# The BuildZoom & Urban Economics Lab Index
# Technical Document

Issi Romem*
BuildZoom

Albert Saiz
Massachusetts Institute of Technology

April 8, 2016

## 1 Introduction

The BuildZoom & Urban Economics Lab index is a joint endeavor of BuildZoom and the Urban Economics Laboratory of the MIT Center for Real Estate. The set of indices leverages BuildZoom's repository of building permit data to track residential permitting activity, and was developed jointly with Professor Albert Saiz of the Center for Real Estate at MIT. Separate indices track permitting for construction of new homes and for remodeling[1] of existing homes both nationally and in individual metropolitan areas.

The indices complement residential building permit statistics published by the U.S. Census. The Census' Building Permits Survey considers only new homes, not existing homes. In addition, whereas the Census figures are obtained from a regular questionnaire asking local governments how many building permits were issued,[2] the BuildZoom & Urban Economics Lab index is compiled from the ground up, based on individual building permit records.

This document briefly describes BuildZoom's building permit repository, and then proceeds to review the indices' underlying methodology, quantify their reliability, and benchmarks indices against data from the U.S. Census.

*Corresponding author. Email: issi@buildzoom.com.

[1]In this document the term remodeling is used broadly to include home improvement, maintenance and repair.

[2]The questionnaire is Form C-404.

## 2 The BuildZoom data repository

BuildZoom's repository of building permit data is collected from numerous building permit issuing authorities. The authorities are primarily building departments in cities and counties, and their geographic domains are referred to herein as *jurisdictions*. The precise nature of the data available differs across jurisdictions, as do the parameters of the collection process, such as the update frequency and the recorded fields. The ingestion of data into the repository includes a standardization process.

The data repository is expanding quickly. At the time of writing the data cover jurisdictions which are home to approximately one third of the U.S. population.[3] The data are geographically distributed such that coverage is available in most metropolitan areas.

New data is typically drawn from jurisdictions at a daily or weekly frequency, although in some jurisdictions the frequency is limited by the relevant authority's data release schedule. The median building permit record in the repository is dated from 2006 and the 10th percentile is dated from 1996. The historical extent of the data differs across jurisdictions.

The data consist of building permits associated with residential and non-residential properties. They include all types of permits, e.g. for electrical and plumbing work as well as structural work, and for work on new structures as well as existing structures. Building permits are classified into a growing number of types based on textual analysis.[4]

## 3 The "Additive Chain" Index Method

In an ideal setting, we would observe all building permits issued in an area dating back from the present, and by adding them up each month we would produce a time series tracking the level of permitting, and qualifying as an index.[5] The index would correspond to the area defined, such as a metro area,[6] a state or the entire U.S., and obtaining indices for different

[3] June 22, 2015.

[4] Building permit classification is currently performed using a mixture of Naive Bayesian classifiers and simple conditional string-matching, however building permit classification is an area of active research.

[5] The index would track construction and remodeling as reflected by the number of permits. Alternatively, we could add up the properties associated with permits issued each month to produce an index that tracks these activities as reflected by the number of *properties* undergoing permitted work, rather than the number of permits. To the extent that more intensive activity requires more permits, the former index would capture both the intensive and extensive margins, whereas the latter would capture only the extensive margin. The ratio of the former index to the latter would convey information regarding the intensive margin.

[6] In this document, metropolitan areas correspond to the standard Core-Based Statistical Areas (CBSAs) used by the Census.

types of activity, e.g. construction of new homes versus remodeling of existing ones, would require us to add up only building permits associated with that type.[7]

In practice, the definition of areas is straightforward and different types of permitted work are feasibly identified, but a substantial challenge emerges because only a partial set of building permits issued is ever observed.[8] This occurs because data is not available for all jurisdictions, and because the periods in which data are observed vary by jurisdiction.[9] Consider the example below, in which the number of permits in Jurisdictions A and B is observed in all months, but in Jurisdiction C it is only observed from March. Adding up the number of permits observed each month yields a nonsensical time series, because it implies a 25% increase from 40 to 50 permits between February and March. The jump obviously follows from the entry of Jurisdiction C's data stream in March, and does not reflect any real increase in activity.

| Permits observed in June | | | | | |
|---|---|---|---|---|---|
| Jurisdiction | Jan | Feb | Mar | Apr | May |
| A | 31 | 30 | 30 | 29 | 30 |
| B | 11 | 10 | 10 | 10 | 11 |
| C | | | 10 | 9 | 9 |
| Observed total | 42 | 40 | **50** | **48** | **50** |

To address the change in the composition of observed jurisdictions in March, we use a simple chaining method, which we refer to as the "additive chain" method. The ratio $\frac{A+B+C}{A+B} = \frac{50}{40}$ in March captures the proportionality, or relative scale, of Jurisdictions A, B and C to A and B alone. Dividing by the ratio adjusts ("chains") the sum of observed permits

[7] Any classification of building permits can be used for this purpose. For example, we could construct an index of solar power system installations, or an index of kitchen remodeling.

[8] We identify permitted work on new and existing homes as follows. We identify building permits corresponding to residential properties using a binary Naive Bayesian classifier, and only permits identified as such contribute to the indices for new home construction and existing home remodeling. We identify building permits involving new construction using a separate binary Naive Bayesian classifier. We identify building permits involving existing structures as those not involving new construction, and not issued within 12 months of a building permit involving new construction on the same property. We identify building permits involving demolition work, which often pre-date new construction work, using a third Naive Bayesian classifier, and for the purpose of the indices we consider them as involving neither new construction nor existing structures.

[9] The absence of data from some jurisdictions begs the question whether and when an observed subset of jurisdictions is sufficient to represent the entire area. We address the matter in detail in Section 4.

from March onwards to the scale of Jurisdictions A and B alone.[10] Instead of introducing a nonsensical shift in March, the chained time series shows a 4% drop from 40 to 38.4.[11] Of course, the scale associated with Jurisdictions A and B alone is of no particular interest, and it is changes in the series - as opposed to levels - which convey information of interest. We obtain the index by adjusting the level of the chained time series to 100 in a selected reference period.[12]

| Permits observed in June | | | | | |
|---|---|---|---|---|---|
| | Jan | Feb | Mar | Apr | May |
| A | 31 | 30 | 30 | 29 | 30 |
| B | 11 | 10 | 10 | 10 | 11 |
| Entry: C | | | **10** | **9** | **9** |
| Sum of A and B | 42 | 40 | 40 | | |
| Sum of A, B and C | | | 50 | 48 | 50 |
| Chained series | 42 | 40 | **40** | **38.4** | **40** |
| Index | 100 | 95.2 | 95.2 | 91.4 | 95.2 |
| % Change | | -5 | 0 | -4 | 4.2 |

The "additive chain" method accounts for the more general case, in which jurisdictions' can also exit the data, and in which multiple entries and exits can occur in the same period.[13] In the extended example below, Jurisdiction C enters at the same time that Jurisdiction D exits. We define the *backward-looking* sum as the sum of observed permits issued by jurisdictions which are also observed in the *previous* period, and similarly the *forward-looking* sum for jurisdictions which are also observed in the *next* period. In the general

[10]Of course, the ratio only reflects the relative scale in the *initial* overlap period (March), and in principle we could incorporate multiple overlap periods to obtain a ratio that is less susceptible to short term fluctuations. However, doing so appears to have little practical significance, especially in areas with a large set of potential jurisdictions, so for the sake of simplicity we obtain the ratio from the initial overlap period alone.

[11]The change is a weighted average of the separate changes in Jurisdictions A, B and C, with March numbers as weights. Arithmetically: $\frac{30}{50} \cdot \frac{29-30}{30} + \frac{10}{50} \cdot \frac{10-10}{10} + \frac{10}{50} \cdot \frac{9-10}{10} = \frac{38.4-40}{40} = -0.04$.

[12]Currently, the reference period is set to January 2006, which is the January most closely approximating the peak of the cycle preceding the recent housing market crisis. In areas in which data is not available as early as January 2006, we use the first month of the index as the reference period.

[13]Because there is substantial variability across jurisdictions in the availability of digitized historical data, new data stream entry is, in fact, the norm. Exiting data streams, on the other hand, are far less common, and typically result from temporary data collection issues.

case, the relative scale of the set of jurisdictions observed in the subsequent period and the set observed in the previous period is captured by the ratio of forward- and backward-looking sums. We chain the series in March by dividing the forward-looking sums from March onwards by the ratio.[14] The additional information from Jurisdiction D affects the level of the index throughout, but it only affects changes in the index during the periods to which the information applies (February and March, but not April and May). Finally, if the composition of observed jurisdictions changes multiple times, we chain iteratively on the outcome of the previous iteration, in chronological order.

| Permits observed in June | | | | | |
|---|---|---|---|---|---|
| | Jan | Feb | Mar | Apr | May |
| A | 31 | 30 | 30 | 29 | 30 |
| B | 11 | 10 | 10 | 10 | 11 |
| Entry: C | | | **10** | **9** | **9** |
| Exit: D | **36** | **34** | **35** | | |
| Backward-looking sum (A, B and D) | 78 | 74 | 75 | | |
| Forward-looking sum (A, B and C) | | | 50 | 48 | 50 |
| Chained series | 78 | 74 | 75 | 72 | 75 |
| Index | 100 | 94.9 | 96.2 | 92.3 | 96.2 |
| % Change | | -5.1 | 1.4 | -4 | 4.2 |

In addition to the steps of the "additive chain" method described, we smooth the published indices using a centered 3-month moving average, after which we adjust for seasonality.[15]

We provide a mathematical representation of the index in Appendix A.1. Before using them, we subject the data from different jurisdictions to a sequence of quality controls, the

[14]Note that periods in which the backward-looking sum equals zero are problematic because then the ratio involves division by zero. This situation is more likely to emerge when the area corresponding to the index is smaller, or when the type of permitted work is very narrowly defined. Missing values of the forward- or backward-looking sums present a similar problem (the first backward-looking sum and the last forward-looking sum of a series are not missing because we specially define them as though the observed set of jurisdictions remains constant across the relevant periods - see Appendix A.1). To address the problem of zero or missing values of backward-looking sums, we scan the jurisdiction-level time series entering the "additive chain" method for gap periods in which no jurisdictions are observed. If one or more such periods exist, we omit any information prior to the last gap period.

[15]We adjust for seasonality using the Census' X-13 ARIMA-SEATS software.

details of which we provide in Appendix A.2. The following section quantifies the reliability of the index, and briefly discusses some alternative approaches that were considered.

# 4 Reliability of the Index

Recall the ideal setting in which all building permits issued in an area dating back from the present are observed. For a sufficiently large area, such as a metro area, it is intuitive that even if records from one neighborhood were lost, the index would still reliably track activity in the area as a whole. The opposite is also intuitive; if all records were lost *except* for those from one neighborhood the resulting index would be unreliable, because the particular neighborhood whose data survived could be quite unrepresentative of the whole metro area.

In practice, intermediate cases prevail, in which BuildZoom's data coverage in most areas is substantial, but incomplete, because it is limited to observed jurisdictions. To assess the degree to which the observed jurisdictions in a metro area represent the whole, we conducted a Monte Carlo simulation. In each metro area, we used the index obtained from the full set of observed jurisdictions as a benchmark, and we produced additional indices from a large number of randomly drawn samples of jurisdictions.[16] At every point of every index we recorded the percent deviation of each sample-based index from the benchmark index, which we refer to henceforth as the error, and we plotted the error distribution against the sample coverage of the benchmark population.[17,18]

Figures 1a and 1b show the simulation results for remodeling of existing homes. Figure 1a reports the median as well as the 75th and 90th percentiles of the error distribution against population coverage, expressed as a percent of the benchmark population, and Figure 1b does the same but expresses population coverage in terms of people.[19]

[16]We drew up to 100 samples without replacement for each metro area $\times$ jurisdiction N-tuple size pair (for the smallest and largest N-tuple sizes, the number of possible samples is less than 100). To simulate indices corresponding to permitting for existing homes we compared 34 benchmark indices to a total of 1,245,864 simulated indices. To simulate indices corresponding to permitting for new homes we compared 35 benchmark indices to a total of 1,321,809 simulated indices. Differences in the numbers corresponding to existing homes and new homes stem from differences in the number of jurisdictions whose time series emerge as valid from the data quality control measures described in Appendix A.2, which we performed separately for existing homes and new homes.

[17]In fact, the errors recorded deviations of *changes* in each sample-based index from contemporaneous *changes* in the benchmark index, as opposed to deviations of *levels*. This is because deviations in levels during all periods are unduly susceptible to noise in the reference period, whereas deviations in the changes are not.

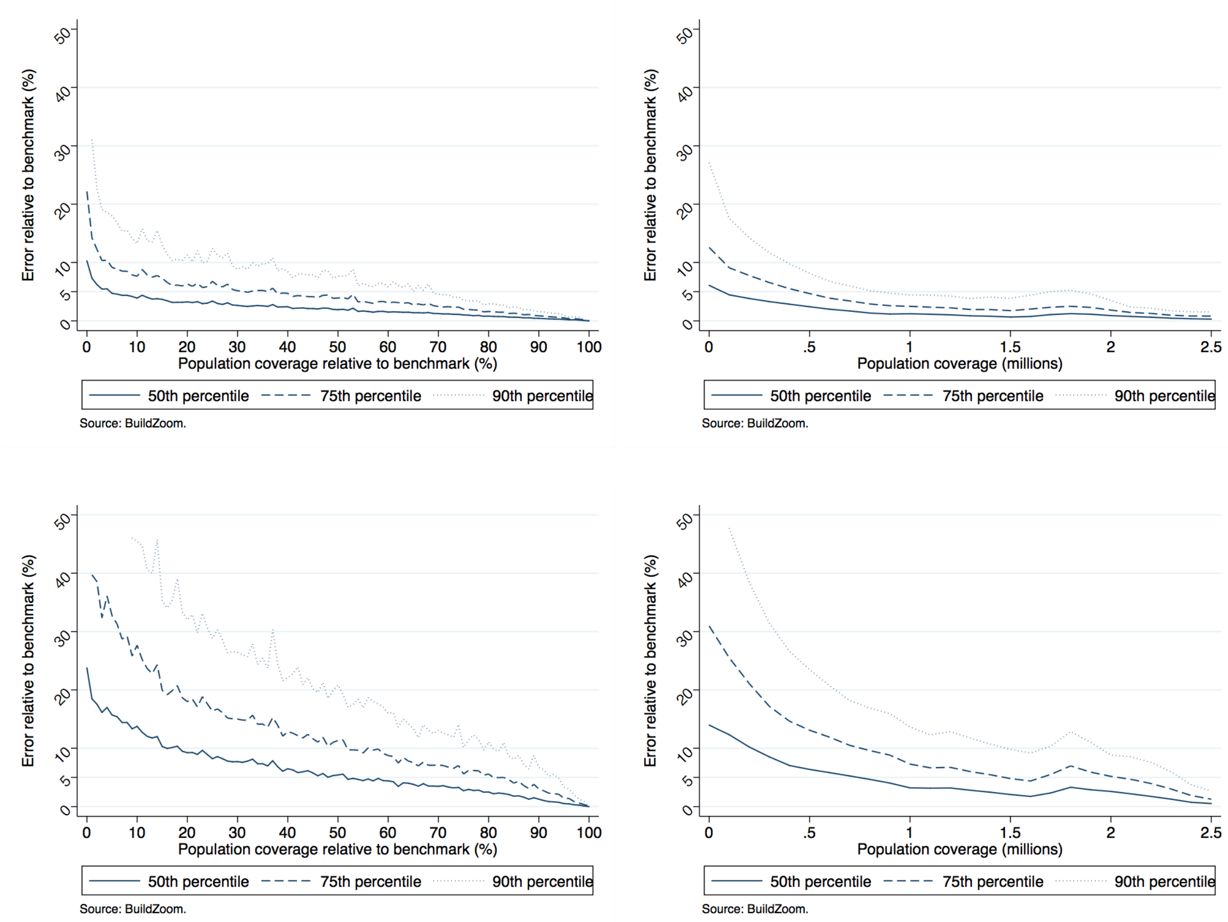
[18]Because of the variation across jurisdictions' coverage periods we observed sample coverage of the benchmark population separately for every point of every index.

[19]The reported distribution is of errors' absolute value.

Figure 1: Simulated error distribution versus benchmark population coverage

(a) Existing homes; coverage in percent

(b) Existing homes; coverage in people




(c) New homes; coverage in percent

(d) New homes; coverage in people

We set two required thresholds for publishing a metro area index:

1. Coverage of at least 30% of the metro area population.

2. Coverage of at least 1 million metro area residents.[20]

The thresholds do not guarantee that every published index will perfectly mimic the index that would have been obtained with full coverage, but they do suggest that in published indexes of remodeling of existing homes large deviations will be uncommon. From Figure 1a we expect that over 90% of data points exceeding the first threshold will have errors below 10%, and we expect that more than 75% will have errors below 5%. From Figure 1a we expect that over 90% of data points exceeding the second threshold will have errors below 5%.[21]

Errors in indexes of permitting for new homes are substantially higher than those corresponding to existing homes, because new construction tends to be less common and more volatile than remodeling work, and because it tends to cluster in certain parts of metropolitan areas, such as the fringe of expanding metros, which makes the selection of observed jurisdictions more pertinent. The results for new home construction are given in Figures 1c and 1d. Nevertheless, for the sake of consistency, the publication of indexes corresponding to new homes is determined by the same thresholds as those corresponding to existing homes.

Note that the sampling method used is such that the reported error rates are conservative. The sampling method used in the simulation, whereby jurisdictions are randomly selected with equal probability, does not accurately capture the way in which jurisdictions are added to the BuildZoom data repository. In order to mimic the actual selection process, we performed a separate simulation in which jurisdictions' probability of selection increased with population. Compared to the previous sample selection regime, population-based sample selection yields error rates that are slightly *lower* at every level of benchmark population coverage.[22,23]

In an attempt to more accurately represent entire metropolitan areas using just the observed jurisdictions, we experimented with a variant of the "additive chain" index method

[20] Note that in metro areas for which 30% of the population exceeds 1 million, i.e. with a total population greater than 3 and 1/3 million, the second threshold is redundant, while in all other metro areas the first threshold is redundant. The second threshold can be thought of as a substitute for the first one in smaller metro areas which is more stringent.

[21] These results hold when each threshold is applied only to metro areas in which the threshold is non-redundant - see footnote 20.

[22] The result is conditional on holding the number of jurisdictions in a sample fixed.

[23] Simulation results with population-based sample selection are available upon request.

in which we weighted permits according to their similarity to the metro area as a whole. In the variant, permits from jurisdictions that are more similar to the metro area as a whole are weighted so as to exert greater influence on the index.[24] As shown in Figures 2a-2d, and despite substantial differences across jurisdictions in their similarity to the metro as a whole, simulations using the similarity-weighted variant produced error rates which are remarkably close to the original. The implication is that the similarity-weighted variant adds complexity without improving upon the original "additive chain" method, which is the one we ultimately use. Although additional refinements of the notion of similarity-weighting may still potentially reduce error rates, the exercise we performed suggests that their benefit is likely to be bounded.[25] The best solution for the problem at hand is to expand the set of observed jurisdictions, and we are aggressively pursuing it.

In addition to the "additive chain" index method and its similarity-weighted variant, we used a third, regression-based index method which yielded remarkably similar indices. The method considers changes over time in the number of permits within individual jurisdictions, and produces an index which is a population-weighted average of the jurisdiction-level changes.[26] This method does not map into a simple (i.e. linear) form of permit weighting,

[24] We observed similarity based on the following list of jurisdiction characteristics, which we deemed relevant ex-ante: (i) share of households with annual income above $100,000; (ii) log vacancy rate; (iii) share of renter households; (iv) share of housing units in multifamily structures; (v) log share of housing units built after 1980. We drew data on the characteristics from the 2009-2013 5-year American Community Survey files. We set jurisdiction weights as the inverse of jurisdictions' Mahalanobis distance to the metro area mean in the characteristic space, using a metro area-specific variance matrix in the Mahalanobis metric.

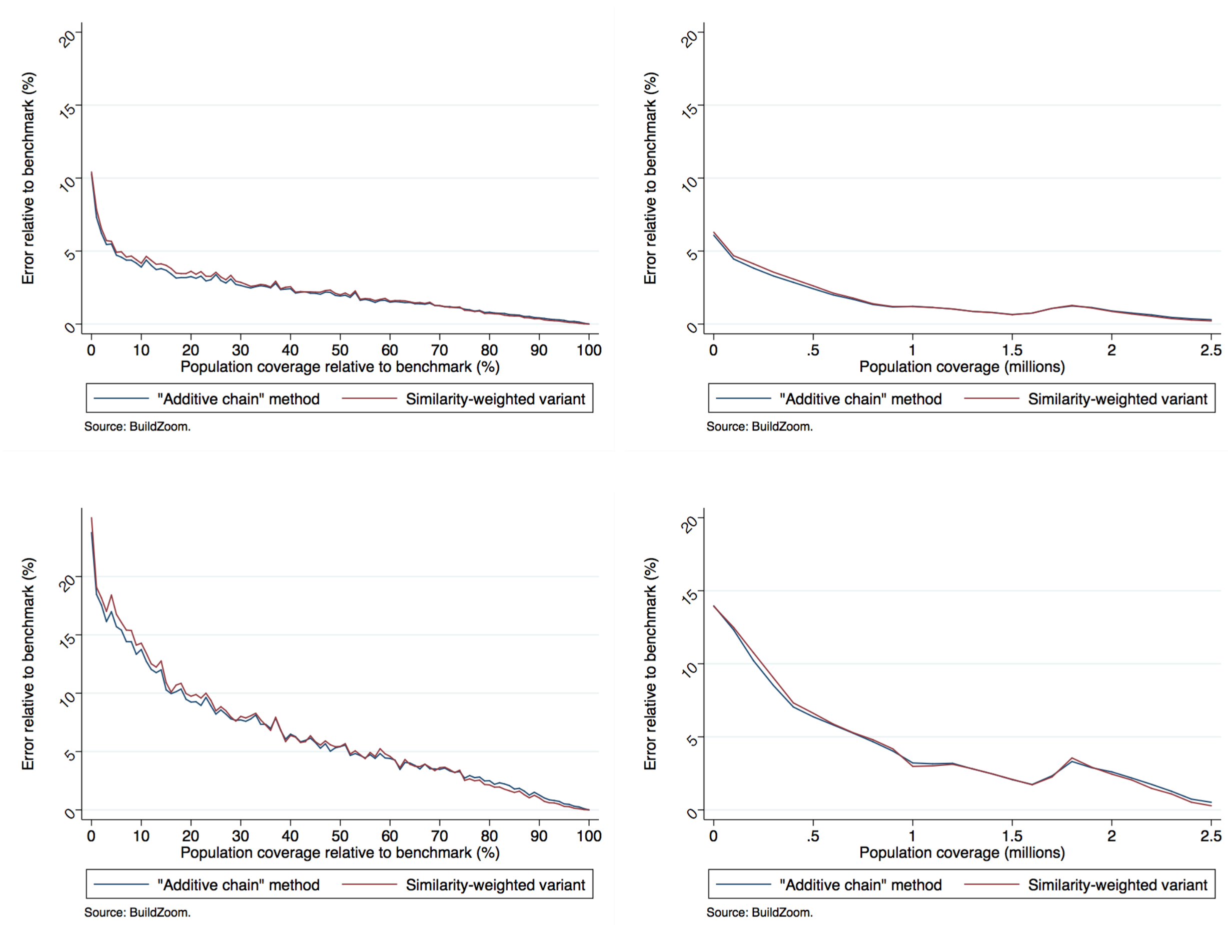
[25] An example of such a refinement is a variant of the "synthetic matching" technique, in which jurisdictions are weighted so that the resulting index mimics the index of the metro area as a whole. Applying refinement is feasible (we applied it) but applying it repeatedly as part of a Monte Carlo simulation is not.

[26] The regression-based method involves running the weighted least squares regression $\log Y_{it} = \phi_t + \theta_i + \epsilon_{it}$, weighted by jurisdiction populations, where $Y_{it}$ is the permit count for jurisdiction $i$ in month $t$, $\phi_t$ is a time fixed effect, $\theta_i$ is a jurisdiction fixed effect and $\epsilon_{it}$ is an error term. Defining $t = 0$ as the reference period and setting it as the omitted period of the time fixed effects, we derived the index as follows. The regression model implies $\log Y_{it} - \log Y_{i0} = \phi_t - \phi_0 + \epsilon_{it} - \epsilon_{i0}$, from which it follows that $Y_{it}/Y_{i0} = e^{\phi_t - \phi_0 + \epsilon_{it} - \epsilon_{i0}}$ and that $E(Y_{it}/Y_{i0}) = e^{\phi_t - \phi_0} E(e^{\epsilon_{it} - \epsilon_{i0}})$. Note that even if $E(\epsilon_{it} - \epsilon_{i0}) = 0$ for all $i, t$, it is generally *not* the case that $E[e^{\epsilon_{it} - \epsilon_{i0}}] = 1$, because $\exp(\cdot)$ is a non-linear transformation. However, assuming that $\epsilon_{it} - \epsilon_{i0} \sim N(\mu, \sigma^2)$ implies that $e^{\epsilon_{it} - \epsilon_{i0}} \sim logN(m, v)$, such that $m = e^{\mu + \sigma^2/2}$ (the assumption that $\epsilon_{it} - \epsilon_{i0} \sim N(\mu, \sigma^2)$ is supported by the fact that the distribution of differences between residuals from the regression corresponding to $\epsilon_{it} - \epsilon_{i0}$ visually resembles a normal distribution). Using the predicted values from regression and assuming that $E(\epsilon_{it}) = 0$ for all $i$ and $t$, we estimated $Y_{it}/Y_{i0}$ as $\widehat{Y_{it}/Y_{i0}} = e^{\widehat{\log Y_{it}} - \widehat{\log Y_{i0}} + s^2/2}$, where $s^2 = var(r_{it} - r_{i0})$ and $r$ denotes residuals from the regression. We obtained an index in levels by setting the value for the reference period to 100, and proportionately scaling the values in the remaining periods, after which we smoothed and adjusted for seasonality in the same manner described in section 3. The estimator $\widehat{Y_{it}/Y_{i0}}$ guarantees that $\widehat{Y_{it}/Y_{i0}} \geq 0$, and its interpretation does *not* rely on log approximations of percentage changes. However, the method as a whole is less transparent than the "additive chain" index method, and it is more sensitive to the treatment of instances in which $Y_{it} = 0$ at the jurisdiction level, whereas the "additive chain" method only encounters difficulty if the sum of permits in *all* observed

Figure 2: Median simulated error using "additive chain" method vs. similarity-weighted variant

(a) Existing homes; coverage in percent

(b) Existing homes; coverage in people




(c) New homes; coverage in percent

(d) New homes; coverage in people

but it does implicitly treat every resident of the area as though he or she were a representative agent of his or her jurisdiction, and it weights each resident equally, so more heavily populated jurisdictions sway the index more forcefully. The fact that the regression-based method yields very similar indices to the "additive chain" method suggests that the path of permitted work captured by the "additive chain" method is robust, in the sense that it is not driven by the methodology.

# 5 Validation

The Census regularly publishes the number of new housing units authorized by building permits. This data provide a benchmark against which we can validate the BuildZoom & Urban Economics Lab index for new home construction. Figure 3a compares the non-seasonally adjusted BuildZoom & Urban Economics Lab index with the corresponding Census numbers. The similarity between the series is striking.

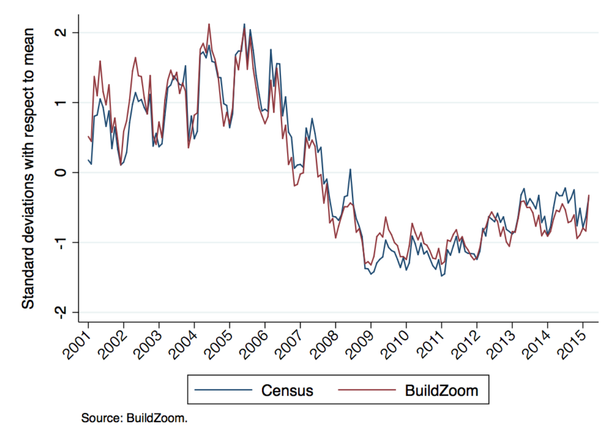
Nevertheless, there is a small alternating gap between the series. Figure 3b presents smoothed versions of the two series which make the gap easier to see. The gap emerges primarily because the series measure slightly different things: the BuildZoom & Urban Economics Lab index reflects the *number of permits* issued, whereas the Census figures correspond to the *number of housing units* authorized. The number of permits required for a construction project is more closely related to the number of structures than the number of housing units. As a result, when new construction shifts towards multi-family projects - or more precisely, when the average number of housing units per permit increases - the Census figures are more positively affected than the BuildZoom & Urban Economics Lab index, and vice versa. Figure 3c plots two additional curves: one showing the gap (on a separate scale), and another showing the share of new construction corresponding to multi-family projects as per the Census. The two curves appear to move roughly in tandem, which suggest that the multi-family share of new homes explains much of the variation over time in the Census to Index gap, especially during and since the recent housing crisis. If it were possible to observe the average number of housing units per permit, that curve would likely align even more closely with the gap curve.

On the whole, comparing the BuildZoom & Urban Economics Lab index for new home construction to the Census numbers of authorized new housing units suggests that the index is, in fact, a reliable measure of residential new construction, and by association that it is a

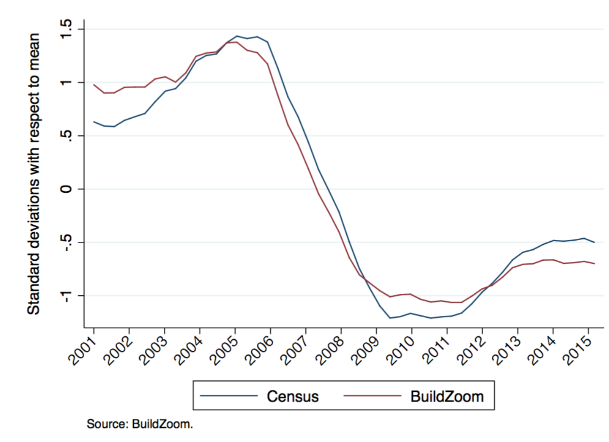
jurisdictions in a period equals 0.

Figure 3: New home construction

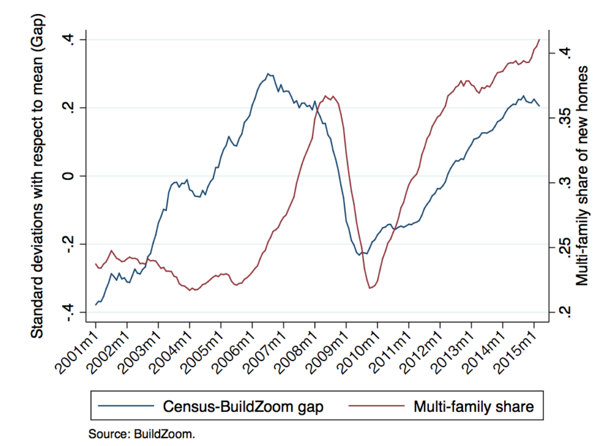
(a) Raw series




(b) Smoothed




(c) Census-BuildZoom gap vs. multi-family share



reliable measure of remodeling of existing homes as well.

# A Appendices

## A.1 Mathematical Representation of the "Additive Chain" Index Method

Let $\mathcal{J}$ be the full set of potentially observable jurisdictions in an area. Let $P_{jt}$ be the number of permits observed in jurisdiction $j \in \mathcal{J}$ in month $t \in \{1, \ldots, T\}$, let $\mathcal{B}_t \subseteq \mathcal{J}$ be the set of backward-looking jurisdictions in month $t$, i.e. $\mathcal{B}_t \equiv \{j : P_{j,t-1} \text{ is observed}\}$, and let $\mathcal{F}_t \subseteq \mathcal{J}$ be the set of forward-looking jurisdictions in month $t$, i.e. $\mathcal{F}_t \equiv \{j : P_{j,t+1} \text{ is observed}\}$. In addition, define the end-cases $\mathcal{B}_1$ and $\mathcal{F}_T$ as $\mathcal{B}_1 \equiv \mathcal{B}_2 \equiv \{j : P_{j,1} \text{ is observed}\}$ and $\mathcal{F}_T \equiv \mathcal{F}_{T-1} \equiv \{j : P_{j,T} \text{ is observed}\}$. Finally, let $t_{ref} \in \{1, \ldots, T\}$ be the reference period. The elements of the chained series are then

$$I_t^* = \frac{\sum_{j \in \mathcal{F}_t} P_{jt}}{\prod_{s=1}^{t} R_s},$$

where $R_t$ is the forward- to backward-looking ratio

$$R_t \equiv \frac{\sum_{j \in \mathcal{F}_t} P_{jt}}{\sum_{j \in \mathcal{B}_t} P_{jt}},$$

and the elements of the Index are

$$I_t = \frac{100 \cdot I_t^*}{I_{t_{ref}}^*}.$$

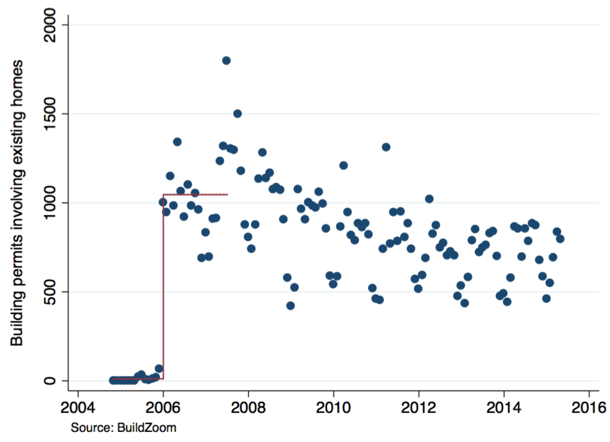
## A.2 Time Series Quality Controls

Prior to the application of the "additive chain" index method, we subject each jurisdiction-specific time series to the following sequence of quality control measures.

1. We unconditionally omit the first month of each time series, as it tends to contain only partial data.

2. Before the full flow of permits is observed, the time series often exhibit a trickle of permits that is obviously incomplete. For example, data for the City of Chicago shown in figure A.1 indicate a typical flow of several thousand permits per month from January

Figure A.1: Permits associated with remodeling of existing homes in Chicago




2006 to the present, however the data also contain a trickle of a single- or double-digit number of permits each month from October 2004 until January 2006, which is obviously incomplete. To address this kind of issue, we use structural break tests to identify the "effective start" of each time series, before which we consider the data unreliable, and we truncate observations before that time.[27] The red line in Figure A.1 indicates the time period determined to be the "effective start" of the time series.

3. We use similar structural break tests to identify any sharp fall in the number of permits that acts as an "effective end" of the series. Such instances are an exception, and usually indicate a temporary deficit in data collection, which renders the data from that time onwards unreliable. We truncate observations following an "effective end" of a series until the deficit is corrected.[28]

4. Notwithstanding the "effective start" or "effective end" of a time series, we use a third

[27]The structural break test is as follows. For every time period, $s$, we run the regression $Y_t = \alpha + \beta D_t(s) + u_t$ on observations from the window $s - 18$ to $s + 18$, where $Y_t$ is the monthly count of relevant permits for the jurisdiction in month $t$, $D_t(s)$ equals $\mathbb{1}\{t \geq s\}$ and $u_t$ is an error term. We identify a single structural break candidate in the period $s^*$ which maximizes the $R^2$ among the set of regressions. If the regression corresponding to $s^*$ yields $\hat{\beta} > 0$ and $\frac{\hat{\alpha}+\hat{\beta}}{\hat{\alpha}} > k_{start}$ for a positive threshold $k_{start}$, we designate the candidate structural break at $s^*$ the "effective start" of the series, and we truncate the observations preceding and including the break. We limit the set of candidate periods to the first 10 years of the time series. The results of the procedure generally align with the results of manually inspecting the "effective start" of the data, so the procedure essentially serves to automate the manual inspection process.

[28]The structural break test identifying the "effective end" of a time series is identical to the one identifying the "effective start," except that $s^*$ is required to yield $\hat{\beta} < 0$ and $\frac{\hat{\alpha}+\hat{\beta}}{\hat{\alpha}} < k_{end}$ for a positive threshold $k_{end}$. We limit the set of candidate periods to the last 2 years of the time series and run the regression on a window $s - 12$ to $s + 12$.

structural break test to automatically flag for manual inspection any additional changes in levels which are unusually sharp.[29]

5. We automatically identify low and high outliers and flag them for manual inspection.[30,31]

6. As described in footnote 8, we do not consider building permits as involving remodeling of existing structures when they were issued within 12 months of a building permit involving new construction on the same property. As a result, the first 12 months of a time series is likely to underestimate the number of permits involving existing structures. For this reason, we omit the first 12 months following the "effective start" of each time series. Purely for the sake of time frame consistency, we also omit the same period with respect to time series of new home construction.[32]

[29]The structural break test is similar to those identifying the "effective start" and "effective end" of a series, but it is not identical. For every time period, $s$, we run the regression $Y_t = \alpha + \beta D_t(s) + \gamma_1 t + \gamma_2 t^2 + \gamma_3 t^3 + \gamma t^4 + u_t$ on observations from the window $s - 60$ to $s + 60$, where $Y_t$ is the monthly count of relevant permits for the jurisdiction in month $t$, $D_t(s)$ equals $\mathbb{1}\{t \geq s\}$, $(t, t^2, t^3, t^4)$ comprise a quartic polynomial time trend, and $u_t$ is an error term. We identify a single structural break candidate in period $s^*$ which maximizes the $R^2$ among the set of regressions. If the regression corresponding to $s^*$ yields a Bonferroni-corrected p-value below 1 percent for the test of the null hypothesis $H_0 : \beta = 0$, then we assign the candidate structural break at $s^*$ for manual inspection. The test indicates all qualifying structural breaks.

[30]We identify month $t^*$ as a low outlier if the level of permits that month is less than 20 percent of the mean level during the interval $t - 24 \leq t \leq t + 24$ (measured in months), and neither the lead month, the lagging month nor the combination of the two exhibit a compensating amount of permits, as per the test immediately described. The idea is that a low outlier may indicate a period during which a permit-issuing authority may have been less than fully functioning, and in which it diverted the incoming flow of permits to the following and/or preceding months. If this a low outlier is not "compensated" by higher months before and/or after, we flag it for manual inspection. The test is as follows: for each potential outlier, we run the regression $Y_t = \alpha + \beta_0 D_{current} + \beta_1 D_{lead} + \beta_2 D_{lag} + u_t$ on observations from the window $t - 24$ to $t + 24$, where $Y_t$ is the monthly count of relevant permits for the jurisdiction in month $t$, $D_{current}$ is an indicator that $\mathbb{1}\{t = t^*\}$, $D_{lead}$ is an indicator that $\mathbb{1}\{t = t^* + 1\}$, $D_{lag}$ is an indicator that $\mathbb{1}\{t = t^* - 1\}$, and $u_t$ is an error term. If $\beta_1 < 0.5 \cdot |\beta_0|$ and $\beta_2 < 0.5 \cdot |\beta_0|$ and $\beta_1 + \beta_2 < 0.5 \cdot |\beta_0|$, then we flag the low outlier for manual inspection.

[31]We identify month $t^*$ as a high outlier and flag it for manual inspection if the level of permits that month is 3 or more standard deviations above the jurisdiction mean during the interval $t - 24 \leq t \leq t + 24$ (measured in months). Standard deviations correspond to this interval as well.

[32]Note that a symmetric problem could emerge with respect to building permits that *precede* new construction, such as demolition permits. Such cases would cause the final months of a time series to overestimate permitting associated with existing homes, but omitting the final months would unreasonably extend the the indices' publication lag. Because such cases primarily involve demolition, and for the purpose of the index alone, we do *not* associate demolition work with either new construction or existing structures.