

Local House Price Comovements

Fischer, Marcel; Füss, Roland; Stehle, Simon

Document Version

Accepted author manuscript

Published in:

Real Estate Economics

DOI:

[10.1111/1540-6229.12331](https://doi.org/10.1111/1540-6229.12331)

Publication date:

2021

License

Unspecified

Citation for published version (APA):

Fischer, M., Füss, R., & Stehle, S. (2021). Local House Price Comovements. *Real Estate Economics*, 49(Supplement S1), 169-198. <https://doi.org/10.1111/1540-6229.12331>

[Link to publication in CBS Research Portal](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please contact us (research.lib@cbs.dk) providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 04. Jul. 2025

Local House Price Comovements

Marcel Fischer*

Roland Füss†

Simon Stehle‡

This version: July 28, 2020

*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, Research Associate at the Centre for European Economic Research (ZEW), Mannheim, Germany, Research Fellow at the Center for Real Estate and Environmental Economics, NTNU Business School, Trondheim, Norway; 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.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/1540-6229.12331

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 price comovement, even within a zip-code, is a highly local and persistent phenomenon. The strength of such 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.

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. These patterns can be explained by 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 de-control (Autor et al., 2014), that spills over from affected to unaffected properties. Similarly, a foreclosure 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, 2020). Even interior renovations, which are generally unobservable, can raise house prices in close proximity after transaction prices of the renovated homes become public (Szumilo, 2019). Despite all the progress that has been made in exploring specific sources of local house price comovements and spillovers, surprisingly little is known about their general dynamics, for instance with respect to the distance between traded homes, the time between trades, and the dependence on the market cycle.

Beyond specific events, comovements should be largely driven through two main channels. First, comovements can be caused by the information channel. Financing decisions as well as negotiations between buyers and sellers are influenced by different anchors for the value of a home, that typically rely on comparable sales. This valuation approach makes use of sales prices of homes that are similar in terms of location, other characteristics, and past sales date. Important anchors include appraised values by independent parties as well as Zillow's estimates for the marked values of homes, the so-called Zestimates. By affect-

ing anchors for the value of a home, past sales prices should indirectly affect current sales prices. In addition, past sales prices can also directly affect present sales prices by affecting both buyers' and sellers' behavior. For instance, 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).

Second, homes in the same neighborhood are of similar size and quality, and 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.

In this paper, we investigate the micro-level price dynamics of homes in urban areas using repeat sales. We evaluate the order of magnitude, the persistence, and the state dependence of these effects. Specifically, we ask whether outperformance or underperformance relative to other homes in a given zip-code is a locally concentrated and persistent phenomenon. For that purpose, we compute excess returns of homes as the difference between individual and county-wide returns over the respective period. We run regressions to explore excess comovements, i.e., whether excess returns on individual homes can be explained by excess returns of nearby recently traded homes – even after controlling for zip-code-year dynamics, as well as transaction specific, locational, and macroeconomic controls. In these regressions, recently traded homes are grouped by their physical and temporal distance to the traded home. This categorization allows us to explore the dynamics of excess comovements, in

particular, whether excess comovements die out with physical and temporal distance.

Our data covers repeat sales on Manhattan Island between 2004 and 2015 from the CoreLogic database. Manhattan Island is well suited 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 on the homepage of the New York City Department of Finance,² implying that information is easily available for all market participants.

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. 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. About 21% of past within-building trades are reflected in present trades. For other trades within 500 feet, the order of

¹<https://www.propertyshark.com/Real-Estate-Reports/2016/12/13/payments-manhattan-now-500k-almost-double-median-sale-price-us/> retrieved on May 18, 2020

²<http://www1.nyc.gov/site/finance/taxes/property-rolling-sales-data.page> retrieved on May 18, 2020

magnitude decreases to 4% and dies out for distances exceeding 1,000 feet. 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.

In addition to the spatial dimension, performance is also persistent on the temporal dimension, i.e., excess comovements exist over longer time horizons. 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. Similar to the generally higher level of correlation in markets with falling prices in stock markets (e.g., Ang and Chen, 2002), as well as housing markets on the aggregate level (e.g., Cotter et al., 2015), our results reveal that local excess comovements are stronger in markets with falling prices. 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. Our results thus suggest a higher heterogeneity in terms of local house price changes when house prices increase.

To further investigate which channel drives excess comovements, we exploit that after a home is sold, information about the trade is not immediately publicly available. More specifically, our data allows us to distinguish between the sales date at which the home was sold, and the recording date, after which information about the trade is publicly available. Splitting our sample of past transactions into recorded and non-recorded trades, we show that excess comovements are substantially stronger with past trades for which information is already publicly available. As the substitution channel would imply no difference in comovements between recorded and non-recorded sales, our results suggest that it is rather the information channel that is driving excess comovements.

Our work contributes to two important strands of literature. First, it contributes to the literature that documents the existence of excess comovements on the index level (e.g., Kallberg et al., 2014; Cotter et al., 2015; Landier et al., 2017; Cohen and Zabel, 2018), by showing that excess comovements also exist in the trading prices of individual homes on the micro level. Second, it contributes to the growing strand of literature investigating 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., 2019), 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, 2020) are important drivers of micro-level house price dynamics. In contrast to the work of Rossi-Hansberg et al. (2010), Campbell et al. (2011), Guerrieri et al. (2013), and Szumilo (2019), which focuses on spillovers related to specific events, we document that, potentially driven by the information channel, prices comove even in the absence of specific events.

This paper proceeds as follows: In Section “Data and Methodology”, we introduce our data as well as our empirical methodology. Section “Empirical Results” presents our results on excess comovements in residential house prices. Section “Robustness Analysis” documents the robustness of our results. The final section concludes. An Online Appendix (currently at the end of this document) provides additional robustness results.

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 is well suited to investigate excess comovements in residential house prices

³<https://www.corelogic.com/solutions/university-data-portal.aspx> retrieved on May 18, 2020

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 should be less affected by information asymmetries. Our period of investigation covers trades from January 2004 to December 2015. To compute realized returns on these trades, we use past trading prices ranging back to 2000.

Data Cleaning

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

⁴According 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

⁵In 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.

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 data set of 39,771 observations. To account for outliers coming e.g. from data errors or physical changes of the property, we remove observations in the third and 97th percentile of the annualized return distribution.⁶ Our final data set then consists of 37,385 observations.

[Figure 1 about here]

Figure 1 visualizes the distribution of these observations geographically.⁷ Unsurprisingly, there are no trades in park areas, such as in Central Park, or industrial areas, such as 207th Street Train Yard Facility in the northern part of Manhattan Island.

[Figure 2 about here]

Figure 2 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 2, 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 layoffs in the financial industry. These layoffs did not occur instantly when house prices on the national level started declining, but with a certain delay.

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

⁶The results are qualitatively robust to the removal of only one or two percent of each tail.

⁷The data for zip-code boundaries was obtained from the US Census Bureau, download link: <http://www2.census.gov/geo/tiger/GENZ2015> retrieved on October 30, 2019

⁸In Subsection “Excess Comovements over Market Cycles”, we use these differences in the general evolution of house prices to investigate whether comovements vary with the phase of the housing market cycle.

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

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

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 Subsection “Excess Comovements over Market Cycles”, are associated with differential excess comovements.

Likewise, by investigating excess returns over different holding periods, our analysis aims at highlighting the importance of the availability of information for residential house prices. That is, we conjecture that only excess returns that become observable via information about the publicly available trading price, affect excess returns of nearby homes, but not changes in value of untraded homes, for whom the new value does not become public due to the absence of a trade. We further test the role of information frictions more formally by comparing excess comovements at the sales and recording date. We exploit that information about the sales price does not become publicly available prior to the recording date.

Control Variables

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

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.

While 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. In addition, we construct a dummy variable indicating whether the seller is a corporation or not.⁹

⁹We 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, HSNB, 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.

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, 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 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.

In a similar fashion, 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. 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.¹¹

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.

¹⁰The geographic coordinates of the New York subway entries are from NYC Open Data (<https://opendata.cityofnewyork.us>) retrieved on May 18, 2020

¹¹We also investigated a setting in which we excluded our controls for the Central Park, New York Stock Exchange, and Times Square. These results are qualitatively identical to those reported in Table 2. They can be found in Table A4 in the Online Appendix.

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.

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 about here]

Table 1 summarizes key properties of our data. As to be expected, the annualized excess return is not significantly different from zero.¹² 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

¹²The small positive value reflects that the market return constructed using the Case-Shiller methodology weights observations unequally.

Manhattan Island are among the most expensive in the U.S. This high average transaction 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.

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 $K - 1$ 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.¹³

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.¹⁴ 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

¹³The total number of neighboring, recently transacted prices ranges from about 200,000 to 246,000 (each property can be a neighboring transaction multiple times) 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.

¹⁴In Section “Robustness Analysis”, we apply the city-block metric to demonstrate the robustness of our results.

every current trade.¹⁵ Figure 3 visualizes our construction of $K = 6$ neighborhoods for a specific property. For every neighborhood k , we define a neighborhood-specific excess return, $\bar{r}_{i,k}^e$ as the average of the observed excess returns in the T days prior to trade i .

To better illustrate this procedure, consider the following example: In the first-order neighborhood ($k = 1$), a home i that is sold at $t = 0$. Suppose that there are four other observed trades in that neighborhood, realized at $t = -181$, $t = -145$, $t = -25$, and $t = 5$, respectively. In our base case, we consider the past $T = 180$ days, such that when computing $\bar{r}_{i,1}^e$, we only consider the second and the third trade, realized at $t = -145$ and $t = -25$, respectively. The first trade lies too far in the past to be included in $\bar{r}_{i,1}^e$. The fourth trade was not yet realized when i was traded, i.e., this information was not yet available and is therefore not included in $\bar{r}_{i,1}^e$. The average excess return for the first-order neighborhood of observation i , $\bar{r}_{i,1}^e$, is then computed as the average of the two excess returns of trades two and three.

[Figure 3 about here]

We employ the following regression setup:¹⁶

$$r_{i,t,t-,z}^e = \alpha_z + \sum_{k=1}^K \rho_k \bar{r}_{i,k}^e + \delta_{a(t),z} - \delta_{a(t-),z} + X_{i,t} \beta + \epsilon_{i,t,t-,z}, \quad (2)$$

¹⁵Intuitively, not only the physical distance, measured in feet, may be important, but also the direction, reflecting that neighborhoods in one direction are more similar/attractive than in another direction. It would have been interesting to investigate this issue in more detail. However, with an average number of observations of 5.3 to 6.5 in the neighborhoods, homes are traded too infrequently to split our samples further up. Simultaneously, whether neighborhoods in one direction are more attractive/similar than in another direction may be challenging to decide for some trades. Yet, we do control for a few specific attractive places in the neighborhoods, namely for the water front, the Central Park, and New York Stock Exchange.

¹⁶Equation (2) can be easily rewritten in spatial econometrics notation because $\bar{r}_{i,k}^e$ 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. In our robustness analysis, we also contrast our main findings with those derived from a spatial autoregressive (SAR) model. The results are given in Table A7 in the Online Appendix.

in which $r_{i,t,t-,z}^e$ 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 zip-code-year fixed effect dummies, for which the subscripts $a(t)$ and $a(t-)$ refer to the years in which the trades took place, respectively. For example, if t corresponds to any day of sale 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 level intercepts.

The excess returns, $\bar{r}_{i,k}^e$, from past trades are (deliberately) based on transactions in close areas suggesting that multicollinearity between these returns may be an issue. Empirically, however, multicollinearity between returns in our neighborhoods is not an issue with correlations among the $\bar{r}_{i,k}^e$ s ranging from 0.047 to only 0.121.

Our specification has several advantages compared to a spatial autoregressive (SAR) model. First, it avoids the use of a large weighting matrix consisting of more than 37,000 individual properties, as it would be the case with a spatiotemporal autoregressive model as, e.g., in Tu et al. (2004). The inclusion of such a highly dimensionally weighting matrix is computationally burdensome and only manageable by either imposing thresholds for the proximity measure or by using a concentrated log likelihood function such as in the specification of a matrix exponential spatial (MESS) model. Second, our specification avoids the use of a log determinant with multiple spatial lags, which would be computationally demanding. Third, the exogenous variables can still be interpreted as marginal effects and not as first round effects without taking into account feedback and spillover effects as in the case of the SAR model. Finally, our model provides a simple way to incorporate the temporal distances as illustrated in Figure 4 and interaction terms to control for market cycles as in Subsection “Excess Comovements over Market Cycles”.

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 on the difference D between t and $t-$: $\text{Var}(\epsilon_{i,t,t-,z}) = \exp(\gamma_1 + \gamma_2 D)$ 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.¹⁷

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

Intuitively, 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. Comovements in residential house prices should then 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. The distance between two homes can be interpreted as a proxy for how well two homes can be substituted with each other. More precisely, the smaller the distance between the two homes, the more a price signal from a previous trade in one home should affect the price of the respective other. The order of magnitude of excess comovements should therefore decrease with increasing

¹⁷Our estimates are qualitatively robust to the homoskedastic case.

distance between homes.

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 as further dimension of excess comovements in local housing markets. Consequently, the order of magnitude of comovement should not only be affected by the spatial distance between homes, but also by the temporal distance of trades. With increasing distance in the temporal sense, prices of neighboring substitutes become less informative for contemporaneous price movements.

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, spatial distance) against timeliness for current market movements (i.e, temporal distance). In other words, if the physical distance between homes is small, agents should be willing to accept a greater time distance for a trade to be used as a reference.

In sum, we test three main 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.

Empirical Results

For our empirical analysis, we need to determine a few parameters for our model introduced in Subsection “Methodology”. Specifically, we need to choose a meaningful number of distinct neighborhoods. In particular, we want to understand whether excess comovements are strongest in the first-order neighborhood and whether they die out in more distant neighbor-

hoods. We set the number of neighborhoods to $K = 6$, which is a reasonable choice between the longest price-relevant distance and the identical sizes of neighborhoods.¹⁸

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 "Robustness Analysis".

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 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.

¹⁸Empirically, 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.

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).¹⁹

Excess Comovements and Spatial Distance. In this section, we put emphasis on the spatial dimension of price comovements. As outlined in Subsection “Methodology”, 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.

[Table 2 about here]

For all specifications in Table 2, the sharpest decline in excess comovements is observed

¹⁹Similarly to Campbell et al. (2011) who use census-tract-year clusters, we cluster standard errors over the zip-code-year level. In Table A2 in our Online Appendix (currently at the end of this document), we show that our results are robust to alternative ways of clustering.

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.²⁰ Second, within 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.²¹

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 gen-

²⁰To 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 buildings should be better substitutes than non-co-op homes. Our results in Table A3 in the Online Appendix (currently at the end of this document) reveal that excess comovements in second to fifth-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.

²¹A 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.

eral 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 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.

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 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 . 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. To facilitate the interpretation of our results, we present our results graphically in Figure 4 for five mutually exclusive time lags and the six neighborhoods from our base case (the regression output is reported in Table A1 in the Online Appendix).

[Figure 4 about here]

Figure 4 reveals 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 – 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 reveals that when spatial distance between properties is small, excess comovements die out slower in the temporal dimension.

Our results from Figure 4 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.

Test on Information Frictions. So far, our results reveal the existence of both economically and statistically significant excess comovements in residential house prices. However, it is important to understand in more detail what causes these excess comovements. As we already argued above, information frictions may play an important role in that regard.

To investigate the role of information frictions more formally, in this section we exploit two distinct dates that relate to each individual transaction: the recording date and the contract date. The contract date is the date, at which the contract was signed. The recording date, on the other hand, is the date, at which the information about the transaction is sent to the county's recorder's office. The difference between these two dates is important, because only information that is registered can be made publicly available. Consequently, market participants can only use the information on sales prices after their respective recording dates. That is, the different information available to the public at the sales and the recording date allow us to directly test to which extent information frictions play an important role for excess comovements in residential house prices. More specifically, if the information channel is in play, excess comovements should be weaker prior to the information about the trade becoming public, i.e., before the recording date. That is, the differences between sales and recording dates allow us to ask whether the actual trade on the sales date, or the information about the trade on the recording date are more important for excess comovements in residential house prices.

To test the importance of the information channel more formally, for every trade i , we

split the trades over the past $T = 180$ days in each neighborhood into two groups. First, we calculate an average excess return from all observations that have been recorded before home i was sold, i.e., trades that have been publicly available before home i was traded. Second, we calculate an average excess return of homes that have been traded, but not yet recorded. Only in the first set, information about the respective other trades is publicly available. Hence, if the information channel is in play, excess comovements should be stronger in the first set.

[Table 3 about here]

Table 3 reports results from regressions using the two different sets of average excess returns in the neighborhoods. The column marked “Recorded sales” reports results for the first set, in which excess returns are computed using all other past trades in the neighborhoods that have been recorded by the date of a home’s sales date. The column “Non-recorded sales” reports results for the second set, in which excess returns are computed using all other past trades in the neighborhoods that have been sold (but not yet recorded) by the date of a home’s sales date. From Table 3, excess comovements are substantially stronger for recorded than non-recorded sales. If information would not play any role, the effects should be of similar size - if ever, effects should be slightly stronger for non-recorded sales, reflecting that the excess returns computed from the non-recorded sales include even more recent trades, namely those trades that have already occurred, but are not yet recorded. Excess comovements for non-recorded sales are largely insignificant, except for the first-order neighborhood, which may reflect that within a building information about sales prices becomes available more quickly via informal channels. Overall, our results in this section reveal that variation in information frictions is important for explaining excess comovements in residential house prices.

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 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 2. 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.²²

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. 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.²³ Formally, we extend our empirical application from Equation

²²Using 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, retrieved on May 18, 2020), 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 Subsection “Robustness of Boom versus Non-Boom”, we document that our results are robust to all of these alternative specifications.

²³It is important to note, that a negative excess return does not necessarily mean a loss for the seller. A negative sign only indicates that the performance of the trade was smaller than the performance of the market.

(2) to:

$$r_{i,t,t-,z}^e = \alpha_z + \sum_{s=1}^2 \sum_{k=1}^K \rho_{k,s} \times \bar{r}_{i,k}^e \times \mathbb{1}_{t \cap s} + \delta_{a(t),z} - \delta_{a(t-),z} + X_{i,t} \beta + \epsilon_{i,t,t-,z}, \quad (3)$$

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 about here]

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 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. Therefore, 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. The observed asymmetry of local house price comovements can be interpreted as further evidence for the information channel, as the substitution channel would imply rather similar comovements for under- and overperforming areas.

A systematic investigation of local over- and underperformance is challenging, because the number of trades in the same building within a shorter time period is typically small. To investigate the persistence of local over- and underperformance, we therefore ask whether locally, on the census tract level, census tracts with local over- or underperformance are more likely to again over- or underperform in the following years. We find that conditional on a positive average excess return in the past year, the probability for a positive average excess return in the current year is 25% higher after controlling for zip-code fixed effects. An extended analysis with three lags of one year each, again controlling for zip-code fixed effects, leaves us with coefficients of 22.5% for one year, 11.7% for two years, which are both significant at the 1% level, and 3.5%, which is insignificant at the 10%, for three years (standard errors are clustered on the zip-code level). Consequently, local over- and underperformance seem to be systematic for up to two years. For three years, we no longer find a systematic effect. Hence, our results suggest that local over- and underperformance is a systematic phenomenon, which might reflect positive or negative attributes of the neighborhoods.

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.

[Table 5 about here]

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.

Accepted Article

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.

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.

Robustness Analysis

This section documents the robustness of our key findings with respect to various assumptions. Subsection “Robustness of Base Case Results” provides evidence for our base case parameter setting, in which we do not distinguish between boom and non-boom periods. Subsection “Robustness of Boom versus Non-Boom” provides results for different definitions of the boom and non-boom periods. Additional robustness results are placed in the Online Appendix (currently at the end of this document). This Online Appendix documents that our results are robust to alternative clustering of standard errors (Table A2), when nearby homes are part of the same cooperative and thereby better substitutes (Table A3), to the omission of controls related to Central Park, Times Square, and the New York Stock Exchange, risk-adjusting returns and restricting the neighborhoods to the same zip-code (Table A4), and accounting for school rezoning, new constructions, and renovations (Table A5). Furthermore, we investigate whether there is more explanatory power in a neighborhood’s excess return if trades in respective lower-order neighborhoods are unavailable (Table A6), and provide estimates of SAR coefficients for comparison (Table A7).

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.

[Table 6 about here]

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.²⁴

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 New York City neighborhoods yields strong comovements of within-neighborhood excess returns from the first- and second-order neighborhoods, but virtually no connection to adjacent neighborhoods, suggesting that in Manhattan, where adjacent neighborhoods are often very heterogeneous, the preferred locations of households have sharp boundaries. This finding suggests that attention to prices of recently transacted homes is especially attracted within the same neighborhoods. This pattern provides further evidence for the information channel rather than the substitution channel, for which no sharp comovement bounds at neighborhood borders should be

²⁴Data 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.

observed.

In Panel D, we restrict the maximum holding period to seven and ten years, respectively.²⁵ 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.²⁶ Further robustness checks can be found in Tables A4 and A5 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

²⁵For 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.

²⁶We 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.

that our key results on comovements in boom and non-boom periods remain robust when using different criteria to determine these two subperiods.

Robustness of Boom versus Non-Boom

So far, 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 index, our non-boom period lasts from November 2005 to February 2013. We further document 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.

[Figure 5 about here]

The left panel of Figure 5 depicts the evolution of real house price indices for Manhattan (dotted line), New York (dashed line), and the entire United States (solid line). As before, 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 5 depicts the number of sold apartments and condominiums on Manhattan Island after removing observations with missing values in sales prices, sale 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.

[Table 7 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 Subsection “Excess Comovements over Market Cycles”, 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 Table 4 that during non-booming periods, excess comovements are stronger for within-building trades. Otherwise, the role of spatial distance is similar for both phases.

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 2004 and 2015. In doing so, we also test information frictions more formally by comparing excess comovements at the sales date of a private transaction and the recording date, when the individual sales price becomes publicly available.

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.

Acknowledgements

The authors would like to thank Konstantin Spring for his excellent research assistance. We are grateful for the helpful comments and suggestions from the editor, Wenlan Qian, two anonymous reviewers, Martin Brown, João Cocco, Christian Riis Flor, Günther Franke, Jens Jackwerth, Winfried Koeniger, Frederic Menninger, Daniel Ruf, Nawid Siassi, Bertram Steininger, Yildiray Yildirim, seminar participants at the Universities of Konstanz, Odense, Regensburg, St. Gallen, the EBS Business School, and the NTNU Trondheim, as well as participants at the 2018 AREUEA National Conference in Washington DC, the 2018 EFMA Annual Meeting in Milan, the 2018 Conference on Decision Sciences in Konstanz, the 2018 ReCapNet Conference of the Centre of European Economic Research (ZEW) in Mannheim, the conference on Housing in the 21st Century at the University of Bonn, and the 2020 American Finance Association Annual Meeting in San Diego.

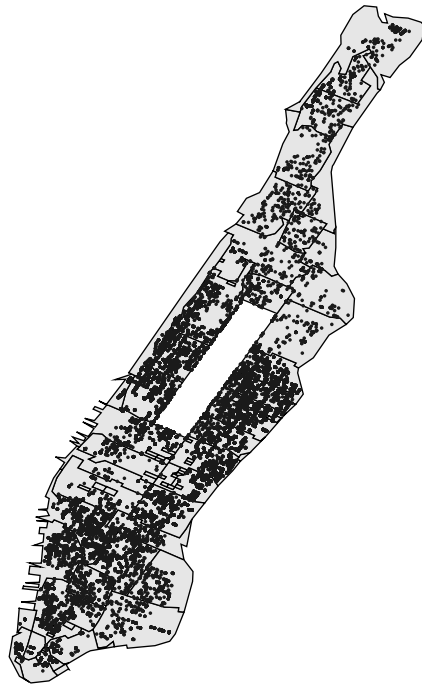
References

- Amromin, G., J. C. Huang, C. Sialm, and E. Zhong, 2018, “Complex Mortgages,” *Review of Finance*, 22(6), 1975–2007.
- Anenberg, E., and E. Kung, 2014, “Estimates of the Size and Source of Price Declines Due to Nearby Foreclosures,” *American Economic Review*, 104(8), 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, 2019, “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=3077951>.
- 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.
- Cohen, J. P., and J. Zabel, 2018, “Local House Price Diffusion,” *Real Estate Economics*, Forthcoming.
- 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.
- 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*, 127(3), 1063–1117.
- 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, 2020, “How Do Foreclosures Exacerbate Housing Downturns?” *Review of Economic Studies*, 87(3), 1331–1364.
- Harding, J. P., E. Rosenblatt, and V. W. Yao, 2009, “The Contagion Effect of Foreclosed Properties,” *Journal of Urban Economics*, 66(3), 164–178.

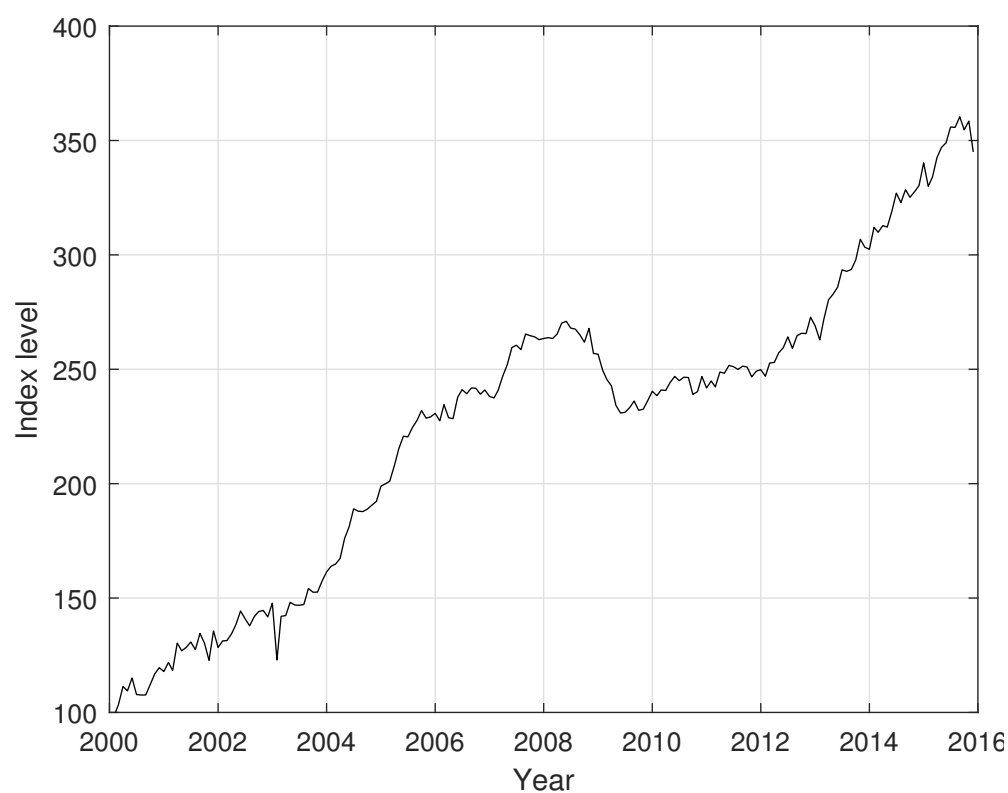
- Kallberg, J. G., C. H. Liu, and P. Pasquariello, 2014, “On the Price Comovement of US Residential Real Estate Markets,” *Real Estate Economics*, 42(1), 71–108.
- Landier, A., D. Sraer, and D. Thesmar, 2017, “Banking Integration and House Price Comovement,” *Journal of Financial Economics*, 125(1), 1–25.
- 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.-i. Sarte, and R. Owens III, 2010, “Housing Externalities,” *Journal of Political Economy*, 118(3), 485–535.
- Szumilo, N., 2019, “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.
- Tu, Y., S.-M. Yu, and H. Sun, 2004, “Transaction-Based Office Price Indexes: a Spatiotemporal Modeling Approach,” *Real Estate Economics*, 32(2), 297–328.

Figure 1
Geographic illustration of the data set



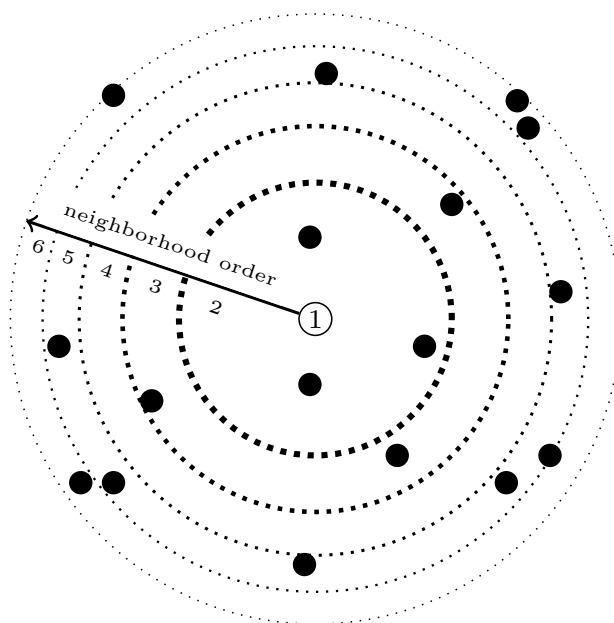
This graph visualizes the geographic distribution of our observations on Manhattan Island. The black lines indicate zip-code boundaries.

Figure 2
Evolution of house prices on Manhattan Island



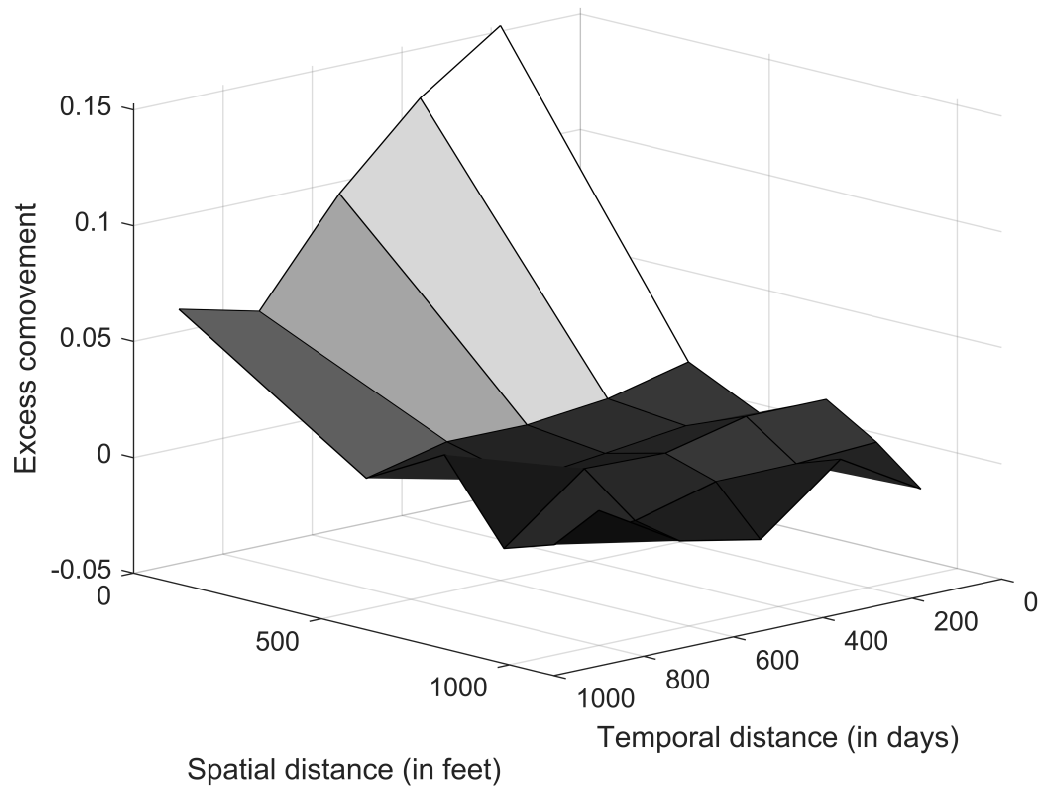
Nominal repeat sales index of Manhattan Island's condominium/apartment market based on our final data set. The index level is normalized to 100 in January 2000.

Figure 3
Construction of neighborhoods



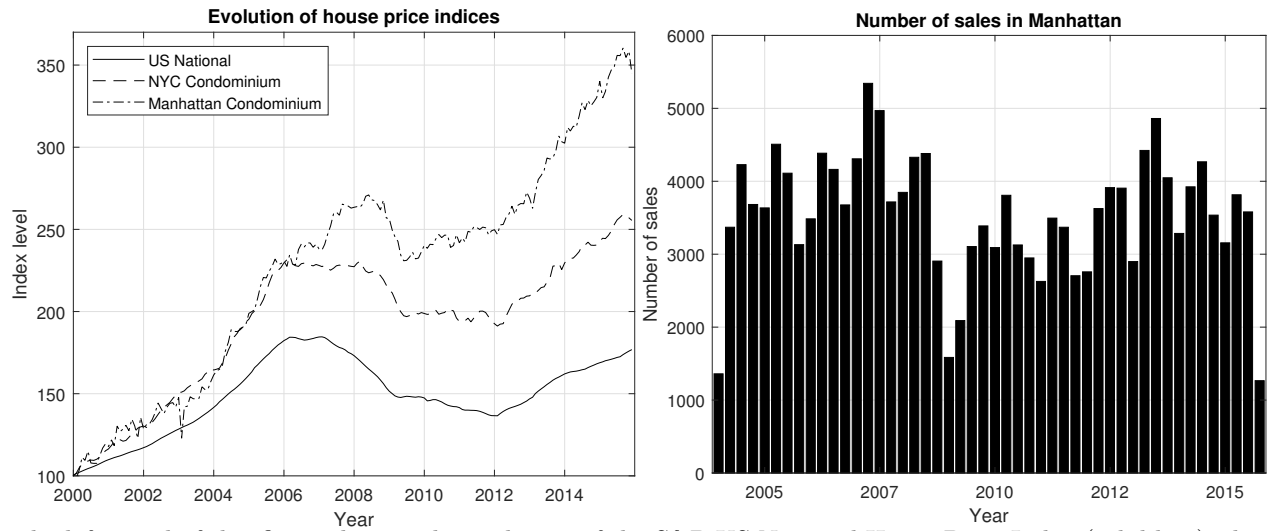
This figure visualizes our construction of neighborhoods. The center symbolizes a trade for a given home. Other trades in the same building are defined as trades in the first-order neighborhood. The dotted circles surrounding the center depict edges of mutually exclusive neighborhoods of orders two to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

Figure 4
Excess comovements over spatial and temporal distance



This figure visualizes excess comovements over spatial distance, measured in feet, and temporal distance, measured in days. The degree of excess comovements as given on the vertical axis is estimated by extending Equation (2) by five temporal lags, each of 180 days, such that 6×5 single excess comovement coefficients are estimated. Table A1 in the Online Appendix provides the exact regression coefficients with corresponding standard errors and significance tests.

Figure 5
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	0.01	0.06
Holding period (in years)	5.41	2.67
Liquidity	168.71	100.86
Central Park view	0.03	0.17
Central Park walking	0.04	0.20
Very close subway	0.02	0.15
Close subway	0.19	0.39
Short distance Times Square	0.003	0.06
Medium distance Times Square	0.01	0.09
Short distance NYSE	0.01	0.11
Medium distance NYSE	0.02	0.14
Waterfront view	0.03	0.16
Waterfront walking distance	0.05	0.22
Dummy one year	0.02	0.13
Dummy two years	0.07	0.26
Price (in mio USD)	1.28	1.30
Seller corporation	0.10	0.30
Buyer corporation	0.16	0.36
Owner-occupied	0.53	0.50
Lagged GDP growth	0.005	0.005
Lagged unemployment growth	-0.007	0.02
Lagged interest change * 10,000	-1.94	300.38

This table provides descriptive statistics of the variables used. *Annualized excess returns* are defined in Equation (1). *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	0.27*** (18.3)	0.25*** (17.3)	0.23*** (16.6)	0.23*** (16.4)	0.21*** (16.7)
Second-order neighborhood	0.10*** (9.7)	0.06*** (6.1)	0.06*** (6.0)	0.06*** (5.7)	0.04*** (4.0)
Third-order neighborhood	0.07*** (7.1)	0.04*** (3.9)	0.04*** (4.2)	0.04*** (4.0)	0.03** (2.9)
Fourth-order neighborhood	0.09*** (9.7)	0.05*** (6.0)	0.05*** (5.8)	0.05*** (5.5)	0.04*** (4.9)
Fifth-order neighborhood	0.07*** (5.9)	0.04*** (3.5)	0.04** (3.1)	0.03** (2.9)	0.02* (2.2)
Sixth-order neighborhood	0.05*** (4.8)	0.02 (1.6)	0.02 (1.6)	0.01 (1.3)	0.002 (0.1)
ln(1+Liquidity)		-0.01*** (-6.7)	-0.002*** (-3.5)	-0.002** (-2.7)	-0.001 (-0.8)
Central Park view		0.01** (2.6)	0.01* (2.4)	0.01* (2.4)	0.01* (2.5)
Central Park walking distance		0.0004 (0.2)	-0.002 (-0.7)	-0.002 (-0.7)	-0.001 (-0.2)
Very close subway		0.003 (1.6)	0.003 (1.5)	0.003 (1.5)	0.003 (1.6)
Close subway		0.003*** (3.5)	0.003*** (4.0)	0.003*** (4.0)	0.003*** (4.3)
Short distance Times Square		-0.01* (-2.3)	-0.01* (-2.0)	-0.005 (-1.9)	-0.01* (-2.5)
Medium distance Times Square		0.002 (0.7)	-0.001 (-0.4)	-0.001 (-0.5)	-0.002 (-0.9)
Short distance NYSE		-0.002 (-0.4)	-0.004 (-1.0)	-0.004 (-1.0)	-0.01* (-2.5)
Medium distance NYSE		-0.01*** (-5.2)	-0.01*** (-5.0)	-0.01*** (-4.7)	-0.02*** (-6.3)
Waterfront view		0.01** (2.8)	0.004 (1.4)	0.004 (1.4)	0.0005 (0.2)
Waterfront walking distance		-0.01*** (-5.3)	-0.005*** (-3.5)	-0.005*** (-3.3)	-0.003 (-1.7)
Dummy one year			0.06*** (10.6)	0.06*** (10.6)	0.05*** (8.7)
Dummy two years			0.03*** (17.4)	0.03*** (17.3)	0.02*** (12.9)
ln(Price)			-0.18*** (-10.9)	-0.18*** (-10.0)	-0.17*** (-11.3)
ln(Price) ² /100			0.62*** (10.7)	0.62*** (9.8)	0.61*** (11.1)
Seller corporation			0.01*** (5.4)	0.01*** (5.5)	0.01*** (5.7)
Buyer corporation			0.01*** (10.2)	0.01*** (10.2)	0.01*** (11.4)
Owner-occupied			-0.001 (-1.9)	-0.001* (-2.4)	-0.001* (-2.4)
Lagged GDP growth				-0.001 (-0.02)	-0.05 (-0.9)
Lagged unemployment growth				0.1*** (4.5)	0.06** (2.6)
Lagged interest change				-0.01 (-0.6)	-0.01 (-1.3)
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). The figures in parentheses report t -statistics derived from robust standard errors that are clustered over the zip-code-year level. ***, **, and * denote significance at the 0.1%, 1%, and 5% levels, respectively.

Table 3
Estimation results, test on information frictions

Variable name	Recorded sales	Non-recorded sales
First-order neighborhood	0.20*** (15.8)	0.09*** (7.1)
Second-order neighborhood	0.04*** (3.9)	0.005 (0.5)
Third-order neighborhood	0.03** (3.1)	0.01 (1.7)
Fourth-order neighborhood	0.04*** (4.4)	0.02* (2.4)
Fifth-order neighborhood	0.02* (2.0)	-0.02 (-1.7)
Sixth-order neighborhood	0.01 (0.5)	-0.001 (-0.1)
Fixed effects	zip-year	
Akaike criterion	-121,796	

This table summarizes the results of a single Maximum Likelihood regression explaining the annualized excess return of repeat sales relative to trades on Manhattan Island. The coefficients describe the comovement of individual trades with average excess returns realized within the past $T = 180$ days in neighborhoods of different orders. For this analysis, the average excess returns from the base case are divided into two categories. The first category are “Recorded sales”, i.e., sales that have been reported to the county’s recorder’s office (and are thus publicly available). The second category contains sales which have already been completed, but have not yet been recorded. 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 Subsection “Control Variables”. Fixed effects are on the zip-code-year level (zip-year). The figures in parentheses report t -statistics derived from robust standard errors that are clustered over the zip-code-year level. ***, **, and * denote significance at the 0.1%, 1%, and 5% levels,

Table 4
Estimation results, cycle dependencies

Neighborhood	Market trend			
	Panel A: Macro		Panel B: Micro	
	Boom	Non-boom	Above avg.	Below avg.
First-order neigh.	0.19*** (10.3)	0.24*** (14.6)	0.17*** (10.4)	0.32*** (14.7)
Second-order neigh.	0.04* (2.3)	0.04*** (3.4)	0.02 (1.6)	0.10*** (4.3)
Third-order neigh.	0.03* (2.1)	0.03* (2.1)	0.03** (2.8)	0.02 (0.07)
Fourth-order neigh.	0.04*** (3.7)	0.04*** (3.3)	0.04*** (3.5)	0.05* (2.2)
Fifth-order neigh.	0.03 (1.7)	0.02 (1.6)	0.03* (2.2)	0.01 (0.5)
Sixth-order neigh.	-0.001 (-0.1)	0.003 (0.3)	0.005 (0.3)	-0.01 (-0.5)
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 Subsection “Control Variables”. 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 null hypothesis that $\rho_{1,m} = \rho_{1,nm}, \dots, \rho_{6,m} = \rho_{6,nm}$, where m (nm) denotes being (not being) in micro or macro phase m , respectively. The figures in parentheses report t -statistics derived from robust standard errors that are clustered over the zip-code-year level. ***, **, 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.	0.10*** (4.5)	0.36*** (12.5)	0.23*** (11.4)	0.26*** (8.6)
Second-order neigh.	-0.004 (-0.2)	0.12*** (4.0)	0.03* (2.2)	0.07* (2.3)
Third-order neigh.	0.04* (2.2)	0.01 (0.4)	0.02 (1.8)	0.02 (0.7)
Fourth-order neigh.	0.03* (2.1)	0.07** (2.8)	0.05** (3.1)	0.02 (0.5)
Fifth-order neigh.	0.03 (1.5)	0.03 (0.9)	0.03 (1.8)	-0.01 (-0.2)
Sixth-order neigh.	-0.01 (-0.2)	0.004 (0.1)	0.01 (0.7)	-0.03 (-0.9)
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 Subsection “Control Variables”. Fixed effects are on the zip-code-year level (zip-year). The figures in parentheses report t -statistics derived from robust standard errors that are clustered over the zip-code-year level. ***, **, 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
<i>Panel A: Base case</i>	0.21*** (16.7)	0.04*** (4.0)	0.03** (2.9)	0.04*** (4.9)	0.02* (2.2)	0.002 (0.1)
<i>Panel B: Varying computation of excess returns in neighborhoods</i>						
$T = 240$	0.22*** (18.6)	0.05*** (4.2)	0.03* (2.4)	0.04*** (4.9)	0.03** (2.8)	-0.003 (-0.2)
$T = 120$	0.19*** (13.6)	0.03** (3.1)	0.03** (3.2)	0.04*** (5.3)	0.02 (1.8)	-0.02 (-1.8)
<i>Panel C: Varying neighborhood definitions</i>						
333 feet	0.21*** (16.7)	0.04*** (3.7)	0.02 (1.6)	0.04*** (4.1)	0.02** (2.6)	0.03** (3.2)
0.1, 0.25 miles	0.21*** (16.6)	0.05*** (5.4)	0.07*** (3.9)			
NYC neighborhoods	0.21*** (16.5)	0.17*** (4.7)	-0.01 (-0.2)			
<i>Panel D: Varying maximum holding period</i>						
Seven years	0.21*** (16.1)	0.03** (2.9)	0.01 (1.5)	0.03** (3.1)	0.01 (0.8)	-0.0001 (-0.01)
Ten years	0.21*** (16.9)	0.04*** (4.1)	0.02** (2.6)	0.04*** (3.8)	0.02* (2.2)	0.003 (0.2)
<i>Panel E: City block and waterfront</i>						
City block metric	0.21*** (16.6)	0.04*** (3.8)	0.04*** (4.4)	0.04*** (3.8)	0.03** (2.8)	0.01 (1.4)
Exclude waterfront obs.	0.21*** (15.9)	0.05*** (4.3)	0.03** (2.9)	0.03*** (3.6)	0.03* (2.2)	-0.003 (-0.3)

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. The figures in parentheses report t -statistics derived from robust standard errors that are clustered over the zip-code-year level. ***, **, and * denote significance at the 0.1%, 1%, and 5% levels, respectively.

Table 7
Robustness, boom versus non-boom

Neighborhood order	Boom						Non-boom					
	First	Second	Third	Fourth	Fifth	Sixth	First	Second	Third	Fourth	Fifth	Sixth
<i>Panel A: Base case</i>												
	0.19*** (10.3)	0.04* (2.3)	0.03* (2.1)	0.04*** (3.7)	0.03 (1.7)	-0.001 (-0.1)	0.24*** (14.6)	0.04*** (3.4)	0.03* (2.1)	0.04*** (3.3)	0.02 (1.6)	0.003 (0.3)
<i>Panel B: Varying non-boom periods</i>												
03/2006 - 04/2012	0.19*** (11.6)	0.04** (3.0)	0.04* (2.6)	0.04*** (3.4)	0.04* (2.4)	0.01 (0.5)	0.25*** (13.2)	0.04** (2.8)	0.02 (1.5)	0.04** (3.2)	0.01 (0.5)	-0.01 (-0.8)
03/2006 - 02/2012	0.19*** (11.8)	0.04** (2.9)	0.04** (2.6)	0.04*** (3.6)	0.04* (2.5)	0.01 (0.5)	0.25*** (13.4)	0.04** (2.9)	0.02 (1.3)	0.04** (3.2)	0.003 (0.2)	-0.01 (-0.8)
10/2008 - 03/2012	0.20*** (14.1)	0.04** (3.0)	0.03** (2.8)	0.04*** (3.9)	0.03* (2.0)	0.004 (0.3)	0.25*** (9.6)	0.06** (3.1)	0.02 (1.0)	0.04* (2.4)	0.02 (1.1)	-0.006 (-0.4)

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 5, 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. The figures in parentheses report t -statistics derived from robust standard errors that are clustered over the zip-code-year level. ***, **, 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)