Heterogeneous Information and Appraisal Smoothing

Authors | Ping Cheng, Zhenguo Lin, and Yingchun Liu

Abstract | This study examines the heterogeneous appraiser behavior and its implication on traditional appraisal smoothing theory. The findings demonstrate that the partial adjustment model is consistent with the traditional appraisal smoothing argument only when all appraisers choose the same smoothing technique. However, if appraiser behavior is heterogeneous and exhibits cross-sectional variation due to the difference in their access to, and interpretation of information, the model actually leads to a mixed outcome: The variance of the appraisal-based returns can be higher or lower than the variance of transaction-based return depending on the degree of such heterogeneity. Contrary to what the traditional appraisal smoothing theory would predict, appraisal-based indices may not suffer any "smoothing" bias. These findings suggest that the traditional appraisal smoothing theory, which fails to consider the heterogeneity of appraiser behaviors, exaggerates the effect of appraisal smoothing.

One of the most extensively documented observations in the real estate literature is that appraisal-based return series (as exemplified by the NCREIF Property Indices) exhibit comparable returns to common stocks, but with much lower volatility. That is, the risk-adjusted returns of commercial real estate are significantly higher than those of common stocks. This is often referred to as the "real estate risk premium puzzle" (Lusht, 1988; Shilling, 2003) as it suggests that real estate rewards investors with additional premium for the same risk. The consistently repeated observation and the desire to solve the "puzzle" have prompted one of the leading theories in real estate: the appraisal smoothing theory.

As articulated by Geltner (1989, 1991) and many others, the foundation of appraisal smoothing theory is that the seemingly "superior" performance of real estate is essentially caused by the rational behavior that appraisers exhibit in their practices. Particularly, it is believed that, in forming an appraisal estimate on a given property, the appraiser tends to not only analyze the current market information, but also gives some consideration to the past market information of the same or similar properties. This is called *partial adjustment behavior*, as the current market information is not given 100% consideration. This behavior

effectively implies that appraisers tend to weight-average the past and present information to form their appraisal opinion. An appraisal-based return series such as the NCREIF Property Index (NPI), therefore, is essentially some sort of "moving average" of the underlying asset performance. As such, it is expected to be "smoothed" and exhibit downward-biased volatility over time. Correcting such bias, it has been argued, will lead to increased variance and reduced risk-adjusted returns for real estate to the levels comparable to that of common stocks, thus solving the "real estate risk premium puzzle."

Based on the assumed appraiser behavior, Geltner (1989, 1991) extended the work of Blundell and Ward (1987) to propose a formal model known as the partial adjustment model, in which the current appraisal return is expressed as a weighted average of the past true property returns (or true property appreciation rates). This model formalizes the smoothing argument and provides the normative foundation for the theory. Over the past twenty years, the partial adjustment model has gained wide popularity among academics, and is often taken as the theoretical launch pad for studies that want to quantify the smoothing effect or develop de-smoothing methodologies.

Despite repeated and often sophisticated elaborations by many studies from various angles, it is worth noting that the whole idea of appraisal smoothing theory is almost entirely motivated to explain one simple fact: the exceptionally low volatility in the only industry standard benchmark of commercial real estate at the time, the NPI. In retrospect, presented in this paper, it appears that we might have leaped to find the cause of a phenomenon before making sure to what extent that phenomenon is actually real. This study examines the root of the appraisal smoothing theory, re-examines the facts and logic behind the partial adjustment model, and presents new empirical evidence to verify the key predictions of the theory. After a brief review of relevant literature in the next section, the rest of the paper presents three investigations of the appraisal smoothing theory:

First, the initial facts that motivated the appraisal smoothing theory are examined. The findings show that the widespread belief of the NPI exhibiting extremely low volatility turns out to be a biased observation that is valid only under the implicit assumption that commercial properties are held for very short periods of time (i.e., a quarter or a year). Under longer and more practical holding periods, however, the volatility of NPI is not nearly as low as previously believed. In other words, had the variance of NPI been examined over more realistic investment horizon, it might not have appeared so low as to prompt the idea of the appraisal smoothing theory. The initial factual ground of the appraisal smoothing theory is thus questionable.

Second, the logic behind the normative foundation (i.e., the partial adjustment model) of the appraisal smoothing theory is re-examined. The findings show that the model supports the traditional appraisal smoothing theory *only* under the assumption of homogeneous appraiser behavior, that is, all appraisers exhibit exactly the same partial adjustment behavior in terms of weighting past versus

present information. Previous literature generally discusses appraisal heterogeneity in the context of appraisal errors in the return or value estimates. The current study analyzes the impact of the other heterogeneity in the model: the choices appraisers make in weighting past versus present information. When such heterogeneity is considered, the model leads to a mixed outcome. The variance of the appraisal-based returns is not always lower than the true variance; rather it could be either higher or lower than the variance of transaction-based return depending on the degree of such heterogeneity.

Third, the central prediction of the appraisal smoothing theory is empirically verified. The prediction is that appraisal-based returns are expected to exhibit low volatility and that, in the absence of appraisal influence, transaction-based returns should exhibit significantly higher volatility. Analyses using two well-respected housing price indices indicate that *appraisal*-based returns exhibit moderately higher (not lower) volatility than the transaction-based returns. This finding contradicts what the smoothing theory would predict, but it is more consistent with our theoretical investigation, which suggests appraisal-based returns may not suffer any ''smoothing'' bias.

Related Literature

After two decades of development, the notion that appraisal causes ''smoothing'' has become so entrenched in the literature that there is rarely any formal investigation into the validity of the facts or logic upon which the appraisal smoothing argument is based. Much of the appraisal smoothing literature simply starts with the presumption that smoothing does exist, and proceeds in identifying the magnitude of the smoothing effect and developing de-smoothing methods for recovering the ''true'' volatility. Although the size of the literature is too large to be reviewed here in detail, a comprehensive discussion on the evolution of the appraisal smoothing theory and de-smoothing methodology can be found in Geltner, MacGregor, and Schwann (2003), in which many prominent studies on the subject are reviewed. The papers reviewed in this section are mainly those that present conflicting evidences and arguments to the traditional appraisal smoothing theory.

Webb, Miles, and Guilkey (1992) constructed a return index using only sold properties, and reported that transaction-based portfolio returns have approximately the same volatility as the appraisal-based portfolios. De Wit (1993) compares in-house versus outside appraisals and find that independent outside appraisals suggest no significant smoothing bias. Hendershott and Kane (1995) develop a benchmark of property fundamental values and report the change of fundamental (market) values exhibits less volatility than normally observed from appraisal-based returns. Brueggeman, Chan, and Thibodeau (1992) use a non-appraisal-based index to show that commercial real estate exhibits much higher risk-adjusted returns than financial assets during the 1972–1991 period, suggesting that the ''good data'' (i.e., transaction-based index) do not exhibit volatility near

the level of stocks. Cho and Megbolugbe (1996) compared an appraisal-based index with a transaction-based index and found no evidence of significantly lower volatility in the appraisal-based index. A similar finding is also reported by Chinloy, Cho, and Megbolugbe (1997), who compared the volatility of separate purchase and appraisal data samples obtained from Fannie Mae and Freddie Mac for 1975–1993, and found the appraisal data to be more volatile than the purchase data. These anecdotal evidences, though important, are not direct challenges to the appraisal smoothing theory, and they are often regarded as exceptions of what the appraisal smoothing theory predicts.

The only direct challenge to the appraisal smoothing theory so far is a study by Lai and Wang (1998), in which the authors take a theoretical approach to examine the conditions and assumptions for the appraisal smoothing theory to hold. The authors point out a subtle inconsistency in Geltner (1989, 1991) in which the partial adjustment model is expressed as appraisal return being a moving average of past *true return*, but is interpreted as appraisal value being a moving average as past *true values* (not returns). They believe that this distinction is important. They argue that the *return*-averaging assumption, though consistent with a lower appraisal variance, is inconsistent with appraisers' actual behavior, because the true market returns are never observable. On the other hand, the *value*-averaging assumption, though somewhat more realistic, is likely to require past true values be approximated by past appraisal values in practice. And they show that, when both the appraisal value and the true value of a property are assumed to be stochastic, the resulted appraisal variance will be higher rather than lower. In a comprehensive review of the appraisal smoothing literature, Geltner, MacGregor, and Schwann (2003) counter the criticism of Lai and Wang (1998) by insisting that the "moving average" process reasonably captures the manner by which appraisers typically use the "comps" in practice. In the discussion of a general version of *the partial adjustment* model (p. 1052, Eq. (20)), they continue to disregard the distinction between the *return*-averaging and *value*-averaging assumption that Lai and Wang (1998) deemed important.

The study of Lai and Wang (1998) is widely acknowledged in the subsequent literature, but few have joined the debate on the issues they raised. Instead, the majority of the research continues to press ahead under the assumption that appraisal smoothing exists, and propositions for various de-smoothing methods keep appearing in leading academic journals. To mention a few of the more prominent ones, in a study of real estate allocation in mixed-asset portfolios, Corgel and deRoos (1999) compared several de-smoothing models proposed by other studies and find that they all achieve the objective of inflating real estate volatility and reducing real estate allocations. But the allocations seem to be rather sensitive to which de-smoothing model used. By rationalizing the appraisal process in a so-called sticky value framework, Brown and Matysiak (2000) argue that, despite appraisal errors being small at individual property levels, it is still necessary to de-smooth the aggregated index, though they concede that it is

unclear which de-smoothing procedure produces the best result. Clayton, Geltner, and Hamilton (2001) note the lack of empirical support and calibration of the partial adjustment model, and attempt to provide such support by conducting a direct empirical test using a sample of Canadian commercial property that contains both appraisal and information. They find strong evidence that appraisers tend to anchor on their previous appraisals to form the current estimates, which is a behavior that the partial adjustment model is based upon. Their finding confirms the result of Diaz and Wolverton (1998), who find the anchoring behavior of appraisers to be significant, a conclusion that was only weakly supported by an earlier similar experiment of Diaz (1997). While Cho, Kawaguchi, and Shilling (2003) continue to refine the Fisher-Geltner-Webb de-smoothing technique (Fisher, Geltner, and Webb, 1994) and apply both the original and revised procedures to commercial property return indexes in the United States, United Kingdom, and Japan, Pagliari, Scherer, and Monopoli (2005) simply use an earlier method proposed by Geltner (1993) to de-smooth the NPI for performance comparison between private and public real estate. Edelstein and Quan (2006), however, suggest that de-smoothing may not be necessary because even if smoothing bias exists at the individual property level, the individual errors may offset each other at the aggregated index level. On the other hand, Wang (2006) continues to argue for the necessity of de-smoothing and proposes a multivariate approach to uncover the "true" real estate volatility without the assumption of a weak form market efficiency that was implicitly made in some early studies. His method suggests the "true" real estate volatility in the U.K. is about 1.5–2 times that of appraisal-based series. Interestingly, also using commercial property data from U.K., Bond and Hwang (2007, p. 378) find that "the level of smoothing in appraisal-based real estate indices is far less than assumed in many academic studies."

It is perhaps fair to say that, despite years of tremendous research effort, little consensus has been established among academics beyond the general belief that appraisal causes "smoothed" return series. There is little agreement with regard to how much smoothing there is in the NPI and what model produces the best de-smoothing outcome. In fact, it is not even clear whether researchers agree on what the best outcome should be. While a market efficiency argument suggests the de-smoothed real estate return should exhibit similar volatility as stocks, Fisher, Geltner, and Webb (1994) argue the reasonable real estate volatility should only be about half that of stocks, considering other risk factors of real estate. The findings of Cho, Kawaguchi, and Shilling (2003) further suggest that the objective of de-smoothing should at least vary across different countries and different property types. As the objective of de-smoothing remains ad hoc and largely subjective, it is not clear that a consensus will ever be reached on the degree of smoothing bias in the NPI (or any appraisal-based real estate index), let alone a practical and generally applicable de-smoothing method. Given this, it is perhaps time to step back and reflect on how we got here in the first place before we move forward.

The Fact that Motivated the Appraisal Smoothing Theory

Despite its rather complex evolution over two decades, the appraisal smoothing theory is mainly motivated by a simple observation: the NPI exhibits exceptionally low volatility (variance) compared to stocks, so low that it cannot be explained by classical pricing models. This observation can be easily confirmed with the publically available NPI today. The quarterly NPI for the period of 1978:Q1–2008:Q4 was obtained for demonstration purposes. During this period, the average quarterly return and standard deviation of the NPI was 2.45% and 1.70%, respectively. On a per-unit-of-return basis, the risk-to-return ratio was 1.7%/2.45% = 0.69. For comparison, the same ratio for the quarterly S&P 500 Index during the same period was 3.69. Clearly, the volatility of the NPI was indeed very low, less than about one-fifth of the S&P 500 on a per-unit-of-return basis. Few people have doubted this fact because, first, this is more than a NPI-only phenomenon. Similar results have been consistently reported from various data sources and over different time periods, some of which date back to the post-World War II era (e.g., Zerbst and Cambon, 1984; Sirmans and Sirmans, 1987). Second, the calculation of mean and variance (or standard deviation) is so simple and straightforward that nobody would doubt anything can go wrong with their computations.

One thing that few people have paid attention to, though, is the implicit investment horizon assumed behind the calculations. Specifically, the mean and standard deviation directly computed from a quarterly (or annual) index of NPI implies a holding period of one quarter (or one year). Since real properties are typically held for multiple quarters or years, does the single-period performance represent the *average* periodic performance of a multi-period investment? In the finance literature, we are used to assuming the answer is ''yes,'' because an efficient market implies that since the stock returns are reasonably independent and identically-distributed (i.i.d.), the holding period is irrelevant. However, this is not the case for real estate and the holding period may actually matter in this sector. To find out whether the real estate performance appears differently under different holding periods, this study conducts a bootstrap-style simulation to examine the quarterly NPI over the 1978:Q1–2008:Q4 period.

First, since the NPI is reported in quarterly returns, the entire series is converted into an index with 1977:Q4 being 100. Taking the index as the original sample, the simulation is run in the following steps: for a given holding-period, say one year, a quarter is randomly picked in which to ''buy'' the portfolio at the index level, then it is ''sold'' at the index level at the end of the holding period (i.e., four quarters later), and the ''return'' of this investment is computed. By repeating this process and randomly buying the market at any time for 100,000 times through re-sampling (with replacement), gives 100,000 simulated annual returns. The average quarterly return and standard deviation is computed for these simulated returns, along with the risk-to-return ratio (or risk per unit of return). The ratios are then obtained for a different holding period, ranging from 1 to 36 quarters (9 years) by repeating the entire simulation process. For comparison, the

Exhibit 1 | Simulated Risk-to-Return Ratios of NPI vs. S&P 500 (1978:Q1–2008:Q4)
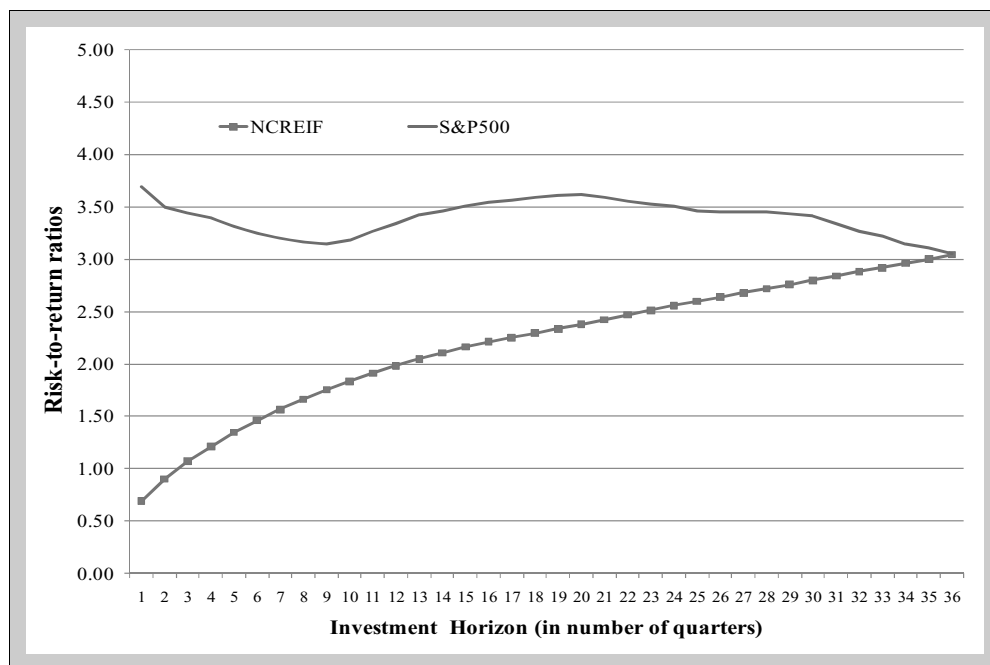
exact same process is applied to the quarterly S&P 500 Index over the same period as the NPI. The simulation results are summarized in Exhibit 1.

The pattern of the risk per unit of return for the NPI is quite different from that of the S&P 500. The ratio of the S&P 500 essentially moves horizontally with modest fluctuation, suggesting it is basically independent of the investment horizon. In sharp contrast, the ratio for the NPI is highly horizon-dependent and consistently increases with the holding period. At short horizons, the ratio is consistent with the "exceptionally low volatility" or significant "risk premium" that is so widely documented in the literature. However, as the holding period becomes more realistic and longer, the difference between NPI and the S&P 500 quickly narrows, and eventually disappears as the holding period approaches 9 years (36 quarters) and beyond (which is not uncommon for commercial real estate). This finding suggests that the conclusion that the NPI exhibits extremely low volatility, which is what motivated the original idea of the appraisal smoothing theory, turns out to be a partial truth and valid only when the holding period is unrealistically short. Had the NPI performance been examined under more realistic investment horizons (e.g., 5 years or longer), there would be a very different perception about the volatility of the NPI, or commercial real estate in general.

But what about the fact that, in Exhibit 1, the risk-to-return ratio of NPI remains below that of the S&P 500 throughout the 36 quarters? Does it suggest appraisal

smoothing bias? If in theory the risk-to-reward ratios of stocks and real estate should be about the same, there are two competing explanations. One is that the volatility gap in Exhibit 1 is caused by what Lusht (1988) called the "non-variance" risk factors: illiquidity, lumpiness, high trading costs, etc. that are unique to real estate. Since these factors must be priced in the marketplace by rational investors, it is possible that when the risks associated with these factors are properly accounted for, the total real estate risk would be close to that of stocks on a per-unit-of-return basis. Essentially, what Lusht (1988) implies is that the real estate pricing model (or the conventional performance metrics) are deficient in that they fail to price the non-variance factors, and he calls for the development of new performance measures and pricing models that are more consistent with real estate fundamentals.

The other explanation would be the appraisal smoothing argument. This explanation basically attributes the volatility gap in Exhibit 1 to a "bad data" problem that is caused by a supposed appraiser behavior. De-smoothing is essentially an effort of inferring good information from "bad" data. Although both explanations can be true to a certain extent, recent studies by Lin and Vandell (2007), Lin and Liu (2008), and Cheng, Lin, and Liu (2010a) have proposed a new real estate risk metric that explicitly integrates illiquidity risk (time-on-market risk) with price volatility. This risk metric, when applied to the NPI, suggests that the risk-adjusted return is almost identical to that of stocks, leaving virtually no room for the appraisal smoothing argument to explain the gap.

These recent studies aside, for those who believe that the volatility gap in Exhibit 1 is entirely caused by appraisal smoothing, it is at least obvious that, since the volatility gap is horizon-dependent, so must be the objective of de-smoothing. This is an important observation as it questions much of the de-smoothing literature that disregards the holding period and attempts to inflate the appraisal-based return volatility by a presumed fixed proportion. For example, the assumption that real estate volatility should be about half that of stocks (Fisher, Geltner, and Webb, 1994) has to be reconsidered because, as shown in Exhibit 1, the NPI's volatility is already more than half that of stocks as soon as the holding period exceeds 3 years (12 quarters). In reality, most real properties are held far longer than 3 years (Cheng, Lin and Liu, 2010b). Had the NPI been examined accordingly, there might be a different perception as to how "low" the NPI volatility really is.

It is worth noting that the importance of the horizon-dependence of real estate performance should not be underestimated. If the highly liquid and information-efficient financial market is characterized by security returns being independent and identically-distributed (i.i.d.), the illiquid and inefficient private real estate market is then characterized by property returns being horizon-dependent. The real estate literature has extensively and repeatedly documented that property prices do not follow the random walk and that returns over time exhibit serial persistence (non-i.i.d.).[1] This is a significant distinction between the two markets and it questions the validity of applying classical theories such as Modern

Portfolio Theory (MPT) and the Capital Asset Pricing Model (CAPM) to real estate investment analysis. Models such as the MPT and CAPM are essentially single-period models in which they assume all assets are to be held for ''one period,'' and the optimal portfolio is one that maximizes the investor's objective over such a single-period horizon. The reality, however, is that assets are often held for multiple periods of time. The validity of a single-period model to multi-period investment is largely based on the studies by Merton (1969), Samuelson (1969), and Fama (1970), among others, which show that an investor's utility maximization over multiple holding periods is essentially indistinguishable from that over a single-period *if* asset returns are independent and identically-distributed (i.i.d.) over time. Without the i.i.d. condition, these classical models are not applicable to real estate. In other words, the ''real estate risk premium puzzle'' can very well be caused by inappropriate application of these single-period pricing models and the traditional performance metrics. This issue, of course, should be investigated more rigorously in separate studies.

Examination of the Partial Adjustment Model

The discussion now turns to the normative foundation of the appraisal smoothing theory: the partial adjustment model. First, the model is shown to only support the appraisal smoothing argument under the implicit assumption of homogeneous appraiser behavior. Second, once the implicit assumption is relaxed to recognize heterogeneous appraisal behavior, the model is shown to be inconsistent with the appraisal smoothing argument and may even support the opposite conclusion.

As proposed in Geltner (1989, 1991), the formal partial adjustment model expresses the appraisal return as a weighted average of the current and past true market returns:

$$\tilde{r}_t^* = \sum_{i=0}^{\infty} w_i \tilde{r}_{t-i} \text{ and } \sum_{i=0}^{\infty} w_i = 1, \tag{1}$$

where $\tilde{r}_{t-i}$ is the true market return in period $t - i(i = 0, 1, 2,...)$, and w_i is the corresponding weight the appraiser places based on his/her knowledge; w_i is also known as the smoothing parameter. Unlike Lai and Wang (1998), who challenge the empirical justification of the model in the first place, the process begins here by taking the partial adjustment model as given. That is, the proposed methodology does not consider whether the model reasonably captures the fundamental appraiser behavior in real life, the true market returns are obtainable, or the inconsistency discerned by Lai and Wang (1998) that the model is expressed as the average of past *returns* but often interpreted as the average of past *values*. Obviously, these are all critical issues that question the validity of the model. But the current study focuses on a different question: let's assume the model is valid;

however, is it really consistent with the appraisal smoothing argument that it is intended to support?

How Does the Model Support the Appraisal Smoothing Argument?

The smoothing effect (lower appraisal volatility) can be easily demonstrated if there are only two periods in the model for mathematical simplicity. In practice, a two-period model is not necessarily unrealistic as most appraisers typically only look back to consider the most recent sales information on the same or comparable properties (market comps). Thus, the appraisal return in period t of Equation (1) can be rewritten as:

$$\tilde{r}_t^* = w\tilde{r}_t + (1 - w)\tilde{r}_{t-1}, \tag{2}$$

where w is the weight placed on the current market information. It tends to be higher if the appraiser is more confident about the current market information, and vice versa.

For simplicity, suppose that $\tilde{r}_t$ and $\tilde{r}_{t-1}$ share the same distribution with mean u and variance σ^2 and the correlation between $\tilde{r}_{t-1}$ and $\tilde{r}_t$ is assumed to be ρ. This assumption will be relaxed later in the paper. For now, the volatility of the appraisal return $\tilde{r}_t^*$ can be expressed as:

$$Var(\tilde{r}_t^*) = Var[w\tilde{r}_t + (1 - w)\tilde{r}_{t-1}] = [w^2 + (1 - w)^2 \\ + 2w(1 - w)\rho]\sigma^2. \tag{3}$$

Given that $\rho \leq 1$ and $0 < w < 1$, then $[w^2 + (1 - w)^2 + 2w(1 - w)\rho] \leq [w^2 + (1 - w)^2 + 2w(1 - w)] = 1$. Thus:

$$Var(\tilde{r}_t^*) \leq \sigma^2. \tag{4}$$

The equality holds only if $\rho = 1$. Therefore, when returns over time are not perfectly correlated, Equation (4) suggests that the variance of the appraisal return $\tilde{r}_t^*$ is always less than the variance of the true market return $\tilde{r}_t$. In other words, appraisal returns always understate (or smooth out) the volatility of true real estate returns, as long as Equation (2) is a reasonable representation of the partial adjustment behavior for all appraisers. This is essentially the normative foundation

for the traditional appraisal smoothing theory. It demonstrates that appraisal causes smoothing.

How Does the Model FAIL to Support the Smoothing Argument?

An issue that has been given little consideration by previous literature, however, is the choice of the ''weights,'' or the smoothing parameters, *w*. In one of the model's original presentations, Geltner (1991) suggests that, although the partial adjustment is an individual behavior, the only difference between the models he presented at the aggregated index level and the disaggregated individual level (Equation (1) and (1a) of that paper) is that the individual appraisal errors are diversified away at the aggregated level, but the weighting parameters remained the same at both levels. This implies an implicit assumption: all appraisers, despite their heterogeneous access to and interpretation of information from various sources, would weigh the past versus present information in exactly the same way. That is, appraisers exhibit homogeneous partial adjustment behavior. Subsequent literature using the partial adjustment model has generally adopted the assumption and treats the parameters as invariant across appraisers. It is under this assumption that the partial adjustment model, as demonstrated above, supports the argument that the appraisal variance is less than that of the true variance.

In reality, even if the ''partial adjustment'' is consistent with the appraisers' behaviors in using past information such as identifying the ''comps,'' there is no precise rule on how such adjustment ''should'' be done. Therefore, it is not hard to imagine that, a certain degree of professional consensus notwithstanding, heterogeneous ''partial adjustment'' exists even when there is no apparent incompetence, negligence, or impropriety on the part of the appraisers. Such heterogeneity implies that the smoothing parameter, although specific to each individual appraiser at the disaggregated level, must exhibit cross-sectional dispersion at the aggregated index or portfolio level and thus introduce another source of variation into the partial adjustment model. The resulted appraisal return index, therefore, is likely to exhibit larger variance than if the parameter is treated as invariant.

It is necessary to note that this additional variation due to heterogeneous smoothing parameter is not caused by some kind of random behavior of the appraisers. It is due to the fact that all appraisers cannot, and indeed should not weigh the past versus present information in exactly the same way because of their differences in access to, and interpretation of information. The different choice of smoothing parameter, therefore, is a result of rational appraiser behavior. Each appraiser chooses the most appropriate weighting parameter based on his/her professional judgment. At the aggregate level, the weights vary across appraisers with certain dispersion. Ignoring such heterogeneity in the past may have led us to exaggerate the smoothing effect on appraisal-based return indices.

Below is a formal analysis that investigates the impact of heterogeneous partial adjustment behavior on the resulted appraisal variance. In the context of the model in Equation (2), this analysis begins by recognizing the cross-sectional variation of the smoothing parameter.

Suppose that w follows a distribution with mean $\overline{w}$ and variance σ_w^2.[2] The variance σ_w^2 essentially captures the variation of the smoothing parameter among appraisers. Based on the conditional variance formula, the variance of the appraisal return $\tilde{r}*$ can be expressed as:

$$Var(\tilde{r}_t^*) = E[Var(w\tilde{r}_t + (1 - w)\tilde{r}_{t-1}|w)]$$
$$+ Var[E[w\tilde{r}_t + (1 - w)\tilde{r}_{t-1}|w]]. \qquad (5)$$

Since $E[w\tilde{r}_t + (1 - w)\tilde{r}_{t-1}|w] = u$, hence $Var[E[w\tilde{r}_{t-1} + (1 - w)\tilde{r}_t|w]] = 0$. Simplifying Equation (5) yields:

$$Var(\tilde{r}_t^*) = \sigma^2 E[w^2 + 2\rho w(1 - w) + (1 - w)^2]. \qquad (6)$$

Note that w varies among different appraisers and it is distributed with mean $\overline{w}$ and variance σ_w^2. Given that $\sigma_w^2 = E[w - E(w)]^2 = E[w^2] - \overline{w}^2$, then $E[w^2] = \sigma_w^2 + \overline{w}^2$. Equation (6) can then be simplified as:

$$Var(\tilde{r}_t^*) = 2(1 - \rho)\sigma_w^2\sigma^2 + [\overline{w}^2 + 2\rho\overline{w}(1 - \overline{w}) + (1 - \overline{w})^2]\sigma^2. \qquad (7)$$

Comparing Equation (7) with Equation (3), it can be seen that the second term of Equation (7) is exactly the right-hand side of Equation (3). In other words, Equation (7) shows that the traditional appraisal smoothing argument (Equation (3)) understates the volatility of appraisal return $\tilde{r}_t^*$ by an amount of $2(1 - \rho)\sigma_w^2\sigma^2$. Given that this amount is always positive when, $\rho < 1$ depending on the magnitude of σ_w^2, it is possible that the first term $2(1 - \rho)\sigma_w^2\sigma^2$ in Equation (7) could be large enough to reverse Equation (4) such that $Var(\tilde{r}_t^*) \geq \sigma^2$. That is, the variance of appraisal returns may actually overstate the true volatility of σ^2. In any event, the positive first term in equation (7) suggests that the smoothing effect should be less than previously believed under the traditional appraisal smoothing theory, which assumes w is invariant and $\sigma_w^2 = 0$. Therefore, once the assumption of homogeneous partial adjustment behavior is relaxed to account for appraisal heterogeneity, the partial adjustment model no longer necessarily supports the appraisal smoothing argument.

To see this point with an example, suppose that there are two types of appraisers with equal numbers in the market and the correlation ρ is assumed to be zero for simplicity. The first type of appraiser has no confidence in the current market information and thus gives 100% weight to the return of last period and 0% weight to the current information. However, the second type of appraiser is 100% confident in the current information and thus will not consider the past information in the valuation process. With this example, the weight w is distributed with mean of $\overline{w} = \frac{1}{2}$ and variance of $\sigma_w^2 = \frac{1}{4}$. Inserting these numbers in Equation (7) with the assumption of, $\rho = 0$ gives:

$$Var(\tilde{r}_t^*) = 2 \times \frac{1}{4} \sigma^2 + \left[\left(\frac{1}{2}\right)^2 + \left(1 - \frac{1}{2}\right)^2 \right] \sigma^2 = \sigma^2.$$

This suggests that after considering cross-sectional variation among appraisers, the variance of the appraisal return may not be reduced at all. However, based on the traditional appraisal smoothing theory (ignoring the heterogeneity of appraisal behaviors and assuming $\sigma_w^2 = 0$), the variance of appraisal return is only $\dfrac{\sigma^2}{2}$.

Heterogeneous Appraisal under Changing Market Conditions

The discussion now turns to another aspect of the heterogeneous appraisal: the time-varying nature of the smoothing parameter. A few studies have recognized that, in weighting past versus present information, some appraisers may choose different weight (w) at different times, or under different market conditions. Matysiak and Wang (1995) suggest that the appraisal smoothing parameter w is likely to be time-varying as appraisers would adapt their valuation process according to market conditions. Brown and Matysiak (1998) applied the Kalman filter approach to study the issue and found that the smoothing parameter varied across market conditions. Their finding was confirmed by Clayton, Geltner, and Hamilton (2001), who examined this issue using a sample of Canadian properties. They found that a greater smoothing effect (more temporal lag bias) is correlated with periods of slow real estate market when fewer s and less current information are available. While it is an empirical matter as to how much the smoothing parameter varies over market conditions, the fact that it varies warrants an extension of the analysis to the setting of changing market conditions.

In the last section, it is assumed that $\tilde{r}_t$ and $\tilde{r}_{t-1}$ share the same mean. In reality, the housing market may experience a good or bad time in the next period when market conditions change. In other words, the expected return in period t may become higher or lower than that of the previous period. Suppose that $u_t - u_{t-1} = \Delta$. When $\Delta > 0$, the market improves over the last period. When $\Delta < 0$,

the opposite is true. The magnitude of Δ reflects the degree of how much the market conditions change. Intuitively, since appraisers are less agreeable on how to respond to dramatic market movements, they tend to either over- or under-react to available information. Therefore, a dramatic change in the market conditions is likely to be associated with a large dispersion of w, that is, σ_w^2 is a function of Δ, i.e., $\sigma_w^2(\Delta)$. When market conditions change, i.e., $\Delta \neq 0$, the second term of Equation (5) becomes:

$$Var(E[w\tilde{r}_t + (1 - w)\tilde{r}_{t-1}|w]) = Var(u_{t-1} + w\Delta)$$
$$= \Delta^2\sigma_w^2(\Delta). \tag{8}$$

As a result, when $\Delta \neq 0$ Equation (5) can be further simplified as:

$$Var(\tilde{r}_t^*) = \Delta^2\sigma_w^2(\Delta) + 2(1 - \rho)\sigma_w^2(\Delta)\sigma^2$$
$$+ [\overline{w}^2 + 2\rho\overline{w}(1 - \overline{w}) + (1 - \overline{w})^2]\sigma^2. \tag{9}$$

Thus Equations (7) and (9) lead to:

$$Var(\tilde{r}_t^*|\Delta > 0) - Var(\tilde{r}_t^*|\Delta = 0) = \Delta^2\sigma_w^2(\Delta)$$
$$+ 2(1 - \rho)\sigma^2[\sigma_w^2(\Delta) - \sigma_w^2]. \tag{10}$$

Since a larger dispersion of w is often associated with the change of market conditions, i.e., $\sigma_w^2(\Delta) > \sigma_w^2$, Equation (10) implies that when the real estate market experiences a downturn or upturn, the appraisal returns are likely to become more volatile.

Another relaxed assumption is that the variances of returns in periods t and $t - 1$ are equal. In reality, the variances may differ, especially when market conditions change. Suppose that $\tilde{r}_i$ is distributed with mean u_i and variance $\sigma_i^2(i = t, t - 1)$, Equation (9) then becomes:

$$Var(\tilde{r}_t^*) = \Delta^2\sigma_w^2(\Delta) + (\sigma_t^2 + \sigma_{t-1}^2 - 2\rho\sigma_t\sigma_{t-1})\sigma_w^2(\Delta)$$
$$+ [\overline{w}^2\sigma_t^2 + 2\rho\overline{w}(1 - \overline{w})\sigma_t\sigma_{t-1} + (1 - \overline{w})^2\sigma_{t-1}^2]. \tag{11}$$

Equation (11) is a general representation of the variance in appraisal-based returns under changing market conditions and heterogeneous appraiser behaviors. In the

traditional appraisal smoothing arguments, which implicitly assume that all appraisals share the same smoothing parameter (i.e., $\sigma_w^2(\Delta) = 0$), Equation (11) becomes:

$$Var(\tilde{r}_t^* | \sigma_w(\Delta) = 0) = \overline{w}^2\sigma_t^2 + 2\rho\overline{w}(1 - \overline{w})\sigma_t\sigma_{t-1}$$
$$+ (1 - \overline{w})^2\sigma_{t-1}^2. \qquad (12)$$

The difference between Equations (11) and (12) is:

$$\Theta = Var(\tilde{r}_t^*) - Var(\tilde{r}_t^* | \sigma_w(\Delta) = 0)$$
$$= (\Delta^2 + \sigma_t^2 + \sigma_{t-1}^2 - 2\rho\sigma_t\sigma_{t-1})\sigma_w^2(\Delta). \qquad (13)$$

Θ essentially captures the effect of heterogeneity of appraisal. Since Θ is always positive,[3] it implies that ignoring the heterogeneity of w understates the volatility of appraisal-based returns and overstates the effect of smoothing. Furthermore, when $\sigma_{t-1} > \sigma_t$, it can been shown from Equation (11) that $Var(\tilde{r}_t^*) > \sigma_t^2$. That is, the appraisal-based returns may exhibit higher, rather than lower, volatility than that of the true returns. Furthermore, given that $\Theta > 0$, $\frac{\partial\Theta}{\partial\sigma_w(\Delta)} > 0$ and $\frac{\partial\Theta}{\partial\Delta} > 0$, the smoothing effect suggested by the traditional appraisal smoothing theory is exaggerated, to the larger extant when the dispersion of smoothing parameter among appraisals is higher and/or when market conditions change dramatically.

The above theoretical analysis can be summarized in three findings. First, the traditional appraisal smoothing argument, which state appraisal must result in "smoothed" returns, is only valid under the implicit assumption that all appraisers exhibit exactly the same partial adjustment behavior and weigh the past versus current information in the same way. Second, however, if one agrees that in reality the appraisers' weighting of past versus current information varies both cross-sectional and over time, the appraisal variance will be higher than what is indicated under the homogeneous assumption. This suggests that the smoothing effect is not as strong as previously believed. Third, and more importantly, the findings show that it is possible that, when the appraisal heterogeneity is high enough, the resulted appraisal variance can actually exceed the true variance. In other words, the partial adjustment model does not definitively support the appraisal smoothing argument, that is, even if we do not question its validity in the first place.

From a certain angle, these findings echo some of the earlier findings by Lai and Wang (1998), in which the authors dispute the validity of the original model in Equation (1). The conclusions here are drawn in a different way. That is, even if the model in Equation (1) is taken as given, the appraisal variance is shown to be

either higher or lower than the true variance depending on the degree of the appraisal heterogeneity. In other words, there is no definitive connection between appraisal and smoothing.

## A Reality Check on What the Appraisal Smoothing Theory Predicts

The above theoretical conclusions suggest that, depending on the degree of appraisal heterogeneity, the partial adjustment model can result in appraisal variance that is either higher or lower than the true variance (or not much different from each other). This contradicts the traditional appraisal smoothing argument, which predicts that appraisal-based returns necessarily exhibit low volatility and that, in the absence of appraisal influence, transaction-based returns should exhibit significantly higher volatility. Empirical data are used next to examine which of these predictions is more consistent with the reality.

The appraisal smoothing theory alleges the volatility of transaction-based real estate returns *should not* be as low as what is observed from appraisal-based returns such as the NPI. However, as surprising as it may be, there is virtually no compelling evidence as to what the transaction-based volatility *should be*. The difficulty is understandable. There is simply not a transaction-based index that is "otherwise comparable" with the appraisal-based NPI.[4] Therefore, studies that attempt to show the existence of smoothing bias often must make critical assumptions about the transaction-based returns and appraiser behavior that causes the bias. This approach effectively implies that the smoothing bias has to be demonstrated based on assumption as oppose to facts. As Edelstein and Quan (2006, p. 43) characterize the problem: "Smoothing is defined as the deviation of an index from one which is never observed; and since it is this deviation between series which gives rise to the problem, it is not surprising smoothing is often demonstrated based on assumptions made about the true series and appraiser methodology and practice. ... Frequently, in the absence of empirical support, the existence of smoothing is largely assumed and unsmoothing techniques amounts to a contrived solution to an assumed problem."

Recently available data in the residential market has made it possible to directly measure the degree of smoothing bias at the aggregate level. Although smoothing has been mainly discussed in the context of commercial real estate, the partial adjustment behavior, the alleged cause for smoothing, is common among both commercial and residential appraisers. In fact, residential appraisers are known to rely more on the so-called market comparable approach, which suggests they tend to give more weight to the past price information. Therefore, if the partial adjustment causes "smoothing," the phenomenon should also be observable from the housing market data.

In late 2004, the then Office of Federal Housing Enterprise Oversight (OFHEO), which later merged into the Federal Housing Finance Agency (FHFA) in October

Exhibit 2 | Historic Housing Price Changes Reported by the FHFA (formerly OFHEO)
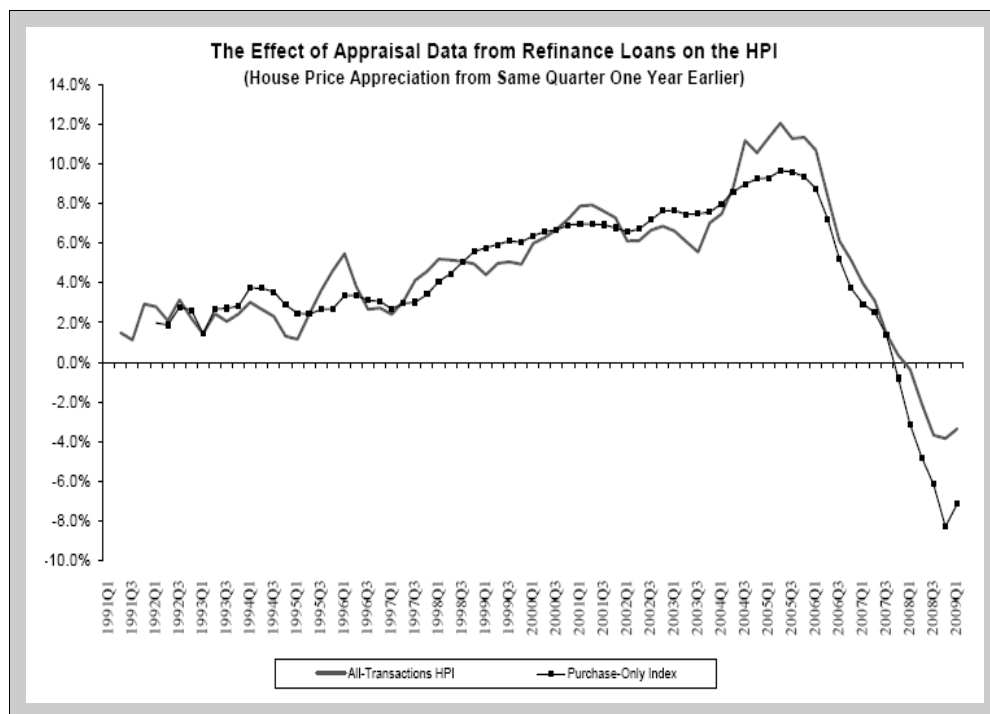
Source: The Federal Housing Finance Agency news release, 2009:Q1.

2008, released a new "purchase-only" Home Price Index (HPI) to supplement its original "All-Transaction Index." Unlike the original index, which uses a large sample mixed with housing sales and refinancing deals (where the property value is the appraisal value instead of sales price), the Purchase-Only Index excludes refinances from the whole sample, thus it is based on the part of the sample that is free of appraisal bias. The portion of the "purchase" subsample averages about 42% of the whole sample over the period of 1991–2008, which implies that the appraisal data (from refinance deals) averages about 58% of the sample during the same period. Such proportion is high enough for an appraisal effect to be noticeable in the All-Transaction Index. Given that both indexes are constructed with the same repeated sales methodology and differ only in whether their samples include appraisal-based values, the difference in the volatilities of the two indexes offers good indication of the appraisal effect.

A quick reality check of the appraisal effect can be found in the official FHFA Quarterly News Release. Exhibit 2 is copied from the 2009:Q1 issue. Exhibit 2 displays the annualized historic returns of the two U.S. national indexes. If the predictions of the appraisal smoothing argument is true, one would expect that the appraisal-based index is more "smoothed" and the "better data" (i.e.,

the transaction-based Purchase-Only Index) should exhibit significantly higher volatility.

Clearly, neither expectation is supported by Exhibit 2. In fact, the opposite seems to be true. The purchase-only index is consistently less volatile and more ''smoothed'' than the appraisal-based All-Transaction Index during the 18-year period, except perhaps in more recent quarters after 2007:Q3 when the financial crisis started, during which time the market is distorted by a disproportionally large number of abnormal sales such as foreclosure and short sales with depressed prices.

The simple fact presented in Exhibit 2 challenges the basic predictions of the traditional appraisal smoothing theory. However, interpretation of Exhibit 2 should be made with caution, because there is another source of noise in the All-Transaction Index besides heterogeneous appraisal behavior. This is the ''mix ratio'' between the purchase versus refinance s in the sample, which may vary over time.[5] When the ratio is relatively stable, the difference in Exhibit 2 can be attributed mainly to the appraisal effect. But when the ratio varies dramatically, the variation may become a significant source of noise. While it is not easy to disentangle the two effects (heterogeneous appraisal and variant mix ratio) in the All-Transaction Index, at least one point is clear: The two indices do not differ significantly, suggesting ''better data'' (i.e., the Purchase-Only Index) does not exhibit significantly higher volatility than the appraisal-based index. To the contrary, the patterns of the two indices are more consistent with the argument that an appraisal-based index does not suffer any smoothing bias at all.

Given the geographic diversity of the housing market, the investigation shown in Exhibit 2 is extended to the regional sub-index level. The FHFA reports sub-indices for the nine census regions (see the Appendix for a map of these regions). Taking the All-Transaction and Purchase-Only Indices for each region over the period of 1991:Q1–2008:Q4, Exhibit 2 is replicated for all nine regions and the results are shown in Exhibit 3. The regional comparison between the two indices exhibit rather a consistent pattern with the national results (Exhibit 2) in at least two aspects. First, the appraisal effect, as indicated by the discrepancy between the All-Transaction and the Purchase-Only Indices, seems to be rather modest in most of the years except the latest market peak and the subsequent downfall. Second, without exceptions, the volatility of the appraisal-based All-Transaction Index seems higher (or less smoothed) than the Purchase-Only Index over time and across most of the regions. While there is some room for interpretation, the case for appraisal causing ''smoothing'' is rather difficult to make in light of these facts.

The historic trends reported in Exhibits 2 and 3 are annual changes of the housing indices. These trends are consistent with the expectation of investors with a one-year investment horizon. In reality, properties are held for various lengths of time. Given the horizon-dependence of the NPI indicated in Exhibit 1, there is good reason to believe residential real estate will exhibit the same characteristics.

Exhibit 3 | Comparison between Appraisal-based and Transaction-based FHFA Index by Regions (1992:Q1–2008:Q4)
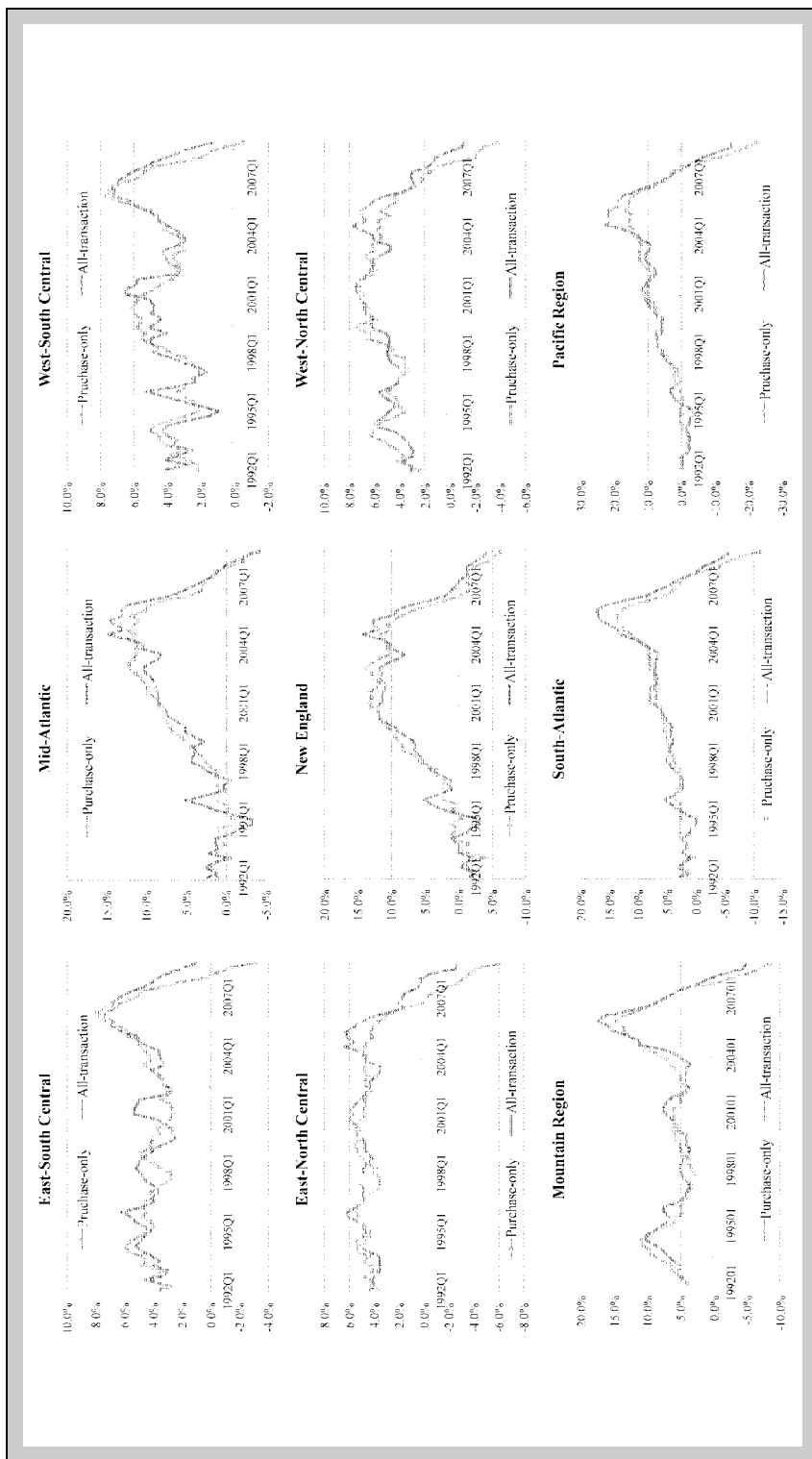
Note: This figure displays the historic annual housing price changes for the nine census regions over the period of 1992:Q1–2008:Q4. It shows that (1) the appraisal effect only causes modest deviation between the two indexes; and (2) the appraisal-based All-Transaction Index seems more (not less) volatile over time and across all regions.
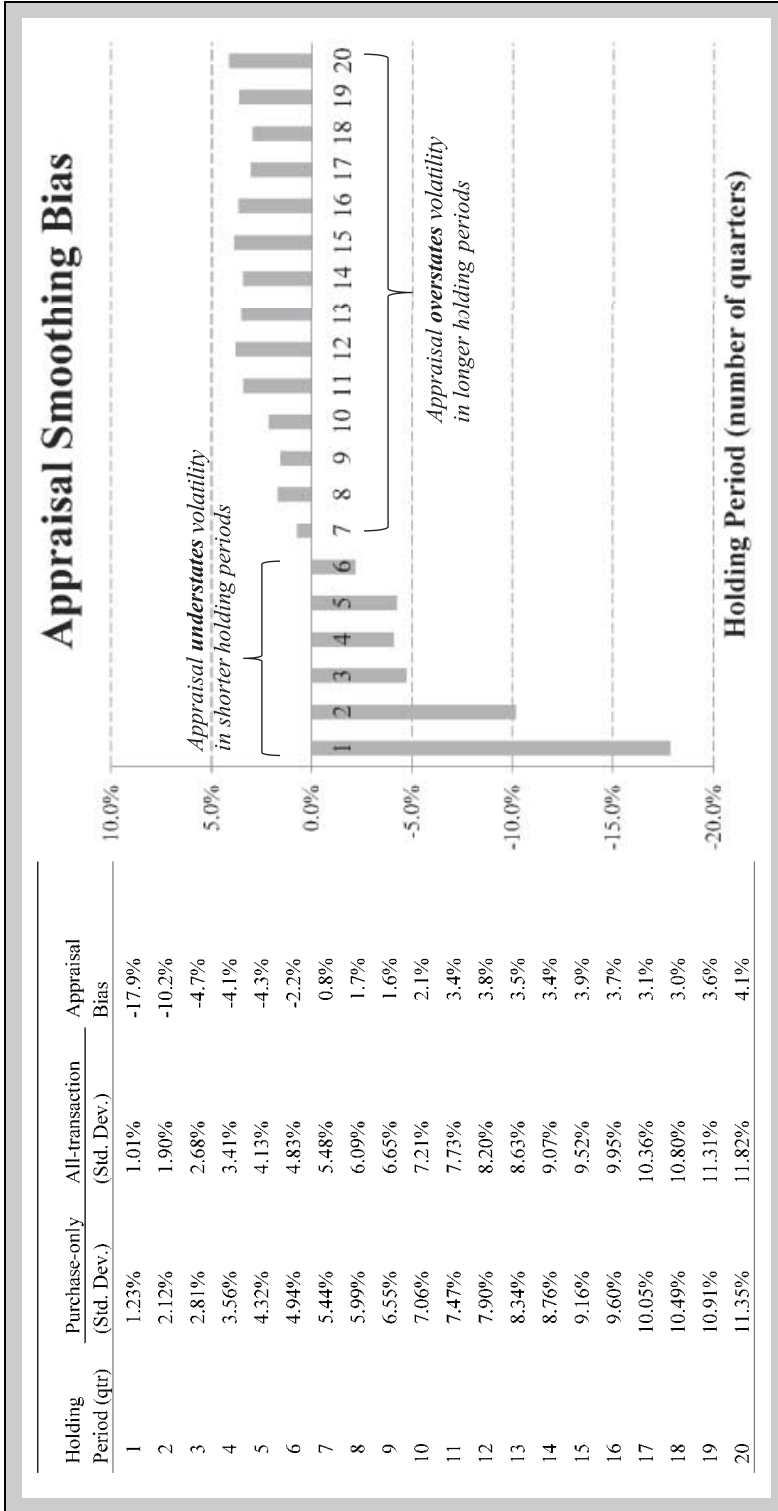
Therefore, the FHFA National Home Price Index is used to examine the appraisal effect over various investment horizons. Since the Purchase-Only Index covers only the period of 1991:Q1–2008:Q4, the range of hypothetical holding periods is limited to be 1–20 quarters (or up to five years). Then the same simulation approach that produced Exhibit 1 is applied to the two FHFA indices. The simulated standard deviations of each index for each holding period are used to compute the smoothing bias, which is measured as the percentage by which the standard deviation of the All-Transaction Index differs from that of the Purchase-Only Index for a given holding period. The results are shown in Exhibit 4.

Exhibit 4 clearly indicates that smoothing bias is also horizon-dependent. "Smoothing" can only be observed under short holding periods (less than 1.5 years). For example, at a holding period of one quarter, the standard deviations of the two indexes (1.23% and 1.01%, respectively), which are the conventional estimates of return volatilities from quarterly indexes, indicate the quarterly All-Transaction Index understates the volatility of the Purchase-Only Index by 17.9%—a significant "smoothing" bias. However, the quarterly return and risk implicitly assume that an investor only holds the index portfolio for a single quarter and sells it immediately, which is not possible in reality. As holding periods becomes more practical (say, two years or longer), the appraisal bias quickly shrinks and turns to positive. That is, the All-Transaction Index actually *overstates* the standard deviations of the Purchase-Only Index, albeit by a modest margin of about 2%–4%. These results are consistent with Exhibits 2 and 3 in that they further suggest that it is a misperception that appraisal-based real estate returns exhibit low volatility. Had the return volatility been examined under more practical holding periods in some research, the traditional appraisal smoothing argument would have appeared ill-supported empirically to begin with.

Some may note that, at a holding period of four quarters (one year), the smoothing bias is still negative in Exhibit 4, which seems to contradict the annual index changes shown in Exhibits 2 and 3. This is because the most recent several quarters during which the market is distorted by rampant foreclosure and short sales are included in the calculations, which causes the Purchase-Only Index to exhibit significantly higher volatility. However, the data period prior to the recent financial crisis is examined (i.e., 1991:Q1–2006:Q4), the standard deviation of the annual Purchase-Only Index is 2.4%, versus 2.8% for the All-Transaction Index. These results are therefore consistent with Exhibits 2 and 3, suggesting the Purchase-Only Index is more "smoothed." In addition, the serial correlation for the annual returns of Purchase-Only Index is 0.73, versus 0.64 for the All-Transaction Index during the same period. This finding contradicts yet another traditional appraisal smoothing argument that predicts an appraisal-based index is supposed to exhibit higher serial correlation than a transaction-based index.[6]

Residential data may not be ideal for examining the traditional appraisal smoothing theory, as the theory is mainly developed for commercial real estate. However, if using past market information is a common behavior among residential and commercial appraisers, the impact of *partial adjustment* should be

Exhibit 4 | Appraisal Effect across Investment Horizons

Appraisal Smoothing Bias

*Appraisal **understates** volatility in shorter holding periods*

*Appraisal **overstates** volatility in longer holding periods*

Holding Period (number of quarters)

Holding Period (qtr)	Purchase-only (Std. Dev.)	All-transaction (Std. Dev.)	Appraisal Bias
1	1.23%	1.01%	-17.9%
2	2.12%	1.90%	-10.2%
3	2.81%	2.68%	-4.7%
4	3.56%	3.41%	-4.1%
5	4.32%	4.13%	-4.3%
6	4.94%	4.83%	-2.2%
7	5.44%	5.48%	0.8%
8	5.99%	6.09%	1.7%
9	6.55%	6.65%	1.6%
10	7.06%	7.21%	2.1%
11	7.47%	7.73%	3.4%
12	7.90%	8.20%	3.8%
13	8.34%	8.63%	3.5%
14	8.76%	9.07%	3.4%
15	9.16%	9.52%	3.9%
16	9.60%	9.95%	3.7%
17	10.05%	10.36%	3.1%
18	10.49%	10.80%	3.0%
19	10.91%	11.31%	3.6%
20	11.35%	11.82%	4.1%

Note: "St. Dev." is the holding period (total) standard deviation. The appraisal bias is measured as the percentage by which the standard deviation of the "All-transaction" index differs from that of the "Purchase-only" index.

observable in both residential and commercial real estate data. Furthermore, recent study suggests that the difference between commercial and residential real estate markets should not be overly exaggerated. For example, Gyourko (2009) has presented some rather comprehensive evidence showing the inextricable connections between the economic and financial fundamentals of commercial and residential real estate. He finds that, among other things, the property value appreciation in owner-occupied houses and commercial properties move together contemporaneously and over time, despite the fact that the two markets are driven by demand from two groups of buyers that differ greatly in their motivations, preferences, and decision criteria. Finally, it is necessary to point out that the theoretical analysis presented here does not depend on any data. So the imperfection of the data should not affect the main conclusions of the analysis.

Conclusion

This study examines the empirical motivation and normative foundation of traditional appraisal smoothing theory. First, it examines the widely reported 'fact' that motivated the appraisal smoothing argument in the first place: the ''exceptionally low volatility'' of the appraisal-based NPI. The findings show that the fact is based on an incomplete observation, which is valid only for unrealistically short holding periods (a year or a quarter). When more realistic and longer holding periods are considered, the NPI's volatility (per unit of return) is not as low as previously believed compared to that of stocks. The fact that the volatility gap between stocks and real estate in Exhibit 1 is horizon-dependent further suggests that it is inappropriate to de-smooth the NPI by inflating its volatility by a presumed percentage.

Second, a formal analysis is conducted to demonstrate that the appraisal smoothing theory is supported by the partial adjustment model only under the assumption of homogeneous appraiser behavior that all the appraisers weigh the past versus current information in exactly the same way. By relaxing the homogenous assumption, the heterogeneous appraisal behavior will result in appraisal variance that is higher than under the homogeneous condition, implying that the appraisal smoothing effect, if it exists, is not as strong as the traditional appraisal smoothing theory suggests. Furthermore, it is possible that the heterogeneity may be to such a degree that the resulted appraisal variance actually exceeds the true variance. That is, appraisal may exaggerate, rather than understate the true variance of real estate returns.

Third, empirical evidence from the residential market is presented, which consistently suggests, historically, the appraisal-based FHFA All-Transaction Index in fact appears to be slightly less (not more) ''smoothed'' than the Purchase-Only Index. This evidence is consistent with the theoretical findings, as they suggest that the appraisal heterogeneity may indeed be such that an appraisal-based index actually exceeds (albeit moderately) the variance of an otherwise comparable transaction-based index. The facts contradict the central prediction of

the traditional appraisal smoothing theory: the transaction-based index should have significantly higher volatility than the appraisal-based index.

The findings of this study have three implications. First, if real estate performance is horizon-dependent and appraiser behaviors are heterogeneous, inferring true market volatility from an appraisal-based index is even more complex than what the de-smoothing literature has attempted before. Given the appraisers' heterogeneous treatment of information in aggregate, and smoothing bias that may exist at the individual property level, the ultimate volatility of the index depends on which influence dominates the other. De-smoothing may be the wrong thing to do because the appraisal-based index may already overstate the true volatility, as the FHFA indices suggest in this paper. Therefore, both influences should be understood and properly controlled before inferring the true market volatility from an appraisal-based index. This suggests an interesting empirical research: How different are appraisers in choosing their smoothing parameters, both cross-sectionally and over time?

Second, if the FHFA indices are indicative of the commercial real estate market, as Gyourko (2009) suggests, the NPI may not suffer much smoothing bias after all. As the longest and most observed market index for commercial real estate, the credibility of the NPI, and to some extent, the credibility of the appraisal industry in general, has been severely undermined by the appraisal smoothing theory in recent years. While many studies often blame the smoothing bias for their nonsensical findings, some researchers have dismissed the NPI as completely useless. This kind of criticism appears premature in light of the findings of this study. Smoothing bias may exist at individual property level, but its impact to the aggregated index is likely to be largely offset by the influence of the appraisal heterogeneity.[7] In the end, the index will not be as smoothed as many had believed, especially when it is viewed under more realistic investment horizons.

Finally, how can the long standing real estate risk premium puzzle be explained? The traditional appraisal smoothing theory can no longer provide a credible explanation, as the theory does not seem to stand up well against simple logic and facts. Therefore, in order to understand why real estate exhibits higher risk-adjusted returns than financial assets (i.e., the real estate risk premium puzzle), it is perhaps useful to reconsider the alternative direction suggested by Lusht (1988).

In his presidential address to the American Real Estate and Urban Economics Association in 1988, Lusht points out the fact that "in real estate markets more than variance affects price." That being the case, he argued: "less variance is not an uncomforting finding, but rather precisely what we would expect in markets where more than variance is priced." These non-variance factors include illiquidity, heterogeneity, high trading costs, etc. His view is concurred by Lai and Wang (1998, p. 532), who state: "a potentially more fruitful way to study why the variance of property returns is lower than that of stock market after adjusting for risk is to examine the unique characteristics of property markets." Until the unique characteristics of real estate is properly captured and priced in a new real

estate risk metric, fair comparison of the performances between real estate and stocks, as well as the proper role of real estate in a mixed-asset portfolio, will remain a major challenge facing academics and practitioners alike.

Appendix

The Nine Census Regions used by OFHEO
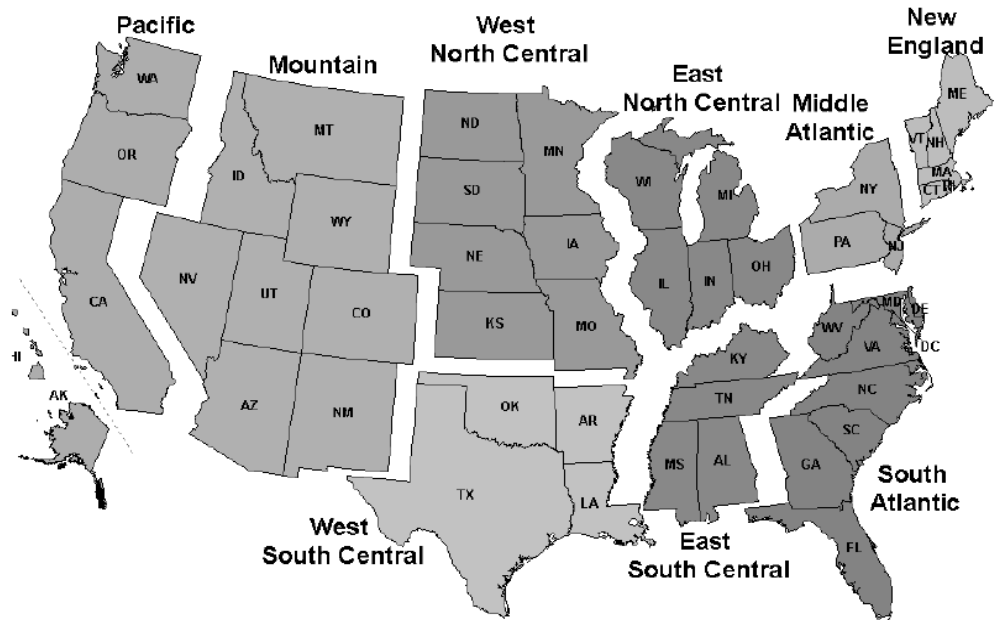
Endnotes

[1] The literature on the subject is too large to be reviewed here fully. A few examples include Case and Shiller (1989), Young, and Graff (1995), Englund, Gordon, and Quigley (1999), and Gao, Lin, and Na (2009), among others.

[2] The analysis does not require the assumption of a specific distribution for w. It only assumes that the mean and variance of w can be estimated.

[3] It can be shown that $\Theta = (\Delta^2 + \sigma_t^2 + \sigma_{t-1}^2 - 2\rho\sigma_t\sigma_{t-1})\sigma_w^2(\Delta) \geq [\Delta^2 + (\sigma_t - \sigma_{t-1})^2]\sigma_w^2(\Delta) \geq 0$.

[4] Recently, however, NCREIF has adopted the MIT transaction-based index. But this new index is constructed with a very different methodology and is not "otherwise comparable" with the traditional NCREIF Property Index in terms of index methodology.

[5] In this regard, the NPI also includes sales when the property is sold. So, like the All-Transaction Index, the NPI is an appraisal-based, but not appraisal-only index, although its "mix ratio" may be quite different.

[6] The authors thank an anonymous referee for suggesting the comparison of serial correlations for the two indices.

[7] For instance, based on an analysis of housing data from Fannie Mae and Freddie Mac over the 1975–1993 period, Chinloy, Cho, and Megbolugbe (1997, p. 97) conclude that ''smoothing is less of a problem than the volatility induced by small sample size and heterogeneous beliefs among appraisers.''

References

Blundell, G. and C. Ward. Property Portfolio Allocation: A Multi-factor Model. *Land Development Studies*, 1987, 4:2, 145–56.

Bond, S. and S. Hwang. Smoothing, Nonsynchronous Appraisal and Cross-sectional Aggregation in Real Estate Price Indices. *Real Estate Economics*, 2007, 35, 349–82.

Breuggeman, W., A. Chan, and T. Thibodeau. Some Additional Evidence on the Performance of Commingled Real Estate Investment Funds: 1972–1991. *Journal of Real Estate Research*, 1992, 7, 433–48.

Brown, G. and G. Matysiak. Valuation Smoothing without Temporal Aggregation. *Journal of Property Research*, 1998, 15:2, 89–105.

——. Sticky Valuations, Aggregate Effects and Property Indices. *Journal of Real Estate Finance and Economics*, 2000, 20:1, 49–66.

Case, K. and R. Shiller. The Efficiency of the Market for Single-Family Homes. *American Economic Review*, 1989, 79, 125–37.

Cheng, P., Z. Lin, and Y. Liu. Illiquidity and Portfolio Risk of Thinly-traded Assets. *Journal of Portfolio Management*, 2010a, 36:2, 126–38.

——. Illiquidity, Cost, and Optimal Holding Period for Real Estate: Theory and Application. *Journal of Housing Economics*, 2010b, 19, 109–118.

Chinloy, P., M. Cho, and I. Megbolugbe. Appraisals, Incentives, and Smoothing. *Journal of Real Estate Finance and Economics*, 1997, 14, 89–111.

Cho, H., Y. Kawaguchi, and J. Shilling. Unsmoothing Commercial Property Returns: A Revision to Fisher-Geltner-Webb's Unsmoothing Methodology. *Journal of Real Estate Finance and Economics*, 2003, 27, 393–405.

Cho, M. and I.F. Megbolugbe. An Empirical Analysis of Property Appraisal and Mortgage Redlining. *Journal of Real Estate Finance and Economics*, 1996, 13:1, 45–55.

Clayton, J., D. Geltner, and S. Hamilton. Smoothing in Commercial Property Appraisals: Evidence from Individual Appraisals. *Real Estate Economics*, 2001, 29:3, 337–60.

Corgel, J. and J. deRoos. Recovery of Real Estate Returns for Portfolio Allocation. *Journal of Real Estate Finance and Economics*, 1999, 18:3, 279–96.

Diaz, J. An Investigation into the Impact of Previous Expert Value Estimates on Appraisal Judgment. *Journal of Real Estate Research*, 1997, 13:1, 57–66.

Diaz, J. and M. Wolverton. A Longitudinal Examination of the Appraisal Smoothing Hypothesis. *Real Estate Economics*, 1998, 26:2, 349–58.

De Wit, D. Smoothing Bias in In-house Appraisal-based Returns of Open-end Real Estate Funds. *Journal of Real Estate Research*, 1993, 8:2, 157–70.

Edelstein, R. and D. Quan. How Does Appraisal Smoothing Bias Real Estate Return Measurement? *Journal of Real Estate Finance and Economics*, 2006, 32:1, 41–60.

Englund, P., T. Gordon, and J. Quigley. The Valuation of Real Capital: A Random Walk Down Kungsgatan. *Journal of Housing Economics*, 1999, 8, 205–16.

Fama, E., Multi-period Consumption-Investment Decisions. *American Economic Review*, 1970, 163–74.

Fisher, J., D. Geltner, and R.B. Webb. Value Indices of Commercial Real Estate: A Comparison of Index Construction Methods. *Journal of Real Estate Finance and Economics*, 1994, 9, 137–64.

Gao, A., Z. Lin, and C. Na. Housing Market Dynamics: Evidence of Mean Reversion and Downward Rigidity. *Journal of Housing Economics*, 2009, 18: 256–66.

Geltner, D. Bias in Appraisal-Based Returns. *Journal of the American Real Estate and Urban Economics Association*, 1989, 17:3, 338–52.

——. Smoothing in Appraisal-Based Returns. *Journal of Real Estate Finance and Economics*, 1991, 4:3, 327–45.

——. Estimating Market Values from Appraised Values without Assuming an Efficient Market. *Journal of Real Estate Research*, 1993, 8:3, 325–45.

Geltner, D., B. MacGregor, and G. Schwann. Appraisal Smoothing and Price Discovery in Real Estate Markets. *Urban Studies*, 2003, 1047–64.

Gyourko, J. Understanding Commercial Real Estate: How Different Is It from Housing? *Journal of Portfolio Management*, Special Real Estate Issue, 2009, 23–37.

Hendershott, P. and E. Kane. U.S. Office Market Values During the Past Decade: How Distorted Have Appraisals Been? *Real Estate Economics*, 1995, 23, 101–16.

Lai, T-Y. and K. Wang. Appraisal Smoothing: The Other Side of the Story. *Real Estate Economics*, 1998, 26, 511–36.

Lin, Z. and Y. Liu. Real Estate Returns and Risk with Heterogeneous Investors. *Real Estate Economics*, 2008, 36, 753–76.

Lin, Z. and K. Vandell. Illiquidity and Pricing Biases in the Real Estate Market. *Real Estate Economics*, 2007, 35, 291–330.

Lusht, K. The Real Estate Pricing Puzzle. *Journal of the American Real Estate and Urban Economics Association*, 1988, 16, 95–104.

Matysiak, G. and P. Wang. Commercial Property Prices and Valuation: Analyzing the Correspondence. *Journal of Property Research*, 1995, 12:1, 181–202.

Merton, R. Lifetime Portfolio Selection under Uncertainty: The Continuous Time Case. *Review of Economics and Statistics*, 1969, 51, 247–57.

Pagliari, J., K. Scherer, and R. Monopoli. Public Versus Private Real Estate Equities: A More Refined, Long-term Comparison. *Real Estate Economics*, 2005, 33, 147–87.

Quan, D., and J. Quigley. Price Formation and the Appraisal Function in Real Estate. *Journal of Real Estate Finance and Economics*, 1991, 4:2, 127–46.

Samuelson, P. Lifetime Portfolio Selection by Dynamic Stochastic Programming. *Review of Economics and Statistics*, 1969, 51, 239–46.

Sirmans, G.S. and G.E. Sirmans. The Historical Perspective of Real Estate Returns. *Journal of Portfolio Management*, 1987, 13:3, 22–31.

Shilling, J. Is There a Risk Premium Puzzle in Real Estate? *Real Estate Economics*, 2003, 31, 501–25.

Young, M. and R. Graff. Real Estate is Not Normal: A Fresh Look at Real Estate Return Distributions. *Journal of Real Estate Finance and Economics*, 1995, 10, 225–59.

Wang, P. Errors in Variables, Links between Variables and Recovery of Volatility Information in Appraisal-based Real Estate Return Indexes. *Real Estate Economics*, 2006, 34, 497–518.

Webb, R.B., M. Miles, and D. Guilkey. Transactions-driven Commercial Real Estate Returns: The Panacea to Asset Allocation Models? *Journal of the American Real Estate and Urban Economics Association*, 1992, 20, 325–57.

Zerbst, R. and B. Cambon. Historical Returns on Real Estate Investments. *Journal of Portfolio Management*, Spring, 1984.

Ping Cheng, Florida Atlantic University, Boca Raton, FL 33431 or pcheng@fau.edu.

Zhenguo Lin, California State University, Fullerton, CA 92834-6848 or zlin@fullerton.edu.

Yingchun Liu, Laval University, Quebec, G1V 0A6, Canada or yingchun.Liu@fsa.ulaval.ca.