Local House Price Comovements*

Marcel Fischer† Roland Füss‡ Simon Stehle§

This version: December 20, 2019

*We are grateful for excellent research assistance from Konstantin Spring and helpful comments and suggestions from Martin Brown, João Cocco, Christian Riis Flor, Günther Franke, Jens Jackwerth, Winfried Koeniger, Frederic Menninger, Daniel Ruf, Nawid Siassi, Bertram Steininger, Yildiray Yildirim, and seminar participants at the Universities of Konstanz, Odense, Regensburg, St. Gallen, the EBS Business School, and the NTNU Trondheim, as well as the 2018 AREUEA National Conference in Washington DC, the 2018 EFMA Annual Meeting in Milan, the 2018 Conference on Decision Sciences in Konstanz, and the 2018 ReCapNet Conference of the Centre of European Economic Research (ZEW) in Mannheim.

†Copenhagen Business School, Department of Finance, Solbjerg Plads 3, DK-2000 Frederiksberg, Denmark and University of Konstanz, Department of Economics, Postbox 147, D-78457 Konstanz, Germany, phone: +49 7531 882645, e-mail: marcel.fischer@uni.kn.

‡Swiss Institute of Banking and Finance (s/bf), University of St.Gallen, Unterer Graben 21, 9000 St. Gallen, Switzerland and Research Associate at the Centre for European Economic Research (ZEW), Mannheim, Germany; Phone: +41 (0)71 224 7055; Email: roland.fuess@unisg.ch

§University of Konstanz, Department of Economics, Postbox 147, D-78457 Konstanz, Germany, phone: +49 7531 883620, e-mail: simon.stehle@uni.kn.

Abstract

Local House Price Comovements

We study the micro-level evolution of residential house prices using data on repeat sales on Manhattan Island from 2004 to 2015. We document that excess price comovement is a highly local and persistent phenomenon. The strength of such excess comovements vanishes with both spatial and temporal distance. Local underperformance is more persistent than local overperformance – particularly when house prices on aggregate level increase.

JEL Classification Codes: R30, R32
Key Words: Housing market, price comovements, urban economics, real estate, repeat sales.

1 Introduction

The recent boom and bust in house prices dramatically illustrates the need for a better understanding of price dynamics of residential homes. Since the pioneering work of Case and Shiller (1989), it is a well-established fact that returns on national and city-wide house price indices are subject to strong auto- and cross-sectional correlation. Two important channels explaining these correlations are comovements and spillovers in residential house prices. Comovements are caused by common underlying factors, such as gradually changing credit conditions (e.g., Chambers et al., 2009; Landvoigt et al., 2015; Amromin et al., 2018). Spillovers, on the other hand, are caused by a trigger, such as gentrification (Guerrieri et al., 2013), or rent decontrol (Autor et al., 2014), that spills over from affected to unaffected properties. For instance, a foreclosure also affects the trading prices of other properties in a neighborhood that do not go through a foreclosure (e.g., Campbell et al., 2011; Gupta, 2019; Guren and McQuade, 2019). Even interior renovations, which are generally unobservable, can raise house prices in close proximity after transaction prices of the renovated homes become public (Szumilo, 2018), suggesting that publicly available trading prices on their own can affect trading prices of other homes.

In this paper, we document the existence of excess comovements, that is, comovements beyond macroeconomic or common factors, in residential house prices on the micro level. We show that even after controlling for price changes on the monthly macro level as well as for zip-code-year based price developments, excess returns are positively related to past excess returns in the nearest surrounding of the traded home. Excess comovements are most pronounced when house prices depreciate.

Our work contributes to two important strands of literature. First, it contributes to the literature that has documented the existence of excess comovements on the index level (e.g., Kallberg et al., 2014), by showing that excess comovements also exist in the trading prices of individual homes on the micro level. Second, it contributes to the literature explaining cross-sectional correlations in house prices on the local level through spill over effects, by documenting that such cross-sectional correlation even exists in the absence of event-specific externalities.

Consistent with the spillover effects documented in, e.g., Campbell et al. (2011), Guerrieri et al. (2013), and Rossi-Hansberg et al. (2010), we show that excess comovements are strongest in the nearest neighborhood — particularly within the same building — and die out quickly with increasing distance between traded homes. Our results are robust to controlling for the evolution of house prices on the borough-level, on a monthly basis, as well as zip-code-year based price movements. In extensive robustness checks, we document our results to withstand other model specifications and parameter choices.

We further document that excess comovements exist over longer time horizons. Even local outperformance dating back as long as 2.5 years contributes to explaining present excess returns. Similar to the generally higher level of correlation in stock returns in markets with falling prices (e.g., Ang and Chen, 2002) and the evidence in Cotter et al. (2015), our results reveal that excess comovements are stronger in markets with falling prices. Local underperformance is particularly persistent when house prices on the aggregate level appreciate. Our results thus suggest a higher heterogeneity in terms of local house price changes when house prices increase.

Apart from the well-documented specific events discussed in the literature, excess comovements should be largely driven through two main channels. First, homes in the same neighborhood share common amenities, such as access to schools, recreational areas, shopping facilities, etc. Hence, ceteris paribus, homes in the same neighborhood should be better substitutes than more distant ones. When house prices in a given neighborhood increase, a potential buyer's budget constraint is more likely to be binding, thus increasing the incentive to search for cheaper homes in the nearest surrounding. This substitution effect should cause price increases in one neighborhood to also affect close-by neighborhoods.

Second, excess comovements can be caused through the information channel. Available information is likely to affect both buyers' and sellers' behavior. For buyers, the market for residential real estate is characterized by an information disadvantage (Coval and Moskowitz, 1999; Garmaise and Moskowitz, 2004; Kurlat and Stroebel, 2015). Information about locally realized sales prices that is not (yet) publicly available is typically easier to access for sellers via private channels, such as mouth-to-mouth propaganda. Thus, buyers have an incentive to use previous sales prices in the neighborhood to reduce the information gap. Simultaneously, sellers and their real estate agents should incorporate past sales prices in their offer prices and during price negotiations – for instance, because they do not want to sell at a worse price than their neighbors. Hence, past price changes in the neighborhood should affect present trading prices via both buyers' and sellers' incentives to use past sales prices as easily available anchors (Murfin and Pratt, 2019). Furthermore, the particularly strong within-building excess comovements are likely to be affected by a second anchoring effect: If a real estate agent has successfully sold a flat in a given building, other households wishing to sell may want to hire the same real estate agent, who likely uses his past realized sales price as an anchor for the new ask price. Ask prices, in turn, are known to affect the level of transaction prices of properties in the neighborhood (Horowitz, 1992; Anenberg, 2016).

We investigate the micro-level price dynamics of homes in urban areas using repeat sales on Manhattan Island between 2004 and 2015. We evaluate the order of magnitude, the persistence, and the state dependence of excess comovements. Manhattan Island seems ideal for investigating excess comovements in residential house prices in many regards. First, Manhattan is a liquid real estate market. Second, Manhattan is densely populated, implying that new constructions are scarce and unlikely to have major price impacts. Third, downpayments in Manhattan are very high and, over the last years, more than 50% of the condominiums and house sales have consistently closed without mortgage financing.[1] This makes buyers less dependent on the lending policy of banks, reducing trading frictions enormously and turns Manhattan into a highly efficient real estate market. As most of the transactions are conducted by a real estate broker, information travels quickly to the buyers who have appointed an agent. Finally, the exact trading prices for all homes are publicly available from the New York City Department of Finance,[2] implying that information is easily available for all market participants.

In contrast to the work of Rossi-Hansberg et al. (2010), Campbell et al. (2011), Guerrieri et al. (2013), and Szumilo (2018), which focuses on spillovers related to specific events, our goal is to document that prices comove even in the absence of specific events and to quantify the order of magnitude of such excess comovements along several dimensions. In our empirical analysis, we focus on excess returns relative to the evolution of house prices on the zip-code-year level, which should remove any event that affects an entire zip-code, such as changes in air quality (Chay and Greenstone, 2005), for instance. We further include variables controlling for the proximity to special areas such as the Central Park. Our results show that excess comovements only exist in the nearest vicinity, dying out quickly with increasing distance. Hence, our results are unlikely to be driven by events that affect larger neighborhoods within a zip-code. For instance, a 6% increase (i.e., one standard deviation) in the annualized excess return of an apartment leads to a 1.3% increase in the expected returns of an apartment located in the same building. This effect decreases by 80% within a 500 feet radius neighborhood.

In addition to the spatial dimension, performance is also persistent on the temporal dimension. For example, within the same building, even conditional on most recent excess returns, 6% of past average excess returns from 2 to 2.5 years ago are reflected in excess returns today. Local underperformance is more persistent than local overperformance, particularly in markets with generally appreciating house prices: Within the same building, 36% of negative excess returns are reflected in today's prices during booming periods, in contrast to only 26% during non-booming states of the aggregate market.

[1] https://www.propertyshark.com/Real-Estate-Reports/2016/12/13/payments-manhattan-now-500k-almost-double-median-sale-price-us/

[2] http://www1.nyc.gov/site/finance/taxes/property-rolling-sales-data.page

Our work contributes to a growing strand of literature investigating micro-level dynamics of house prices. This literature demonstrates that local events, such as gentrification (Guerrieri et al., 2013), urban revitalization (Rossi-Hansberg et al., 2010), air pollution (Chay and Greenstone, 2005), legislative amendment (Autor et al., 2014), unnatural deaths (Bhattacharya et al., 2017), the Low Income Housing Tax Credit (Diamond and McQuade, 2019), and foreclosures (Harding et al., 2009; Campbell et al., 2011; Anenberg and Kung, 2014; Gerardi et al., 2015; Gupta, 2019; Guren and McQuade, 2019) are important drivers of micro-level house price dynamics. Our work contributes to this line of research by documenting that prices comove not only in the presence of specific events, but also in their absence. We further quantify how the order of magnitude of such excess price comovements changes with the distance between traded homes.

Our work further contributes to the literature on excess comovements by showing that such comovements not only exists on the index level (e.g., Kallberg et al., 2014), but also in the prices of individual homes. In our empirical analysis, we use individual property returns in excess of Manhattan Island's monthly repeat sales index as endogenous variable as well as yearly zip-code specific deviations of annualized excess returns from this local index as control variables in order to account for local and neighborhood-specific events.

This paper proceeds as follows: Section 2 explains the existence of comovements based on a stylized model from which we derive our main hypotheses. In Section 3, we introduce our data. Section 4 presents our results on excess comovements in residential house prices. Section 5 documents the robustness of our results. Finally, Section 6 concludes.

2 A Simple Model

In this section, we motivate general excess comovements in residential house prices. In a stylized market microstructure model for the housing market, we demonstrate why prices should comove with future transactions in the neighborhood.

We consider a set of homes, $H_1, H_2, \ldots, H_n$ that only differ by their location. We denote the physical distance between two homes H_i and H_j by $D_{i,j}$. We assume that households typically have a preference for a certain location of their homes. This preference could both reflect the neighborhood's facilities, such as good schools, restaurants, shops, as well as social ties, such as other family members or friends living in the neighborhood. A home outside the preferred location is a substitute for the home at the preferred location, because both homes provide households with the same housing services. The prices P_i and P_j of the two homes H_i and H_j should therefore be positively correlated:

$$P_i = f\left(P_j\right) \text{ with } \frac{\partial f}{\partial P_j} > 0. \tag{1}$$

That is, comovements in residential house prices reflect that households react to price increases for homes in a given neighborhood by purchasing substituting homes in close-by neighborhoods, thus causing price increases in these neighborhoods. For instance, an increase in house prices in a potential buyer's preferred neighborhood increases the likelihood that his budget constraint is binding and he thus has to search for alternatives in the surrounding area of the preferred neighborhood.

The distance $D_{i,j}$ between two homes can be interpreted as a proxy for how well two homes H_i and H_j can be substituted with each other. The distance is, among others, a good proxy for the commuting costs and time it takes to get from H_i to the amenities at H_j. The smaller the distance $D_{i,j}$ between two homes H_i and H_j, the better they proxy for each other. That is, we can refine our model to

$$P_i = g\left(P_j, D_{i,j}\right) \text{ with } \frac{\partial g}{\partial P_j} > 0 \text{ and } \frac{\partial g^2}{\partial P_j \partial D_{i,j}} < 0. \tag{2}$$

In other words, the evolution of two homes' prices in the same neighborhood should be positively correlated. More precisely, the smaller the distance D_{ij} between two homes, the more a price signal from a previous trade should affect the price of H_i.

The adjustment of prices due to substitution should not take place instantaneously, as search for houses is time consuming. Thus, the gradual adjustment of prices brings up temporal distance $T_{i,j}$ as further dimension of excess comovements in local housing markets. Consequently, the price, P_i should thus not only be affected by the spatial distance between traded homes, but also by the temporal distance $T_{i,j}$. With increasing distance in the temporal sense, prices of neighboring substitutes become less informative for contemporaneous price movements:

$$P_i = h(P_j, D_{i,j}, T_{i,j}) \text{ with } \frac{\partial h^2}{\partial P_j \partial T_{i,j}} < 0. \tag{3}$$

In order to find the best price estimate, i.e., the fair market value, for a home in a given location, agents on both the seller and buyer side have to trade off substitutability (i.e, physical distance $D_{i,j}$) against timeliness for current market movements (i.e, temporal distance $T_{i,j}$). In other words, if the physical distance between i and j is small, agents should be willing to accept a greater time distance for home H_j to enter the price estimation:

$$\partial \left(\frac{\partial h^2}{\partial P_j \partial D_{i,j}} \right) \bigg/ \partial T_{i,j} < 0 \tag{4}$$

Our simple model in this section makes three predictions. First, positive excess returns from nearby homes should lead to positive excess returns for a given home. Second, the strength of these excess comovements should die out with increasing spatial and temporal distance between traded homes. Third, when physical distance is small, excess comovements should be more persistent in the time dimension. We test these model predictions in Section 4, after introducing our data in Section 3.

3 Data and Methodology

Our data is from the CoreLogic database, which covers 99.9% of the U.S. population.[3] We focus on repeat sales in urban areas using data from Manhattan Island, New York City. Manhattan Island seems ideal to investigate excess comovements in residential house prices in several regards. First, given that Manhattan is generally perceived as a very attractive place to live, the market for real estate is liquid and foreclosures are rare.[4] Second, compared to more rural areas, Manhattan Island is densely populated and space for new buildings is therefore extremely scarce. This severely limits the amount of new construction and the price impact of new buildings on existing places. Third, the exact prices of all trades are publicly available at the New York City Department of Finance's homepage. That is, information about actual trading prices of adjacent homes is easily available for all market participants and our results are less affected by information asymmetries. Our data spans the time period from January 2004 to December 2015, plus past sales prices on the most recent previous transactions dating back to 2000.

3.1 DATA CLEANING

We consider repeat sales of condominiums and apartments in order to compute returns on investments for such places. Initially, our dataset of the market on Manhattan Island consists of 43,466 observations, covering the period from January 2004 to December 2015. The dates for the most recent prior sale dates range back until January 2000, allowing for longer holding periods even at the beginning of the sample. The removal of observations

[3] https://www.corelogic.com/solutions/university-data-portal.aspx

[4] According to RealtyTrac.com, only one in every 12,410 trades in New York City relates to a foreclosure. See http://www.realtytrac.com/statsandtrends/foreclosuretrends/ny/new-york-county/new-york/ as of April 2019

that are not classified as resales (e.g., subdivisions) or for which information about the date of the transaction, the current or most recent preceding sales price (prior sales price) is not available leaves us with 42,301 observations.[5] The removal of duplicates with identical sales prices, prior sales prices, transaction dates, and geographic coordinates leaves our sample with 41,905 observations. Following Landvoigt et al. (2015), we remove speculative trades with holding periods of less than 180 days, leaving us with 41,283 observations.[6] Finally, similar to Campbell et al. (2011), for every year we remove outliers with current or prior sales prices in the first and 99th percentile, respectively, leading to a cleaned data set of 39,771 observations. To account for data errors or physical changes in a property, we follow Standard & Poor's (S&P Dow Jones Indices, 2017) in removing outliers. More specifically, we remove observations in the third and 97th percentile of the annualized return distribution.[7] Our final data set then consists of 37,385 observations.

Figure 1 summarizes the evolution of residential house prices in our cleaned data set using a repeat sales index (Case and Shiller, 1989) constructed on a monthly basis for the time period from January 2000 to December 2015. Similar to house prices on the national level, from Figure 1, Manhattan Island experienced a significant boom during the 2000s, with prices more than doubling from 2000 to 2006. Thereafter, house prices did not show a clear trend until they sharply declined in late 2008 – later than on the national level.[8] This relatively late decline may reflect that layoffs in the financial industry and their implications for house prices on Manhattan Island did not occur instantly when house prices on the national level started declining, but with a certain delay.

[Figure 1 about here]

3.2 EXCESS RETURNS

The repeat sales in our data differ along two important dimensions that make a direct comparison of returns difficult. First, the lengths of the time intervals between two trades may differ substantially. Second, returns depend crucially on the phase of the housing market cycle. To control for these two effects, we compute annualized market-adjusted

[5] In our data, the characteristics of repeat-sales and non-repeat-sales are remarkably similar, indicating that the removal of non-repeat-sales does not leave us with a non-representative sample. For instance, the average trading price of our repeat-sales is USD 1.41 million (in January 2015 dollars), whereas it is USD 1.42 million for the non-repeat sales. The similarity of both subsamples further suggests that the non-repeat sales do not constitute a systematic, confounding factor in our analysis.

[6] Similarly, in the construction of the S&P 500 Case-Shiller house price index, observations with holding periods of less than six months are removed (S&P Dow Jones Indices, 2017).

[7] The results are qualitatively robust to the removal of only one or two percent of each tail.

[8] In Section 4.2, we use these differences in the general evolution of house prices to investigate whether comovements vary with the phase of the housing market cycle.

excess returns, $r_{t,t-}$, for properties traded at month t and previously traded at month $t-$ as follows:

$$r_{t,t-} = \left(\frac{P_t}{P_{t-}}\right)^{\frac{1}{y(t,t-)}} - \left(\frac{C_t}{C_{t-}}\right)^{\frac{1}{y(t,t-)}}, \tag{5}$$

in which P_t and P_{t-} denote the present and prior trading prices of the property in months t and $t-$, respectively. $y\,(t, t-)$ is the time distance in years between the two trades, and C_t and C_{t-} denote the index levels of the Manhattan Island repeat-sales price index constructed as in Case and Shiller (1989) from our cleaned data in months t and $t-$, respectively. By subtracting the index return, we remove aggregate effects that systematically affect house prices, such as inflation, seasonal effects, and the phases of the housing market cycle at the moments of the two trading dates, as well as other common specific events, such as the September 11 attacks. Simultaneously, excess returns allow us to distinguish between states with local over- and underperformance, which, as we demonstrate in Section 4.2, are associated with differential excess comovements.

3.3 CONTROL VARIABLES

Our control variables can be broadly split into three different categories: (1) transaction-specific, (2) locational, and (3) macro-financial control variables.

3.3.a Transaction-Specific Variables

In our data cleaning procedure, we remove transactions with holding periods of less than 180 days, which are likely to be speculative trades. Short holding periods may be targeted at larger renovations during that period, aiming at substantially increasing a property's value. To account for these possible effects, we include mutually exclusive dummy variables for holding periods of less than one and less than two years, respectively.

The results in Landvoigt et al. (2015) document that during the recent housing market boom, housing returns varied substantially between homes in different price segments in a nonlinear fashion. To account for this effect, we control for the log of the inflation-adjusted prior sales price (in January 2015 dollars) as well as its square. Whereas private investors profit from both their home as a durable consumption good and from house price appreciations, corporations should place higher emphasis on earning higher returns on their investments. To control for these effects, we include two dummies for whether a property is sold or bought by a corporation and a dummy for whether a home is bought to become an owner-occupied home. Transactions in which the buyer is a corporation or the home is bought to serve as an owner-occupied home are already marked in our database. We further construct a dummy variable indicating whether the seller is a corporation or not.[9]

3.3.b Locational Variables

The location of a residential home is one of the key factors determining its price (e.g., Can, 1990; Case and Mayer, 1996). To control for possible changes in the pricing of location-specific factors, we control for the view on the Central Park and the waterfront as well as the walking-distance to these two amenities. We further control for distance to Times Square, the New York Stock Exchange, and the nearest entry to the subway.[10] More specifically, we include mutually exclusive dummies for a view and a walking-distance to Central Park if the beeline does not exceed 100 feet and the city block walking-distance does not exceed 500 feet, respectively. In similar fashion, we include a waterfront-view-dummy if a home has a direct view on the water surrounding Manhattan Island; i.e., if the home is separated from water only by a road, a park, or both, but not by a building. We further include a walking-distance dummy, if the city block distance to the waterfront does not exceed 500 feet. To account for easy access to the subway system, we include two dummies: a dummy for very close distances to the nearest entry for city block walking-distances of less than 100 feet and a dummy for close distances of 100 to less than 500 feet. For Times Square and the New York Stock Exchange we include two dummies for short walking-distance and medium walking-distance if the city-block distance is less than 1,000 feet and 1,000 to less than 2,000 feet, respectively.

Guerrieri et al. (2013) document substantial differences in house price growth across neighborhoods. To account for these differences, we proceed similar to Campbell et al. (2011), who use census-tract-year dummies, and control for zip-code-year fixed effects in the current and the prior year of trade of the home. To attain a reasonable number of observations per zip-code (at least 1,000 observations), we have to cluster a few adjacent zip-codes. A detailed overview over the clustered zip-codes can be found in the Appendix. To control for the impact of liquidity in the local housing market on transaction prices (Caplin and Leahy, 2011), we control for the log of one plus the number of trades in the past 180 days on the zip-code level.

[9] We define a seller as a corporation if the seller's name contains keywords such as ACQUI, ASSOC, AVENUE, BANK, BOARD, CORP, CREDITOR, EQUIT, ESTATE, FUND, HDFC, HLDGS, HOLDING, HOUSING, HSNG, INC, INVEST, L*L*C, LLC, LP, LTD, OWNER, PARTNER, PLC, PORTFOLIO, PROP, QUATAR, REALTY, STREET, TRUST, or *LP, in which * signifies blank spaces. A manual comparison of more than 3,000 observations did not indicate any missing words.

[10] The geographic coordinates of the New York subway entries are from NYC Open Data (https://opendata.cityofnewyork.us)

3.3.c Macro-Financial Variables

To control for changes in the macroeconomic environment, we include the seasonally-adjusted real growth rate of the GDP relative to the previous quarter with a lag of one period from the US Bureau of Economic Analysis, the seasonally-adjusted monthly growth rate of the unemployment rate in New York City from the Bureau of Labor Statistics, and the percentage change in the average fixed mortgage lending rate from the Federal Housing Finance Board.

Since the pioneering work of Case and Shiller (1989), it is known that residential house prices exhibit a significant degree of autocorrelation. To explain price movements, it is therefore important to control for this persistence. Our analysis focuses on explaining excess returns rather than raw returns, thus removing the systematic autocorrelation.

Table 1 summarizes key properties of our data. As to be expected, the annualized excess return is not significantly different from zero.[11] The average holding period is only about 5.5 years, indicating that Manhattan is a fairly liquid market for residential homes. About 9% of properties are even resold within up to 2 years, which may, among others, reflect institutional investors' activities that account for about 15% of purchases and 10% of sales. Yet, the majority of trades (53%) still represents sales of owner-occupied places. With an average prior trading price of USD 1.279 million (inflation-adjusted to 2015 prices), prices on Manhattan Island are among the most expensive in the U.S. This high average trading price suggests that prices should be largely determined by location. In contrast, renovations or a new kitchen should have a lower impact on the trading price, advocating the repeat sales approach. Likewise, the short average holding period provides additional support for the repeat sales approach.

[Table 1 about here]

3.4 METHODOLOGY

The goal of our work is to investigate how recent past excess returns in residential house prices comove with present returns of homes in the neighborhood. We further ask whether the strength of these effects varies with the stage of the housing market cycle and whether excess comovements vanish with temporal distance between two trades. For that purpose, we define K mutually exclusive neighborhoods for each observed trade. We refer to trades with coinciding geographic coordinates, i.e., trades in the same building, as the first-order neighborhood throughout. Additionally, we draw $K-1$ circles around each observed trade. We want to end up with the same expected number of observations in each of these $K-1$

[11] The small positive value reflects that the market return constructed using the Case-Shiller methodology weights observations unequally.

circles to make sure that, on average, liquidity is the same in all circles and the average excess returns from all of these neighborhoods are thus estimated with the same precision. We therefore draw the circles such that the area inside each of them is identical.[12]

The first of the $K-1$ circles, also referred to as the second-order neighborhood throughout, is characterized by a maximum distance of 500 feet, roughly corresponding to two blocks.[13] The borders of the other circles, which we refer to as third-, fourth-, fifth-, and sixth-order neighborhoods thus lie at 707, 866, 1,000, and 1,118 feet, respectively, leaving us with on average 5.3 to 6.5 historical trades in the second- to sixth-order neighborhood for every current trade. Figure 2 visualizes our construction of $K = 6$ neighborhoods for a specific property. For every neighborhood k, we define a neighborhood-specific excess return, $\bar{r}^e_{i,k}$ as the average of the observed excess returns in the T days prior to trade i.

[Figure 2 about here]

We employ the following regression setup:[14]

$$r^e_{i,t,t-,z} = \alpha_z + \sum_{k=1}^{K} \rho_k \bar{r}^e_{i,k} + \delta_{a(t),z} - \delta_{a(t-),z} + X_{i,t}\beta + \epsilon_{i,t,t-,z}, \tag{6}$$

in which $r^e_{i,t,t-,z}$ is the annualized excess return on property i in zip-code z realized between time $t-$ and t, $\delta_{a(t),z}$ and $\delta_{a(t-),z}$ are the zip-code-z specific deviations of the annualized excess returns from the Manhattan Island wide index. The subscripts $a(t)$ and $a(t-)$ refer to the years in which the transactions took place, respectively. For example, if t corresponds to any sale date in the year 2010, $a(t) = 2010$. $X_{i,t}$ is a vector of control variables, $\epsilon_{i,t,t-,z}$ is a normally-distributed error term, and α_z reflect zip-code means.

The precision of our estimate for the annualized excess return is generally increasing with the length of the time interval between two trades. Intuitively, when the two trades occur within a relatively short time period, small deviations in observed trading prices of individual properties and short-term fluctuations in the local house price index lead to significant amplifications when being annualized. Hence, annualized excess returns tend to be subject to higher variation when two trades occur within a relatively short time period. To account for this phenomenon in our analysis, we allow the variance of $\epsilon_{i,t,t-,z}$ to depend

[12] The total number of observed past trading prices ranges from about 200,000 to 246,000 in the $K-1 = 5$ circles of our empirical analysis, i.e., the numbers of observations among our defined neighborhoods are roughly of equal size.

[13] In Section 5, we demonstrate the robustness of our results to applying the city-block metric.

[14] Equation (6) can be easily rewritten in spatial econometrics notation because $\bar{r}^e_{i,k}$ reflects the k-th spatial lag. Nevertheless, under the assumption of homoskedastic error terms, OLS is applicable, since we account for the time-directionality in constructing the spatial weights.

on the difference D between t and $t-$: $\mathrm{Var}\left(\epsilon_{i,t,t-,z}\right) = \exp\left(\gamma_1 + \gamma_2 D\right)$ where γ_1 and γ_2 are regression-endogenously determined coefficients. We use the exponential function to ensure positivity of the variance in the optimization process.[15]

Our goal is to explore whether recent past excess returns in the neighborhood comove with present excess returns, i.e., whether the ρ_ks are different from zero and, if so, whether price comovements decay with increasing distance, i.e., whether $|\rho_1| > \cdots > |\rho_K|$.

4 Empirical Results

For our empirical analysis, we need to determine a few parameters for our model introduced in Section 3.4. Specifically, we need to choose the number of distinct neighborhoods that we want to consider. In particular, we want to understand whether excess comovements are strongest in the first-order neighborhood and whether they die out in more distant neighborhoods. We therefore set the number of neighborhoods to $K = 6$.[16]

We further need to choose the maximum number of days prior to our trade, T, such that trades on other properties should reasonably have the potential of affecting a home's price. Hence, we consider the persistence of price comovements not only in the spatial, but also in the temporal sense. The choice of T is driven by a tradeoff between two opposing objectives. On the one hand, we want to estimate price comovements as precisely as possible, suggesting that we should use as much past data as possible. On the other hand, the precision can be reduced by using outdated observations that may have little informational content for present prices, among others, because the information is already incorporated in more recent prices. We set $T = 180$ for three main reasons. First, gathering information in the housing market costs more time than for example gathering information about the stock market. Second, finding a buyer for a given home typically takes time. Third, our choice of about half a year provides us with a reasonable number of observations to estimate effects with good precision. We document the robustness of our results to the choice of T in Section 5.

4.1 EXCESS COMOVEMENTS

In this section, we provide empirical evidence on the existence and strength of excess comovements in residential housing markets. Table 2 summarizes the results of five Maximum Likelihood regressions explaining the annualized excess returns of repeat sales relative to

[15] Our estimates are qualitatively robust to the homoskedastic case.

[16] Empirically, it turns out that a larger number of neighborhoods does not further contribute significantly to explaining house prices, while a smaller number does not allow us to fully capture the decay in the comovement magnitude with increasing distance.

trades on Manhattan Island. The first-order neighborhood relates to trades in the same building. Second-, third-, fourth-, fifth-, and sixth-order neighborhoods are less than 500, 500 to 707, 707 to less than 866, 866 to less than 1,000, and 1,000 to less than 1,118 feet. Our choice of distances from the traded homes is motivated by the goal to build neighborhoods of identical sizes in order to end up with similar numbers of traded homes in every neighborhood. Locational controls are our measure for liquidity, as well as dummies indicating Central Park view, Central Park walking distance, a very close subway station, a close subway station, short distance to Times Square, medium distance to Times Square, short distance to the NYSE, medium distance to the NYSE, waterfront view, and waterfront walking distance. Transaction-specific controls include two dummy variables indicating that a resale took place within one year, or between one and two years, respectively, log inflation-adjusted prior sale price and its square, two dummies indicating whether the seller or buyer of the property is a corporation, and a dummy indicating whether the property is owner-occupied. Macro-financial controls are lagged GDP growth, lagged unemployment growth, and lagged percentage interest rate change. Fixed effects are on the zip-code level (zip) or the zip-code-year level (zip-year).[17]

4.1.a Excess Comovements and Spatial Distance

In this section, we put emphasis on the spatial dimension of price comovements. From Section 2, prices should exhibit excess comovements and these comovements should decay in magnitude as the distance between homes increases. From Table 2, the coefficients for neighborhoods one to five are all positive and significant. That is, Table 2 confirms the existence of excess comovements in regular sales. From the first- to the sixth-order neighborhood the coefficients generally decrease, indicating that comovements decrease with increasing distance between two traded homes. Coefficients are monotonically decreasing in space, except for the third-order neighborhood, for which the strength of comovements is of slightly lower magnitude.

For all specifications in Table 2, the sharpest decline in excess comovements is observed for the transition from the first- to the second-order-neighborhood, for which coefficients drop by at least 63%. This result should be mainly driven through two channels. First, trades within the same building should be among the closest substitutes.[18] Second, within

[17] Similarly to Campbell et al. (2011) who use census-tract-year clusters, we cluster standard errors over the zip-code-year level. In Table 8 in our online appendix (currently at the end of this document), we show that our results are robust to alternative ways of clustering.

[18] To investigate the substitution channel in more detail, we also explored a setting in which we identified co-operatives. Two dwellings within a given co-op tend to be more homogeneous than two random dwellings. Co-ops often span over several buildings or even an entire building block. Hence, co-op homes in adjacent

the first-order neighborhood both the transmission of information via informal channels, such as chats among neighbors, but also active search for information, should be most intense.[19]

[Table 2 about here]

Local price movements should generally be driven by location-specific events. It is therefore important to control for them. A comparison of columns (1) and (2) in Table 2 reveals that after including our locational controls and controlling for zip-code fixed effects, the coefficients generally decrease, but remain highly significant for the first five neighborhoods. That is, even after controlling for location-specific events, there is still a strong informational content in excess house price movements in the closest neighborhoods.

However, our results also reveal that the coefficient for the most remote, the sixth-order neighborhood, becomes close to zero and insignificant. In other words, our local controls and zip-code fixed effects already capture local price trends quite well. Furthermore, the insignificance of the coefficient for the sixth-order neighborhood in column (2) points to two conclusions: First, beyond general price trends, the sixth-order neighborhood no longer contains price information. Second, the reduction in the coefficients for the first- to fifth-order neighborhood largely reflects the removal of the location- and zip-code-specific events. The locational controls and the zip-code fixed effects thus should not only capture the general price movement in the sixth-, but also in the first- to fifth-order neighborhoods very well. Changes in our coefficients in the transition from column (2) to (3), in which we add transaction-specific controls, are rather small. In the transition from column (3) to (4), in which we include macro-financial controls, these changes are even smaller, indicating that the excess returns, which our work builds on, already capture the effects of macroeconomic events very well.

Column (5) reports the estimates for our full specification. Compared to the model presented in column (4), we include zip-code-year fixed effects as opposed to zip-code fixed effects. From column (5), 21% of the increase in the annualized excess return in the first-order neighborhood are reflected in future home prices. For example, a one standard deviation increase in the annualized excess return in the first-order neighborhood, i.e., an increase in the annualized excess return by about 6%, leads to an increase in the expected annualized

buildings should be better substitutes than non-co-op homes. Our results in Table 9 in the online appendix (currently at the end of this document) reveal that excess comovements in second to fourth-order neighborhood from co-op homes are stronger and within-building excess comovements are weaker, since the excess comovements in the adjacent buildings already capture some of the effects.

[19] A further reason for stronger comovements within the same building could be rent stabilization, which applies to entire buildings when built before 1974 containing more than six units. Rent stabilization should have only little effect on price comovements in our work, since we use repeat sales and the stabilization should be priced into the initial purchase.

excess return of future home prices of about 1.3%. For an average holding period of about five years, the expected excess return is then about 6.5%. For the second- to fourth-order neighborhoods these effects are around 80% weaker than for the first-order neighborhood. Here, a one standard deviation increase in the second- to fourth-order neighborhood's excess return leads to an increase in the expected future excess return after the typical holding period of five years of 0.9 to 1.2 percent, respectively. With past excess returns of identical signs in the first- to fifth-order neighborhoods, the effects accumulate and expected future excess returns can be even higher. For example, a one standard deviation increase in all five neighborhoods leads to an increase in the expected future excess return of around 11% over five years.

4.1.b ***Excess Comovements and Temporal Distance***

Our results in the previous section document that excess comovements exist in the spatial dimension, but are dying out with increasing distance between traded properties. In this section, we turn to the second prediction from section 2, and ask whether in addition to the spatial dimension, excess comovements also exist over longer time horizons, i.e., above systematic autocorrelation. Specifically, we investigate whether adding more lagged excess returns from previous periods has additional predictive power for present excess returns, and, if so, whether the predictive power is decaying with increasing temporal distance. More technically, instead of investigating the informational content of only the most recent $T = 180$ days, we analyze the excess comovements of prices from multiple lags of intervals of length T.

Table 3 summarizes our baseline results when choosing a total of five mutually exclusive lags. Accordingly, to be included in lag 1, a neighboring trade should have been settled in the most recent 180 days prior to the respective sale, for lag 2 during the most recent 181 to 360 days, etc. As for all following tables, we only show results for our most advanced specification, including all sets of control variables, as well as zip-code-year fixed effect dummies. In order to keep the amount of coefficients to be estimated at a tractable amount, we unify second- and third-order, as well as fourth- and fifth-order neighborhoods. Coefficients for the sixth-order neighborhood remain insignificant in all our specifications. We therefore exclude the sixth-order neighborhood throughout our analysis in this section.[20]

The results in Table 3 reveal that excess comovements in residential house prices do not only exist in the spatial dimension, but also over longer time horizons. Lagged excess returns extending beyond the first lag have a strong predictive power for present excess returns

[20] We report results when keeping the number of neighborhoods at $K = 6$ in the online appendix in Table 10 (currently at the end of this document).

– particularly the first-order neighborhood. These effects are decreasing with increasing temporal distance and are dying out completely in all but the first-order neighborhood. Compared to our results from Table 2 with only one temporal lag of excess returns, the point estimates for the first-order neighborhood are smaller, reflecting that the additional lags are already picking up some of the effects. The persistence of comovements in the first-order neighborhood confirms the third prediction from our model: When spatial distance between properties is small, excess comovements die out slower in the temporal dimension.

[Table 3 about here]

Our results in Table 3 suggest implications for the efficiency of local housing markets: Even conditional on the most recent, price movements from greater temporal distances are reflected in excess returns today to a both economically and statistically significant extent. Our results thus indicate that on the local level, information is processed very slowly, suggesting an explanation for the well-documented autocorrelation of residential house prices on the macro level.

4.2 EXCESS COMOVEMENTS OVER MARKET CYCLES

Having demonstrated the local nature of excess comovements, our next step is to ask whether the order of magnitude of excess comovements varies with phases of the housing market cycle, i.e., whether the strength of excess comovements and the distance, over which they are measurable, differs with the stage of the housing market cycle at the time of the sale; i.e., whether it differs between good states with generally increasing house prices and other stages of the housing market cycle. We differentiate between aggregate (macro) and local (micro) market trends. For the macro perspective, we define boom and non-boom periods ex-post using our price index for Manhattan Island from Figure 1. According to this index, the boom in the early 2000s ends in October 2005, and house prices start booming again in March 2013. We therefore define the period from November 2005 to February 2013 as the non-boom period and the remaining months as the boom period.[21]

Taking on the local, micro perspective, we define local positive (negative) markets according to the sign of the average excess returns in each neighborhood from the past T days.

[21] Using the publicly available S&P CoreLogic Case-Shiller New York City condominium index (download link https://us.spindices.com/documents/additionalinfo/20170926-589149/589149_cs-condoindices-0926.xls?force_download=true), we identify a non-boom period between February 2006 and April 2012. Using the S&P Case-Shiller National home price index, we identify a period from March 2006 to March 2012. Similarly, we characterize our non-boom period using a purely liquidity-based approach building on the number of observed trades. In Section 5.2, we document that our results are robust to all of these alternative specifications.

Consequently, the k-th order neighborhood is above (below) average markets when we observe a positive (negative) average excess return for the past $T = 180$ days. Put differently, we ask whether positive and negative information from past excess returns affect future excess returns asymmetrically.[22] Formally, we extend our empirical application from Equation (6) to:

$$r^e_{i,t,t-,z} = \alpha_z + \sum_{s=1}^{2}\sum_{k=1}^{K} \rho_{k,s} \times \bar{r}^e_{i,k} \times \mathbb{1}_{t\cap s} + \delta_{a(t),z} - \delta_{a(t-),z} + X_{i,t}\beta + \epsilon_{i,t,t-,z}, \tag{7}$$

in which the two stages of the cycle are defined by $s \in \{\text{boom,non-boom}\}$ for the macro, and $s \in \{\text{above average, below average}\}$ for the micro trend, and $\mathbb{1}_{t\cap s}$ is an indicator function that equals one if the housing market is in stage s at time t.

Table 4 summarizes the results of two separate regressions explaining the annualized excess return of repeat sales relative to trades on Manhattan Island in boom and non-boom periods for the macro panel (Panel A: Macro) and positive (negative) trends on the micro level (Panel B: Micro). For both regressions, the full set of control variables from Section 3.3 as well as zip-code-year fixed effect dummies are used.

From Panel A of Table 4, coefficients are generally smaller in booming stages of the housing market cycle than in other stages. That is, consistent with Cotter et al. (2015), excess comovements in residential house prices seem to be stronger in markets with falling prices. Similarly, from Panel B, coefficients for above-average excess returns are generally smaller than for below-average excess returns. That is, excess comovements in residential house prices seem to be stronger for negative deviations from the market than for positive deviations. For instance, about a third of a negative excess return in the first-order neighborhood is reflected in the excess return of an existing trade, whereas only less than 17% of a positive excess return is.

[Table 4 about here]

Having documented differences in how excess returns comove with past ones in positive and negative states of the housing market cycle on both the macro and micro level in Table 4, we next investigate these two effects jointly. In particular, we want to shed light on whether the generally stronger negative persistence in bad market environments is further amplified by a negative macro trend or not.

[22] It is important to note, that a negative excess return does not necessarily mean a loss for the seller. A negative sign simply indicates that the performance of the trade was smaller than the performance of the market.

Table 5 presents results from a single regression for four mutually exclusive sets of variables, which we select according to the phase of the housing market cycle on the macro and the micro level as in Table 4. Irrespective of the stage of the housing market cycle on the macro level, underperformance on the micro level is generally more persistent than local overperformance - particularly in the first- and second-order neighborhoods.

Excess comovements are strongest for below-average returns on the micro level in booming stages of the housing market cycle. That is, local underperformance is most persistent when house prices are generally increasing. Simultaneously, excess comovements from above-average local returns in a booming market are relatively weak. In a non-booming stage of the housing market cycle, differences between the strength of comovements induced by local above- and below-average performance are much weaker.

[Table 5 about here]

In sum, our results in Table 5 thus suggest that there is more heterogeneity in terms of the local evolution of house prices in booming stages of the housing market cycle than in non-booming ones.

5 Robustness Analysis

This section documents the robustness of our key findings with respect to various assumptions. Section 5.1 provides evidence for our base case parameter setting, in which we do not distinguish between boom and non-boom periods. Section 5.2 provides results for different definitions of the boom and non-boom periods. Additional robustness results are placed in an online appendix (currently at the end of this document).

5.1 ROBUSTNESS OF BASE CASE RESULTS

With our results in Table 2, we demonstrate that excess comovements in residential house prices are a highly local phenomenon. In this section, we demonstrate the robustness of our results with regard to four key dimensions and report these results in Table 6. To simplify the comparison with our base-case results, we repeat the results from Table 2 in Panel A of Table 6.

In Panel B of Table 6, we allow for a different number of past days used to compute average excess returns in the neighborhoods. In our base case parameter setting, we used the past $T = 180$ days, which we consider a good tradeoff between the two opposing goals of having a reasonably larger number of observations and very recent up-to-date observations.

In Panel B, we explore the cases in which we set $T = 120$ or $T = 240$ days. Our results for these two cases demonstrate the robustness of our key findings that effects are strongest in the same building, i.e., the first-order neighborhood, remain significant in the second- to fourth-order neighborhood, and fade out for higher-order neighborhoods. Similarly, the point estimates for the strength of comovement in the various neighborhoods are of a very similar order of magnitude.

In Panel C, we vary the definitions of the neighborhoods. In our base case parameter setting, the second-order neighborhood was characterized by a maximum distance from the traded home of not more than 500 feet, roughly corresponding to two blocks. Here, we report results when shrinking this distance measure by two thirds, i.e., to 333 feet. Again, the borders of the higher-order neighborhoods are defined such that the area is the same as in the second-order neighborhood. We also depict results for the case, in which the neighborhoods are defined as in Campbell et al. (2011), i.e., a maximum distance of 0.1 miles, corresponding to 528 feet, for the second-order and 0.25 miles, corresponding to 1,320 feet, for the third-order neighborhood. As in Campbell et al. (2011), we do not account for neighborhoods of a higher order. Finally, we depict results for the case in which the second-order neighborhood is defined by the City of New York (e.g., Chinatown, Lower East Side, etc.) and the third-order neighborhood consists of the corresponding neighborhoods adjacent to the second-order neighborhood.[23]

Our results in Panel C again document the robustness of our key finding that comovements are strongest in the first-order neighborhood. With smaller second- to sixth-order neighborhoods for the former case, results remain significant even in the sixth-order neighborhood, reflecting that the maximum distance of a trade in this neighborhood is 744 feet, corresponding to a trade in the fourth neighborhood in our base-case parameter setting. A more narrow definition of neighborhoods again suffers from the problem of relatively small numbers of historical trades in each of the neighborhoods, which, among others, leads to the coefficient for the third-order neighborhood being insignificant. For instance, the number of historical trades in this neighborhood decreases by about 65% compared to our base-case parameter setting with wider neighborhoods. Using the definition of City of New York neighborhoods yields strong comovements of within-neighborhood excess returns from the first- and second-order neighborhoods, but virtually no connection to adjacent neighborhoods,

[23]Data on the neighborhoods is obtained from New York City Open Data: https://data.cityofnewyork.us/City-Government/Neighborhood-Tabulation-Areas/cpf4-rkhq (retrieved on July 13, 2018), which provides a shape file defining the neighborhoods. The shape file includes a "miscellaneous" area that consists of several dispersed areas, such as parks, cemeteries, etc., that are not related to a particular neighborhood. A few trades in our data fall into this area, but are only a few feet away from the nearest non-"miscellaneous" neighborhood. We assign these observations to the nearest non-"miscellaneous" neighborhood.

suggesting that in Manhattan, where adjacent neighborhoods are often very heterogeneous, the preferred locations of households have sharp boundaries.

[Table 6 about here]

In Panel D, we restrict the maximum holding period to seven and ten years, respectively.[24] Further restricting the maximum holding period to less than seven years leads to such a strong decline in the number of observations that it no longer provides a representative picture of market movements and – due to the lack of this information – predicts largely insignificant effects. Specifically, reducing the maximum holding period to six years removes more than a third of all trades and the information contained in these trades.

In Panel E, we change the distance measure used in the definition of our neighborhoods from the Euclidean to the city-block metric and ask whether our results are affected by excluding observations for which the waterfront lies within at least the sixth-order neighborhood. Intuitively, for such observations, the area covered by higher-order neighborhoods may be smaller than that of smaller-order neighborhoods giving rise to potentially significantly different numbers of past trades in the different neighborhoods. Our results for both cases confirm our key findings that prices comove the strongest in the first-order neighborhood and fade out for the most distant neighborhoods.[25] Further robustness checks can be found in Table 11 in our online appendix (currently at the end of this document).

Overall, our results in this section confirm the robustness of our key findings on house price comovements to various assumptions in our base case parameter setting, in which we do not split the sample into boom and non-boom periods. We next proceed to demonstrate that our key results on comovements in boom and non-boom periods remain robust when using different criteria to determine these two subperiods.

5.2 ROBUSTNESS OF BOOM VERSUS NON-BOOM

In Section 4.2, we defined boom and non-boom periods based on our Manhattan Condominium index, constructed using the Case-Shiller methodology (Case and Shiller, 1989). Using this database, our non-boom period lasted from November 2005 to February 2013. We

[24] For shorter holding periods, larger reconstructions and major changes in the neighborhood should be less likely. That is, the repeat-sales approach should yield particularly precise estimates.

[25] We apply the city block metric to proxy commuting distance between properties. This is possible by exploiting the geometric design of Manhattan. We therefore shift the coordinates of the properties in our sample such that the streets approximately align with the lines of longitude and latitude. More precisely, we shift the coordinates (after standardizing) by 35 degrees counterclockwise around the south-east corner of the Central Park. Again, we construct six neighborhoods (of cubic form due to the metric) around each property. To ease comparison with our base case results, each cube encompasses the same area as our base-case circles.

further documented that excess comovements in the first-order neighborhood are stronger during non-boom periods and weaker during boom periods and that the other estimates are of comparable magnitude. In this section, we use alternative definitions for the boom and non-boom periods, using different house price indices and a liquidity measure.

The left panel of Figure 3 depicts the evolution of real house price indices for Manhattan (dotted line), New York (dashed line), and the entire United States (solid line). Similar to our proceeds from Section 4.2, we define the beginning of a non-boom period as the month in which a previously sharp incline in house prices ends. Likewise, the end of a non-boom period is the month in which a new sharp incline in house prices begins. That is, for the NYC Condominium Index, the non-boom period is March 2006 to April 2012 and for the US National House Price Index, this period is March 2006 to February 2012. The right panel in Figure 3 depicts the number of sold apartments and condominiums on Manhattan Island after removing observations with missing values in sales prices, sales dates, and duplicates. From this panel, the number of sales declined from 4,381 to 2,906 trades in October 2008 and did not recover systematically before March 2012. As an additional definition for our non-boom period, we therefore use the time period October 2008 to March 2012 as a liquidity-based definition of our non-boom period.

[Figure 3 about here]

Table 7 summarizes our results for the different definitions of the non-boom period. For ease of comparison, the results from Table 4 are repeated in Panel A of Table 7. Consistent with our key findings from Section 4.2, our robustness results with different definitions of the non-boom period confirm that during non-boom periods, excess comovements with the first-order neighborhood are stronger. Irrespective of the exact definition of our non-boom period, point estimates for our coefficients are very similar. Our results in Table 7 thus confirm the finding from Section 4.2 that during non-booming periods, excess comovements are stronger for within-building trades. Otherwise, the role of spatial distance is similar for both phases.

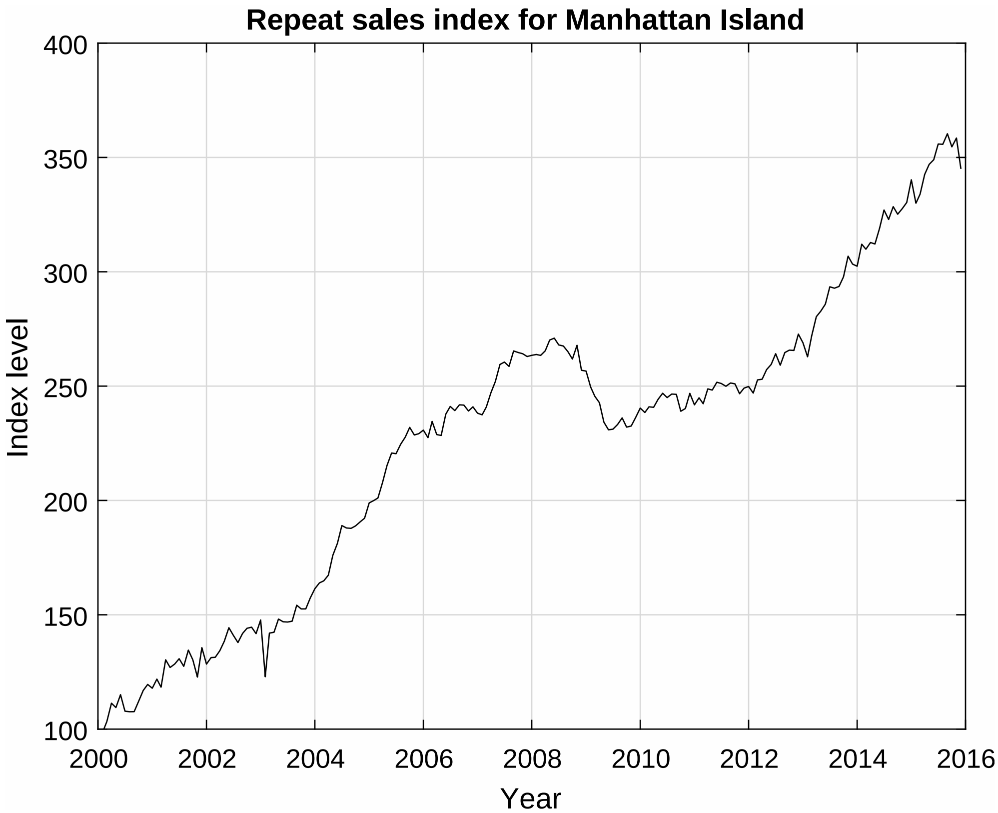
[Table 7 about here]

6 Conclusion

The housing market boom and bust of the early 2000s highlights the importance for a better understanding of the evolution of residential house prices. We contribute to this challenging endeavor by exploring the micro-level evolution of residential house prices, using data from trades on Manhattan Island between 2000 and 2015.

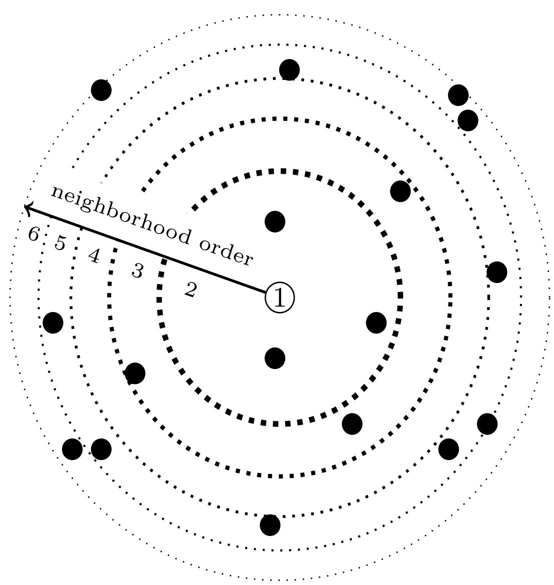
We document that even after controlling for monthly aggregate market movements and zip-code-year based price movements, excess comovements in residential house prices are a highly persistent local phenomenon. The strength of these excess comovements vanishes with the distance between traded homes. In addition to these spatial excess comovements, excess comovements in residential house prices also have a persistent temporal dimension. Unlike in stock markets, house prices seem to adjust slowly to new information, and even price movements from more than two years ago still have a significant impact on present price movements. Moreover, local underperformance is more persistent than local overperformance. This phenomenon is particularly strong when house prices on the aggregate level appreciate.

Figure 1
Evolution of house prices on Manhattan Island



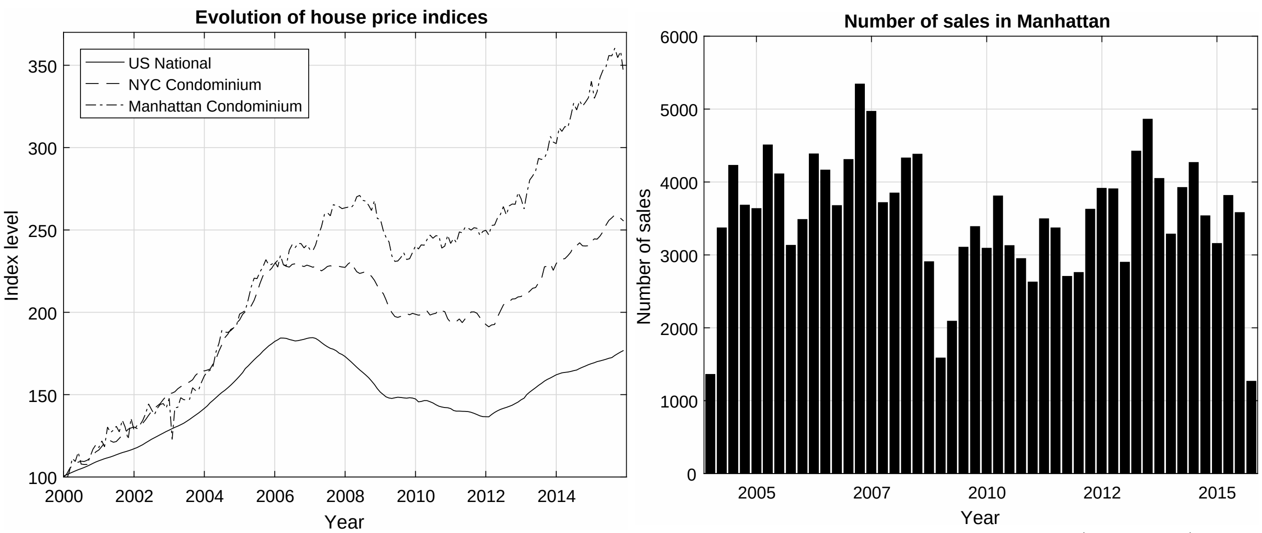
Nominal repeat sales index of Manhattan Island's condominium/apartment market based on our final data set. The index level is normalized to 100 in January 2000.

Figure 2
Construction of neighborhoods



This figure visualizes our construction of neighborhoods. The center symbolizes a trade for a given home. Other trades in the same building are defined as trades in the first-order neighborhood. The dotted circles surrounding the center depict edges of mutually exclusive neighborhoods of orders two to six.

Figure 3
Identification of non-booming periods



The left panel of this figure depicts the evolution of the S&P US National House Price Index (solid line), the S&P Case-Shiller Condominium Index for New York City (dashed line) and the Manhattan Condominium Index (dotted line) constructed using the methodology of Case and Shiller (1989). Index levels are normalized to 100 in January 2000. The right panel depicts the absolute number of sales of apartments and condominiums on Manhattan Island from the first quarter of 2004 to the fourth quarter of 2015 after removing observations with missing values in sales prices, sales dates, and duplicates.

Table 1
Summary statistics

Variable name	Mean	Standard deviation
Annualized excess return	0.007	0.056
Holding period (in years)	5.407	2.666
Liquidity	168.709	100.86
Central Park view	0.029	0.167
Central Park walking	0.042	0.200
Very close subway	0.024	0.154
Close subway	0.185	0.389
Short distance Times Square	0.003	0.056
Medium distance Times Square	0.007	0.085
Short distance NYSE	0.013	0.114
Medium distance NYSE	0.019	0.135
Waterfront view	0.027	0.162
Waterfront walking distance	0.050	0.218
Dummy one year	0.018	0.134
Dummy two years	0.074	0.261
Price (in mio USD)	1.279	1.301
Seller corporation	0.103	0.304
Buyer corporation	0.156	0.363
Owner-occupied	0.531	0.499
Lagged GDP growth	0.005	0.005
Lagged unemployment growth	-0.007	0.018
Lagged interest change * 10,000	-1.935	300.379

This table provides descriptive statistics of the variables used. *Annualized excess returns* are defined in Equation (5). *Holding period (in years)* is the number of years between two trades of a given residential home. *Liquidity* is the number of sales during the past 180 days in the respective zip-code. *Central Park view* and *Central Park walking* are two dummies indicating whether a home has a view on the Central Park (distance of less than 100 feet beeline) and the city-block distance to the nearest entrance is less than 500 feet, respectively. *Very close subway* and *Close subway* are two mutually exclusive dummies indicating whether the city-block distance to the nearest subway entrance is less than 100 feet or 100 to less than 500 feet, respectively. *Short distance Times Square / NYSE* and *Medium distance Times Square / NYSE* are mutually exclusive dummies for whether the city-block distance to Times Squares / NYSE is less than 1,000 feet or 1,000 to less than 2,000 feet, respectively. *Water front view* is a dummy indicating whether a home has direct view on the water surrounding Manhattan Island. *Waterfront walking distance* is a dummy indicating whether the city-block distance to the waterfront does not exceed 500 feet. *Dummy one year* and *Dummy two years* are indicators for holding periods of one and two years, respectively. *Price (in mio USD)* is the most recent available prior trading price of the home CPI-adjusted to January 2015 dollars. *Seller/Buyer* corporation is a dummy indicating whether the seller/buyer is a corporation. *Owner-occupied* is a dummy indicating whether the buyer is the new inhabitant. *Lagged GDP growth* is the previous quarter's U.S. GDP growth. *Lagged unemployment growth* is the previous month's New York City wide unemployment growth rate. *Lagged interest change* is the percentage change of the average fixed mortgage lending rate in the month prior to the sale.

Table 2
Estimation results, base case

Variable name	(1)	(2)	(3)	(4)	(5)
First-order neighborhood	0.273***	0.254***	0.231***	0.229***	0.210***
	(0.015)	(0.015)	(0.014)	(0.014)	(0.013)
Second-order neighborhood	0.100***	0.064***	0.060***	0.056***	0.041***
	(0.010)	(0.011)	(0.010)	(0.010)	(0.010)
Third-order neighborhood	0.075***	0.043***	0.042***	0.039***	0.029**
	(0.011)	(0.011)	(0.010)	(0.010)	(0.010)
Fourth-order neighborhood	0.090***	0.053***	0.050***	0.048***	0.042***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Fifth-order neighborhood	0.074***	0.045**	0.037**	0.035**	0.025*
	(0.012)	(0.013)	(0.012)	(0.012)	(0.011)
Sixth-order neighborhood	0.051***	0.018	0.017	0.014	0.002
	(0.011)	(0.011)	(0.010)	(0.011)	(0.012)
ln(1+Liquidity)		-0.005***	-0.002***	-0.002**	-0.001
		(0.001)	(0.001)	(0.001)	(0.001)
Central Park view		0.006**	0.006*	0.006*	0.007*
		(0.002)	(0.002)	(0.002)	(0.003)
Central Park walking distance		0.000	-0.002	-0.002	-0.001
		(0.002)	(0.002)	(0.002)	(0.003)
Very close subway		0.003	0.003	0.003	0.003
		(0.002)	(0.002)	(0.002)	(0.002)
Close subway		0.003***	0.003***	0.003***	0.003***
		(0.001)	(0.001)	(0.001)	(0.001)
Short distance Times Square		-0.006*	-0.005*	-0.005	-0.007*
		(0.002)	(0.002)	(0.002)	(0.003)
Medium distance Times Square		0.002	-0.001	-0.001	-0.002
		(0.002)	(0.002)	(0.002)	(0.002)
Short distance NYSE		-0.002	-0.004	-0.004	-0.012*
		(0.004)	(0.004)	(0.004)	(0.005)
Medium distance NYSE		-0.011***	-0.010***	-0.010***	-0.016***
		(0.002)	(0.002)	(0.002)	(0.003)
Waterfront view		0.006**	0.004	0.004	0.000
		(0.002)	(0.003)	(0.003)	(0.003)
Waterfront walking distance		-0.008***	-0.005***	-0.005***	-0.003
		(0.001)	(0.001)	(0.001)	(0.002)
Dummy one year			0.060***	0.060***	0.050***
			(0.006)	(0.006)	(0.006)
Dummy two years			0.034***	0.033***	0.024***
			(0.002)	(0.002)	(0.002)
ln(Price)			-0.176***	-0.177***	-0.175***
			(0.017)	(0.018)	(0.017)
$\ln(\text{Price})^2/100$			0.618***	0.621***	0.612***
			(0.063)	(0.064)	(0.059)
Seller corporation			0.010***	0.010***	0.010***
			(0.002)	(0.002)	(0.002)
Buyer corporation			0.012***	0.012***	0.013***
			(0.001)	(0.001)	(0.001)
Owner-occupied			-0.001	-0.001*	-0.001*
			(0.001)	(0.001)	(0.001)
Lagged GDP growth				-0.001	-0.052
				(0.061)	(0.058)
Lagged unemployment growth				0.096***	0.061**
				(0.021)	(0.023)
Lagged interest change				-0.005	-0.013
				(0.009)	(0.010)
Fixed effects	no	zip	zip	zip	zip-year
Akaike criterion	-117,834	-118,571	-120,729	-120,771	-121,762

This table summarizes the results of Maximum Likelihood regressions explaining the annualized excess return of repeat sales relative to trades on Manhattan Island. The first-order neighborhood relates to trades in the same building. Second-, third-, fourth-, fifth-, and sixth-order neighborhoods have distances to the traded home of less than 500 feet, 500 to less than 707 feet, 707 to less than 866 feet, 866 to less than 1,000 feet, 1,000 to less than 1,118 feet, respectively. For further variable descriptions see Table 1. Fixed effects are on the zip-code level (zip) or the zip-code-year level (zip-year). Heteroskedasticity-robust standard errors are clustered over the zip-code-year level and reported in parentheses. ***, **, and * denote significance at the 0.1%, 1%, and 5% levels, respectively.

Table 3
Estimation results, temporal distance

Variable name	lag 1	lag 2	lag 3	lag 4	lag 5
First-order neigh.	0.153***	0.129***	0.096***	0.051***	0.060***
	(0.011)	(0.011)	(0.009)	(0.011)	(0.010)
Second & third-order neigh.	0.039***	0.025*	0.020	0.022*	0.014
	(0.012)	(0.010)	(0.012)	(0.010)	(0.010)
Fourth & fifth-order neigh.	0.026*	0.021*	0.020*	0.022*	-0.010
	(0.011)	(0.010)	(0.010)	(0.011)	(0.010)
Akaike criterion			-122,600		

This table summarizes the results of Maximum Likelihood regressions explaining the annualized excess return of repeat sales on Manhattan Island. The first-order neighborhood relates to trades in the same building. Second-, third-, fourth-, fifth-, and sixth-order neighborhoods have distances to the traded home of less than 500 feet, 500 to less than 707 feet, 707 to less than 866 feet, 866 to less than 1,000 feet, 1,000 to less than 1,118 feet, respectively. For the lags, mutually exclusive time intervals of 180 days are set for which the average excess returns are calculated, e.g., for lag 1 the sales in the most recent 180 days are used. The regression includes all control variables as in column (5) of Table 6. Fixed effects are on the zip-code level (zip) or the zip-code-year level (zip-year). Heteroskedasticity-robust standard errors are clustered over the zip-code-year level and reported in parentheses. ***, **, and * denote significance at the 0.1%, 1%, and 5% levels, respectively.

Table 4
Estimation results, cycle dependencies

	Market trend			
	Panel A: Macro		Panel B: Micro	
Neighborhood	Boom	Non-boom	Above avg.	Below avg.
First-order neigh.	0.186***	0.239***	0.166***	0.323***
	(0.018)	(0.016)	(0.016)	(0.022)
Second-order neigh.	0.039*	0.043***	0.020	0.096***
	(0.017)	(0.013)	(0.013)	(0.022)
Third-order neigh.	0.033*	0.026*	0.032**	0.017
	(0.015)	(0.012)	(0.011)	(0.025)
Fourth-order neigh.	0.044***	0.040***	0.040***	0.046*
	(0.012)	(0.012)	(0.011)	(0.021)
Fifth-order neigh.	0.030	0.020	0.029*	0.012
	(0.018)	(0.013)	(0.013)	(0.023)
Sixth-order neigh.	0.000	0.003	0.005	-0.011
	(0.020)	(0.013)	(0.015)	(0.022)
LR test (p-value)	0.000		0.000	
Akaike criterion	-121,763		-121,819	

This table summarizes Maximum Likelihood regression results on two separate regressions explaining the annualized excess return of repeat sales relative to trades on Manhattan Island. The two regressions depict explanatory power of neighboring excess returns conditional on being in a specific phase of a macro, and a micro cycle. For the macro cycle, we define a boom (January 2004 to October 2005, and March 2015 to December 2015) and a non-boom (November 2005 to February 2013) period. For the micro cycle, we define an above (below) average local market by a positive (negative) average excess return in the past $T = 180$ days. The first-order neighborhood relates to trades in the same building. Second-, third-, fourth-, fifth-, and sixth-order neighborhoods have distances to the traded home of less than 500 feet, 500 to less than 707 feet, 707 to less than 866 feet, 866 to less than 1,000 feet, 1,000 to less than 1,118 feet, respectively. The locational, transaction-specific and macro-financial control variables used for the two regressions are defined in Section 3.3. Fixed effects are on the zip-code-year level. The Likelihood Ratio test (LR test) is a test of joint equality of neighborhood coefficients, i.e. under the hypothesis that $\rho_{1,m} = \rho_{1,nm}, \ldots, \rho_{6,m} = \rho_{6,nm}$, where m (nm) denotes being (not being) in micro or macro phase m, respectively. Heteroskedasticity-robust standard errors are clustered over the zip-code-year level and reported in parentheses. ***, **, and * denote significance at the 0.1%, 1%, and 5% levels, respectively.

Table 5
Estimation results, cross-cycle dependencies

Macro level	Boom		Non-boom	
Micro level	Above avg.	Below avg.	Above avg.	Below avg.
First-order neigh.	0.102***	0.363***	0.232***	0.258***
	(0.023)	(0.029)	(0.020)	(0.030)
Second-order neigh.	-0.004	0.123***	0.035*	0.067*
	(0.021)	(0.031)	(0.016)	(0.029)
Third-order neigh.	0.040*	0.013	0.025	0.022
	(0.018)	(0.035)	(0.014)	(0.032)
Fourth-order neigh.	0.035*	0.068**	0.044**	0.016
	(0.017)	(0.025)	(0.015)	(0.033)
Fifth-order neigh.	0.031	0.027	0.025	-0.004
	(0.021)	(0.031)	(0.014)	(0.032)
Sixth-order neigh.	-0.005	0.004	0.011	-0.032
	(0.027)	(0.028)	(0.016)	(0.034)
Akaike criterion		-121,763		

This table summarizes Maximum Likelihood regression results for a single regression explaining the annualized excess return of repeat sales relative to trades on Manhattan Island. The average past excess returns within neighborhoods are divided according to the current state of the macro and micro cycle in which each transaction was settled. The macro state describes booming and non-booming periods of the Manhattan Island market, where the non-booming period is set to November 2005 to February 2013. Above (below) average returns on the micro level are defined as average excess returns in a neighborhood from the past $T = 180$ days being positive (negative). The first-order neighborhood relates to trades in the same building. Second-, third-, fourth-, fifth-, and sixth-order neighborhoods have distances to the traded home of less than 500 feet, 500 to less than 707 feet, 707 to less than 866 feet, 866 to less than 1,000 feet, 1,000 to less than 1,118 feet, respectively. The locational, transaction-specific and macro-financial control variables used in this regression are defined in Section 3.3. Fixed effects are on the zip-code-year level (zip-year). Heteroskedasticity-robust standard errors are clustered over the zip-code-year level and reported in parentheses. ***, **, and * denote significance at the 0.1%, 1%, and 5% levels, respectively.

Table 6
Robustness, base case

Neighborhood order	First	Second	Third	Fourth	Fifth	Sixth
Panel A: Base case	0.210***	0.041***	0.029**	0.042***	0.025*	0.002
	(0.013)	(0.010)	(0.010)	(0.009)	(0.011)	(0.012)
Panel B: Varying computation of excess returns in neighborhoods						
$T = 240$	0.222***	0.047***	0.027*	0.042***	0.031**	-0.003
	(0.012)	(0.011)	(0.011)	(0.009)	(0.011)	(0.014)
$T = 120$	0.186***	0.030**	0.032***	0.042***	0.019	-0.017
	(0.014)	(0.010)	(0.010)	(0.008)	(0.010)	(0.009)
Panel C: Varying neighborhood definitions						
333 feet	0.211***	0.037***	0.015	0.037***	0.023**	0.029***
	(0.013)	(0.010)	(0.009)	(0.009)	(0.009)	(0.009)
0.1, 0.25 miles	0.210***	0.053***	0.072***			
	(0.013)	(0.010)	(0.018)			
NYC neighborhoods	0.211***	0.170***	-0.009			
	(0.013)	(0.036)	(0.052)			
Panel D: Varying maximum holding period						
Seven years	0.206***	0.03**	0.014	0.034**	0.009	0.000
	(0.013)	(0.010)	(0.009)	(0.011)	(0.012)	(0.010)
Ten years	0.210***	0.039***	0.024**	0.036***	0.022*	0.003
	(0.012)	(0.010)	(0.009)	(0.009)	(0.010)	(0.010)
Panel E: City block and waterfront						
City block metric	0.210***	0.039***	0.040***	0.041***	0.025**	0.014
	(0.013)	(0.010)	(0.009)	(0.011)	(0.009)	(0.010)
Exclude waterfront obs.	0.208***	0.046***	0.032**	0.034***	0.025*	-0.003
	(0.013)	(0.011)	(0.011)	(0.009)	(0.011)	(0.013)

This table documents the robustness of our key results with respect to various assumptions. Panel B presents results when varying the definition of T, the maximum number of past days used to compute average excess returns in the neighborhood. Panel C presents results for different neighborhood definitions. In the row "333 feet", the second-order neighborhood is defined by a maximum distance of 333 feet. The subsequent neighborhoods are defined such that the area within each neighborhood is the same as in the second-order, yielding borders of 470, 576, 666, and 744 feet. In the row marked "0.1, 0.25 miles", the second- and third-order neighborhoods are defined by maximum distances of 0.1 and 0.25 miles from the traded home, i.e., 528 and 1,320 feet, respectively. "NYC Neighborhoods" depicts results for the case in which the second-order neighborhood is the neighborhood as defined by the City of New York (e.g., Chinatown, Lower East Side, etc.) and the third-order neighborhood are the neighborhoods adjacent to the second-order neighborhood. In Panel D, observations with a holding period of more than seven or ten years, respectively, are excluded. Panel E shows results for a change in the distance measure to the city block metric, and when excluding observations for which the waterfront lies within at least the sixth-order neighborhood (i.e., 1,118 feet). All regressions include the entire set of controls: locational, transaction-specific, and macro-financial. Fixed effects are on the zip-code-year level. Heteroskedasticity-robust standard errors are clustered over the zip-code-year level and reported in parentheses. ***, **, and * denote significance at the 0.1%, 1%, and 5% levels, respectively.

Table 7
Robustness, boom versus non-boom

	Boom						Non-boom					
Neighborhood order	First	Second	Third	Fourth	Fifth	Sixth	First	Second	Third	Fourth	Fifth	Sixth
Panel A: Base case												
	0.186***	0.039*	0.033*	0.044***	0.030	0.000	0.239***	0.043***	0.026*	0.040***	0.020	0.003
	(0.018)	(0.017)	(0.015)	(0.012)	(0.018)	(0.020)	(0.016)	(0.013)	(0.012)	(0.012)	(0.013)	(0.013)
Panel B: Varying non-boom periods												
03/2006 - 04/2012	0.189***	0.044**	0.036**	0.039***	0.036*	0.011	0.246***	0.038**	0.019	0.043**	0.007	-0.009
	(0.016)	(0.015)	(0.014)	(0.012)	(0.015)	(0.018)	(0.019)	(0.013)	(0.013)	(0.014)	(0.014)	(0.012)
03/2006 - 02/2012	0.187***	0.043**	0.036**	0.040***	0.038*	0.01	0.254***	0.040**	0.018	0.041***	0.003	-0.009
	(0.016)	(0.014)	(0.014)	(0.011)	(0.015)	(0.018)	(0.019)	(0.014)	(0.013)	(0.013)	(0.014)	(0.012)
10/2008 - 03/2012	0.199***	0.037**	0.033**	0.042***	0.026*	0.004	0.251***	0.055**	0.016	0.042*	0.020	-0.006
	(0.014)	(0.012)	(0.012)	(0.011)	(0.013)	(0.014)	(0.026)	(0.018)	(0.016)	(0.017)	(0.019)	(0.016)

This table documents the robustness of our key findings with respect to various ways to define boom and non-boom periods in Panel B. Panel A repeats the results for our base-case parameter setting in which boom and non-boom periods (November 2005 to February 2013) are defined using our Manhattan Condominium index, based on Case and Shiller (1989). The row "03/2006 - 04/2012" depicts results for the case in which the non-boom period is set to its counterpart in the S&P/CS NYC Condominium index, i.e., the time period March 2006 to April 2012; the row "03/2006 - 02/2012" for the case in which the non-boom period is set to its counterpart in the S&P/CS US National Home Price index, i.e., the time period from March 2006 to February 2012. The row "10/2008 - 03/2012" reports results when defining the non-boom period using the liquidity-dry-up-period from Figure 3, i.e., October 2008 to March 2012. All regressions include the entire set of controls: locational, transaction-specific, and macro-financial. Fixed effects are on the zip-code-year level. Heteroskedasticity-robust standard errors are clustered over the zip-code-year level and reported in parentheses. ***, **, and * denote significance at the 0.1%, 1%, and 5% levels, respectively.

Appendix

CLUSTERING OF ZIP-CODES

The zip-codes have been clustered the following way:

- 10001 & 10011 (Chelsea and Clinton)
- 10002 & 10003 & 10009 (Lower East Side)
- 10004 & 10005 & 10006 & 10007 & 10038 & 10280 & 10282 (Lower Manhattan)
- 10012 & 10013 (Greenwich Village/Lower Manhattan)
- 10017 & 10163 (Gramercy Park and Murray Hill)
- 10018 & 10019 & 10036 & 10129 (Chelsea and Clinton)
- 10023 & 10069 (Upper West Side)
- 10026 & 10027 & 10030 & 10037 & 10039 (Central Harlem)
- 10029 & 10035 & 10128 (East Harlem, 10128 is Upper East)
- 10031 & 10032 & 10033 & 10034 & 10040 (Inwood and Washington Heights)

ONLINE APPENDIX

Table 8
Alternatively clustered standard errors

Neighborhood order	Zip-code-year		Two-way		Zip-code		Year	
	s.e.	t-stat	s.e.	t-stat	s.e.	t-stat	s.e.	t-stat
First-order neigh.	0.013	16.662	0.021	10.023	0.019	11.079	0.015	14.033
Second-order neigh.	0.010	4.011	0.010	4.153	0.012	3.461	0.008	5.191
Third-order neigh.	0.010	2.892	0.010	2.880	0.011	2.618	0.009	3.200
Fourth-order neigh.	0.009	4.848	0.008	5.203	0.011	3.784	0.005	8.326
Fifth-order neigh.	0.011	2.223	0.010	2.463	0.012	2.053	0.008	3.079
Sixth-order neigh.	0.012	0.107	0.014	0.089	0.014	0.089	0.011	0.114

This table presents standard errors (s.e.) and corresponding t-statistics (t-stat) for different ways of clustering. The estimated model is our base case, presented in column (5) of Table 2. Column "Zip-code-year" serves for ease of comparison and presents results when clustering over zip-code-year level as used throughout the paper. Column "Two way" shows results when applying the two-way clustering by Cameron et al. (2011) with zip-code and year dimension. Columns "Zip-code" and "Year" depict standard errors and t-statistics when clustering over zip-code and year, respectively.

Table 9
Evidence on substitution: co-operatives

Neighborhood	Base case returns	Co-operative returns
First-order neigh.	0.23***	-0.064*
	(0.018)	(0.026)
Second-order neigh.	0.024*	0.093***
	(0.012)	(0.021)
Third-order neigh.	0.019	0.048*
	(0.012)	(0.022)
Fourth-order neigh.	0.031***	0.07**
	(0.01)	(0.023)
Fifth-order neigh.	0.018	0.037
	(0.012)	(0.021)
Sixth-order neigh.	-0.001	0.013
	(0.013)	(0.022)
AIC	-121812.08	

This table documents results for a single regression explaining annualized excess returns of homes. The column "Base case returns" depicts estimates for the average excess returns in each neighborhood as for the base case results in Table 2. The column "Co-operative returns" shows estimates for the comovement of average excess returns of co-operatives with the subsequent excess returns of single co-operative units. We define a building as co-operative if its land use is flagged as co-operative. All regressions include the entire set of controls: locational, transaction-specific, and macro-financial. Fixed effects are on the zip-code-year level. Heteroskedasticity-robust standard errors are clustered over the zip-code-year level and reported in parentheses. ***, **, and * denote significance at the 0.1%, 1%, and 5% levels, respectively.

Table 10
Estimation results, time dimension with six neighborhoods

Neighborhood	lag 1	lag 2	lag 3	lag 4	lag 5
First-order neigh.	0.153***	0.129***	0.095***	0.051***	0.06***
	(0.011)	(0.011)	(0.009)	(0.011)	(0.01)
Second-order neigh.	0.027**	0.018*	0.015	0.015	0.007
	(0.01)	(0.008)	(0.009)	(0.008)	(0.009)
Third-order neigh.	0.011	0.015	0.01	0.007	0.024*
	(0.009)	(0.009)	(0.01)	(0.009)	(0.01)
Fourth-order neigh.	0.025**	0.025**	0.018*	0.017*	-0.009
	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)
Fifth-order neigh.	0.012	0.011	0.01	0.001	-0.003
	(0.011)	(0.008)	(0.009)	(0.009)	(0.009)
Sixth-order neigh.	-0.003	0.017	-0.011	-0.004	0.018*
	(0.011)	(0.009)	(0.009)	(0.008)	(0.008)
AIC			-122573.75		

This table documents the robustness of our results on the time dimension. In contrast to the base case results reported in Table 3, we report results when keeping the neighborhoods of order one to six separated. The lags are defined in mutually exclusive intervals of $T = 180$ days. The regression includes the entire set of controls: locational, transaction-specific, and macro-financial. Fixed effects are on the zip-code-year level. Heteroskedasticity-robust standard errors are clustered over the zip-code-year level and reported in parentheses. ***, **, and * denote significance at the 0.1%, 1%, and 5% levels, respectively.

Table 11
Further robustness checks, base case

Neighborhood	(1) Risk-adjusted	(2) Zip-code restriction
First-order neigh.	0.194***	0.211***
	(0.013)	(0.013)
Second-order neigh.	0.034***	0.039***
	(0.01)	(0.01)
Third-order neigh.	0.029***	0.026**
	(0.009)	(0.01)
Fourth-order neigh.	0.037***	0.037***
	(0.008)	(0.008)
Fifth-order neigh.	0.025*	0.027*
	(0.01)	(0.012)
Sixth-order neigh.	0.003	-0.006
	(0.011)	(0.012)
AIC	89440.209	-121752.97

This table documents further robustness checks on our key results. Column (1) presents the results when risk-adjusting the excess returns by the annualized standard deviation of the monthly market return between the respective prior sale date and the second sale date. In column (2), excess returns of neighbors are restricted to be in the same zip-code as the corresponding observation. All regressions include the entire set of controls: locational, transaction-specific, and macro-financial. Fixed effects are on the zip-code-year level. Heteroskedasticity-robust standard errors are clustered over the zip-code-year level and reported in parentheses. ***, **, and * denote significance at the 0.1%, 1%, and 5% levels, respectively.

References

Amromin, G., J. C. Huang, C. Sialm, and E. Zhong, 2018, "Complex Mortgages," *Review of Finance*, 22(6), 1975–2007.

Anenberg, E., 2016, "Information Frictions and Housing Market Dynamics," *International Economic Review*, 57(4), 1149–1479.

Anenberg, E., and E. Kung, 2014, "Estimates of the Size and Source of Price Declines Due to Nearby Foreclosures," *American Economic Review*, 104(8), August, 2527–2551.

Ang, A., and J. Chen, 2002, "Asymmetric Correlations of Equity Portfolios," *Journal of Financial Economics*, 63(3), 443 – 494.

Autor, D. H., C. J. Palmer, and P. A. Pathak, 2014, "Housing Market Spillovers: Evidence from the End of Rent Control in Cambridge, Massachusetts," *Journal of Political Economy*, 122(3), 661–717.

Bhattacharya, U., D. J. Huang, and K. M. Nielsen, 2017, "Spillovers in Asset Prices: The Curious Case of Haunted Houses," working paper, 8th Miami Behavioral Finance Conference 2017, Available at SSRN: `https://ssrn.com/abstract=307795`.

Cameron, A. C., J. B. Gelbach, and D. L. Miller, 2011, "Robust Inference With Multiway Clustering," *Journal of Business and Economic Statistics*, 29(2), 238–249.

Campbell, J. Y., S. Giglio, and P. Pathak, 2011, "Forced Sales and House Prices," *American Economic Review*, 101(5), 2108–2131.

Can, A., 1990, "The Measurement of Neighborhood Dynamics in Urban House Prices," *Economic Geography*, 66(3), 254–272.

Caplin, A., and J. Leahy, 2011, "Trading Frictions and House Price Dynamics," *Journal of Money, Credit and Banking*, 43(7), 283–303.

Case, K. E., and C. J. Mayer, 1996, "Housing Price Dynamics within a Metropolitan Area," *Regional Science and Urban Economics*, 26(3), 387–407.

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

Chambers, M., C. Garriga, and D. E. Schlagenhauf, 2009, "Accounting for Changes in the Homeownership Rate," *International Economic Review*, 50(3), 677–726.

Chay, K. Y., and M. Greenstone, 2005, "Does Air Quality Matter? Evidence from the Housing Market," *Journal of Political Economy*, 113(2), 376–424.

Cotter, J., S. Gabriel, and R. Roll, 2015, "Can Housing Risk Be Diversified? A Cautionary Tale from the Housing Boom and Bust," *Review of Financial Studies*, 38(3), 913–936.

Coval, J., and T. Moskowitz, 1999, "Home Bias at Home – Local Equity Preference in Domestic Portfolios," *Journal of Finance*, 54(6), 2045–2073.

Diamond, R., and T. McQuade, 2019, "Who Wants Affordable Housing in Their Backyard? An Equilibrium Analysis of Low-Income Property Development," *Journal of Political Economy, forthcoming.*

Garmaise, M., and T. Moskowitz, 2004, "Confronting Information Asymmetries: Evidence from Real Estate Markets," *Review of Financial Studies*, 17(2), 405–437.

Gerardi, K., E. Rosenblatt, P. S. Willen, and V. Yao, 2015, "Foreclosure Externalities: New Evidence," *Journal of Urban Economics*, 87, 42–56.

Guerrieri, V., D. Hartley, and E. Hurst, 2013, "Endogenous Gentrification and Housing Price Dynamics," *Journal of Public Economics*, 100, 45–60.

Gupta, A., 2019, "Foreclosure Contagion and the Neighborhood Spillover Effects of Mortgage Defaults," *Journal of Finance*, 74(5), 2249–2301.

Guren, A. M., and T. J. McQuade, 2019, "How Do Foreclosures Exacerbate Housing Downturns?" Working Paper 26216, National Bureau of Economic Research, August.

Harding, J. P., E. Rosenblatt, and V. W. Yao, 2009, "The Contagion Effect of Foreclosed Properties," *Journal of Urban Economics*, 66(3), 164–178.

Horowitz, J. L., 1992, "The Role of the List Price in Housing Markets: Theory and an Econometric Model," *Journal of Applied Econometrics*, 7(2), 115–129.

Kallberg, J. G., C. H. Liu, and P. Pasquariello, 2014, "On the Price Comovement of US Residential Real Estate Markets," *Real Estate Economics*, 42(1), 71–108.

Kurlat, P., and J. Stroebel, 2015, "Testing for Information Asymmetries in Real Estate Markets," *Review of Financial Studies*, 28(8), 2429–2461.

Landvoigt, T., M. Piazzesi, and M. Schneider, 2015, "The Housing Market(s) of San Diego," *American Economic Review*, 105(4), 1371–1407.

Murfin, J., and R. Pratt, 2019, "Comparables Pricing," *Review of Financial Studies*, 32(2), 688–737.

Rossi-Hansberg, E., P.-ï. Sarte, and R. Owens III, 2010, "Housing Externalities," *Journal of Political Economy*, 118(3), 485–535.

S&P Dow Jones Indices, 2017, "S&P CoreLogic Case-Shiller Home Price Indices Methodology," Available at `http://us.spindices.com/documents/methodologies/methodology-sp-cs-home-price-indices.pdf?force_download=true`.

Szumilo, N., 2018, "Prices of Peers: Identifying Endogenous Price Effects in the Housing Market," working paper, Available at SSRN: `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3098073`.