To Measure Globally, Aggregate Locally: Home Price Index Consistency in the Presence of Asymmetric Sales and Appreciation

Thom Malone*[1] and Christian L. Redfearn†[2]

[1]Graduate School of Design, Harvard University
[2]Sol Price School of Public Policy, University of Southern California

December 2017

DRAFT ONLY - Do Not Cite Without Permission of the Author

Abstract

This paper highlights the role of housing submarkets in the construction of aggregate home price indexes. Submarkets within metropolitan areas exhibit asymmetric price appreciation and sales volumes. This produces biased aggregate house price indexes because because the sample they are based on is representative neither of the stock nor its appreciation. Our proposed solution is to estimate indexes by submarket and aggregate them by the share of stock they represent. Over the last housing cycle "local-pooling" indexes consistently find a lower peak, less severe crash, and stronger recovery when compared with conventional "global" indexes. The typical indexes appear to have oversampled lower priced homes relative to their share of the stock while these same submarkets rose and fell in price to a greater extent than the rest of the market. For the case of the Los Angeles metropolitan area, these indexes overstated swings in the value of the housing stock by as much as $500B dollars.

*tmalone@gsd.harvard.edu
†redfearn@usc.edu, corresponding author. Thanks to HomeUnion for their support with the Los Angeles data.

1 Introduction

Aggregate house price indexes are derived from a very small amount of information relative to the asset they represent. In the U.S. for example, only about seven percent of the existing owner-occupied housing stock sells in any given year. Challenged by a heterogeneous and infrequently traded good, researchers turn to commonly-used statistical tools to construct quality-adjusted housing price indexes. These weight all sales evenly, implicitly assuming that they are random draws from a common distribution. This assumption, however, is rarely considered critically in the construction of indexes – despite substantial evidence of distinct housing submarkets and significant idiosyncratic price movements across space and time [1]. The submarkets within a metro area do share a common labor market and other shared fundamentals that suggest pooling observations is appropriate. But if these data are not representative of the underlying stock and there is submarket variation in price appreciation these indexes will be biased.

We refer to indexes that first pool observations and then apply statistical techniques as "global-pooling" indexes [2]. These create an index by gathering all the observations within a metropolitan area and then applying statistical techniques to control for variation in the quality of housing. This is global in the sense that all observations are included in a joint estimation with common parameters.

This paper contrasts "global-pooling" with "local-pooling." That is, observations are first pooled within submarkets, indexes are constructed locally, and then aggregate indexes are constructed by weighting the local indexes by the share of the stock they represent. An obvious cost to this approach is the noise associated with smaller samples. Offsetting this is the fact that in urban areas changes in the location premium drive most of the movement in house prices.[3] Our approach to local pooling puts a premium on measuring changes in land and location. This is because housing and amenities are more homogeneous at smaller geographic levels like census tracts or zip codes. Moreover, local pooling allows for greater flexibility in estimating hedonic characteristics.

[1] (Bourassa, Hamelink, Hoesli, and MacGregor 1999), (Bourassa, Hoesli, and Peng 2003), (Clapp and Wang 2006), (Bhattacharjee, Castro, Maiti, and Marques 2016), (Keskin and Watkins 2017) are but a few examples on the topic.

[2] Commonly cited examples of these include S&P/Case-Shiller, LPS, FHFA, FNC, Core-Logic, Radar Logic, Clear Capital, and Zillow home value indexes.

[3] Indeed, we learned from the last housing cycle that we didn't experience a housing bubble, we experienced a land bubble. (see Bostic, Longhofer, and Redfearn (2007))

The underlying economic framework for this approach is straightforward. Tiebout long ago formalized an explanation for diverse public goods within a unified housing market. It's easy to extend the same intuition for local private goods – restaurants, retail, etc. These two sets of characteristics define the neighborhoods that are capitalized into home prices. Shocks to different types of households that occupy these varying neighborhoods allow for clustered housing submarket behavior: both prices and transactions can deviate systematically from the rest of other submarkets for extended periods of time. Other neighborhoods in the MSA are substitutes but imperfectly so.

For instance, the increase in the availability of subprime loans in the early 2000s resulted in lower-end households having greater access to credit. In turn both prices and sale volumes in these submarkets rose relative to other submarkets. In our results, it appears that these dynamics lead to an overestimate of aggregate house prices during the housing boom and an overestimate of the decline after the bust. These sorts of local shocks are common within metropolitan areas in which some particular innovation is manifest within a submarket. The result is that the sample of house sales may not reflect the stock as whole, yielding aggregate indexes that are biased.

We use two disparate data sets to explore these issues. The first is 3.2 million sales from DataQuick for the greater Los Angeles metropolitan area, covering all arm's-length house sales from 2000 to 2015. The second is nineteen years of Swedish housing transactions, numbering over one million raw sales. The two data sets provide an excellent opportunity to examine the impact that order of aggregation has on estimated prices as measured by the most commonly used indexes: median, repeat-sale, and hedonic. Both data sets cover periods of pronounced housing price changes, but with different contexts behind them. The Swedish data afford us a chance to examine variation nationally across markedly different metropolitan areas, while the Los Angeles data allow us to look very locally because each sale is geocoded.

We provide evidence of pervasive idiosyncratic price movements and asymmetric selection within housing markets. In some cases relative prices rise as sale propensities rise yielding over-representation of "winners" and an upward bias in measured aggregate prices; in other cases, the opposite holds. This means that local dynamics are at odds with the assumptions required for global-pooling indexes to recover aggregate price levels for the stock as opposed to the sample of sold dwellings. In the case of Los Angeles locally-pooled indexes show the peak of the 2000s boom is overestimated, and the bust was not nearly as large as in the global indexes suggest. We also see a stronger recovery in local indexes. These differences can be traced to fast-appreciating low end homes being over-sampled in the 2004-2008 period when a heavier proportion of loans in the market were subprime and the housing in the central city being consistently under-sampled. These differences are large enough that they shift the estimated value of housing wealth by $500 billion dollars.

The paper proceeds first by addressing formally the order of aggregation in the presence of idiosyncratic price movement. The data and stylized facts showing asymmetric price appreciation and nonrandom sampling are then discussed in detail in Section 3. Both the Swedish and Los Angeles results are presented in Section 4. In Section 5, we outline our interpretation of these results and discuss extensions.

2 Representativeness & Order of Aggregation

This section outlines the essential differences between aggregate price indexes that employ local- and global-pooling in their construction. Global-pooling indexes are those whose first step in construction is to pool all available observations within the geographic area of interest into a single data set. For this class of index construction methods, controlling for factors that may be confounded with price appreciation is left to the econometrician. For example, a hedonic index controls for quality variation explicitly by pricing it and removing it from the estimate of aggregate prices; repeat-sales indexes difference observations on identical dwellings to remove quality variation from prices.

This intuition for "global pooling" is based on the premise that the stochastic process that guides prices is the same for all dwellings and measured appreciation for any one dwelling is some mixture of true aggregate appreciation and noise. In this case appealing to the Law of Large Numbers is appropriate: the sample mean will tend to the true mean as the sample size increases. Another way of describing this premise is as an assumption of representativeness with regard to appreciation. So long as this is true any sample of homes could yield consistent estimates of aggregate prices.[4]

[4] In this case, "could" refers the usual caveats associated with index construction. Hedonic indexes, for example, typically depend on assumptions regarding functional form and the stability of attribute prices across time and space; repeat-sale indexes rely on strong assumptions about time-invariant attribute bundles and prices. There is an extensive literature on these indexes and their maintained assumptions. See especially JREFE 14(1-2), 1997 and

There are many plausible reasons why the house price appreciation could vary across housing units for reasons other than differences in the physical structures of houses. (Landvoigt, Piazzesi, and Schneider 2015) shows that house price appreciation in San Diego in the first half of the naughts was greater for houses with initially lower prices, presumably because of the expansion in 'sub-prime' mortgages during the time period. Works such as (Genesove and Mayer 2001) and (Anenberg 2011) find evidence that sales prices can differ according to the loan-to-value ratios on the property due to 'loss aversion' (the tendency for people to strongly prefer avoiding losses to realizing gains). Similarly, (Tenreyro, Bracke, et al. 2017) use data on house sales in England and Wales to show that sale prices are 'anchored' by the price the current owner paid.

Other studies find different rates of appreciation due to demand side factors such as buyer income (Case and Mayer 1996) and age (Ortalo-Magné and Rady 1999). The Alonso-Muth-Mills model of a monocentric city also indicates gives reason believe that appreciation could be different depending on how far a dwelling is from the city center (see (Bogin, Doerner, and Larson 2016) for empirical evidence of this.). The elasticity of housing supply could also vary within a metropolitan area, meaning that even if demand side shifters were homogeneous across all buyers, they would have different effects depending on where the house is located.

Evidence of distinct local price paths is provided in two papers on index construction techniques by Meese and Wallace (1991) and Goetzmann and Spiegel (1997). The former paper attempts to develop an alternative method for measuring aggregate prices that is free from many of the strong assumptions associated with the hedonic and repeat-sale approaches (assumptions they find violated). The latter paper presents local index construction methods. Both find significant idiosyncratic price movement at the local level (municipalities and zip codes, respectively).

One solution for the global-pooling indexes is to control for non-random selection. The process, analogous to labor market research by Hausman and Wise (1977), involves modeling the selection process and constructing an additional regressor that captures the likelihood of inclusion in the observed sample. In principle, this technique (or non-parametric variant thereof, e.g. Newey, Powell, and Walker (1990)) will produce price index estimates free of sample selection problems. Only a handful of applications of this type of approach have been undertaken in the housing

AREUEA Journal 19(3) 1991.

literature, most notably Gatzlaff and Haurin (1997), Englund, Quigley, and Redfearn (1999a), and (Hwang and Quigley 2004). More recently (Mason and Pryce 2011) has used a fractional probit regression to adapt this approach when only aggregated data housing submarkets is available This difficultly in this approach is that both prices and the selection process must be modeled accurately.

Hedonic and repeat-sales indexes are employed because the case has been well made that housing is a heterogeneous good and that quality variation across time and space can mislead users of less-sophisticated indexes. However, the accuracy of these more sophisticated indexes depends crucially on the representativeness of the sample. Typically, representativeness is assumed without verification. Without more careful consideration of the sample, it is not clear whether the parameter estimates from these techniques are generalizable. If, in fact, selection is non-random and price appreciation is asymmetric, these indexes will recover only the quality-controlled price indexes for the observed sample of sold homes but not for the entire stock.

Once asymmetric price levels are introduced, the appeal to the Law of Large Numbers does not hold unconditionally. To see this consider the following stylized model of a unified housing market. In it there are two submarkets, A and B, with populations N^A and N^B. The stock of dwellings is $N = N^A + N^B$. For the simplicity assume that these populations are constant over time. Furthermore, assume that dwelling quality is uniform and time-invariant. The issue to be highlighted in this paper is representativeness with regard to both selection and appreciation; attribute heterogeneity, while more realistic, serves only to complicate the model without adding any insight. As such, the usual conception of house prices as quotient of house prices and dwelling quality,

(1) $V_{it} = P_t Q_{it} + \varepsilon_{it}$,

is simplified. Here, V_{it} is the observed sale price, P_t is the aggregate price level, and ε_{it} is variation in price due to buyer and seller heterogeneity. i and t index dwelling and time, respectively. Q_{it} is the dwelling quality and, for the purposes of this paper, is assumed to be equal to one. In this homogeneous world, pricing for individual dwellings in the two submarkets is given by

(2) $V_{it} = \psi^A_{it} + \varepsilon_{it} \quad \forall\, i \in A$

(3) $V_{it} = \psi^B_{it} + \varepsilon_{it} \quad \forall\, i \in B$

where $\psi_{it}^A \sim (P_t^A, \sigma_{A,t}^2)$ and $\psi_{it}^B \sim (P_t^B, \sigma_{A,t}^2)$. This allows prices within each submarket to move around a common (within-submarket) trend. Aggregate prices then are a function of price movements within each submarket. True aggregate prices are defined as

$$(4) \quad P_t = \frac{N^A P_t^A + N^B P_t^B}{N^A + N^B}$$

Selection into the sample of observed dwelling sales may also asymmetric across submarkets. In this case, π_t^A and π_t^B are the sale probability is submarkets, A and B, respectively. The observed sample is then

$$(5) \quad n_t = n_t^A + n_t^B$$

where $n_t^A = \pi_t^A N^A$ and $n_t^B = \pi_t^B N^B$, the subsamples from each of the submarkets. Recall that quality-control issues have been taken off the table by assuming a homogeneous housing stock. Estimated aggregate prices are given by

$$(6) \quad \widehat{P}_t = \frac{1}{n_t} \sum_{i \in n_t} V_{it} = \frac{1}{n_t^A + n_t^B} \left(\sum_{i \in n_t^A} \psi_{it}^A + \sum_{i \in n_t^B} \psi_{it}^B + \sum_{i \in n_t} \varepsilon_{it} \right)$$

The error term ε has expectation zero, so

$$(7) \quad E[\widehat{P}_t] = \frac{n_t^A P_t^A + n_t^B P_t^B}{n_t^A + n_t^B}$$

If the estimate $\widehat{P}_t$ is to be unbiased, $E[\widehat{P}_t]$ must equal P_t, or

$$(8) \quad \frac{N^A P_t^A + N^B P_t^B}{N^A + N^B} = \frac{n_t^A P_t^A + n_t^B P_t^B}{n_t^A + n_t^B}$$

Let S_t^A and S_t^B be the share of the stock represented by the number of dwellings in submarkets A and B, respectively. Let s_t^A and s_t^B be the respective shares of each within the observed sample of sold homes. This yields a slightly different expression of the same equality in Equation (8):

$$(9) \quad S_t^A P_t^A + S_t^B P_t^B = s_t^A P_t^A + s_t^B P_t^B$$

This equality holds under one of two conditions. First, recognize that $S_t^A = 1 - S_t^B$ and $s_t^A = 1 - s_t^B$. Substituting and rearranging obtains

$$(10) \quad P_t^A (S_t^A - s_t^A) = P_t^B ((1 - s_t^A) - (1 - S_t^A))$$

or

$$(11)\quad P_t^A = P_t^B$$

So, any sample - regardless of sample selection - generates a consistent estimate of aggregate prices when price levels among the submarkets have the same first moment. This is important but not surprising. Of greater interest is under what conditions the estimated aggregate price is consistent when price levels are asymmetric across submarkets. Consider a rearrangement of Equation (9):

$$(12)\quad P_t^A(S_t^A - s_t^A) = P_t^B(s_t^B - S_t^B)$$

If $P_t^A \neq P_t^B$, the equality can hold only if both quantities inside parentheses equal zero - that is, if the submarket shares of stock and sales are identical. In other words, if appreciation is not symmetric across submarkets, the sample must be representative of the stock for global-pooling indexes to be consistent.

This stylized example of housing submarkets highlights a feature of housing markets often overlooked in empirical research: the extent to which the sample is representative with regard to appreciation. In other words, by globally pooling observations, bias may increase with the addition of more observations. So long as there are asymmetric price levels across submarkets, more observations from an overrepresented submarket will simply bias aggregate prices to those of the submarket.

A potential solution to the selection problem that doesn't require as strong an assumption of representative appreciation is a "locally-pooled" index. The intuition for local pooling is straightforward and based loosely on the notion of submarkets. While conceptually easy, submarkets are notoriously difficult to define precisely in practice. Rothemberg, Butler, Galster, and Pitkin (1991) dedicate an entire book to the effort, while Goodman and Thibodeau (1998) may be more representative of ongoing research in the area. In it, the authors define roughly a submarket to include dwellings with identical per unit price of housing services and bundles that are approximate substitutes.

The role of submarkets in sample selection is clear. In a unitary housing market, where all dwellings are perfect substitutes, no two dwellings could have different prices. Where there is relative scarcity, either in location or attribute bundles, differences in price can arise. See **?** for a well-developed theoretical model of such internal dynamics. Briefly, examples could include a shock to immigration and changes in the dynamics of "starter" homes, shocks to equities causing asymmetric changes in high-end home prices, or white-collar layoffs resulting in a decrease in price for middle-class housing. At some level, dwellings within a metropolitan area share common fundamentals due to the competition for land. This competition for land may place bounds on the divergence between any two submarkets, but the bounds may be wide with significant idiosyncratic dynamics across all submarkets.

These local dynamics pose problems for the "globally" pooled indexes because the techniques used to construct them are not well equipped to handle asymmetric appreciation or non-representative samples. The solution taken in this paper is to pool observations locally, and then aggregate - not by the share of observations but by the share of the housing stock in the particular area.

Local pooling can be illustrated by extending the same stylized submarket example from above. Under this approach, estimated aggregate price levels are

$$(13) \quad \widehat{P}_t = S^A \widehat{P}_t^A + S^B \widehat{P}_t^B = \frac{S_t^A}{n_t^A} \sum_{i \in n_t^A} (\psi_{it}^A + \varepsilon_{it}) + \frac{S_t^B}{n_t^B} \sum_{i \in n_t^B} (\psi_{it}^B + \varepsilon_{it})$$

with expectation

$$(14) \quad E[\widehat{P}_t] = S_t^A P_t^A + S_t^B P_t^B$$

This right hand side of this expression is, in fact, the definition of true aggregate price: local pooling of observations yields an unbiased estimate of aggregate prices in the presence of asymmetric prices across submarkets.

Of course, the tradeoff made to obtain an unbiased estimate is the noise incurred by smaller samples within each submarket. As ever with submarkets, definition is part art, part science. For this work, the question is not what constitutes a submarket, rather it is whether or not moving toward local pooling is appropriate. Dividing the sample into submarkets unambiguously results in fewer degrees of freedom at the local level. However, one mitigating factor is the extent to which local heterogeneity is less than global heterogeneity, which if true reduces the need for quality-controlled indexes. Moreover, if the number of indexes is large, even noisy local indexes will produce a consistent estimate of aggregate prices.

The estimator for aggregate price level given in Equation (13) can be generalized for M submarkets:

$$(15)\quad \widehat{P}_t = \frac{S_t^A}{n_t^A}\sum_{i\in n_t^A}(\psi_{it}^A + \varepsilon_{it}) + \frac{S_t^B}{n_t^B}\sum_{i\in n_t^B}(\psi_{it}^B + \varepsilon_{it}) + \ldots + \frac{S_t^M}{n_t^M}\sum_{i\in n_t^M}(\psi_{it}^M + \varepsilon_{it})$$

The empirical measures of local prices will be measured with error,

$$(16)\quad \widehat{P}_t = S_t^A(P_t^A + \delta_t^A) + S_t^B(P_t^B + \delta_t^B) + \ldots + S_t^M(P_t^M + \delta_t^M)$$

where δ_t^X is measurement error in submarket X. However, as long as local measurement error is noise (i.e., $\delta_t^X \sim (0, \sigma_\delta^2)$), the aggregate estimator remains unbiased:

$$(17)\quad E[\widehat{P}_t] = S_t^A P_t^A + S_t^B P_t^B \ldots + S_t^M P_t^M + \sum_{X\in A..M}\delta_t^X$$

The summation of the errors will approach zero as the number of submarkets increases.[5] And, therefore, the right-hand side of this expression is once more the definition of the true aggregate price.

3 Data and Motivations

In this section we provide information on the data we used and provide evidence that the two assumptions required for global pooling (that is, symmetric price appreciation across submarkets and submarket sales shares are the same as submarket stock shares). The first subsection shows asymmetric appreciation among submarkets in Sweden from 1981 - 1999. The second subsection shows that neither assumption holds for Los Angeles over the period from 2000 - 2015.

3.1 Sweden

The data used in this subsection consists of 1,000,0000 housing sales in Sweden from 1981 to 1999. These observations have been compiled by Statistics Sweden from two sources: tax assessment records that contain detailed physical characteristics of the dwelling and sales records that contain the date of sale, location, and transaction price. Dwellings are uniquely identified so that multiple sales of a single unit can be distinguished from those dwellings that sell only once during the sample

[5] Isn't the logical conclusion then to treat every dwelling as its own submarket? No, because omitted dwellings (those that do not sell) represent omitted submarkets. Local pooling, in principle, requires all submarkets be represented in the sample.

period. The tax assessment records include detailed characteristics of the property which will be used in constructing hedonic indexes, as well as enforcing the constant quality assumption implied in the repeat sales index. The data also contain crude information about the buyer and seller so that obvious non-market transactions can be removed[6].

Of particular use for this research are the geographic variables. Each observation is identified by region, county, municipality, and parish. For the purposes of this paper, submarkets are defined as municipalities - the rough equivalent of U.S. Census tracts. Some municipalities are quite small. In order to ensure price level estimates in each municipality for each quarter, only municipalities with at least three observations in each quarter over the 19 year period are included. Both the locally- and globally-pooled indexes are based on these data.

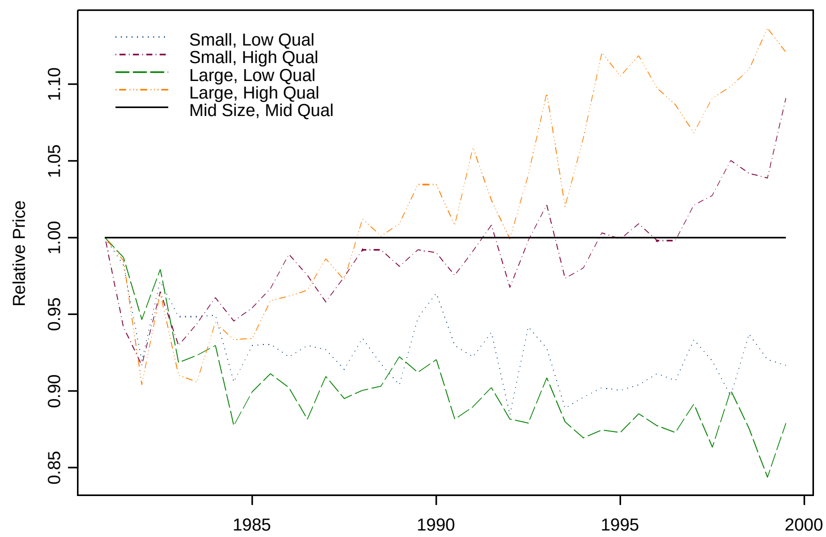
Recall the discussion above regarding the conditions under which global-pooling indexes were able to recover population parameters from the sample of sold dwellings. There were two: either the submarkets share uniform price processes or the relative proportions of submarket sales in the observed sample reflect those of the housing stock. We use two crude means to identify submarkets; again the focus is on representativeness. The first is illustrative of submarkets based on attribute bundles, the second is based on location. Both approaches result in the same interpretation and both should serve to demonstrate the difference between global- and local-pooling.

Consider the following *ad hoc* taxonomy of housing markets. Dwellings are divided into five submarkets based on the results of a hedonic regression. The regression includes characteristics related to size (lot size, living area, number of garages) and to quality (kitchen quality, the presence of a recroom, sauna, tiled bath, etc). The regression results allow the dwellings to be ranked by their overall size and quality (by interacting the respective regression coefficients and attribute values). From these indexes, the sample can be divided into submarkets. For the purposes of illustration, the following discussion uses five submarkets: small/low-quality, small/high-quality, large/low-quality, large/high-quality, and medium size and medium quality (referred to as "mid/mid").

For Region 1, the greater Stockholm area, Figure 1

[6]The data are described more completely in Englund, Quigley, and Redfearn (1998) and Englund, Quigley, and Redfearn (1999b). Since these have been published, the data have been expanded so that they now span nineteen years, from 1981 to 1999, and approach 1,000,000 observations throughout Sweden.

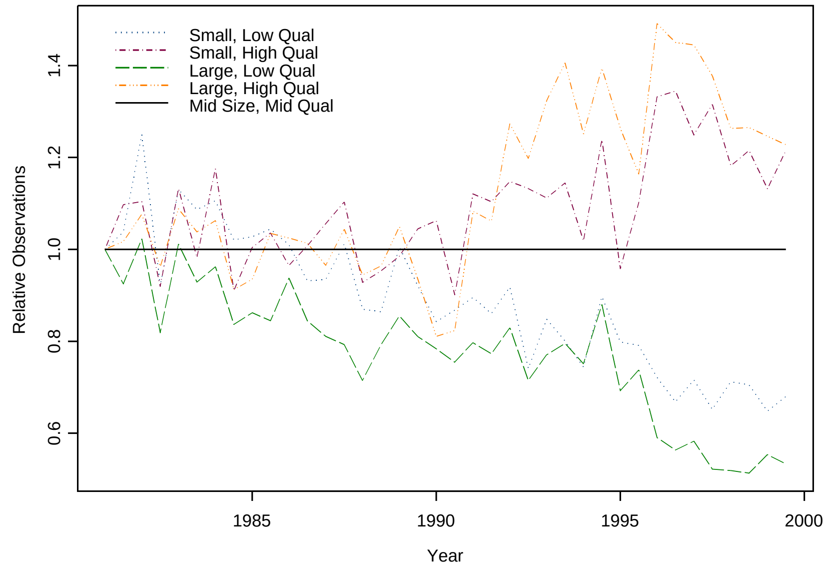
Figure 1: Relative Price Appreciation by Submarket - Stockholm



demonstrates the violation of the first condition under which the global-pooling indexes may be consistent. It shows the relative price levels of the first four submarkets relative to the medium-size/medium-quality submarket based on hedonic price indexes constructed using only the sales from each submarket. It is clear from this figure that the submarkets do not share identical price processes. Standard errors are not shown, but the null hypothesis that they do share the same processes can be rejected at the one-percent level.

Given asymmetric price processes, global-pooling indexes can still be consistent if submarket shares in the observed sample match their shares in the housing stock. Figure 2 demonstrates

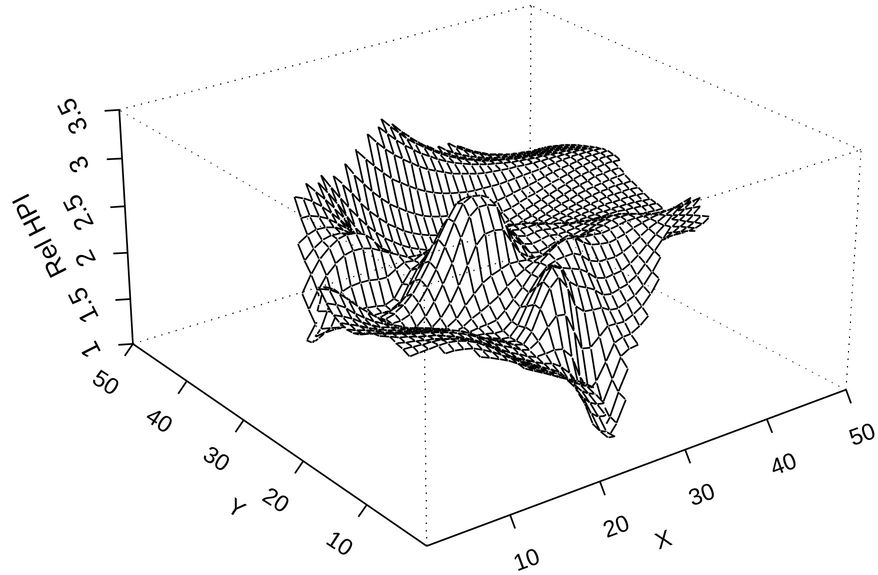
Figure 2: Relative Composition of Observed Sample - Stockholm



that for the case of Stockholm, this condition does not hold. For example, large/high-quality dwellings appreciated 15 percent more over the sample period at the same time their share of the observed sales grew by *forty percent*. On the other hand, large/low-quality dwellings lagged in appreciation as their share of observed sales fell. Granted, these are *ad hoc* submarkets, but the figures do point to inferential problems for any global-pooling index: over time, the observed sample is not representative of the stock. So, while the global-pooling indexes may be consistent estimates of the sample parameters, they may be biased estimators of the population parameters.

This partition by bundles of dwelling attributes ignores the most important single attribute of housing: location. The five submarkets are aspatial within a region of Sweden that contains qualitatively distinct submarkets across its core, suburbs, and periphery. Figures 3 and 4 address both prices and sample selection in a spatial context. The first two of these show relative median prices and sample shares across the Stockholm region comparing the surface of median prices in 1990 to a baseline of 1981. The upper panel of Figure 3 shows a relative higher appreciation in median price in the city center and to northern part of the region. The lower panel relates the analog for sale observations (not stock). It shows that relatively more observations from West of the center are found in the sample of sold dwellings in 1990 than in 1981. The panels in Figure 4 repeat the exercise for the same region nine years later, comparing prices and sample selection between 1981 and 1999.

Figure 3: Relative Price Appreciation & Sales, Stockholm 1981 to 1990



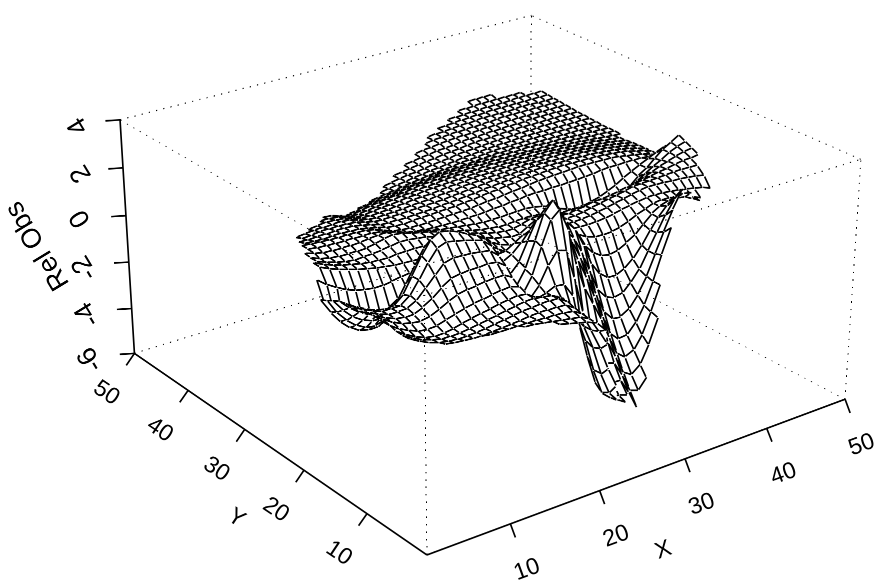
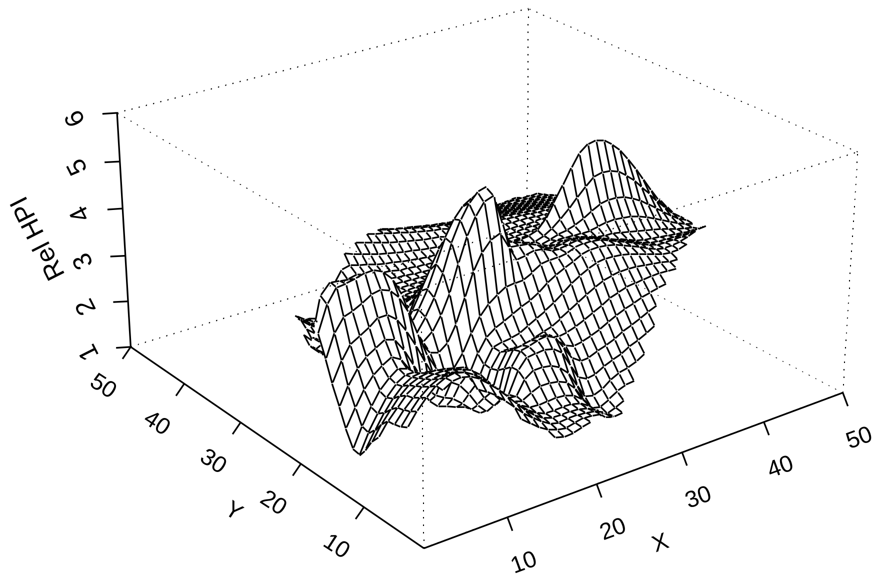
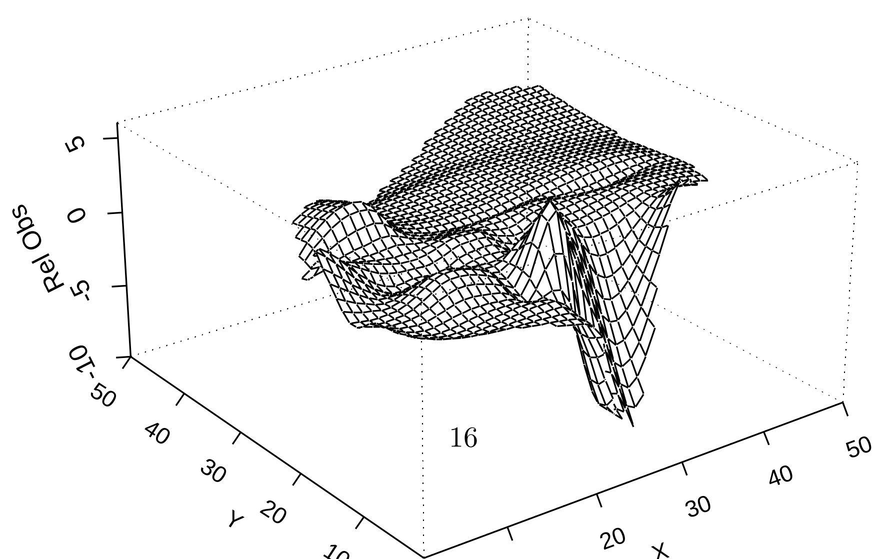


Figure 4: Relative Price Appreciation & Sales, Stockholm 1981 to 1999





Again, the center shows relatively higher appreciation, but by 1999 areas to the east and west of the center gained more than the rest as well. The sample selection picture for 1999 shows non-uniformity in the sample.

These simple metrics of asymmetry with regard to price levels and sample selection cast doubt on the ability of global-pooling indexes to recover population parameters. This finding is robust to submarket definition: essentially any meaningful partition of the data yields idiosyncratic price and selection paths. Many definitions of submarkets were explored with little difference in interpretation.

3.2 Los Angeles

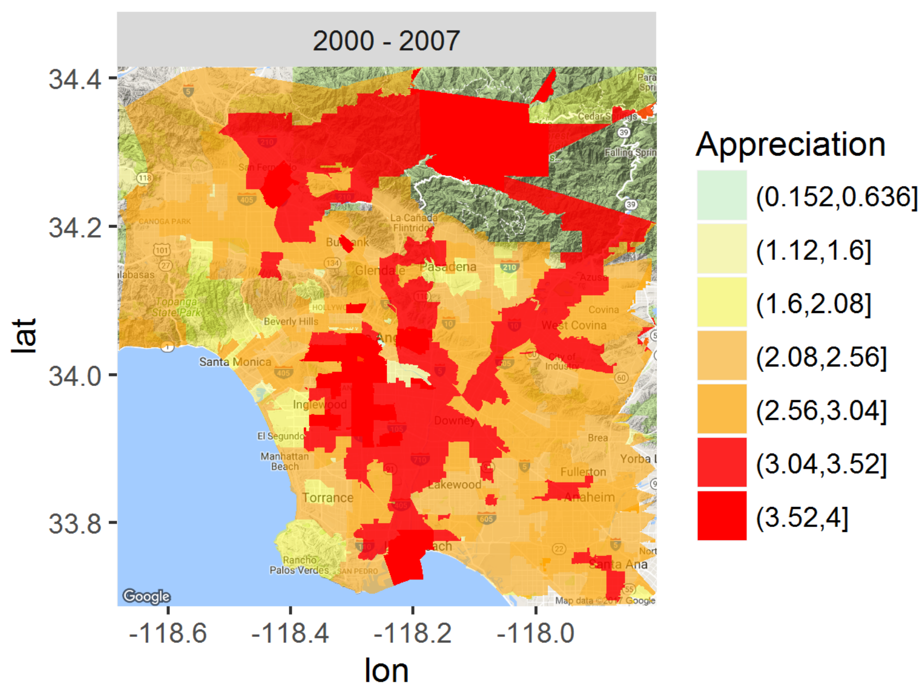
For the Los Angeles metropolitan area, we use data consisting of individual house sales complied from public records by the real estate services firm Dataquick(DQ). It has since been purchased by private data vending company CoreLogic and is available in products from that company called RealQuest and ListSource. Dataquick themselves compiled the data from tax assessors records, and as such their national data covers over 97 percent of all real estate transactions in the United States in the sample period. Our sample covers Los Angeles County, Orange County, Ventura County, Riverside County, and San Bernardino County, from the first quarter of 2000 to the fourth quarter of 2015.

We filter the DQ data in several ways to remove data that either irrelevant or clearly an error. Notably, we record house price as missing if it is more than $10 million. This effectively removes several observations that have sale prices that are implausibly high. Similarly, we also code observations as missing if there bedroom to bathroom ratio (or vice versa) of over 5. We also code the square-footage statistic as missing if the house has a positive number of bedrooms or bathrooms and 0 for square-footage, since this is logically impossible.

To compare sales data to the universe of housing stock we use data from the 2000 and 2010 Census taken from two sources, the Neighborhood Change Database (NCDB), and public use microsamples (IPUMS). The NCDB is aggregated to the tract level, and gives us information on the distribution of housing in each tract. In particular, we take the number of housing units of all types in the tract, the number of bedrooms, bathrooms, total rooms of all types, date the house was moved in to, decade the house was built in, self-reported house price, and whether the house is owner-occupied or rented. We also use the NCDB for sociodemographic information on the neighborhood, namely the racial composition, average family income, and proportion of adults with a college degree in the neighborhood. The microdata gives us more detailed information on the characteristics of housing structures, since it is not aggregated, but does not give us the same level of geographic specificity as each observation is only identified at the public use microarea (PUMA) level.

The L.A data has a specific advantage in that each observation is geocoded with specific coordinates. Figures 5 to 8 map the price appreciation across Los Angeles during throughout the boom, bust, and recovery. The underlying data for the figure are median house prices calculated by zip code. The same figure can be produced using quality controlled indexes, but this is left out for the sake of brevity. Brighter red areas represent higher price growth over the period[7].

Figure 5: House Price Appreciation Across Los Angeles



The maps makes it clear that price appreciation was no even across space over the time period,

[7]The specific figure shown is the ratio of prices in the future year Q4 to prices in the base year Q1. For instance, Figure 5 shows the median price in the zipcode in the fourth quarter of 2007 divided by the median price in the first quarter of 2000.

thus violating the first assumption of global pooling. Figure 5 shows that during the boom price appreciation was strongest south of downtown L.A extending through to downtown Long Beach, and in the San Fernando Valley area at the top corner of the map. However, even if it not to the same extreme degree, price appreciation was strong throughout all of the L.A. Even the areas that experienced the slowest appreciation (upper income areas like Santa Monica, Beverly Hills, and Manhattan beach area) still had the median sale price nearly double.

Figure 6: House Price Appreciation Across Los Angeles

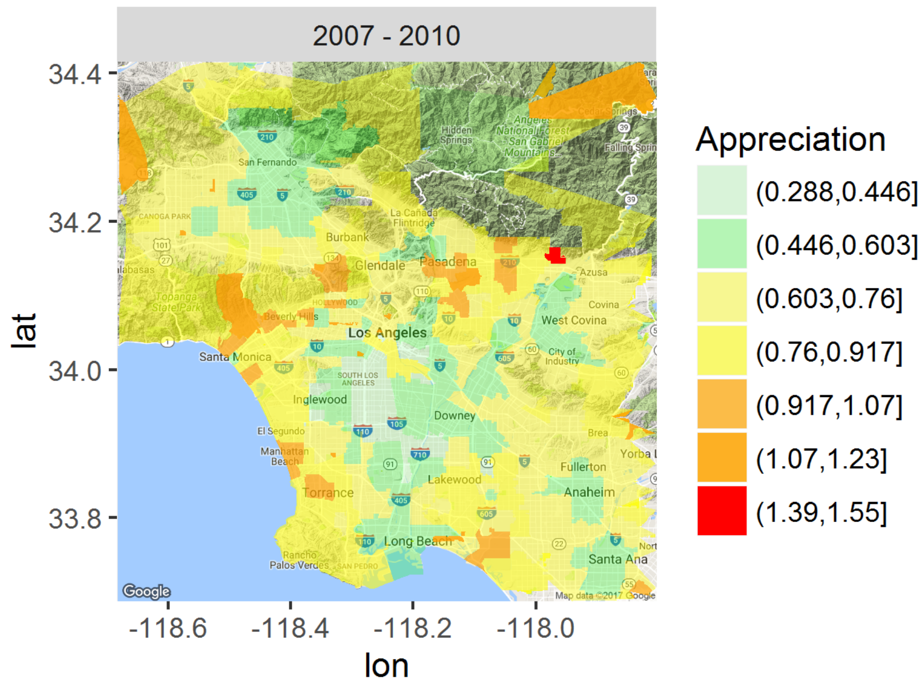


Figure 6 shows nearly the inverse of the previous figure. Those areas that experienced the greatest gains during the boom saw extreme declines in their sale prices. South Los Angeles appears to have been hit the hardest with prices declining by 50 percent or more. Interestingly we also see that the upper income areas that saw the smallest amounts of appreciation in the boom managed to maintain their sale prices (Santa Monica, Beverly Hills, and Manhattan beach are all shaded orange).

Figure 7: House Price Appreciation Across Los Angeles

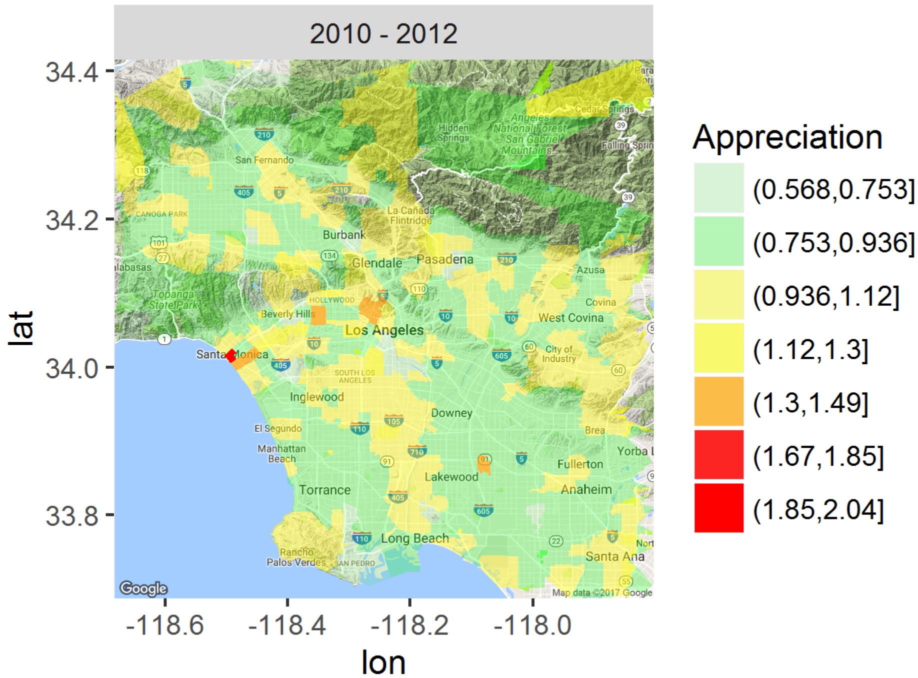


Figure 7 shows that during the stagnant period after the bust the same upper income areas as before managed to maintain values, but so did the downtown and San Fernando Valley areas that experienced the biggest crash. Some areas also continued to decline, most notably those around downtown.

Figure 8: House Price Appreciation Across Los Angeles

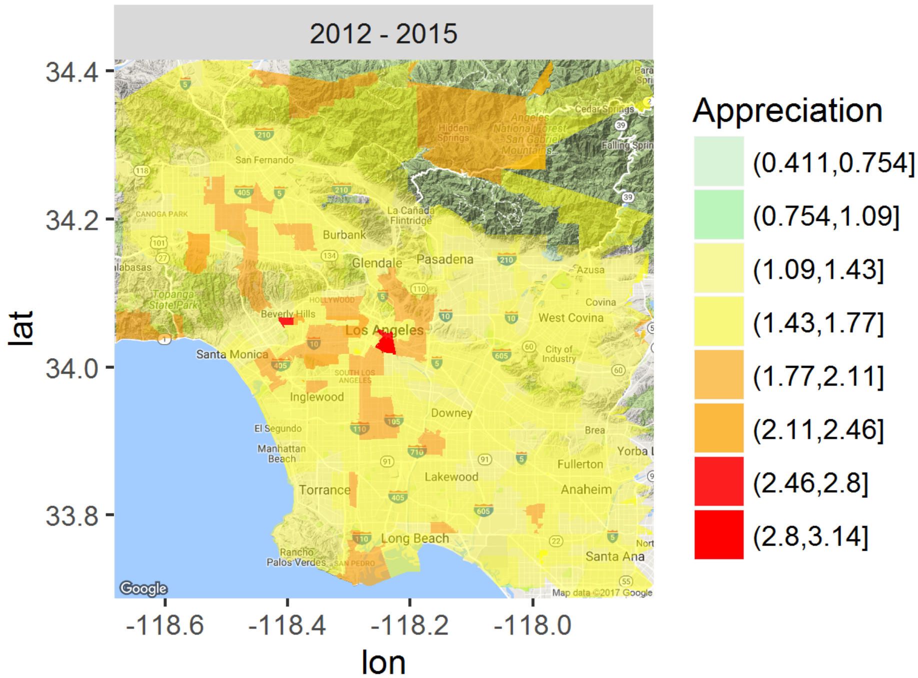


Figure 8 shows that the recovery has been driven by housing in the downtown Los Angeles area in the center of the map, slightly north of the area where the boom and bust were focused. That said, most of the map is shaded yellow, indicating thtat moderate price increases have occured throughout the city.

Taken together, Figures 5 - 8 tell a very interesting story about how the housing boom and bust played out in Los Angeles. Appreciation is clearly asymmetric with the boom and bust been driven driven by the rise and fall of house prices in the area South of downtown and the San Fernando Valley, while the recovery is driven by housing in Downtown area. House prices in upper income areas like Beverly Hills or Manhattan Beach on the other hand, experienced a less extreme cycle. The housing in these structures varies in several important ways. The areas where the boom and bust were concentrated are typically smaller lots, with smaller houses, lower prices, and less amenities.

The trends seen in the maps can be further verified by looking at some local indexes over the time 2000 - 2015 period. Figures 9, 10, and 12, show house price appreciation by structure type, lot size, and price in a base year respectively. Combined they seem to show that appreciation was greatest amongst smaller single family units, in denser downtown areas that were closer to the bottom of the house price hierarchy. However, the degree of asymmetry has also varied greatly across time.

Figure 9: House Price Appreciation by Structure Type, 2000 - 2015

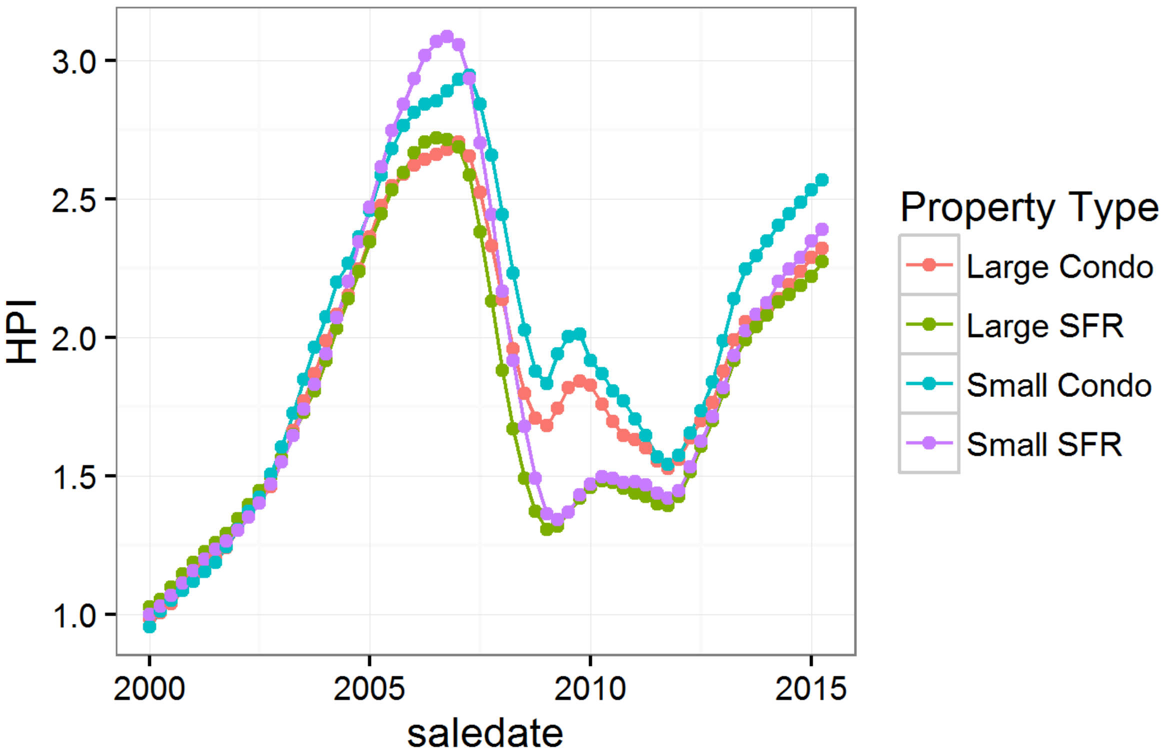


Figure 10: House Price Appreciation by Lot Size, 2000 - 2015

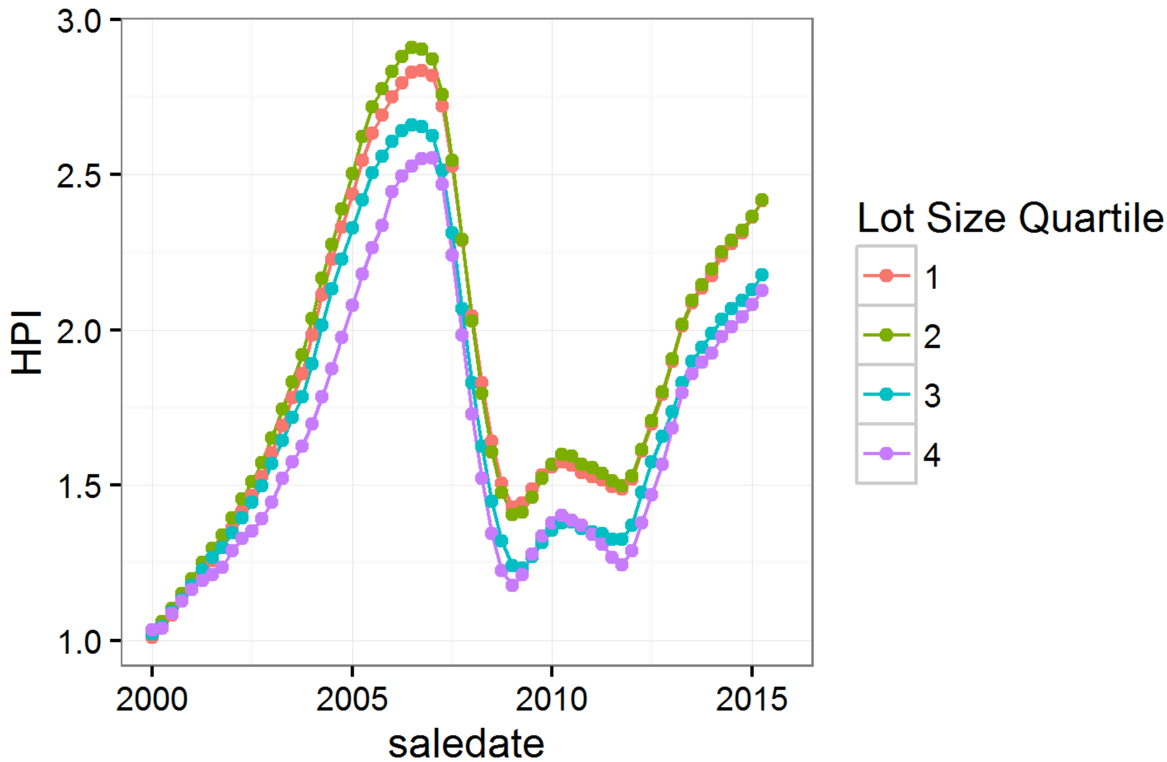


Figure 11: House Price Appreciation by 2000 Quartile, 2000 - 2015

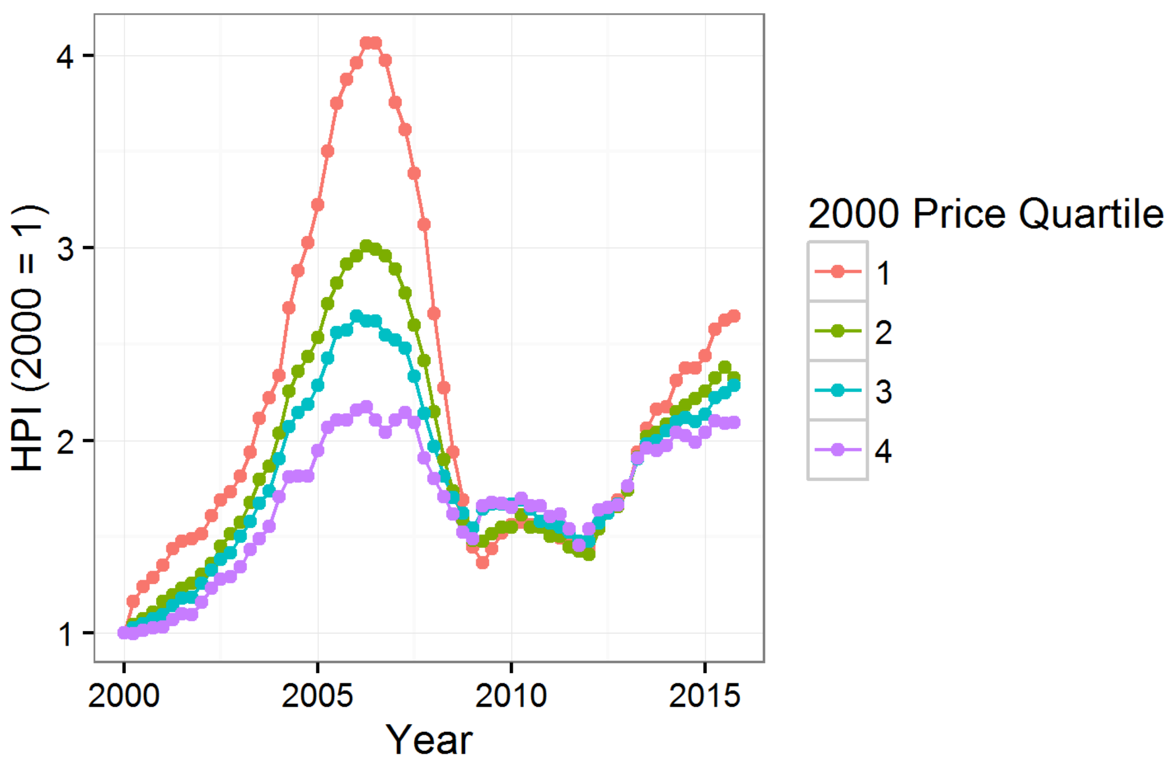
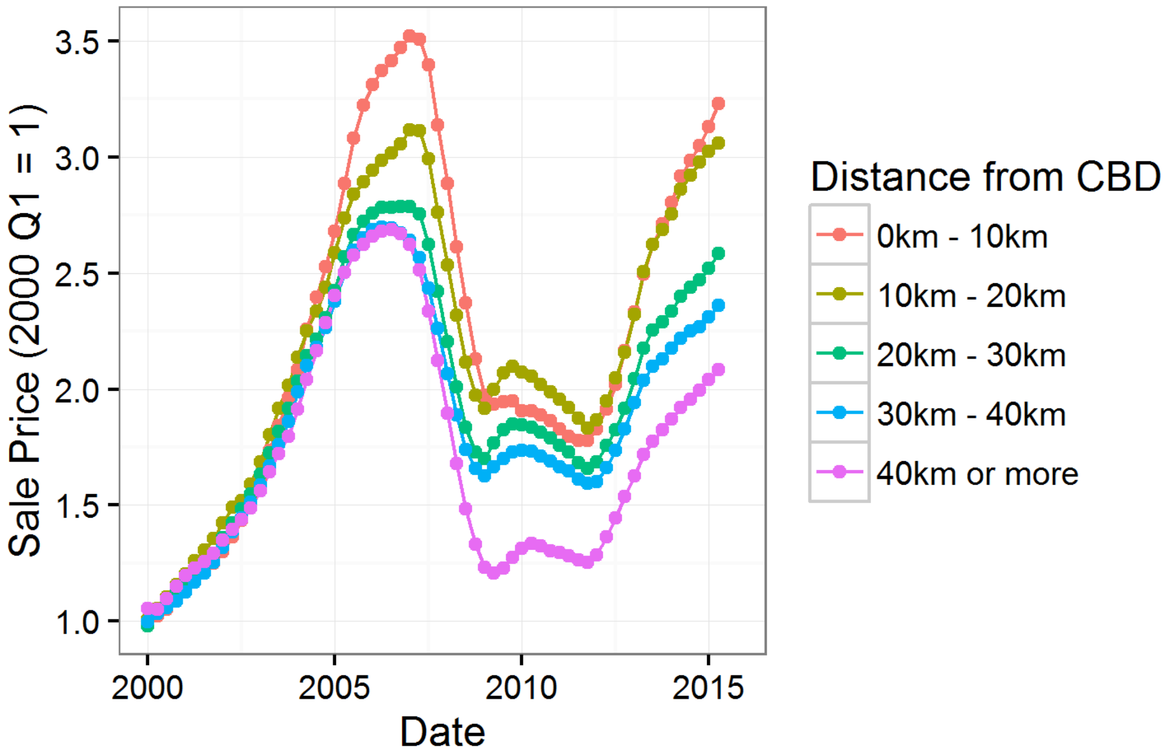


Figure 12: House Price Appreciation by Distance from City Center, 2000 - 2015



In figure 9 we divide sales into 4 types; large/small single family units and large/small condominiums. Where a unit is counted as large if it has 3 or more bedrooms. Each line represents the median price for these units. The plot shows similar patterns of appreciation up until 2005, declines post 2007, and rises post 2012 regardless of the structure. However, it is notable that the peak of small single family homes in 2006 and 2007 is substantially higher than any other type of structure with a maximum value more than 3 times that of in the year 2000. None of the other types went above 2.7 in that period. This likely reflects the nature of the housing boom that was driven by the expansion of subprime mortgages that were be disproportionately used to purchase smaller units.

Figure 10 shows median house prices of units divided into quartiles based on their lot size where 1 is used to denote the smallest lots and 4 the largest. It is clear from this that the smallest 50 percent of lots experienced the most aggressive price increases during the housing boom. Interestingly, the indexes do not converge as in figure 9, but instead seem to main their order throughout the recovery.

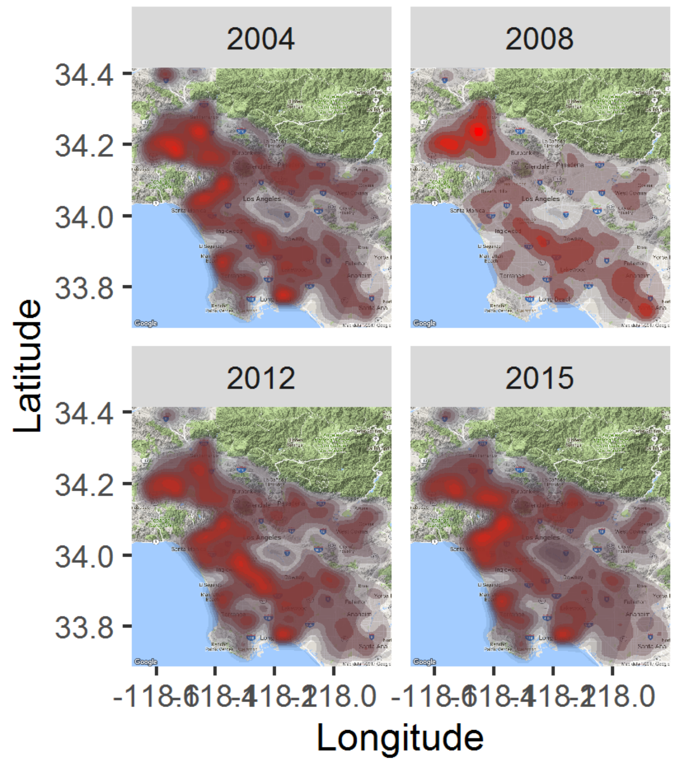
The same story is found in figure 11 that shows house price appreciation by their 2000 quartile.[8] This is done by taking the median sale price each zip code in 2000 then allocating every sale in the data a quartile based on this price. In this we see a similar pattern to the other figures except that the price of the lowest quartile (denoted as quartile 1 in the figure) shows the strongest price appreciation during the boom and bust and the 4 quartiles realigning during the recovery. It is also greatly exaggerated compared with other figures, consistent with the narrative that the boom was driven by smaller, cheaper units that received a demand shock due to improved access to credit amongst lower income borrowers.

Finally, we look at appreciation when houses are grouped by their distance from downtown Los Angeles. This is seen is figure 12 that shows house price appreciation has consistently been stronger the closer to the CBD but the extent to which this is true varies across time. It is clear that appreciation during the boom was fairly symmetric across submarkets, diverged greatly at the peak of the boom, was symmetric during the crash and stagnant period, and has again begun to diverge in the recovery. Clearly the recovery has been much stronger in areas closer to the CBD than other areas. This possibly reflects the trend of urban revival and gentrification in downtown areas that has been widely documented and that housing is more inelastic in areas closer to CBDs(Glaeser, Gyourko, and Saks 2006).

While the exhibits thus far show that there is asymmetric price appreciation amongst submarkets we know from section 2 that this does not mean that globally pooled indexes will be biased. For this to be the case it must also be that submarket shares are different from their shares of the underlying stock. The last set of figures in this section provide evidence that the distribution of housing sales across space shifted frequently through the boom, bust, and recovery. These shifts mean that different areas and types of housing are over represented relative to their share of the stock in global indexes.

[8] The same figure can be reproduced using quartiles from any year during the sample period

Figure 13: Density of Housing Sales in Los Angeles, 2004, 2008, 2012, and 2015



First, we show heat maps of the quantity of house sales in 2004, 2008, 2012, and 2015 in figure 13. Brighter, and more densely red spots show where house sales were most frequent in that year.

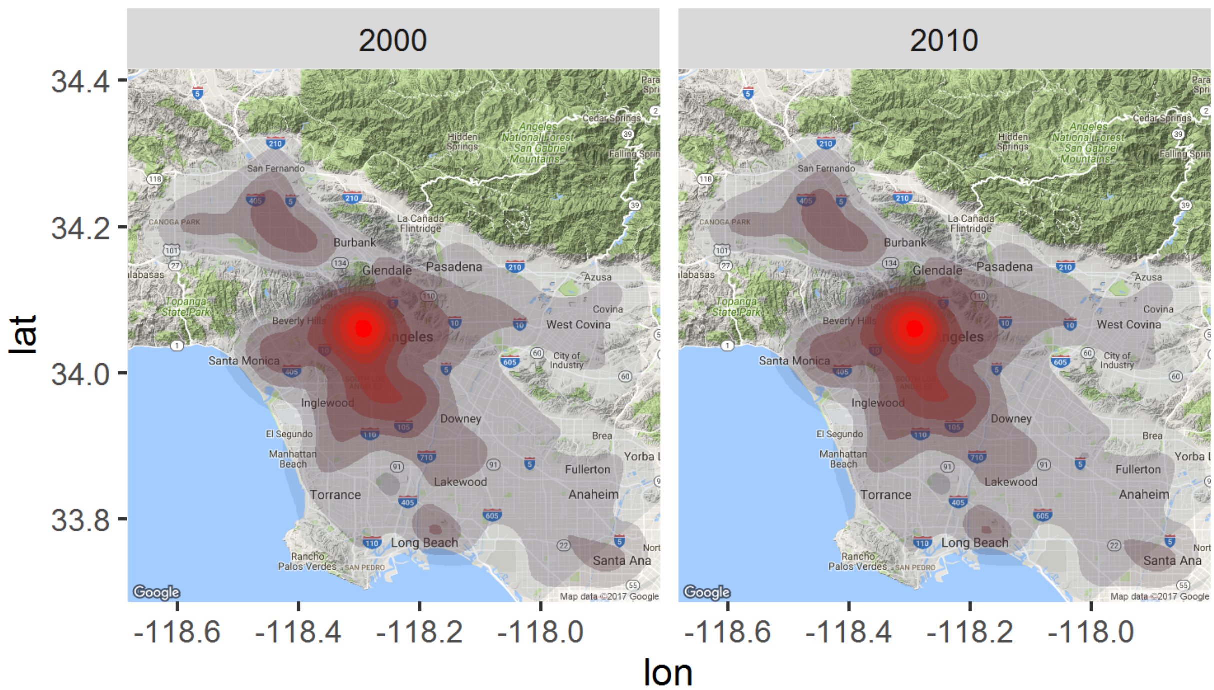
The 2004 panel shows the bulk of housing sales to be concentrated in downtown areas. The San Fernando Valley had the most sales, followed by areas such as downtown Los Angeles, and Downtown Long Beach. This is not too surprising, given that these are all very densely populated areas.

Comparing the 2004 panel with the 2008 panel, it is clear that just after the peak of the boom all house sales in L.A were concentrated in a San Fernando Valley(the top left corner of the map), and to a lesser extent South Central Los Angeles. The panel also shows less sales overall. The 2012 panel shows that sales were stronger during this period, and were mostly concentrated in South Central area, and the San Fernando Valley, similar to in 2004. The 2015 panel, however, shows that most sales occurred slightly north of this area, closer to downtown Los Angeles.

Figure 13 shows similar patterns to those seen in figure 5 to 8. Indicating a positive correlation between price appreciation and sales quantities. This is important for globally pooled indexes because, generally speaking, this means that fast appreciating neighborhoods may take up a greater share of observations than slow appreciating units.

That the distribution of housing sales shifts over time means that its representativeness of the underlying stock changes over time too. This can be assessed by looking at the distribution of the housing stock. This is shown in figure 14. Since we are using data from the Census to get our housing stock numbers, the panels show only 2000 and 2010, though the spatial distribution of the housing stock is virtually unchanged between the two years.

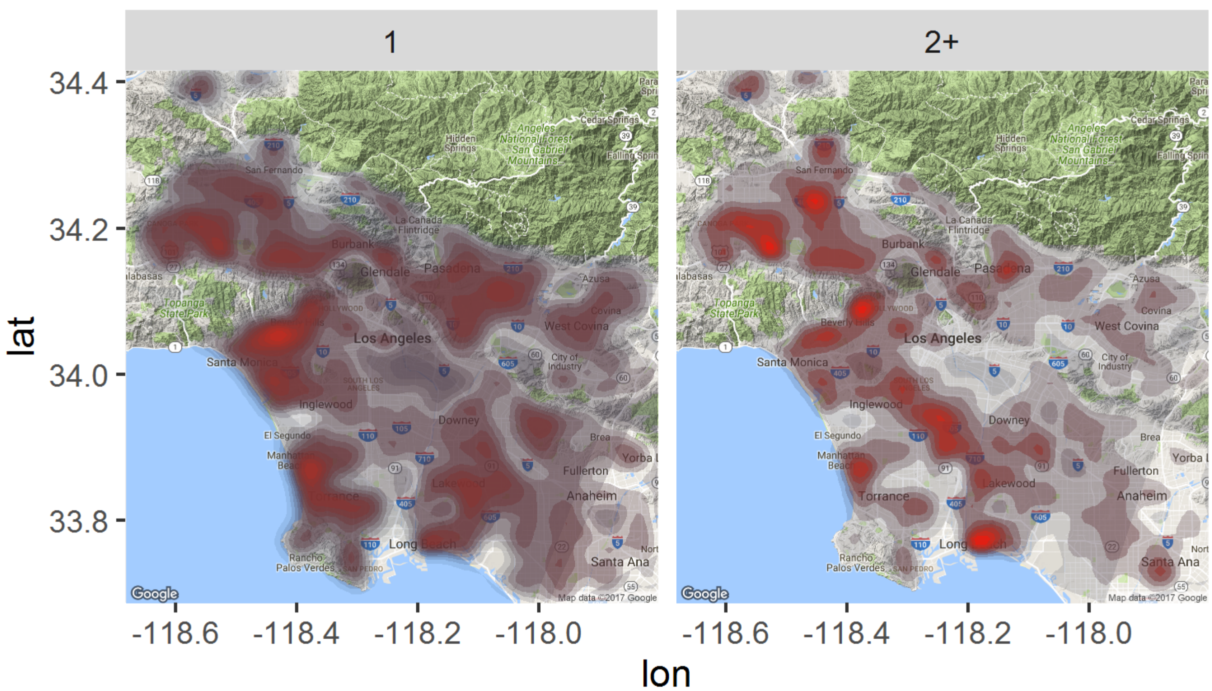
Figure 14: Density of Housing Stock, 2000 and 2010



The differences between it the maps of sales and the maps of the stock are stark. The South Central L.A hot spot has completely disappeared, and the San Fernando Valley and Downtown Long Beach have dulled. Furthermore, a new hot spot has emerged in Downtown L.A, showing the density of housing stock is very monocentric. So, despite representing the bulk of the housing stock of L.A county, housing in Downtown is not traded very frequently, and is consistently underrepresented in global indexes. This means that our locally pooled index in the next section will downweight many of areas where sales and appreciation were the strongest in the period, and instead more heavily weight the housing close to downtown. Since we have already seen that appreciation is greatest close to the CBD this means most globally pooled indexes should be underestimates.

There are other differences between the housing stock and housing sales, and ways that housing sales vary across time that are important to point out. The first is the difference between repeat sales and single sales. To investigate this we again look at heat maps of the sale locations, but make a different map for houses by that sold more than once in the decade. This is shown in 15. From this it is clear that the single sales are much more evenly spread out than repeat sales, the large hot spot in South L.A only appears in the map of homes that sold more than once. The other, smaller, hot spots at Downtown Long Beach, Manhattan Beach, and in the San Fernando are also duller in the left hand panel of the figure.

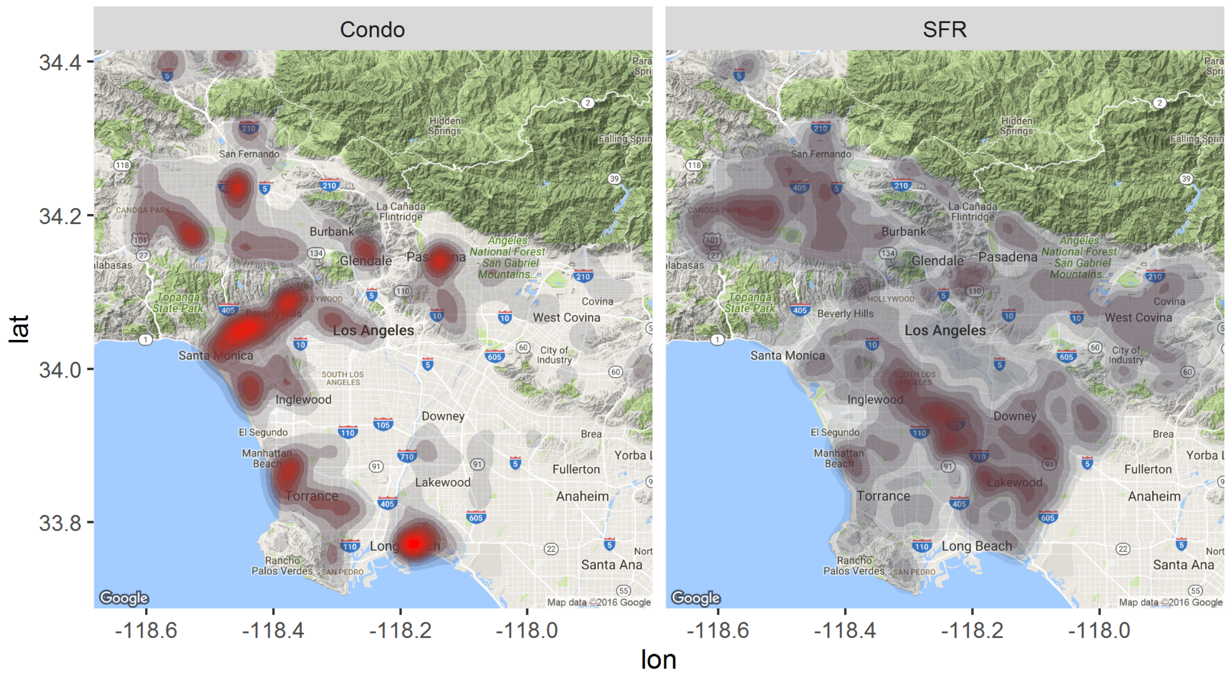
Figure 15: Density of Sales by Number of Times Sold, 2000 - 2015



Sale type also varies greatly across space. Figure 16, replicates figure **??** twice, once using only Condo's, another time using only SFR. This shows that all of the 'hot spots', with the exception of downtown L.A, consist almost entirely of condo sales. Property types are clearly not randomly dispersed across space, especially for Condo's. So, the 'typical' house sale depends on the location.

This is important because many popular indexes consist of only single family units, but doing so means that submarkets where other housing types make up most of the stock are underrepresented.

Figure 16: Density of Housing Sales in L.A County by Property Type, 2000 - 2015



4 Empirics & Preliminary Results

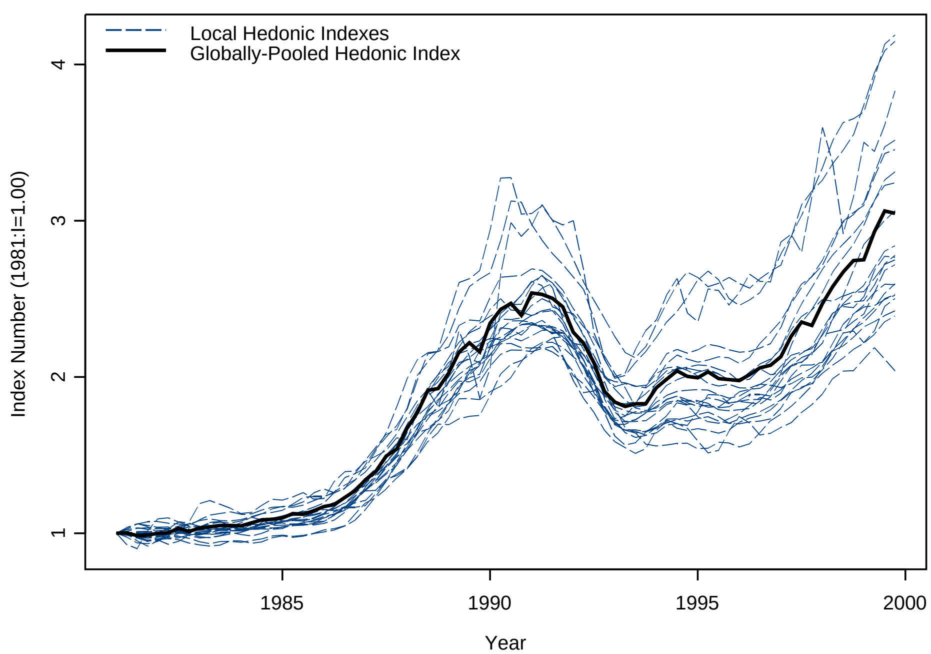
The empirical work centers on the most commonly cited housing price indexes, these are the median, repeat-sale, and hedonic price indexes. There are numerous variations on each, but the goal of this section is to test for differences between the local- and global-pooling approaches and therefore only the typical formulation of each index is employed. Mean and median price indexes are common in the business press, but have fallen from favor in academic work as they do not control for quality differences.

The theoretical model discussed earlier derived two conditions under which globally-pooled sample observations could be used to recover the population parameters. The first condition required that all of the observations could be drawn from the same stochastic price process, in which case any sample would be representative with regard to appreciation. The second condition required that submarket shares of observations were equal to submarket shares of the dwelling stock. The previous section suggests that there is good reason to doubt that either section is true.

In this section we apply the local pooling solution suggested earlier. This method creates individual indexes for each submarket, then taking the weighted mean of these indexes where the weights are that submarket share of the housing stock. Defining a submarket is, of course, more art than science, and providing a comprehensive answer to the question 'what is a submarket?' is beyond the scope of this paper. In light of this, we use geographic areas as our submarkets (specifically, we use zip codes), as there is a general consensus that locations within a city define a submarket of some kind in that everyone who lives there has access to a common set of local amenities.

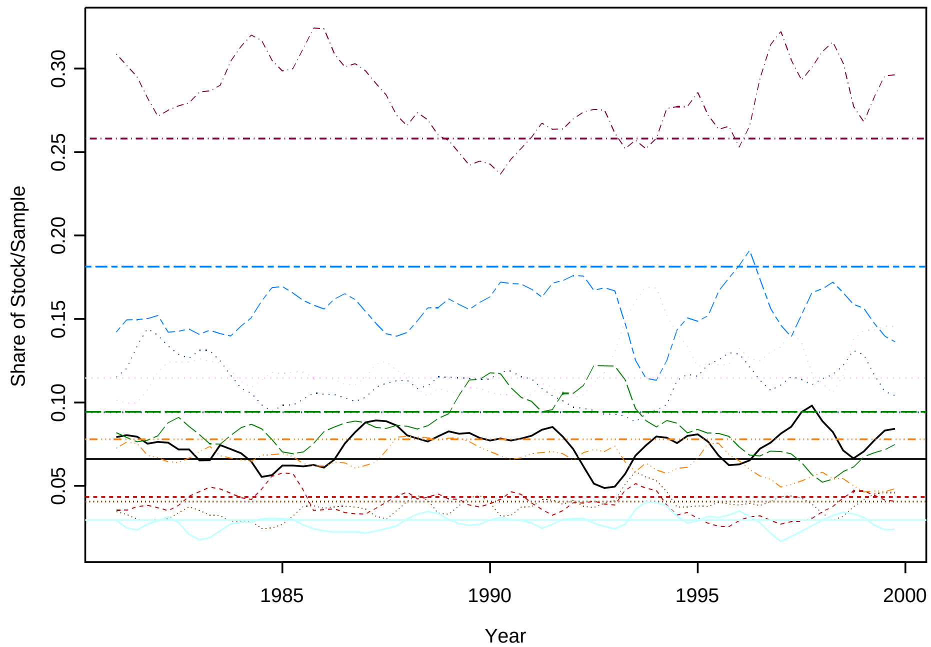
Figure 17 suggests that the first condition does not hold; Figure 18 suggests that the second does not.

Figure 17: Local Price Indexes and Globally Pooled Aggregate Prices for Stockholm



The first of these two figures shows both the set of municipal housing price indexes as well as the globally-pooled hedonic price index for the Stockholm region. These local indexes are significantly

Figure 18: Local Share of Aggregate Stock vs. Local Share of Observed Sales



different from one another - they do not share the same price process. (This is typical of all eight regions.) The figure is suggestive of a non-representative sample as most of the local indexes fall below the aggregate index. It is not possible to determine from the figure, since stock weighting may yield the same aggregate price index. But it is easy to see from this figure how non-random selection could lead to bias in measured aggregate prices: consider a sample of sold dwellings from only the top half of the index distribution. Of course, such a sample is not found in practice, but as the second figure shows, municipal shares of observations can differ substantially from the municipality's share of the region's housing stock. In it, the time-varying lines are the municipal share of observed sales from period to period, while the horizontal lines are the municipal shares of the housing stock.[9] When the stock-share line is above the sample-share line, the municipality is under-represented in that period.[10]

[9]To be clear, the dwelling stock is calculated by totaling any dwelling that sells at least once during the 19-year sample period. As such, the stock is understated by the amount of dwellings that do not sell at all. While imperfect, this definition is sufficient for the purposes of examining the difference between the time-varying weights implied by fluctuating sales and the fixed weights implied by the share of dwellings sold over the sample period.

[10]What is also clear from the figure is that a few of the municipalities dominate the stock and sale shares. This

In order to examine differences between indexes based on locally- and globally-pooled samples, a simple set of comparisons were constructed. First, four globally-pooled indexes were constructed for each of eight regions in Sweden. These indexes are constructed as typically implemented in practice. In particular, the repeat sale indexes are built using the weighting approach developed in Case, Shiller, et al. (1987).[11] The hedonic regressions include eighteen measures of housing quality (including lot size, living area, age, and dummies for one- or two-car garage, sauna, fireplace, tiled bath, roof type, quality of insulation, and quality of kitchen, as well as dummies for each municipality).

Local indexes are constructed analogously. For each municipality, mean, median, repeat sale, and hedonic indexes are estimated. The hedonic specification at the municipal level obviously does not include the municipal dummies found in the global hedonic specification. Aggregate price indexes are built by weighting the local indexes by the respective municipality's share of the dwelling stock.

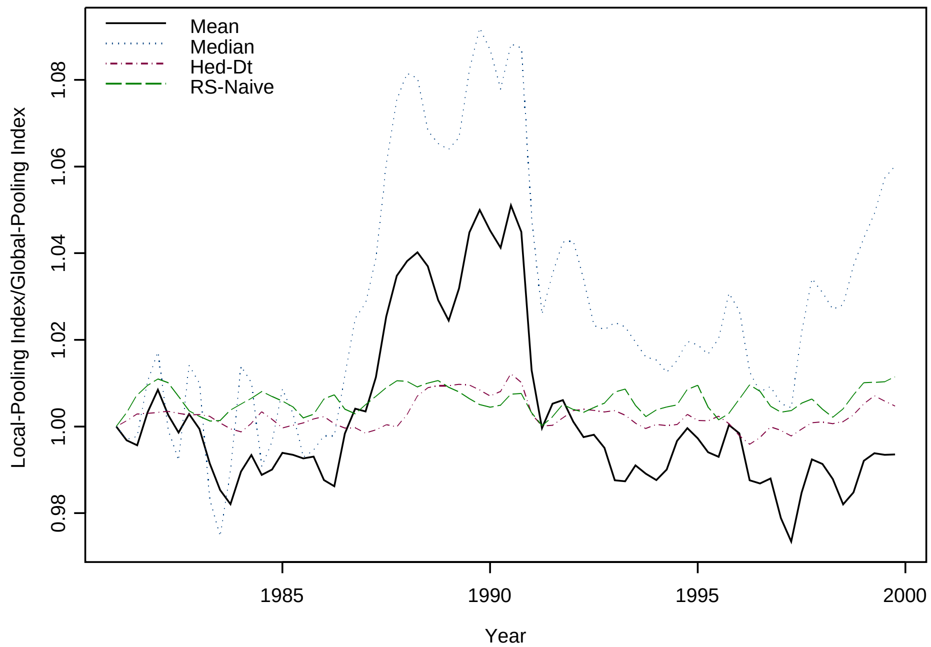
The first question is whether or not the order of aggregation matters. Figure 19 shows the relative difference between local and global pooling in the Malmö region for the four indexes: mean, median, repeat sale, and hedonic. It indicates that meaningful differences exist only for mean and median indexes. This pattern is consistent across all eight regions as indicated in Table 1. It reports the mean standard deviation of the index relatives. As in the figure, the table supports the notion that the order of aggregation is relevant only for the simple indexes. This is somewhat surprising given the wide dispersion of local prices and sample shares displayed in Figures 17 and 18. The modest difference between the locally- and globally-pooled repeat sale and hedonic indexes is likely due to the dominance of a handful of municipalities in the sample. In all the regions, there are one to three central municipalities from which a significant minority of observations are drawn. In fact, an example of this can be seen in Figure 2, where two municipalities contain 45 percent of the stock[12].

may explain the lack of decisive findings as to the relative levels of aggregate prices as measured using locally- and globally-pooled indexes. This is discussed in more detail below.

[11] Two forms of the auxiliary regression were explored, one with and the other without squared time between sales. The results presented here are based on just the linear term; that is, squared residuals are regressed on a constant and time between sales only. Differences in the subsequent weighted regressions were negligible.

[12] To date, two levels of geographical aggregation have been examined: parish and municipality. The first of these created many problems with parishes too small to support repeat sale indexes; the second, presented here, appears to be too broad a geographical area. Initial results from the parish-level analysis showed greater difference between

Figure 19: Locally-Pooled Indexes Relative to Globally-Pooled Indexes - Malmö



The second question is whether the set of indexes measure the same phenomenon. That is, are the indexes interchangeable? Much has been made of the differences in underlying samples and maintained hypotheses among the different indexes. **?** find problems with both repeat sale and hedonic indexes sufficient to suggest use of a median index in their stead. Advocated as well in their paper is the use of a non-parametric estimator that is free of the assumptions required of the quality-controlled indexes. In this paper, as in all housing price index research, there is no "truth" against which to measure index construction methods. The aggregate index based on locally-pooled hedonic regressions is used as the benchmark. It is the most flexible and employs the most information from market transactions of the indexes examined.

Figure 20 shows several aggregate housing price indexes relative the locally-pooled hedonic index, again for the Malmö region. Apparent in the figure is the wide variability of the global mean and median indexes. As is well known, these indexes are subject to quality variation. The repeat sales indexes, both locally- and globally-pooled, drift higher relative to the benchmark. This

the locally- and globally-pooled repeat sale and hedonic indexes.

Table 1: Comparison of Indexes Locally-Pooled Relative to Globally-Pooled
Average Percent Difference, Standard Deviation in Parentheses

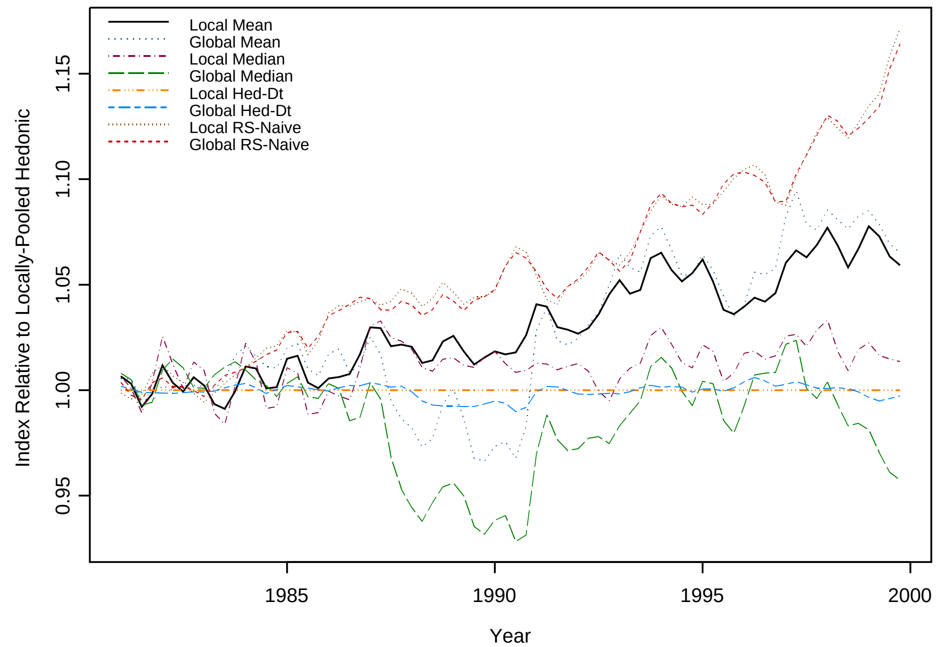
Index	Region							
	I	II	III	IV	V	VI	VII	VIII
Mean	-0.70	-0.94	4.17	0.38	2.06	2.79	-2.35	2.71
	(1.8)	(2.0)	(1.7)	(2.2)	(2.0)	(1.7)	(2.8)	(3.2)
Median	0.75	0.77	8.77	2.94	3.56	7.59	7.74	6.28
	(1.4)	(2.6)	(3.0)	(3.3)	(2.5)	(3.0)	(4.3)	(3.4)
Repeat Sale	-0.92	0.19	1.99	0.95	1.77	0.69	0.05	-0.34
	(0.4)	(0.4)	(0.6)	(0.4)	(0.6)	(0.7)	(0.8)	(0.9)
Hedonic	-1.01	0.01	-0.86	0.19	0.08	-2.76	-1.20	-0.38
	(0.4)	(0.6)	(0.4)	(0.5)	(0.6)	(0.6)	(0.8)	(0.8)

phenomenon has been documented before in Englund, Quigley, and Redfearn (1999b) in which these differences with the hedonic indexes are attributed to unmeasured quality change in the sample dwellings.

The locally-pooled median index shows remarkably similar behavior vis-á-vis the benchmark locally-pooled hedonic price index and is a significant improvement relative to its globally-pooled counterpart. The locally-pooled median is particularly close to the benchmark despite the absence of control for the quality differences apparently manifest in the globally-pooled median index. The basis for this is two-fold. First, local quality variation is less extensive than global variation and therefore quality-control is less of an issue locally. In particular, locally pooled indexes successfully capture local land values which are a dominant component of observed dwelling prices. Second, as shown in Equations 15, 16, and 17, locally-pooled indexes are consistent estimators of aggregate prices so long as measurement error at the local level is not correlated across municipalities.

These results are consistent across the eight regions, as reported in Table 2. The table reports mean and standard deviations for the index relatives - the percent difference between the indexes and the benchmark locally-pooled hedonic index. In general, the locally-pooled median exhibits smaller differences than the others and is more stable. The broad similarity of the locally-pooled median and hedonic indexes points to a potential solution to a central problem in index choice. Researchers interested in aggregate prices are often forced to choose between median and repeat

Figure 20: Aggregate Price Indexes Relative to Locally-Pooled Hedonic Index - Malmö



sale indexes as data are typically unavailable to estimate hedonic indexes, hence the popularity and extensive usage of the OFHEO and NAR indexes. The results presented in Table 2 suggest that the benchmark index is recoverable with a simple set of sales data: sale price, date of sale, *and* dwelling location (even crudely identified).

The availability of minimal sets of sales data implies that unbiased aggregate housing price indexes may be constructed without having to make the assumptions required of the repeat sales index or be subject to the quality variation that haunts the mean and median as useful measures of aggregate prices.

4.1 Los Angeles

Figure 21 shows the local indexes for all zip codes within the 5 counties in the greater L.A area in color, and the globally pooled index for the entire of L.A in black. The left panel shows indexes based on a hedonic regression, the center panel shows the median, and the right most panel shows a weighted repeat sales index constructed using the Case-Shiller method. The extreme dispersion

Table 2: Comparison of Index Performance Relative to Locally-Pooled Hedonic
Average Percent Difference, Standard Deviation in Parentheses

Index	Region							
	I	II	III	IV	V	VI	VII	VIII
Global Mean	4.92	0.66	-5.01	3.72	3.08	-4.61	4.81	-3.88
	(5.0)	(3.4)	(2.8)	(4.0)	(4.3)	(3.0)	(6.4)	(4.4)
Global Median	0.34	-2.80	-10.55	-0.81	-0.12	-9.28	-1.85	-9.70
	(1.8)	(2.8)	(3.3)	(3.0)	(2.5)	(3.1)	(6.0)	(3.5)
Global RS	-3.55	0.95	1.08	3.53	-7.32	3.51	5.13	2.83
	(2.1)	(2.5)	(3.6)	(4.2)	(2.6)	(3.8)	(5.5)	(3.3)
Global Hedonic	1.02	-0.01	0.87	-0.19	-0.07	2.84	1.22	0.39
	(0.4)	(0.6)	(0.4)	(0.5)	(0.6)	(0.7)	(0.8)	(0.8)
Local Mean	4.11	-0.34	-1.08	4.06	5.14	-1.98	2.20	-1.40
	(3.4)	(2.0)	(2.0)	(2.7)	(3.6)	(2.3)	(4.0)	(2.1)
Local Median	1.08	-2.11	-2.78	2.02	3.40	-2.46	5.57	-4.12
	(1.6)	(1.5)	(2.3)	(1.8)	(2.0)	(2.2)	(4.5)	(2.5)
Local RS	-4.43	1.15	3.09	4.52	-5.67	4.22	5.17	2.48
	(2.1)	(2.5)	(3.5)	(4.3)	(2.7)	(3.9)	(5.2)	(3.4)

of the individual zip code indexes around the globally pooled index make it clear that the extent and timing of price appreciation and depreciation vary greatly within the metro area. Combined with the information we know from the previous section that looking solely at sales can mean that certain submarkets will be overrepresented in a globally pooled index this justifies the use of a locally pooled index for the metro area.

Figure 21: Locally Pooled Indexes and Globally Pooled for Los Angeles, 2000-2015

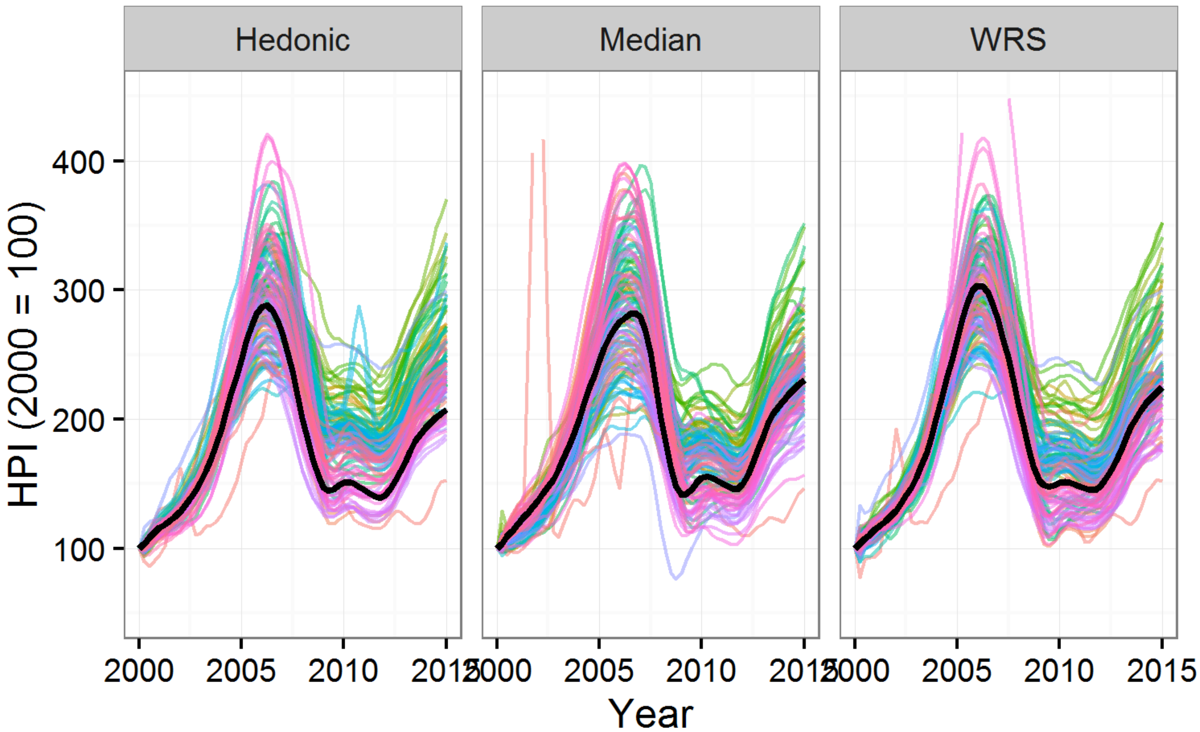
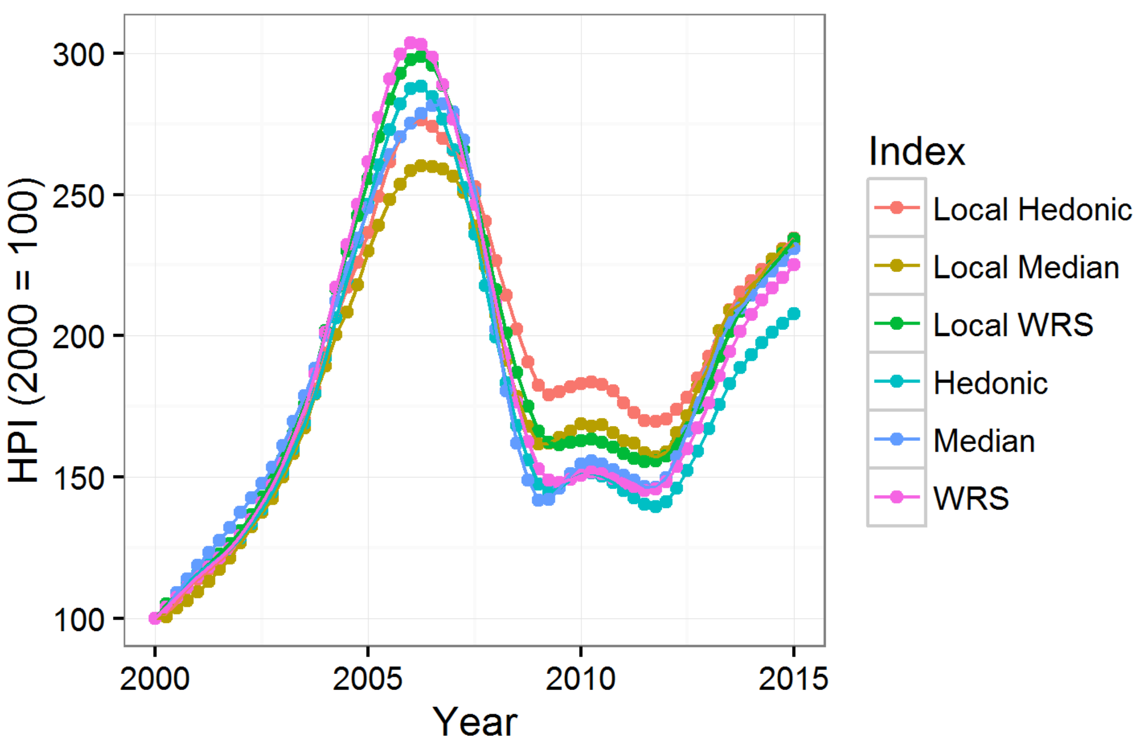


Figure 22: Local Indexes vs Globally Pooled Index for L.A, 2000-2015



Another interesting note about the submarket indexes is the great degree of similarity across methods. The average absolute difference between the two is only 4 points, and the average correlation between the two is over 0.96. This most likely indicates that housing structure is more homogeneous in quality at local levels, so there is less need to control for differences in quality at these levels.

Figure 23: Globally Pooled Indexes relative to Locally Polled Indexes for Los Angeles, 2000-2015

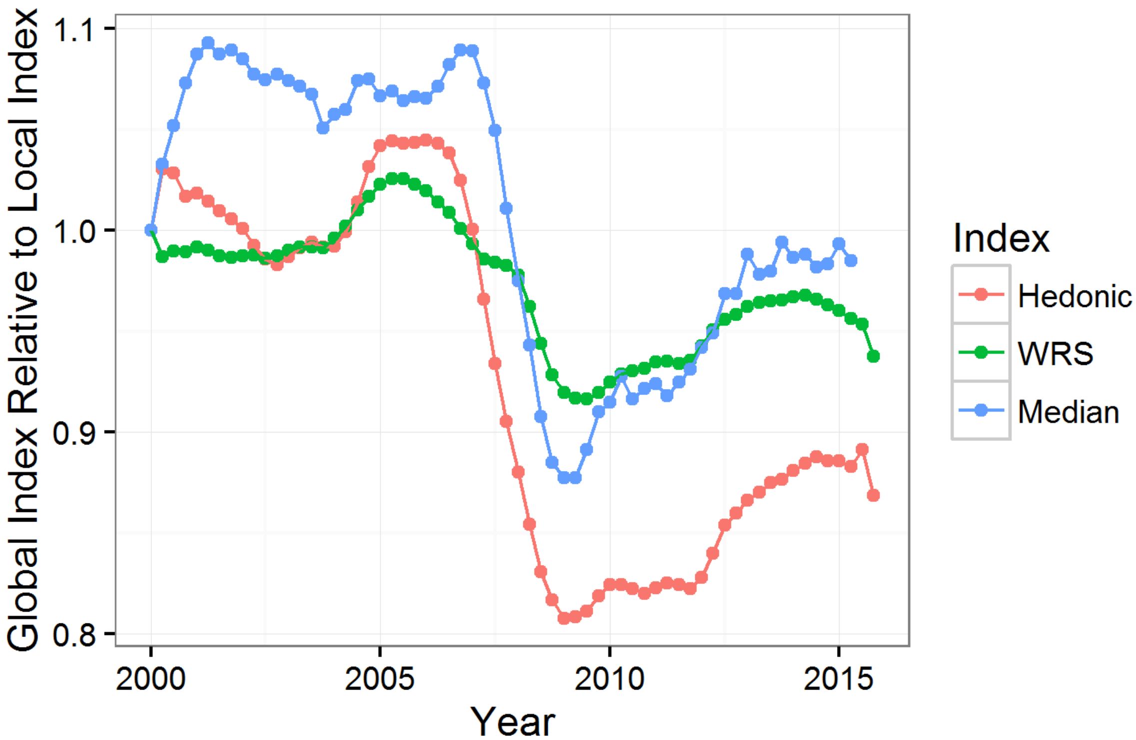


Figure 22 shows the globally pooled index against the median and hedonic indexes shown in the black lines in figure 21, and a weighted repeat sales index. Figure 23 shows the ratio of the globally pooled indexes to the relative indexes. The locally pooled indexes are each the weighted mean of the colored lines in figure 22, where the weights are that submarkets share of the overall housing stock in 2000 [13]. The figures show several features that we might expect given the evidence in the previous section.

First, all the indexes show a great degree of symmetry in the during both the boom, indicative of the great symmetry of price appreciation across submarkets seen in the previous section.

[13] We also used the 2010 housing stock as weights and found the same results.

Second, the locally pooled indexes have a lower peak for the boom than all the others, particularly the repeat sales index. This is not surprising, the evidence in Section 3 showed indicated that fast appreciating areas were overrepresented relative to the housing stock during the 2000 - 2008 period. The repeat sales index shows the highest peak for the boom, indicative of its own selection problems found seen in Figure 15 where we saw that repeat sales were heavily concentrated in South Central L.A during the sample period.

The overstatement of the boom by globally pooled indexes relative to locally pooled ones is most likely due to the demand shock caused by the expansion of credit to lower income individuals through subprime loans. This demand shock was concentrated in the lower priced markets, because access to credit for buyers in higher priced submarkets remained relatively unchanged over the period. The result is that lower priced, smaller homes became overrepresented relative to both earlier years and to its share of the overall housing stock.

Third, the locally pooled index also find a less severe crash and a stronger recovery than globally pooled indexes. This is probably because the global indexes underrepresent downtown L.A, where price appreciation has been the strongest. Future versions of this paper will explore the implications of these findings for such things mortgage default and calculating the value of the housing stock.

5 Conclusions

The goal of this paper was to examine the relative performance of indexes using an alternate order of aggregation. Typical construction of aggregate indexes begins by pooling all relevant observations - a process we call "global pooling." Global pooling implicitly assumes that observation are representative with regard to price level and the dwelling stock. These assumptions are typically made without verification. The first issue the paper addresses is representativeness and the potential bias that arises from non-representative samples.

In a stylized example, two conditions were derived under which global pooling indexes could recover the population parameters from a sample of observed dwellings. First, appreciation could be symmetric across submarkets; second, selection into the sample could be random. In other words, in the presence of asymmetric appreciation, the sample needed to be representative of the dwelling stock. In practice, both conditions were shown to be violated in the case of both the

Swedish and Los Angeles housing markets covered in our data.

An alternative approach to aggregating all observations is to pool observations within smaller levels of geography, construct local indexes at this level, and then build aggregate price indexes by weighting the local indexes by their share of the stock rather than by their share of observed sales. This has the effect of reducing the bias induced by non-representative samples. Whereas an additional observation from an over-represented submarket biases global pooling indexes toward prices in that submarket, the additional observation improves local price estimates and, in turn, aggregate price estimates in the locally-pooled indexes.

Empirically, the difference between local- and global-pooling indexes was greatest for the mean and median indexes, but only very modest for the repeat sale and hedonic indexes. The results are preliminary but surprising given the wide variation in price appreciation at the local level and in their share of the observed sample over time. In particular, the Los Angeles results are compelling in that the locally pooled index outperforms the popular FHFA and Case-Shiller indexes by detecting the start of both the bust and the recovery first. Furthermore, it suggests that these indexes overstated the severity of the housing crash and the understated the strength of the recovery.

One potentially important result in comparing the estimates of aggregate prices produced by the various indexes is that the locally-pooled mean and median indexes track the estimates of the "best" index. In particular, the locally-pooled median index tracks closely the locally-pooled hedonic index (which was chosen as the benchmark due to its flexibility across submarkets and its use of the most information of all the indexes). This similarity offers the potential for an unbiased aggregate price index free of the assumptions implicit in the repeat sales using a minimum of transaction data. More work is required to buttress these results, but these initial results are promising.

References

Anenberg, E. (2011): "Loss aversion, equity constraints and seller behavior in the real estate market," *Regional Science and Urban Economics*, 41(1), 67–76.

Bhattacharjee, A., E. Castro, T. Maiti, and J. Marques (2016): "Endogenous spatial regression and delineation of submarkets: a new framework with application to housing markets," *Journal of Applied Econometrics*, 31(1), 32–57.

Bogin, A. N., W. M. Doerner, and W. D. Larson (2016): "Local house price dynamics: New indices and stylized facts," Discussion paper, Federal Housing Finance Agency.

Bostic, R. W., S. D. Longhofer, and C. L. Redfearn (2007): "Land leverage: decomposing home price dynamics," *Real Estate Economics*, 35(2), 183–208.

Bourassa, S. C., F. Hamelink, M. Hoesli, and B. D. MacGregor (1999): "Defining housing submarkets," *Journal of Housing Economics*, 8(2), 160–183.

Bourassa, S. C., M. Hoesli, and V. S. Peng (2003): "Do housing submarkets really matter?," *Journal of Housing Economics*, 12(1), 12–28.

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., R. J. Shiller, et al. (1987): "Prices of single-family homes since 1970: new indexes for four cities," *New England Economic Review*, (Sep), 45–56.

Clapp, J. M., and Y. Wang (2006): "Defining neighborhood boundaries: Are census tracts obsolete?," *Journal of Urban Economics*, 59(2), 259–284.

Englund, P., J. M. Quigley, and C. L. Redfearn (1998): "Improved Price Indexes for Real Estate: Measuring the Course of Swedish Housing Prices," *Journal of Urban Economics*, 44(2), 171–96.

——— (1999a): "Do Housing Transactions Provide Misleading Evidence About the Course of Housing Values?," *unpublished manuscript*, pp. 1–26.

——— (1999b): "The Choice of Methodology for Computing Price Indexes: Comparisons of Temporal Aggregation and Sample Definition," *Journal of Real Estate Finance and Economics*, 19(3), 91–112.

Gatzlaff, D. H., and D. R. Haurin (1997): "Sample Selection Bias and Repeat-Sales Index Estimates," *Journal of Real Estate Finance and Economics*, pp. 33–50.

Genesove, D., and C. Mayer (2001): "Loss aversion and seller behavior: Evidence from the housing market," *The Quarterly Journal of Economics*, 116(4), 1233–1260.

Glaeser, E. L., J. Gyourko, and R. E. Saks (2006): "Urban growth and housing supply," *Journal of Economic Geography*, 6(1), 71–89.

Goetzmann, W. N., and M. Spiegel (1997): "A Spatial Model of Housing Returns and Neighborhood Substitutability," *Journal of Real Estate Finance and Economics*, 14(1-2), 11–31.

Goodman, A. C., and T. G. Thibodeau (1998): "Housing Market Segmentation," *Journal of Housing Economics*, 7(2), 121–143.

Hausman, J. A., and D. A. Wise (1977): "Social Experimentation, Truncated Distributions, and Efficient Estimation," *Econometrica*, 45(4), 919–38.

Hwang, M., and J. M. Quigley (2004): "Selectivity, quality adjustment and mean reversion in the measurement of house values," *The Journal of Real Estate Finance and Economics*, 28(2), 161–178.

Keskin, B., and C. Watkins (2017): "Defining spatial housing submarkets: Exploring the case for expert delineated boundaries," *Urban Studies*, 54(6), 1446–1462.

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

Mason, P., and G. Pryce (2011): "Controlling for transactions bias in regional house price indices," *Housing Studies*, 26(5), 639–660.

Meese, R., and N. Wallace (1991): "Nonparametric Estimation of Dynamic Hedonic Price Models and the Construction of Residential Housing Price Indices," *American Real Estate and Urban Economic Association Journal*, 19(3), 308–332.

Newey, W. K., J. L. Powell, and J. R. Walker (1990): "Semiparametric Estimation of Selection Models: Some Empirical Results," *American Economic Review*, 80(2), 324–28.

Ortalo-Magné, F., and S. Rady (1999): "Boom In, Bust Out: Young Households and the Housing Price Cycle," *European Economic Review*, 43, 755–766.

Rothemberg, J., R. V. Butler, G. C. Galster, and J. R. Pitkin (1991): *The Maze of Urban Housing Markets: Theory, Evidence, and Policy*. University of Chicago Press: Chicago.

Tenreyro, S., P. Bracke, et al. (2017): "History Dependence in the Housing Market," in *2017 Meeting Papers*, no. 423. Society for Economic Dynamics.