[Abstract] This paper identifies a critical issue in the Weighted Repeated Sales (WRS) method – the omission of market risk in the weight estimation model specified by Case and Shiller (1989). It demonstrates that the omission of market risk is conceptually unjustified. Through extensive examination of the real estate market risk, the study not only proposes a modified WRS model that is empirically supported, but also contributes to the broad discussion on the holding period dependence of real estate risk. It also shows that the Case-Shiller weighting method is likely to be mis-specified in nine of the ten cities where the Case-Shiller metro indices are “tradable” with housing options and futures contracts listed on the Chicago Mercantile Exchange. Using a large sample of repeated sales from Washington DC area, the original repeated sales method of Bailey, Muth, and Nourse (1963) and the Case-Shiller method are compared against the modified WRS, and the results indicate that market risk plays an important role in the index estimation.
Market Risk Factor and the Weighted Repeated Sales Method

[Abstract] This paper identifies a critical issue in the Weighted Repeated Sales (WRS) method – the omission of market risk in the weight estimation model specified by Case and Shiller (1989). It demonstrates that the omission of market risk is conceptually unjustified. Through extensive examination of the real estate market risk, the study not only proposes a modified WRS model that is empirically supported, but also contributes to the broad discussion on the holding period dependence of real estate risk. It also shows that the Case-Shiller weighting method is likely to be mis-specified in nine of the ten cities where the Case-Shiller metro indices are “tradable” with housing options and futures contracts listed on the Chicago Mercantile Exchange. Using a large sample of repeated sales from Washington DC area, the original repeated sales method of Bailey, Muth, and Nourse (1963) and the Case-Shiller method are compared against the modified WRS, and the results indicate that market risk plays an important role in the index estimation.

1. Introduction

The Weighted Repeated Sales (WRS) regression, popularized by Case and Shiller (1987, 1989), is the methodology behind the highly influential S&P/Case-Shiller Home Price Indices. Since its inception, the Case-Shiller Home Price Indices have been widely regarded by economists, mortgage lenders, securities issuers, insurers, rating agencies, and policy makers as the most authoritative and reliable benchmark for monitoring home price movements in the U.S. In mid-2006, the Chicago Mercantile Exchange (CME) began to trade housing options and futures contracts that are based on the “tradable” Case-Shiller home price indices for ten U.S. metropolitan areas. This further enhances the Indices’ status as the predominant benchmark in the U.S. housing market.

The WRS is regarded as a revolutionary improvement over the original repeated sales method pioneered by Bailey, Muth, and Nourse (1963). The key innovation of Case and Shiller is the modification of the original BMN method into a three-step weighted least square
procedure, which takes into account the different time intervals between paired (repeated) sales. Case and Shiller argue that paired sales with longer time intervals between sales should be given less weight than those with shorter time intervals in the regression (hence the name Weighted Repeated Sales regression) and they contend that the resultant index is sensitive to the weighting estimate. The real innovation of WRS, therefore, is the specification of a linear model to obtain the necessary regression weights, which are then used for the weighted regression in the third step.

While the idea of a weighted regression is conceptually correct, Case and Shiller’s linear model for estimating the weights suffers from an important deficiency which undermines the validity of the WRS methodology that is now widely trusted. As our analysis will show, Case-Shiller’s weighting specification (the step two of WRS) omits the market risk factor. Such omission is unjustified in light of their own empirical findings in the same study (Case and Shiller (1989)) that elaborates the WRS method. Details of this self-contradicting logic are revealed in Section 2 of this paper.

A method for producing market index should not ignore the market risk itself. But to properly incorporate this risk requires an understanding of the time-varying nature of market risk. Traditionally, under the Efficient Market Hypothesis, it is widely accepted that security asset returns are independent and identically-distributed (i.i.d.) over time. Case and Shiller (1989) conclude that the real estate market is not efficient because property returns are clearly not i.i.d. But if not i.i.d, what is it? Case and Shiller did not try to answer this question, instead they simply asserted that market risk has no bearing on the weight estimation. The current study tackles this question directly. In Section 3, we present extensive empirical evidence on the relationship between real estate market risk and the holding time (or the time interval between
paired sales). The findings are presented in what we called the *risk lines*. They are direct observations from a wide range of real estate market and submarket indices without resorting to complex statistical manipulations. This part of the findings makes a separate contribution to the literature because the issue is of broad importance. It is widely documented in the literature that real estate returns are not i.i.d. over time, the current study now reveals what they actually are. Future research in real estate investment need not (and indeed should not) continue to rely on the i.i.d. assumption. Conversely, all the classical finance theories (such as the MPT and CAPM) that depend on the i.i.d. assumption should not be naively applied to real estate investment and portfolio analysis. Deviation from the traditional finance paradigm is the imperative first step toward establishing the real estate investment discipline.

For the task at hand, knowledge of the market risk factor provides a solid empirical foundation for us to propose in Section 4 an alternative weight specification that remedies the deficiency in Case and Shiller’s original model. Furthermore, unlike previous studies that assume one model fits all markets, we highlight the notion that proper model specification should be market-specific. For example, we find that the east and west coastal cities are better fitted with one model specification while the inland cities tend to fit a different one. This conclusion is also consistent with the volatility clustering pattern found in Miles (2008) and several others. That being said, we find that, for nine out of ten metropolitan areas where the Case-Shiller Indices serve as the base for housing-related futures and options, the current linear specification is mis-specified, which are likely to result in less accurate price indices.

Finally, we provide an empirical comparison between the original BMN method, the Case-Shiller method, and our alternative method. The WRS was embraced by many primarily because it is conceptually superior. Few have examined whether the Case-Shiller method makes
an empirical difference over the original BMN method. Using a large sample of repeated sales from Washington DC metropolitan statistical area, we first show that market risk clearly affects index outcome – the index differences are smaller when the market is stable but become bigger when the market is more volatile. Furthermore, the results suggest that the Case-Shiller index offers slight improvements over the BMN index when the market is stable, but less so than the BMN index due to the fact that its misspecified model does not weight correctly (i.e. overweighting noisy data and underweighting accurate information) during periods when the market is volatile.

2. Review of the Weighted Repeated Sales Procedure

According to Case and Shiller (1989), the WRS method is based on the assumption that the log price $P_{i,t}$ of the $i$th house at time $t$ can be viewed to have three components:

$$P_{i,t} = C_t + H_{i,t} + N_{i,t}$$

(1)

where $C$ is the log of the citywide level of housing prices at a particular time, which captures the location or market effect on price. $H$ captures the impact of property-specific factors on price and is assumed to follow the Gaussian random walk. $N$ is an identically distributed normal noise term (which has zero mean and variance $\sigma_N^2$ ). It is further assumed that $C, H, N$ are uncorrelated.\(^3\) Equation (1) is consistent with the common notion that the price of an individual property is affected by three factors – the location (market conditions), the property (asset specific characteristics), and the random trading noises. What is interesting, though, is that Case

\(^3\) While these assumptions may be debatable to some, we take the Case-Shiller’s assumptions as-is and focus on the second step of WRS, which is the heart of their innovation.
and Shiller were clear about the distributional form of the $H$ and $N$ factors, but they did not mention the same for the market factor $C$ and, as explained below, they went on to ignore the factor all together in the step two of WRS.

The three-step WRS procedure was described in Case and Shiller (1989) as follows:

“In the first step, the BMN procedure was followed exactly, and a vector of regression residuals was calculated. In the second step, the squared residuals in the first step regression were regressed on a constant and the time interval between sales. The constant term was the estimate of $\sigma^2_N$ and the slope term was the estimate of $\sigma^2_H$. In the third step, a generalized least squares regression (a weighted regression) was run by first dividing each of the observations in step-one regression by the square root of fitted value of step-two regression and running the regression again.” (pages 126-127)

Reading the second step above, one may naturally ask: What about the variance of $C$, or $\sigma^2_C$? Why is there no mentioning of the city-wide price risk factor (or the market risk factor) in the above three steps?

To explain the question, suppose that a house was bought at time $s$ and then sold at time $t$. To follow Case and Shiller’s assumption in Equation (1), the first step of the WRS regression is to compute the dependent variable, which is defined as the difference between the paired log prices at time $s$ and $t$. That is,

$$P_t - P_s = (P^C_t - P^C_s) + (P^H_t - P^H_s) + (P^N_t - P^N_s)$$

(2)
Note that the log price differences are essentially the returns over the time interval \((t - s)\) on a continuous compounding basis. For convenience, we denote the return with a general symbol \(R\), and rewrite Equation (2) as:

\[
R_{t,s} = R_{t,s}^C + R_{t,s}^H + R_{t,s}^N
\]  

(3)

Where \(R_{t,s}\) denotes the total return of a house that was bought at time \(s\) and then sold at time \(t\). \(R_{t,s}^C\) and \(R_{t,s}^H\) capture the market (at city level) and property-specific return effects, respectively, and \(R_{t,s}^N\) is purely a random noise.

Mathematically, under Case and Shiller’s assumption that the three effects \(C, H,\) and \(N\) are uncorrelated, the variance of the paired log price change from time \(s\) to \(t\), \(R_{t,s}\), can be obtained as

\[
Var(R_{t,s}) = Var(R_{t,s}^C) + Var(R_{t,s}^H) + Var(R_{t,s}^N)
\]  

(4)

This implies that the variance (risk) of individual property’s return over the interval \((t - s)\) is determined by three factors – the market risk \(Var(R_{t,s}^C)\), the property-specific risk \(Var(R_{t,s}^H)\), and a random noise \(Var(R_{t,s}^N)\). In addition, Case and Shiller (1989) assume that \(R_{t,s}^H\) follows the Gaussian random walk and \(R_{t,s}^N\) reflects a trading noise, we can further simply Equation (4) as

\[
Var(R_{t,s}) = Var(R_{t,s}^C) + (t - s)\sigma_H^2 + \sigma_N^2
\]  

(4’)

However, the second step of WRS, as quoted above from Case and Shiller (1989), only estimates the following model:

\[
Var(R_{t,s}) = (t - s)\sigma_H^2 + \sigma_N^2
\]  

(5)

Comparing Equation (5) with Equation (4’), we can see that \(Var(R_{t,s}^C)\), the market risk factor, is noticeably omitted.
Case and Shiller (1989) are aware of such omission, but they provide no theoretical justification, partly because the appropriate distributional form of the market factor $C$ was unknown at the time. So instead, they conducted an empirical analysis that led them to conclude that “individual (house) prices are not so heavily influenced by the aggregated market price” (page 127). This point is also reiterated in the technical document “S&P/Case-Shiller Home Price Indices - Index Methodology”. On page 24, the document states that “over longer time intervals, the price changes for an individual home are more likely to be caused by factors other than market forces”. In other words, market condition changes have no impact on individual home prices! Such a claim is surprising. If market risk factor indeed has no impact on individual home prices, one has to wonder why we should estimate home price indices to begin with. Furthermore, the common wisdom of “location, location, and location” in real estate investment speaks in volume that the city-wide location factor (the $C$) should have a big impact on the values of individual properties. Leaving out the factor $C$ in Equation (4’) and relying on Equation (5) for the step-two weight estimation is, therefore, likely to cause biased weight estimates that will adversely affect the final index.

The omission of the market risk factor was never questioned by all subsequent studies that propose alternative weighting specifications. Whereas Case and Shiller (1989) ignore the $C$ component (citywide factor) in the step two of WRS and assumes the $H$ component (property-specific factor) to be Gaussian random such that $Var(R_{i,t}^H)$ increases linearly over time, Abraham and Schauman (1991) suspect the linear pattern may not be “forever”, so they added a time-square variable to Equation (5) without any explanation. Callhoun (1996) argues that the second

---

4 The document is available at http://www2.standardandpoors.com/portal/site/sp/en/us/page.topic/indices_csmahp/0,0,0,0,0,0,0,0,0,2,1,0,0,0,0,0.html
5 For example, Bourassa et al (2010) and numerous others documented that housing prices exhibit strong spatial-dependence.
moment of $H$ should take a quadratic functional form. Both of these studies are silent on the market factor $C$, and neither provides evidence as to why the $H$ factor should not be Gaussian random. A more recent study by Graddy et al (2012) continues to ignore the location factor $C$. They argue that the $H$ factor should not be assumed to follow a random walk and propose a so-called flexible GLS that does not require the i.i.d. assumption in the second step of regression. However, the practical application of their suggestion seems limited as the authors acknowledge that their method will not change the coefficient estimates of the third step of WRS, which result in home price index. Rather it only improves the standard error of the estimated coefficients, which are hardly concerned by typical users of the Case-Shiller Home Price Indices. The current paper differs from previous research in that we do not dwell on the $H$ factor and argue whether it is a Gaussian random. Rather our focus is on the more important factor, the omitted location factor $C$.

While the omission of location factor $C$ is obviously problematic, some may argue that, for practical purposes, the Case-Shiller specification of Equation (5) may still be valid, because $\sigma_H^2$ is nothing more than a regression coefficient of the time interval. If one can assume that the housing returns from the market are independent and identically-distributed (i.i.d.) over time, the market risk $\text{Var}(R_{t,s}^C)$ will increase linearly with the length of time interval $(t - s)$, just like $\text{Var}(R_{t,s}^H)$. Then the slope term in Equation (5) can be interpreted as to capturing both the market and house-specific impacts $[\sigma_C^2 + \sigma_H^2]$. That is, same regression, but different interpretation of the slope term. However, such an interpretation is precisely what is rejected by the main findings of Case and Shiller (1989).

---

6 Graddy et al (2012) assumed real estate returns to follow processes such as Moving Average (MA), Autoregressive (AR), or Autoregressive Moving Average (ARMA), but they provide no theoretical or empirical reason as to why these processes are valid for real estate.
Case and Shiller (1989) tests the market efficiency for single-family houses in several major U.S. cities. Their main conclusion is that real estate market is not efficient, because the citywide market returns over time are correlated, as opposed to being independent and identically-distributed (i.i.d.). This finding, which is widely cited in the literature, implies that the market risk factor, $Var(R_{c,s}^C)$, cannot be assumed to increase linearly with the length of time interval like $Var(R_{h,h}^H)$ does. The slope term of the step-two WRS (Equation (5)), therefore, cannot be interpreted as estimate for $[\sigma_c^2 + \sigma_h^2]$.

This leads us to an interesting question: If the market risk factor $Var(R_{c,s}^C)$ can neither be omitted for mathematical reason or common sense, nor can it be assumed to increase linearly over time (as under the assumption of i.i.d.), then what is the appropriate way to handle the market risk factor and what is the proper weighting specification for the second step of WRS? To answer these questions, we must first find out how the market risk factor is affected by the time elapsed between repeated sales (or the holding period). This is the focus of the next section.

3. The Holding Period Dependence of Real Estate Risk

As discussed before, mainstream finance theories generally accept that, in an efficient market, asset returns are independent and identically-distributed (i.i.d.). The variance of returns (risk) is therefore time-invariant. Under this assumption, the risk of holding an asset for one period is the same as the per-period risk of holding the asset over multiple periods. Numerous studies including Case and Shiller (1989) have discovered that this is not the case with the real estate market, because real estate returns exhibit strong serial correlations and are not i.i.d.\footnote{In fact, Cabrera et al (2011) finds that even securitized real estate violate the i.i.d. condition by exhibiting strong return predictabilities.}

7
More recently, Anglin and Wiebe (2013) demonstrate the inefficiency of real estate market by showing that individual sellers can affect the house selling price through listing price strategies, that is, even small sellers are not merely price takers. The inefficiency of real estate market implies violation of the i.i.d. condition. But if real estate returns are not i.i.d., what are they? As it turns out, an alternative return distribution has emerged from several recent studies including Lin and Liu (2008) and Cheng et al (2013), which formally show that the holding-period real estate risk (the standard deviation of real estate returns) *increases linearly with time*. That is, 
\[ \sigma_{\tau} = \tau \sigma \] 
where \( \tau \) is the holding period, and \( \sigma \) is the single-period risk. This is in sharp contrast with the i.i.d. assumption that suggests that real estate risk increases with the square-root of time, i.e. \( \sigma_{\tau} = \sqrt{\tau \sigma} \).

This section presents a series of empirical examination of the above conjecture by examining a wide range of real estate market data. We obtain market risk by computing the variances of returns of major market indices and sub-indices. For the housing market, we gather the OFHEO all-transactions Home Price Index and the OFHEO Purchase-only Home Price Index.\(^8\) The former is the original OFHEO index that is constructed based on a mix of housing sales and refinancing transactions (where the property value is appraisal value instead of sale price). The OFHEO purchase-only index is a newer index that starts in 1991. Since it uses only sales price, it is free of the influence of appraisal bias. In addition to national indices, we also obtain the regional OFHEO sub-indices. The classification of the nine regions is displayed in the Appendix 1. The second group of housing index is the S&P/Case-Shiller Home Price Indices, which contains sub-indices for 20 major U.S. Metropolitan Areas, as well as aggregated national

---

8 In late 2008, the Office of Federal Housing Enterprise Oversight (OFHEO) was merged into the new Federal Housing Finance Agency (FHFA) and the indices have been renamed as the FHFA Home Price Indices. We retain the OFHEO name here only because it is familiar to most readers.
index. For commercial real estate, we collect the NCREIF Property Index and sub-indices for major property sectors (office, retail, industrial, apartment, etc). For comparison purposes, we also collect major stock market indices including S&P500 composite, Dow Jones Industry, and the NASDAQ Composite. Lastly, we include the public real estate index – the NAREIT composite for equity real estate investment trusts (REITs). A list of these indices is shown in Table 1.

**Table 1. List of Market Indices and Sub-indices to be Examined**

<table>
<thead>
<tr>
<th>Index Name</th>
<th>Time span</th>
</tr>
</thead>
<tbody>
<tr>
<td>OFHEO Home Price Indices:</td>
<td></td>
</tr>
<tr>
<td>National</td>
<td>1975Q1 - 2008Q3</td>
</tr>
<tr>
<td>Regional</td>
<td>1975Q1 - 2008Q4</td>
</tr>
<tr>
<td>OFHEO Purchase-only Home Price Indices:</td>
<td></td>
</tr>
<tr>
<td>National</td>
<td>1991Q1 - 2008Q3</td>
</tr>
<tr>
<td>Regional</td>
<td>1991Q1 - 2008Q4</td>
</tr>
<tr>
<td>S&amp;P/Case-Shiller Home Price Indices</td>
<td></td>
</tr>
<tr>
<td>National</td>
<td>1987Q1 - 2008Q3</td>
</tr>
<tr>
<td>20 Metropolitan Indices</td>
<td>1987Q1 - 2008Q3*</td>
</tr>
<tr>
<td>NCREIF Property Index (commercial real estate)</td>
<td></td>
</tr>
<tr>
<td>National</td>
<td>1977Q4 - 2008Q4</td>
</tr>
<tr>
<td>By property type</td>
<td>1977Q4 - 2008Q4</td>
</tr>
<tr>
<td>S&amp;P500 Composite</td>
<td>1974Q1 - 2008Q4</td>
</tr>
<tr>
<td>Dow Jones Industry</td>
<td>1974Q1 - 2008Q4</td>
</tr>
<tr>
<td>NASDAQ Composite</td>
<td>1974Q1 - 2008Q4</td>
</tr>
<tr>
<td>NAREIT equity REITs</td>
<td>1975Q1 - 2008Q4</td>
</tr>
</tbody>
</table>

* Most metro indices start at 1987Q1, a few started around 1990Q1, Dallas starts at 2000Q1.

Note that these indices are intended to be representative, rather than exhaustive, for our purpose. The indices begin at different times and are reported for different time intervals. Some are available daily, others only quarterly. For easy calculation and comparison, we chose
quarterly indices for all the above time series data. All the indices are obtained from public sources.

To see how real estate market risk changes over time (i.e. holding period), first, we use each quarterly index to calculate the return indices for various holding periods ranging from one quarter to 20 quarters (5 years). We then compute the standard deviation for each return series, which we denote with $\sigma_\tau$, where $\tau = 1, 2, 3, \ldots 20$, and $\sigma_\tau$ is the total standard deviation for holding-period $\tau$.

![Figure 1. Risk Lines of Major National Markets](image)

For easy comparison, we standardize $\sigma_\tau$ by computing $\sigma_\tau / \sigma_1$, so that the risk of a single-quarter holding period is scaled to a unit ($\sigma_1 / \sigma_1 = 1$), and we plot the standardized numbers in Figure 1. Then, given the price risk for a single-quarter holding period being 1, we also plot the multi-period risk increasing along two alternative paths: one is the path of i.i.d.
(\sigma_\tau = \sqrt{\tau \sigma_1}), the other is the alternative linear path of \sigma_\tau = \tau \sigma_1, which implies that the variance increases with the square of holding time, i.e. \sigma_\tau^2 = \tau^2 \sigma_1^2.

Figure 1 displays the standardized risk for major asset indices at the national level. For lack of a better term, we tentatively call these lines the risk lines. Each line in Figure 1 depicts the relationship between an asset’s risk and the investment terms (or holding periods). Each point (except the point (1,1)) on a curve represents the ratio between the risk of holding the asset for certain multiple periods and the risk of holding it for a single period. Each risk line provides a side-by-side comparison of investment risks of holding a particular asset over various holding-periods. Different risk lines illustrate how the risks of various asset classes change over the lengths of their holding-periods. It should be noted that these lines do not indicate the magnitudes of risk. They only indicate how risks change over time relative to other asset markets.

3.1. Evidence from the national markets

Three observations can be made from Figure 1: First, there is a clear diversion between real estates and stocks (including NAREITs). While the security assets are all reasonably close to the i.i.d. path (especially over shorter holding periods), the real estate indices are all far away from that path. To the extent that asset returns over time in an efficient market closely resemble i.i.d., the real estate market cannot be said being efficient. This confirms the finding of Case and Shiller (1989) from a different angle. Second, the theoretically derived linear path sets the upper bound for all the real estate lines. That is, although the real estate lines are away from the i.i.d., they are not all close to the linear path with a slope of 1. Rather, their slopes are somewhat less than 1. In comparison, the stock lines are much closer to the i.i.d. path, although with some drift.
away from it when holding-periods exceed 10 quarters. Third, the real estate lines exhibit strong linear patterns, indicating that the standard deviations of real estate assets increase linearly over time, thus their variances will increase with the square of time. If this linear pattern persists across various markets and submarkets, it would suggest that the proper specification for Equation (5) should include a quadratic term of the time interval \((t - s)\).

With regard to the OFHEO and Case-Shiller housing indices, it is necessary to note that the repeated sales method is only directly applied to samples at regional or metropolitan levels. Their national level indices are aggregated from the regional or metropolitan indices using certain weighting scheme that is subject to occasional revision. Given the heterogeneous nature of local real estate markets, the pattern of the national level indices may or may not be indicative of the behavior of the local markets. Therefore, we examine these curves of individual regional and metropolitan markets in the following sections.

3.2. Evidence from the OFHEO Regional Indices

The OFHEO regional indices are reported for the nine census regions displayed in Appendix 1. For brevity, we only present the results from the Purchase-only index to avoid any potential bias due to the appraisal data in the original All-transaction index. By replicating Figure 1 for the OFHEO regular regional indices, we find that the nine regions can be separated into two groups by their risk lines: group A consists of four of the nine regions, whose risk lines exhibit strong linear trends and are closer to the linear paths. Group B consists five of the remaining nine regions, whose risk lines exhibit somewhat weaker linear trends and are located relatively closer to the i.i.d. path. We plot these two groups in Figures 2A and 2B.

---

9 However, we did examine the All-transaction by regions as well and the results are actually somewhat stronger than the Purchase-only index with 6 regions in Group A and only 3 in Group B (closer to i.i.d.). This result is available upon request.
Figure 2A. Risk Lines of OFHEO Purchase-only Regional Indices

The linear path: $\sigma_\tau = \tau \sigma_1$

The i.i.d. path: $\sigma_\tau = \sqrt{\tau} \sigma_1$

Scaled standard deviations

Holding-periods (number of quarters)

Figure 2B. Risk Lines of OFHEO Purchase-only Regional Indices

The linear path: $\sigma_\tau = \tau \sigma_1$

The i.i.d. path: $\sigma_\tau = \sqrt{\tau} \sigma_1$
Figure 2A contains all the coastal regions (New England, Mid-Atlantic, South-Atlantic, and the Pacific). The risk lines of all these regions are strongly linear and far away from the i.i.d. path, suggesting the variances of the return indices for coastal markets increase with the square of time intervals. On the other hand, all the inland regions, as shown in Figure 2B, tend to exhibit risk lines that are less linear and much closer to the i.i.d. path. These results suggest that, according to the OFHEO purchase-only indices, Case-Shiller’s step-two WRS specification (Equation (5)) is reasonable for the inland regions, but remain mis-specified for the coastal markets, and the costal markets arguably are the more critical or interesting housing markets.

3.3. Evidence from the S&P/Case-Shiller Home Price Indices

Now we take the same approach to examine the S&P/Case-Shiller home price sub-indices at metropolitan level. The S&P/Case-Shiller indices cover 20 major U.S. metropolitan areas. Computing the risk lines for these metropolitan sub-indices reveals, once again, two separate groups. Group A, as shown in Figure 3A, contains 11 of the 20 cities whose risk lines are strongly linear and are far away from the i.i.d. path. With the exception of Denver and Las Vegas, all these cities are major coastal cities, and they all represent critical housing markets. For these cities, Case and Shiller’s weight model specification (Equation (5)) is mis-specified. More importantly to note is that, these cities include nine of the ten cities (except Chicago) that have housing options and futures contracts traded on the Chicago Mercantile Exchange. The fact that these city level indices are estimated based on mis-specified weighting schemes may have a significant impact on the pricing of these housing derivative securities.
Figure 3A. Risk Lines of S&P/Case-Shiller Metropolitan Sub-indices

The linear path: $\frac{\sigma_r}{\sigma_1} = 1$

The i.i.d. path: $\frac{\sigma_r}{\sigma_1} = \sqrt{r}$

Figure 3B. Risk Lines of S&P/Case-Shiller Metropolitan Sub-indices

The linear path: $\frac{\sigma_r}{\sigma_1} = 1$

The i.i.d. path: $\frac{\sigma_r}{\sigma_1} = \sqrt{r}$
Figure 3B displays the risk lines of the remaining nine cities in the Case-Shiller 20-city group. Although relatively speaking, the risk lines of these cities are closer to the i.i.d. path, most of them are actually not as close as the S&P500 does. For example, it perhaps can be argued that Portland and Detroit are both linear and distant enough from the i.i.d. path, and Phoenix and Dallas are closer to the linear path than to the i.i.d. path over shorter holding periods (under 10 quarters). As for Chicago and Seattle, it is perhaps a matter of opinion as to whether they are reasonably close to the i.i.d.. In the end, there are only three cities that are more clearly close to the i.i.d. path – Atlanta, Charlotte, and Cleveland.

3.4. Evidence from the Commercial Real Estate Market

Although the focus of this paper is on the housing market, it is perhaps also interesting to examine the risk lines of the commercial real estate markets to see similarities or differences between the two kinds of markets. The National Council of Real Estate Investment Fiduciaries (NCREIF) Property Index is widely regarded as the industrial standard for commercial real estate benchmark. Besides national composites, the NCREIF releases sub-indices by property types. We replicate Figure 1 for the four major property types, as these sub-indices have sufficiently long time series.

Figure 4 displays the risk lines of the NCREIF commercial real estate indices. Without exception, these indices all exhibit strong linear trends and are very close to the linear path. Compared with the housing markets, commercial properties of any type are further away from the i.i.d. path than residential properties. This is consistent with the notion that the commercial
property markets are less efficient than the residential markets because they are generally less frequently traded.\footnote{In the past many attribute the risk lines of NCREIF to autocorrelation due to appraisal smoothing. Recent work by Cheng, Lin, and Liu (2011), however, shows that the appraisal smoothing theory offers no valid explanation of the risk lines observed in Figure 7, which confirms previous findings of Lai and Wang (1998) and De Wit (1993).}

Judging mainly from the two transaction-based indices – the OFHEO purchase-only and the S&P/Case-Shiller indices, it is perhaps not surprising that housing markets exhibit significant heterogeneity in that each market possesses a distinct risk line. But there does appear to be a discernable pattern, that is, the risk lines of housing markets in the coastal regions or cities tend to exhibit strong linear trends and often lie closer to the linear path. In comparison, the inland regions and cities are more likely to exhibit risk lines that are relatively closer to the i.i.d. path with weaker linear trends. This suggests that a one-size-fits-all weighting specification cannot
properly capture the market risk factor for all markets; rather, the specification should be market-
specific.

4. The Appropriate Weight Specification for the Step-two of WRS

The findings so far can be summarized as follows: First, the omission of market risk
(location factor \(C\)) in the step-two of WRS by Case and Shiller (1989), as well as subsequent
studies, is conceptually unjustified. Since the location factor is not trivial in real estate, it can
adversely affect the estimates for the home price indices in step-three of WRS. Second, based on
the theoretical and empirical evidences in Section 3, proper weighting specification for step-two
of WRS should be market-specific. For markets where the risk lines are reasonably close to the
i.i.d., it can be argued that Case-Shiller’s original specification (Equation (5)) is reasonably
accurate for practical purpose. On the other hand, for markets where risk lines exhibit strong
linear trends, which suggest the standard deviations of the returns increase linearly (along
different slopes) over time, their corresponding variances should increase with the square of the
time interval between paired sales. For these markets, the appropriate weighting specification
should be:

\[
Var(R_{i,t}) \approx (t - s)^2 \theta + (t - s)\sigma^2_H + \sigma^2_N
\]  

(6)

Where \(\theta\) is the regression coefficient for the square of time interval between paired-sales. While
the interpretation of \(\theta\) is a little more complex, it suffice to say that \(\theta\) is closely related to both
the slope of the risk line (\(\beta_{RC}\)) and the market risk of a single period (\(\sigma^2_C\)), that is,

\(\theta = f(\beta_{RC}, \sigma^2_C)\). All other symbols are as defined before. Table 2 provides a summary of these
two groups of markets and attempts to shed some lights on their economic importance. Based on
Figure 2A, 2B, 3A, and 3B, we list these markets in the two groups in Table 2.
### Table 2. Housing Market Shares and Classifications

#### I. The OFHEO Nine Regions:

<table>
<thead>
<tr>
<th>Group A regions</th>
<th>Housing shares (%)</th>
<th>Group B regions</th>
<th>Housing shares (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Middle Atlantic</td>
<td>11.6%</td>
<td>East North Central</td>
<td>16.6%</td>
</tr>
<tr>
<td>South Atlantic</td>
<td>20.3%</td>
<td>East South Central</td>
<td>6.5%</td>
</tr>
<tr>
<td>New England</td>
<td>4.5%</td>
<td>West North Central</td>
<td>7.7%</td>
</tr>
<tr>
<td>Pacific</td>
<td>14.0%</td>
<td>West South Central</td>
<td>11.5%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mountain</td>
<td>7.3%</td>
</tr>
<tr>
<td><strong>Total housing shares</strong></td>
<td><strong>50.4%</strong></td>
<td></td>
<td><strong>49.6%</strong></td>
</tr>
</tbody>
</table>

#### II. The S&P/Case-Shiller 20 Metropolitan areas:

<table>
<thead>
<tr>
<th>Group A metro areas</th>
<th>Housing shares (%)</th>
<th>Group B metro areas</th>
<th>Housing shares (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boston*</td>
<td>4.0%</td>
<td>Atlanta</td>
<td>6.4%</td>
</tr>
<tr>
<td>Denver*</td>
<td>3.0%</td>
<td>Charlotte</td>
<td>2.2%</td>
</tr>
<tr>
<td>Las Vegas*</td>
<td>2.1%</td>
<td>Chicago*</td>
<td>9.2%</td>
</tr>
<tr>
<td>Los Angeles*</td>
<td>10.5%</td>
<td>Cleveland</td>
<td>0.1%</td>
</tr>
<tr>
<td>Miami*</td>
<td>5.3%</td>
<td>Detroit</td>
<td>5.9%</td>
</tr>
<tr>
<td>Minneapolis</td>
<td>3.9%</td>
<td>Dallas</td>
<td>6.6%</td>
</tr>
<tr>
<td>New York*</td>
<td>13.2%</td>
<td>Phoenix</td>
<td>4.9%</td>
</tr>
<tr>
<td>San Diego*</td>
<td>2.8%</td>
<td>Portland</td>
<td>2.5%</td>
</tr>
<tr>
<td>San Francisco*</td>
<td>4.1%</td>
<td>Seattle</td>
<td>3.7%</td>
</tr>
<tr>
<td>Tampa</td>
<td>3.4%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Washington, DC*</td>
<td>6.0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total housing shares</strong></td>
<td><strong>58.4%</strong></td>
<td></td>
<td><strong>41.6%</strong></td>
</tr>
</tbody>
</table>

Note: Housing shares are obtained from the U.S. Census Bureau as of Q4 of 2008. Group A are markets where Case-Shiller's original weight specification is more biased than areas in Group B. A (*) indicates the metro index is the base for housing options and futures traded on the CME.

The housing shares are the housing stocks in individual markets as percentages of the total housing stocks among all the reported markets. Regions or metro areas of group A are markets that exhibit strong linear risk lines (Figure 2A and 3A), for which Equation (6) is more applicable for the step-two of WRS. Regions or metro areas of group B are the rest of the markets where Case-Shiller’s original specification (Equation (5)) can be argued as reasonable only for practical purposes.
As panel I of Table 2 shows, Groups A includes all the coastal regions and group B includes all the inland regions. Although the total housing stocks are about evenly split between the two groups, it is perhaps safe to say that the regions in Group A are more populous and economically vital to the nation. They are perhaps more closely watched by economists and policy makers as well.

Panel II of Table 2 shows bigger differentiations between the two groups. Group A includes 11 of the 20 metropolitan areas covered by the S&P/Case-Shiller Home Price Indices. Most markets of this group are large and critical coastal cities, which represent 58.4% of the total housing stock, compared to 41.6% in group B. The biggest difference, however, is that group A includes nine of the ten cities where Case-Shiller indices are “tradable” with housing options and futures contracts listed on the Chicago Mercantile Exchange. For these nine cities, the accuracy of the Case-Shiller indices is of vital importance to the pricing of the derivative securities that are based on the indexes. Unfortunately, it is for these cities that Case-Shiller’s original specification is less than appropriate.

5. A Comparison between index methods

Market indices are used to track the market movement of certain assets. But the question of whether one index method is better than another is not an empirical one, in the sense that one cannot empirically conclude which index tracks the “true” market movement better because the “true” market movement is unknown. The reason that Case and Shiller’s WRS method is embraced by many as being superior to the original BMN method is mainly because of its conceptual soundness. The current study has shown that, while Case and Shiller’s weighted regression makes conceptual sense, their weighting specification suffers from a conceptual
deficiency itself – it omits market risk factor based on the false claim that market force has no impact on individual property value. Given what is found on market risk factor in Sections 3 and 4, we propose an alternative weight model that remedies the deficiency in the Case-Shiller’s weighting specification, and thus improves its conceptual soundness. In other words, methodologically speaking, our modified WRS is superior to the Case-Shiller WRS for the same reason that the latter is superior to the original BMN method – it is more conceptually sound and mathematically correct.

Therefore, the empirical question is not really “which method is better?”, but rather “does it make a difference?” and whether the difference necessitates the change in practice. To see the difference, ideally, we would want to use the same dataset that is used to construct the current S&P/Case-Shiller Home Price Indices, to run the repeated-sales regression under three weighting schemes – the BMN method, the Case-Shiller method, and our method, and to measure the difference between the results. The difficulty, obviously, is that the dataset for the S&P/Case-Shiller Home Price Indices is proprietary and not available to the public. As the cost of acquiring large samples of repeated sales at a national scale is prohibitive, we use a sample of data from the Washington DC area for demonstration purpose.

The data we used is from a major mortgage institution that collects data from sources such as Fannie Mae, Freddie Mac, DataQuick, etc. The Washington DC sample contains over 1.5 million pairs of repeat residential property sales during 1990 - 2011. This sample is large enough for us to exclude refinance transactions with appraisal values so that only the purchase transactions are used for the analysis. The numbers of paired sales vary from year to year, averaging about 65,000 pairs per year.
Using the sample we first construct three quarterly repeated-sale indices using three different weighting schemes – the original BMN method, the Case-Shiller method, and our modified WRS method. Then we compute the quarterly returns for each index. Two comparisons are made. First, we compare the standard errors of the regression coefficients. As Graddy et al (2012) suggests, smaller standard errors suggest tighter estimates and a more efficient model. Our results indicate that the standard error of Case-Shiller method is on average 10.5 basis points higher than that of our modified WRS. A simple t-test indicates such difference is statistically significant at the 1% level (Table 3), suggesting that our model is more efficient in reducing the estimation errors of the indices.

Table 3. Standard Errors of Regression Coefficients: Case-Shiller vs. Ours

<table>
<thead>
<tr>
<th></th>
<th>Case-Shiller</th>
<th>Our method</th>
<th>Difference</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean std. err.</td>
<td>45.8</td>
<td>35.3</td>
<td>10.5</td>
<td>0.0095</td>
</tr>
</tbody>
</table>

Second, using our modified WRS as the baseline, we compute the index differences between our method and the other two methods. The differences essentially reflect the impact of different weighting schemes. The results (in basis points) are displayed in Figure 5. In addition, in order to see how the differences correlate with market conditions, we plot the overall metropolitan housing index during the same period. Overall, the magnitudes of the differences range from a few basis points to nearly 150 basis points for both the BMN and Case-Shiller indices. The average differences are 20 and 25 basis points for BMN and Case-Shiller methods, respectively.

Closer inspection of Figure 5 also reveals: First, market condition affects the index performances. Prior to 2001, when the market is relatively stable, the differences are small and
stable for both BMN and Case-Shiller indices. But when the market moves rapidly during the booming years and the subsequent bust, the differences became much bigger. The reason is simple: neither BMN nor Case-Shiller index captures the impact of market risk. So when the market is stable and its impact on property value is small (e.g. prior to 2001), ignoring such impact does not make a big difference in the resulted indices. But when the market becomes more volatile, its impact on all property values becomes much more pronounced. Ignoring such impact will cause the resulted indices to differ significantly.

**Figure 5. Comparison of weighting schemes on index accuracy**

<table>
<thead>
<tr>
<th></th>
<th>Prior-2001</th>
<th></th>
<th>Post-2001</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMN</td>
<td>15.78</td>
<td>12.99</td>
<td>0.12</td>
</tr>
<tr>
<td>Case-Shiller</td>
<td>14.39</td>
<td>9.36</td>
<td>0.23</td>
</tr>
</tbody>
</table>

Second, when comparing the differences before and after 2001 (see the table in Figure 5), the results suggest that, prior to 2001 when the market was fairly stable, the average difference of Case-Shiller index is 14.39 basis points, which is modestly closer to the base line (our modified
than the average differences of the BMN index (15.78). However, after 2001 when the
market moves rapidly, the average deviation of Case-Shiller index is much larger than that of
BMN index (34.06 vs. 22.14), that is, rather than improving the BMN result, the Case-Shiller
method seems to have done the opposite. This perhaps suggests that, while a weighted
regression is conceptually superior, not weighting correctly (i. e. overweighting noisy data and
underweighting accurate information) may be worse than not weighting at all.

6. Conclusions

Case and Shiller (1989) is an influential study that established the Weighted Repeated
Sales (WRS) method behind what has become the most trusted housing price index – the
S&P/Case-Shiller Home Price Indices. The key innovation of Case and Shiller is the
specification of weighting scheme which turns the original BMN repeated sales method into a
three-step weighted regression. This improvement is conceptually sound, but the weight
estimating model they specified in the step-two of WRS suffers from an important deficiency –
the omission of the market risk factor. Through formal analysis, we demonstrate that the
omission of market risk is mathematically and conceptually unjustified, and we propose an
alternative weight model which properly incorporates the market risk factor.

Unlike some previous studies that propose alternative models based on arbitrary
assumptions, our proposition is based on extensive theoretical and empirical analyses. The
findings make a separate contribution to the literature as they address a broadly important
question: “if real estate returns are not i.i.d., what are they?” By analyzing a simple model of
real estate transaction process, we show that the standard deviation of real estate returns
increases linearly with holding time, as opposed to the square-root of time under the i.i.d.
assumption. We then examine a wide range of real estate market indices and sub-indices with what we call risk lines (Figures 1 to 4) and show that, broadly speaking, empirical data are much more consistent with our linear path than with the i.i.d. path. The linear pattern of standard deviation suggests that variance of real estate return is a strict convex function of the holding period, rather than just a linear function under the i.i.d. assumption. These findings not only place the quadratic weight model (Equation (6)) on solid empirical ground, but also have broad implications. Future research in real estate investment need not (and indeed should not) continue to rely on the i.i.d. assumption. Conversely, all the classical finance theories (such as the MPT and CAPM) that depend on the i.i.d. assumption should not be naively applied to real estate investment and portfolio analysis. Real estate research must take the notion of “real estate is different” more seriously and explore beyond the traditional finance paradigm.

To see whether the methodological modification makes a difference in the resulted indices, we use a large sample of repeated-sales from Washington DC area and construct three repeated sales indices using the original BMN method, the Case-Shiller method, and our modified WRS method, respectively. Our comparison suggests that market risk clearly affects index performances – the differences are smaller when the market is stable but become bigger when the market is more volatile. Capturing market risk factor also yields tighter index estimates and improves the index efficiency.

As a final note, it is necessary to point out that, the subject of the repeated sales regression is a complex one, both theoretically and, perhaps more so, empirically. This paper is not intended to address all aspects of the issues surrounding the methodology. We accept Case and Shiller’s method as-is, and focus only on the conceptual and empirical problems with their model specification in the second step of WRS. The general idea of giving different weights to
different observed paired sales according to the time elapsed between the sales remains to their credit. Our contribution is on how to properly capturing the impact of market risk factor in this important market index method.
References


Appendix 1. The Nine Census Regions used by OFHEO