	(0.000)
LWR2 – 1	-0.021	0.031	0.010
	(0.113)	(0.086)	(0.142)
LWR2 – 2	-0.045	0.048	0.003
	(0.116)	(0.061)	(0.125)
TOWNFE	-0.642	0.340	-0.302
	(0.000)	(0.000)	(0.000)
TOWNFE - 1	0.030	0.081	0.111
	(0.085)	(0.038)	(0.074)
TOWNFE - 2	0.035	0.064	0.099
	(0.089)	(0.048)	(0.079)
Boston Fed	-0.263	0.140	-0.124
	(0.008)	(0.060)	(0.069)

p-values in parentheses (calculated using bootstrapped standard errors for LWR1, LWR2, TOWNFE, and Boston Fed, calculated based on 100 randomly drawn split samples for LWR1-1, LWR1-2, LWR2-1, LWR2-2, TOWNFE-1, and TOWNFE-2)

Table 4 - Long Run Impacts from a 1 SD Increase in Fundamentals: Town-Level Diffusion Model

	Fundamental	Persistence	Price Spillover	Total
Estimator	Population Growth Rate			
LWR1	-0.001	-0.001	-0.004	-0.006
LWR2	0.049	0.000	-0.002	0.048
TOWNFE	0.051	0.002	-0.013*	0.040
	Population Growth Rate Spillover			
LWR1	-0.077	-0.004	-0.001	-0.082
LWR2	0.008	0.002	0.002	0.012
TOWNFE	-0.126***	-0.006	-0.001	-0.132*
	Employment Growth Rate			
LWR1	0.002	0.000	0.001	0.003
LWR2	0.045	-0.002	0.004	0.047
TOWNFE	0.027	0.001	0.001	0.029
	Employment Growth Rate Spillover			
LWR1	0.044	0.001	0.000	0.045
LWR2	0.081**	-0.002	0.003	0.082**
TOWNFE	0.014	0.001	0.002	0.017
	Single Family Permits			
LWR1	0.016	0.000	0.001	0.016
LWR2	-0.001	-0.004	0.005	0.000
TOWNFE	-0.021	-0.001	-0.001	-0.022
	Single Family Permits Spillover			
LWR1	0.049	-0.001	0.002	0.049
LWR2	0.067**	-0.008	0.001	0.060*
TOWNFE	0.049**	0.000	0.000	0.049**
	Unemployment Rate Change			
LWR1	-0.132	-0.006	0.005	-0.133
LWR2	-0.292	0.000	0.027	-0.265
TOWNFE	-0.230**	-0.010	0.024**	-0.217**
	Unemployment Rate Change Spillover			
LWR1	0.240*	0.007	-0.003	0.244*
LWR2	0.418**	0.001	-0.016	0.403**
TOWNFE	0.342***	0.014	-0.016*	0.340*

Table 4 - Continued

	Fundamental	Persistence	Price Spillover	Total
		Test Score Change		
LWR1	0.038	0.001	0.002	0.041
LWR2	0.021	0.000	-0.001	0.020
TOWNFE	0.017	0.000	0.001	0.019
		Test Score Change Spillover		
LWR1	0.036	0.001	0.002	0.039
LWR2	-0.010	0.000	0.002	-0.008
TOWNFE	0.012	0.000	0.003	0.015
		Crime Change		
LWR1	-0.043	-0.001	-0.004	-0.048
LWR2	-0.045	0.000	-0.004	-0.049
TOWNFE	-0.086**	-0.004	-0.003	-0.092*
		Crime Change Spillover		
LWR1	-0.121***	-0.004	-0.002	-0.127*
LWR2	-0.060	0.000	-0.004	-0.064
TOWNFE	-0.065**	-0.003	-0.003	-0.071**

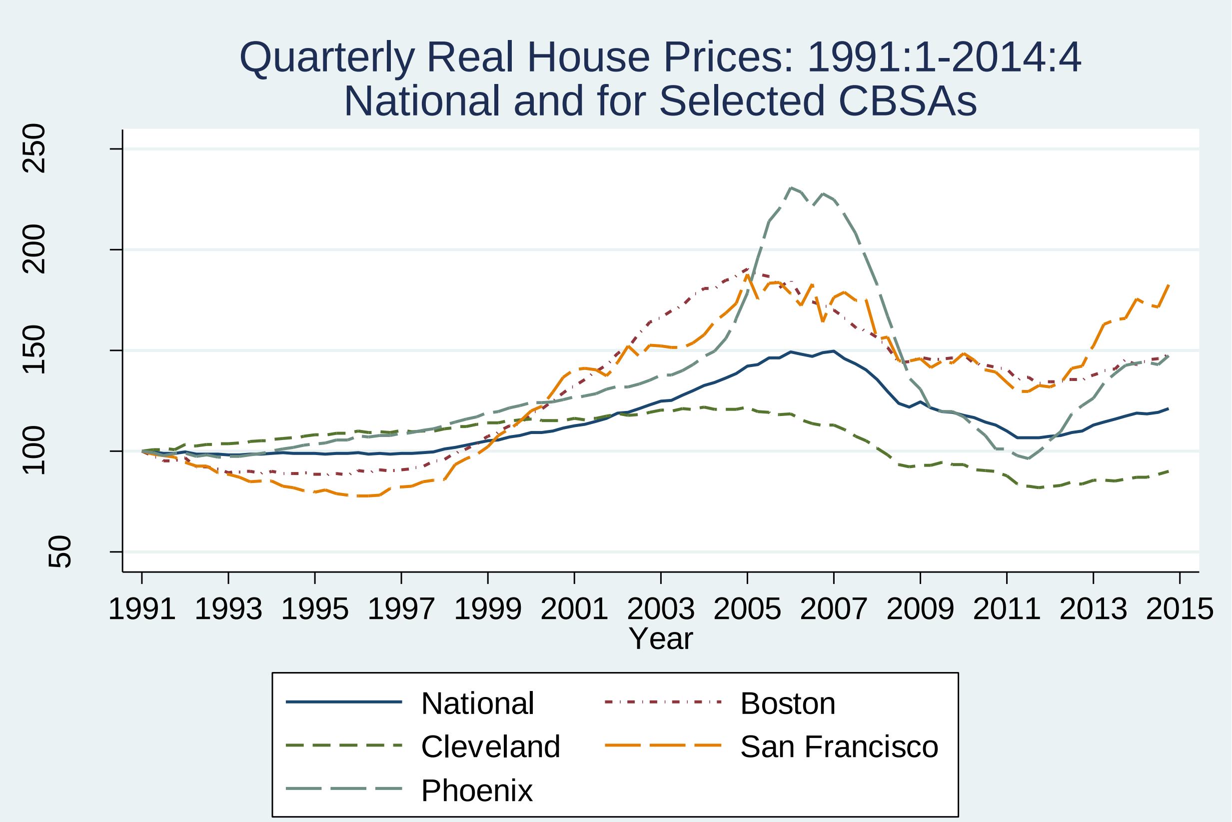
Notes: standard errors are bootstrapped,
*,**,***; p-value < 0.1, <0.05, <0.01, respectively

Table 5 – Long Run Impacts from a 1 SD Increase in the Market Growth Rate: Tract-Level Diffusion Model

Model	Persistence	Spillover-Tract	Spillover-Town	Total
LWR1	-0.868	0.756	0.061	-0.051
	(0.000)	(0.000)	(0.001)	(0.168)
LWR1 – 1	0.047	0.220	0.040	0.307
	(0.597)	(0.096)	(0.293)	(0.009)
LWR1 – 2	0.040	0.274	0.054	0.368
	(0.653)	(0.016)	(0.072)	(0.000)
LWR2	-0.027	0.489	0.139	0.600
	(0.477)	(0.000)	(0.000)	(0.000)
LWR2 – 1	-0.097	0.366	0.061	0.330
	(0.039)	(0.000)	(0.002)	(0.000)
LWR2 – 2	-0.107	0.391	0.065	0.349
	(0.048)	(0.000)	(0.002)	(0.000)
TOWNFE	-1.122	0.614	0.014	-0.493
	(0.000)	(0.000)	(0.162)	(0.000)
TOWNFE - 1	-0.074	0.170	0.013	0.109
	(0.217)	(0.000)	(0.386)	(0.047)
TOWNFE - 2	-0.087	0.202	0.017	0.131
	(0.107)	(0.000)	(0.157)	(0.022)

p-values in parentheses (calculated using bootstrapped standard errors for LWR1, LWR2, TOWNFE, and Boston Fed, calculated based on 100 randomly drawn split samples for LWR1-1, LWR1-2, LWR2-1, LWR2-2, TOWNFE-1, and TOWNFE-2)

Figure 1



Sources: FHFA

Figure 2

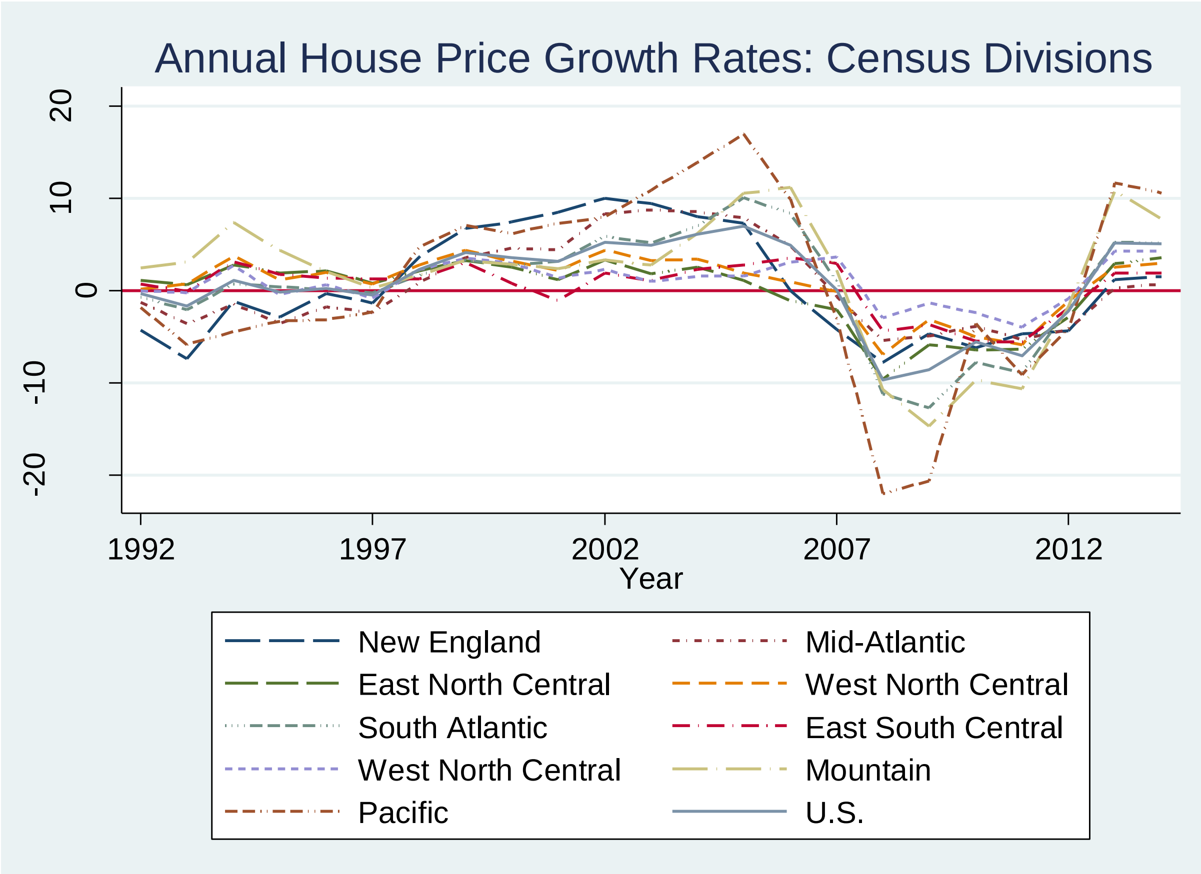
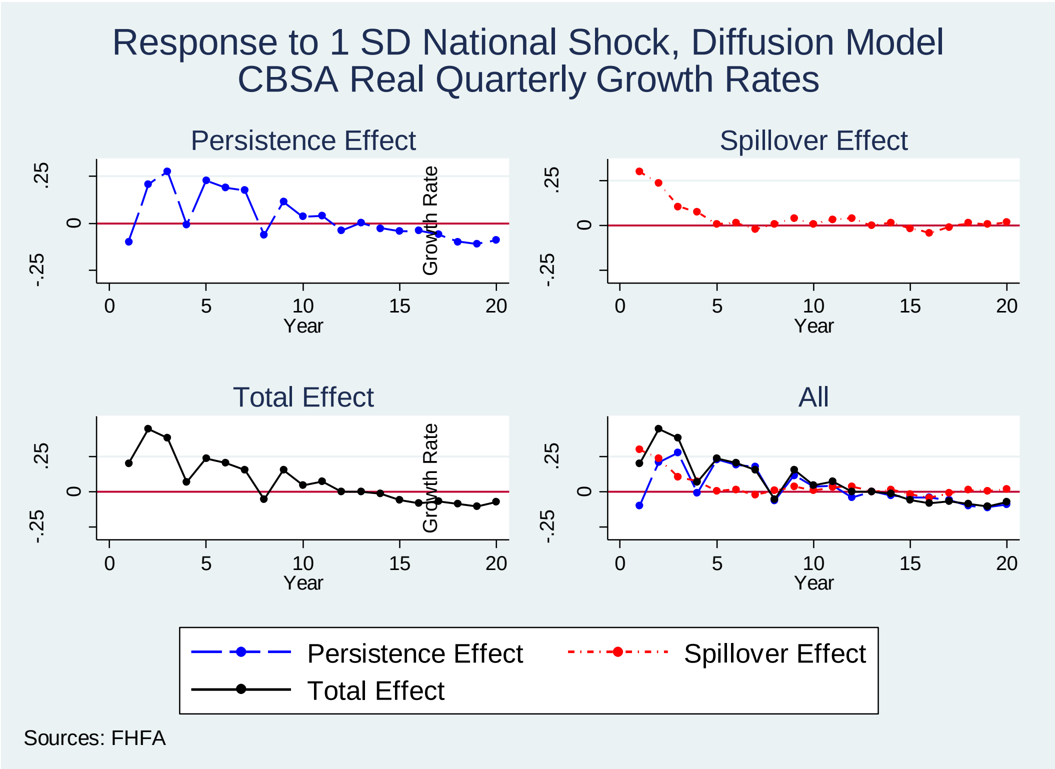
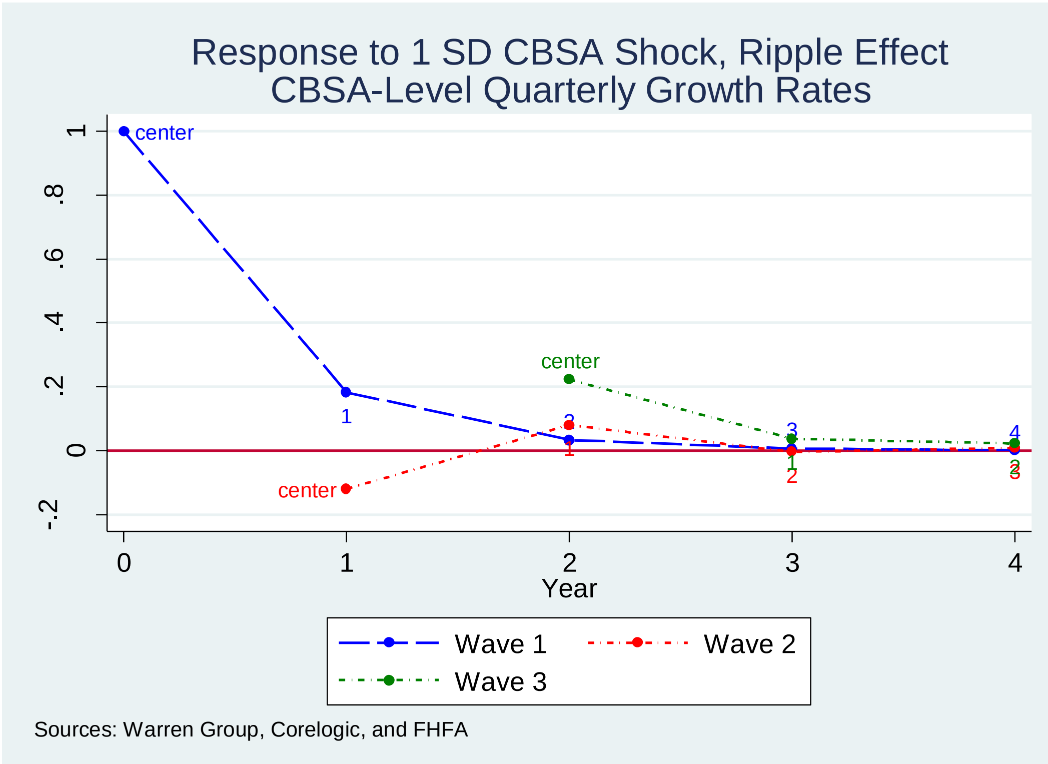

Annual House Price Growth Rates: Census Divisions
20
10
0
-10
-20
1992
1997
2002
2007
2012
Year
New England
Mid-Atlantic
East North Central
West North Central
South Atlantic
East South Central
West North Central
Mountain
Pacific
U.S.

Figure 3



Response to 1 SD National Shock, Diffusion Model
CBSA Real Quarterly Growth Rates
Persistence Effect
Spillover Effect
Total Effect
All
Growth Rate
Year
.25
0
-.25
0
5
10
15
20
Persistence Effect
Spillover Effect
Total Effect
Sources: FHFA

Figure 4



Response to 1 SD CBSA Shock, Ripple Effect
CBSA-Level Quarterly Growth Rates
-.2
0
.2
.4
.6
.8
1
0
1
2
3
4
Year
center
center
center
Wave 1
Wave 2
Wave 3
Sources: Warren Group, Corelogic, and FHFA

Figure 5a

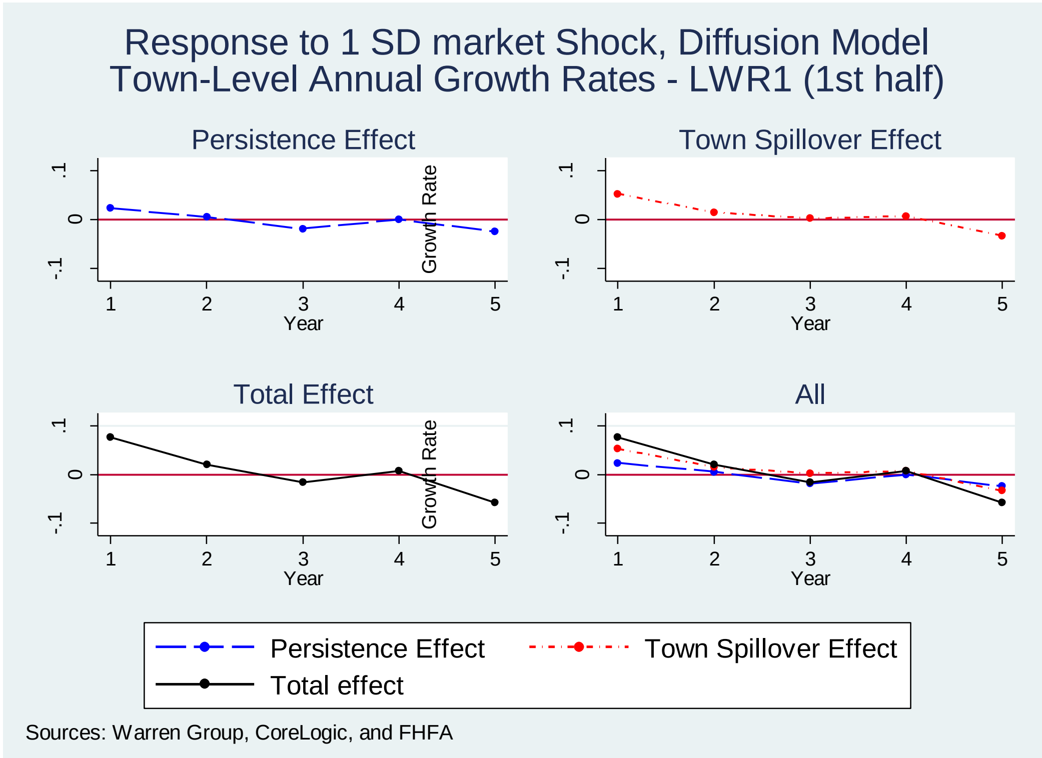


Figure 5b

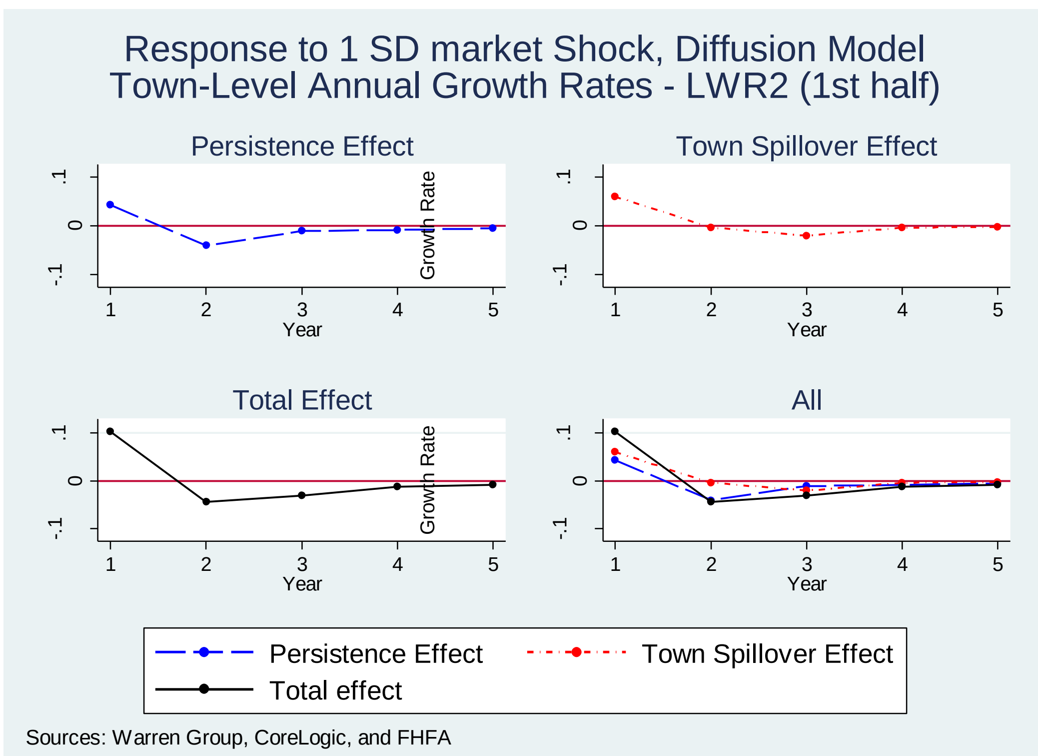


Figure 6a

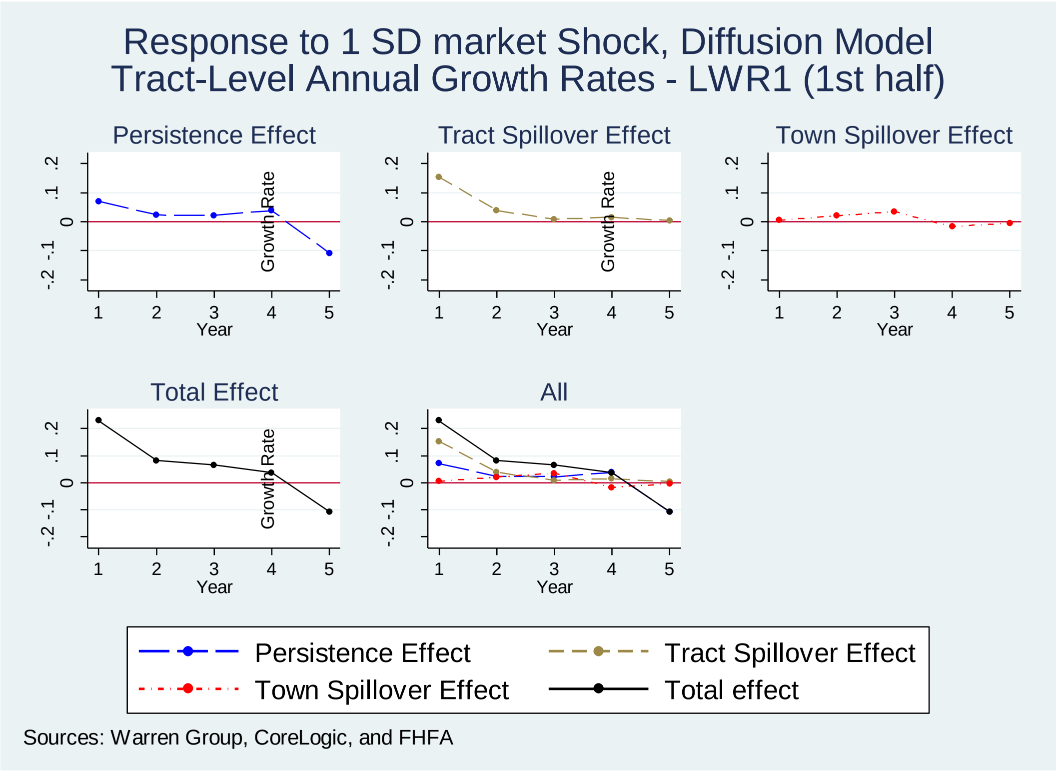


Figure 6b

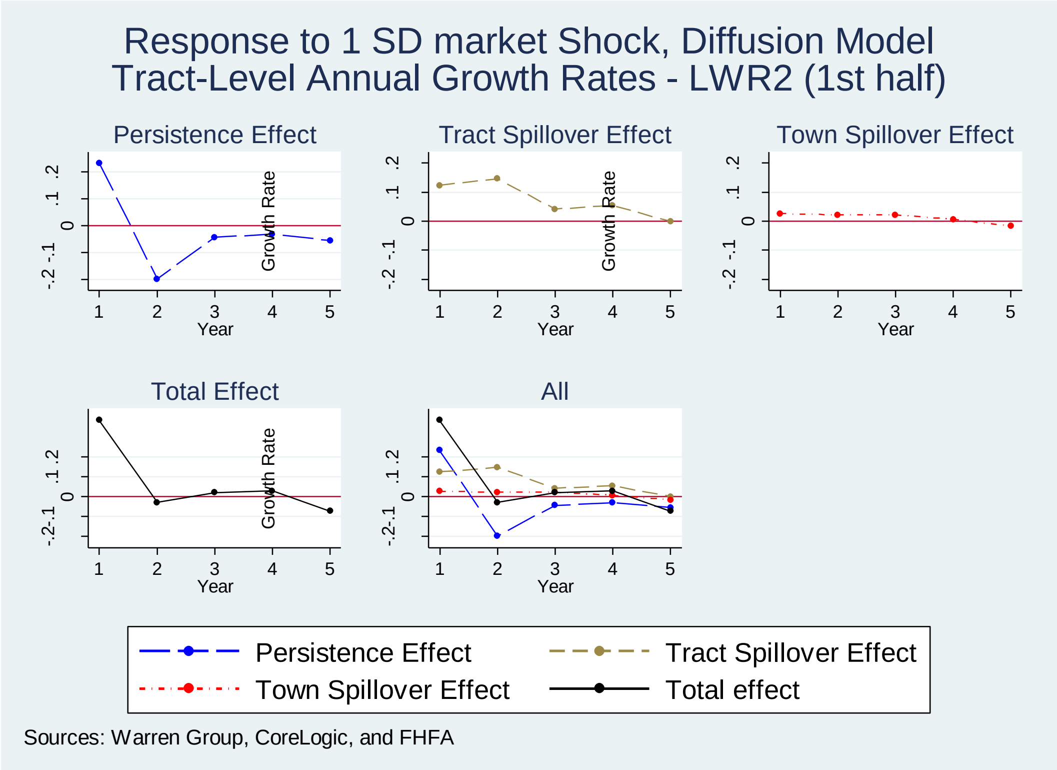


Table A1 - Summary Statistics for Hedonic Variables

Variable	Mean	Std.Dev.	Minimum	Maximum
Real Sale Price ($2012 in thousands)	406.95	238.04	27.34	2097.04
House Age	43.96	35.12	0.00	200.00
Number of Bedrooms	3.31	0.82	1.00	10.00
Number of Bathrooms	1.71	0.74	1.00	10.00
Number of Half Baths	0.61	0.54	0.00	8.00
Total Number of Rooms	7.01	1.61	3.00	23.00
Living Area (100s sf)	19.65	8.54	5.00	79.99
Lot Size (acres)	0.62	0.78	0.10	10.00
Number of Observations	639,859			

Table A2: Town-level Diffusion Equation Results

	Dependent Variable (Growth Rate)			
Variables	LWR1	LWR2	Town FE	Boston Fed
Own Lags	(1)	(2)	(3)	(4)
L1	-0.491	-0.367	-0.281	-0.031
	(0.042)**	(0.042)**	(0.035)**	(0.040)
L2	-0.197	-0.121	-0.072	0.018
	(0.046)**	(0.040)**	(0.033)*	(0.034)
L3	-0.073	-0.078	-0.022	0.099
	(0.035)*	(0.029)**	(0.025)	(0.032)**
L4	-0.115	-0.057	-0.005	-0.068
	(0.063)	(0.028)*	(0.020)	(0.027)*
L5	-0.090	-0.058	-0.055	-0.167
	(0.024)**	(0.025)*	(0.026)*	(0.023)**
Adjacent Lags				
L1	0.011	0.006	-0.003	0.017
	(0.011)	(0.013)	(0.015)	(0.029)
L2	0.014	0.026	0.045	0.040
	(0.027)	(0.019)	(0.016)**	(0.021)
L3	0.023	0.004	0.039	0.052
	(0.014)	(0.010)	(0.010)**	(0.016)**
L4	0.009	0.019	0.010	-0.075
	(0.026)	(0.006)**	(0.008)	(0.025)**
L5	0.009	-0.005	-0.033	-0.007
	(0.018)	(0.008)	(0.007)**	(0.014)
Constant	0.044	0.042	0.023	0.007
	(0.003)**	(0.002)**	(0.002)**	(0.004)
R-squared	0.20	0.12	0.08	0.04
Observations	2,272	2,272	2,272	2,272
Number of dorcodes	142	142	142	142

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure A1

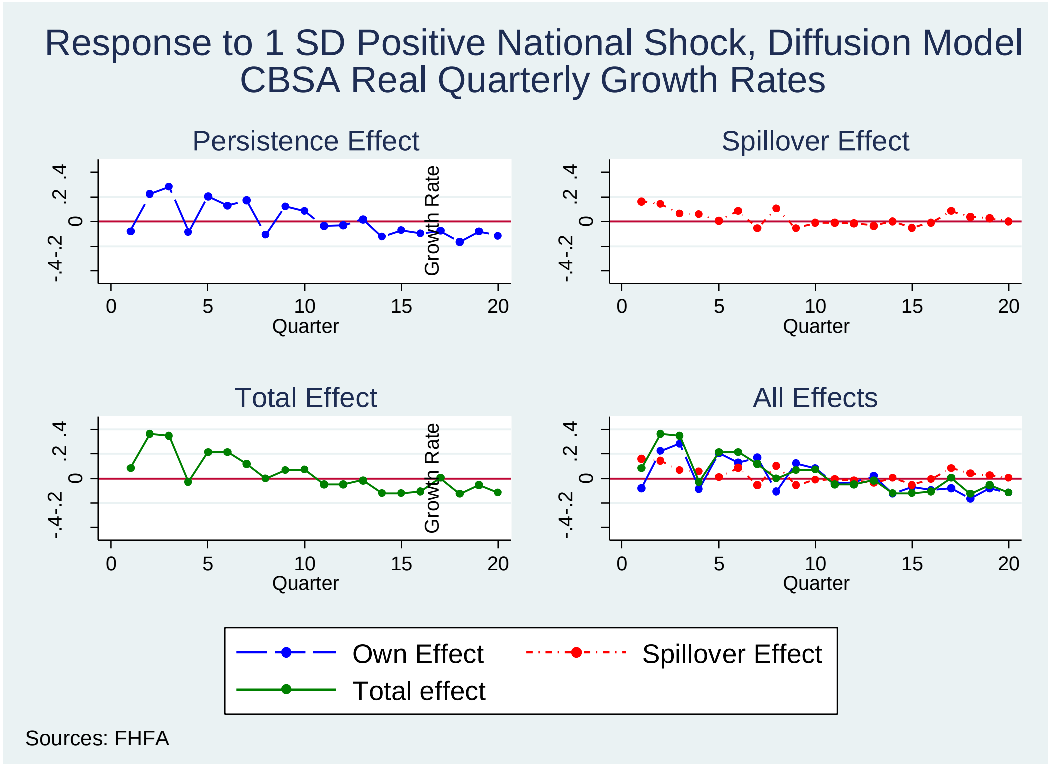


Figure A2