Random Disaggregate Appraisal Error in Commercial Property: Evidence from the Russell-NCREIF Database

*David M. Geltner**
*Richard A. Graff***
*Michael S. Young****

Abstract. **This paper examines the magnitude of random disaggregate appraisal valuation error in institutional-grade commercial property. Unlike previous transactions-based studies of appraisal error, we use a much larger database that is not restricted to sold properties, and we employ a methodology that focuses on appraisal error rather than the difference between transaction price and previous appraised value. Our model gives a point estimate of 11.07% for the standard error of appraisals in the Russell-NCREIF database, with a robust range of 6% to 13%.**

Introduction

It has long been recognized that commercial real estate assets are not fungible. Consequently, in contrast to stock and bond markets, market transactions do not provide an accurate proxy for asset valuation and investment reporting within the commercial real estate sector.

Although appraisal values are widely accepted as the best substitute available, the difference between appraisal and market value has been a subject of concern to institutional investors. Much of this concern has focused on the aggregate value of many properties combined in a portfolio or index. At that level a major focus of research has been the effect of appraisals in causing smoothing and lagging in the capital gains component over time in indices of property returns such as the Russell-NCREIF Property Index. A number of studies have addressed the smoothing issue at both the theoretical and empirical level.[1] At the disaggregate (i.e., individual property) level, most of the recent literature has focused on theoretical issues, such as optimal appraisal technique under various assumptions.[2] There has been little empirical study of disaggregate appraisal error. This paper is an attempt to address that gap in the literature.

Appraisal error is a rather different phenomenon at the disaggregate and aggregate levels. In addition to the smoothing component observed in aggregate data, disaggregate appraisal error contains a purely random error component that diversifies out of the aggregate data. At the disaggregate level, however, the purely random

*Finance Department, College of Business Administration, University of Cincinnati, Ohio 45221-0195.
**Electrum Partners, 400 North Michigan Avenue, Suite 1200, Chicago, Illinois 60611.
***The RREEF Funds, 650 California Street, Suite 1800, San Francisco, California 94108.
Date Revised—June 1994; Accepted—July 1994.

component of appraisal error is important, as appraisals are primarily used at this level to help obtain an accurate estimate of the market value of a specific property.[3] In particular, knowledge about the typical magnitude of the random component of disaggregate appraisal error can aid in judging the reliability of the appraised values of individual properties. This is important, for example, in determining whether there is a significant difference between a review appraisal and the first appraisal, as well as determining how an offer price should be related to an appraised value.[4]

Knowledge about the magnitude of the random component of disaggregate appraisal error also helps in drawing empirically based conclusions about the "heterogeneity" of real estate, i.e., the degree to which investment risk for each property is idiosyncratic (independent of investment risk for other properties) and therefore can be diversified away by investing in a portfolio of sufficiently many properties.[5] Some past studies have indicated that real estate is very heterogeneous, much more so than the stock market. For example, using appraisal-based returns Hartzell-Hekman-Miles (1986) found that over 80% of the variance in the returns to the average individual property could be eliminated in a broadly diversified portfolio of properties. This compares to the elimination of only a little over 50% of the variance in the average individual stock in a diversified portfolio of stocks. However, what matters to investors is variance in actual market value, which is what is measured in the case of stock returns, but not in the case of appraisal-based returns. Appraisal-based returns contain the random disaggregate appraisal error component that adds variance to disaggregate returns, but not to aggregate returns, thus exaggerating real estate heterogeneity. To gauge the extent of this exaggeration we must know the typical magnitude of the random component of disaggregate appraisal error.

The present study seeks to contribute to our knowledge about random disaggregate appraisal error by analyzing disaggregate returns in the Russell-NCREIF Property database. By employing some plausible assumptions and sensitivity analysis around those assumptions, we believe this large and unique database offers some evidence and insights concerning the magnitude of disaggregate appraisal error for a large class of regularly appraised commercial property.[6]

The paper is organized in four subsequent sections. The next section discusses evidence from previous empirical studies and some conceptual issues regarding the measurement of random disaggregate appraisal error. The second section presents the conceptual model and analytical approach we take in this paper and presents our empirical estimate of error magnitude in the historical Russell-NCREIF database. A third section discusses the implications of our findings regarding the heterogeneity issue. A final section summarizes and concludes the paper.

Previous Empirical Studies Relevant to the Study of Disaggregate Appraisal Error

Considering the importance of disaggregate appraisal error of commercial properties, it is surprising that there has been relatively little empirical analysis of this phenomenon. The studies by Miles et al. (1991) and Webb (1994) in which transaction prices are compared to recent appraised values for a sample of sold properties are the most relevant studies to date. These studies find the average absolute difference between the appraised value and the transaction price to be approximately 10%.[7] A similar study by Blundell and Ward (1993) on commercial property in England found even larger differences—18% standard error.

However, there are some problems with this approach if we are trying to infer from it the average magnitude of the random component of disaggregate appraisal error. For one thing, a sample of transacted properties may not be representative. Another problem is that the difference between appraised value and transaction price may be due not only to appraisal error but to changes in market value occurring during the time between the last appraisal and the transaction. Furthermore, observed appraisal error may be due to both the random disaggregate component and the systematic smoothing and lagging observed in aggregate returns. More fundamentally, even regarding the random disaggregate component, there is a conceptual problem in attempting to make inferences about the magnitude of appraisal error using these transaction-based studies, due to the fact that the notions of "transaction price" and "market value" cease to coincide in the case of private real estate markets.

This distinction together with its significance can be seen by comparing the precise definitions of the two terms. "Market value" is defined as the "most likely" transaction price, or more properly, the *mean* of the distribution of ex ante possible transaction prices. This may be thought of as the mean of the distribution of prices that informed individuals would be willing to pay or accept for the property.[8] In relatively less liquid markets such as real estate, it is generally not reasonable to assume that each transaction will occur at a transaction price that exactly equals the mean of this ex ante distribution. Once a property is sold, the observed transaction price is best viewed as a single random drawing from the ex ante distribution. The transaction price will therefore generally differ from the ex ante mean, and therefore from the market value of the property.

Of course, appraised values also do not equal market values of properties. Apart from smoothing and timing issues, random disaggregate appraisal error might be caused by an inability of the appraiser to perfectly adjust for differences between the subject property and comparable properties, by an inability to perfectly match the market's assumptions regarding expected future rents or discount rate, etc. Some suggestion of the perceived magnitude of disaggregate appraisal error is revealed in Damodaran and Liu's (1993) report that when REITs are occasionally subjected to "appraisal audits" (i.e., the REIT hires an independent fee appraiser to check the REIT's own internal valuations), the REIT's internal valuation is "not confirmed" (and then modified) *only* if the external appraisal differs by more than *10%* from the internal valuation. This "10 percent rule" is apparently ad hoc, not based on empirical analysis.[9]

In reference to the previous transactions-based empirical studies by Miles et al. (1991) and Webb (1994), the point here is that disaggregate appraisal error should not be the difference between the appraised value and the subsequent transaction price of a subject property, but rather the difference between the appraised value and the market value of that property. If appraised value equals market value plus a random drawing from an appraisal "error" distribution, and transaction price equals market value plus a random drawing from a transaction "noise" distribution, then the Miles–Webb and Blundell–Ward approach measures the difference between these two random drawings, not the difference between either drawing and market value. If appraisal error is independent of transaction noise, then the dispersion (i.e., the average absolute magnitude or average squared deviation) of the differences between appraised values and transaction prices will be greater than the dispersion in either random variable alone.

For example, ignoring smoothing and timing differences for the moment, suppose both appraised value and transaction price are drawn from distributions with the same mean (equal to the current market value of the property). If the standard deviation of the appraisal error distribution is 5% of the market value (variance equals .0025), and the standard deviation of the transaction price distribution is 10% of the market value (variance equals .01), then the standard deviation of the difference between the transaction price and the appraised value will be 11.2% ($=\sqrt{(.0025+.01)}$) of the market value provided the appraisal error is independent of the transaction noise.

In some cases the appraiser may know in advance the exact or approximate transaction price at which a property he is appraising will soon be sold (e.g., the deal is already "done" or "in the works" at the time the appraisal is made).[10] To the extent this occurs, the appraisal error and the transaction noise random variable would not be independent, but would in fact be positively correlated. This would cause the dispersion in the distribution of observed differences between appraised values and transaction prices to be less than that indicated in the above example. Indeed, such positive correlation could conceivably cause the price difference dispersion found in transactions-based studies to be *less than* the dispersion in either the appraisal error or the transaction noise distributions alone.[11]

In summary, previous empirical studies of the difference between appraised value and subsequent transaction price suffer from a variety of problems if we attempt to use them to quantify the typical magnitude of the random component of disaggregate appraisal error. Some of these problems could cause the price difference dispersion found in transaction-based studies to exceed the dispersion in the random disaggregate appraisal error distribution (e.g., smoothing, timing, independence of appraisal error and transaction noise). Other problems could cause bias in the other direction (e.g., non-representative easy-to-appraise transaction sample, positive correlation between appraisal error and transaction noise), allowing the possibility that random appraisal error is greater than the price difference error found in the transaction-based studies. Thus, it would seem desirable to obtain some empirical evidence about the typical magnitude of random disaggregate appraisal error using approaches different from the previous transaction-based studies.

Estimating Disaggregate Appraisal Error

This section presents and applies a conceptual model to estimate the magnitude of random disaggregate appraisal error in the Russell-NCREIF database, given the constraints imposed by "masking" in the publicly available data.[12]

The basic motivation behind the procedure used in this paper is to recognize that purely random appraisal error is one source, but not the only source, of the dispersion we can observe in the disaggregate level Russell-NCREIF returns. We attempt to quantify, and remove, the dispersion caused by the other sources, thereby exposing the dispersion in the random appraisal errors. In essence, our procedure combines a model of appraisal error with a market model of true real estate returns. By subtracting out the (observable) aggregate return, we are left with disaggregate "residuals" that are composed of statistically independent components, one of which is the random disaggregate appraisal error. In this context our appraisal error model allows us to quantify (after the application of some judgment) the magnitude of the other additive components of the observed residual variance.

We begin by considering the difference between the appraisal value and the market value of property p at time t. This difference ("appraisal error") can be considered to arise from two components: smoothing, and purely random error. We represent these two sources of error by the following model that relates appraised value and the market value at the disaggregate level:

$$V^*_t(p) = \eta_t(p) + \alpha V_t(p) + (1-\alpha)V^*_{t-1}(p),$$

where: $V^*_t(p)$ is the (log of) appraised value of property p at time t; and $V_t(p)$ is the (log of) market value (mean of the ex ante transaction price distribution) of property p at that time. Alternatively, expressed in returns (first differences) rather than levels:

$$r^*_t(p) = [\eta_t(p) - \eta_{t-1}(p)] + \alpha r_t(p) + (1-\alpha)r^*_{t-1}(p), \tag{1}$$

where: $r^*_t(p)$ is the appraisal-based appreciation return for property p between $t-1$ and t, and $r_t(p)$ is the corresponding "true" or market value-based return.[13]

In equation (1), $\eta_t(p)$ represents the purely random component of appraisal error, while the α parameter (where $0<\alpha<1$) applied to the other terms on the right-hand side captures the effect of smoothing at the disaggregate level. (The smaller the α, the greater the smoothing.) The reason such smoothing would typically occur in appraised values is discussed in previous articles.[14] As the random valuation error component is, by definition, idiosyncratic and purely random, the mean of $\eta_t(p)$ equals zero and the realization of $\eta_t(p)$ is uncorrelated both across properties and across time. It is this purely random error component within the historical Russell-NCREIF sample that we seek to quantify in the present paper.

Next, we note that the market value-based return for property p during year t can be expanded into two components: the aggregate or "market-index" component common across all properties within the market sector in which p is classified, and the idiosyncratic component unique to property p individually:[15]

$$r_t(p) = m_t + e_t(p). \tag{2}$$

Here, m_t is the mean or aggregate component (reflecting systematic and market sector responses), and $e_t(p)$ is the market value-based idiosyncratic or unsystematic component of property p's true market value-based return in year t. This second component reflects news that uniquely affects property p (such as the discovery of a leak in its roof), or news that is offset in other properties within the same market sector as p (such as news that one of property p's tenants is moving to a competitive property within the same market sector). By definition, $e_t(p)$ has a zero mean and is uncorrelated (or perhaps negatively correlated) across properties: $\mathrm{COV}[e_t(i), e_t(j)] \leq 0$, such that $e_t(p)$ diversifies out of the aggregate return.[16]

Combining (1) and (2) and expanding, we see that the disaggregate appraisal-based return observations can be expressed as in equation (3):

$$r^*_t(p) = +\alpha \sum_{i=0}^{\infty} (1-\alpha)^{i} \, m_{t-i} + \alpha \sum_{i=0}^{\infty} (1-\alpha)^{i} \, e_{t-i}(p)$$

$$+ \sum_{i=0}^{\infty} (1-\alpha)^{i} [\eta_{t-i}(p) - \eta_{t-i-1}(p)] \, . \qquad (3)$$

Now, the smoothed aggregate return component represented by the first summation on the right-hand side of (3) can be observed in the aggregate appraisal-based return for the market sector:[17]

$$r^*_t = \alpha \sum_{i=0}^{\infty} (1-\alpha)^{i} \, m_{t-i} \, , \qquad (4)$$

where r^*_t is the aggregate appraisal-based return for the market sector (i.e., the appraisal-based index return for office properties, retail properties, etc.). Subtracting (4) from (3) we obtain the observed disaggregate appraisal-based "residuals," labelled $\varepsilon_t(p)$:

$$\varepsilon_t(p) = \alpha \sum_{i=0}^{\infty} (1-\alpha)^{i} \, e_{t-i}(p) + \sum_{i=0}^{\infty} (1-\alpha)^{i} [\eta_{t-i}(p) - \eta_{t-i-1}(p)] \, . \qquad (5)$$

Finally, we note that all of the random variables on the right-hand side of (5) are statistically independent, by definition. Thus, the observable variance of the appraisal-based residuals is given by equation (6):[18]

$$\text{VAR}[\varepsilon_t(p)] = \left(\frac{\alpha}{2-\alpha}\right) \text{VAR}[e_t(p)] + \left(\frac{2}{2-\alpha}\right) \text{VAR}[\eta_t(p)] \, . \qquad (6)$$

Therefore, the dispersion of the purely random component of the disaggregate appraisal error, expressed as the variance of this error component in the pooled data, is given by equation (7):[19]

$$\text{VAR}[\eta_t(p)] = (1-\alpha/2)\text{VAR}[\varepsilon_t(p)] - (\alpha/2)\text{VAR}[e_t(p)] \, . \qquad (7)$$

In order to quantify $\text{VAR}[\eta_t(p)]$, equation (7) makes it clear that we must quantify two factors in addition to the observable value of $\text{VAR}[\varepsilon_t(p)]$. These are: the disaggregate smoothing factor, α, and the volatility in the market value-based idiosyncratic return component, $\text{VAR}[e_t(p)]$. Due to data masking, neither of these parameters can be estimated from our sample. However, we propose to assign plausible values to these parameters within the Russell-NCREIF database and then to employ sensitivity analysis to gain additional insight.

Consider first the disaggregate smoothing factor, α. Geltner (1993a) suggests that the model we are using here with a value of $\alpha = 1/2$ captures well the effects of disaggregate smoothing in annual frequency returns in the Russell-NCREIF database.[20] We therefore propose to use as our most plausible estimate of the disaggregate smoothing factor, the value: $\alpha = .50$, and then check for sensitivity within the range from $\alpha = .33$ to $\alpha = .67$.[21]

Next, consider the volatility in the true idiosyncratic return component within our sample, $\mathrm{VAR}[e_t(p)]$. This moment can be quantified in two steps. First, we consider the likely true (unsmoothed) volatility in the aggregate Russell-NCREIF properties (i.e., a "fully diversified" portfolio as represented by the All-Property Index). This represents the systematic risk component in the property returns. Then, by making assumptions about how heterogeneous the Russell-NCREIF properties actually are, we can estimate the magnitude of the unsystematic risk component (the idiosyncratic return volatility).

Regarding the first step, a number of recent studies suggest that the annual aggregate return volatility of the Russell-NCREIF All-Property Index is around 10%.[22] Now recall from equation (2) that the idiosyncratic return components are not taken with respect to a fully diversified portfolio such as the All-Property Index, but rather with respect to a portfolio within each property-type market sector (the residual observations in equation (5) are computed *within* market sectors, that is, relative to market sector mean returns). Volatility of property market sectors may exceed that of the fully diversified portfolio as a whole, due to the effect of diversification across property types. The Russell-NCREIF Index publishes subindices by property type that allow us to estimate this additional volatility. The average volatility in the annual returns of the Russell-NCREIF property-type subindices is approximately 1.2 times the volatility of the All-Property Index during the 1980–92 period examined here. Thus, we estimate aggregate real estate market sector volatility to be 12%.

The second step in estimating the idiosyncratic return volatility that we need to quantify in equation (7) is to make an assumption regarding the true heterogeneity of the properties in the Russell-NCREIF database. The assumption we employ is that the Russell-NCREIF properties have approximately the same degree of overall heterogeneity as the stock market. Studies such as McEnally and Boardman (1979) have found that the variance of the average individual stock is two to three times the variance of an equally weighted portfolio of all stocks. Thus, if real estate with a fully diversified true volatility of 10% (variance = .01) has heterogeneity like the stock market, then the variance of the average individual property would be between .02 and .03. This implies average individual property volatility between 14.1% and 17.3% per year.[23] After subtracting the .0144 systematic variance (implied by our previous 12% estimate of aggregate market sector volatility), this suggests that .01 would be a reasonable value for our "most likely" estimate of $\mathrm{VAR}[e_t(p)]$ in equation (7).[24] As a sensitivity range, we propose idiosyncratic return variances ranging from $\mathrm{VAR}[e_t(p)] = .005$ to $\mathrm{VAR}[e_t(p)] = .015$.[25]

Substituting the above-described parameter value assumptions into equation (7), we have our estimate of the standard deviation of the random disaggregate appraisal error given, for each market sector, by equation (8):

$$SD[\eta_t(p)] = \sqrt{\{(.75)\mathrm{VAR}[\varepsilon_t(p)] - (.0025)\}}\,, \tag{8}$$

where $\mathrm{VAR}[\varepsilon_t(p)]$ is the empirically observed variance in the disaggregate appraisal-based appreciation return residuals for the market sector. Out lower bound estimate would be given by (8a):

$$SD\,[\eta_t(p)] = \sqrt{\{(.67)\text{VAR}[\varepsilon_t(p)] - (.0050)\}}\,, \tag{8a}$$

and our upper bound by (8b):

$$SD\,[\eta_t(p)] = \sqrt{\{(.83)\text{VAR}[\varepsilon_t(p)] - (.0008)\}}\,. \tag{8b}$$

In obtaining our estimate of the historical magnitude of random disaggregate appraisal error in the Russell-NCREIF database, we have applied the above procedure to the entire Russell-NCREIF property sample, not just the sold-property sample (which is much smaller, and proprietary).[26] Our raw data thus consists of the disaggregate empirical observations of the appraisal-based appreciation returns, $r^*_t(p)$, for each of the hundreds of properties (p) in the database, in each year from 1980 through 1992.[27]

In this historical database the pooled sample variance in the observed residuals, VAR[$\varepsilon_t(p)$], equals .019665. Substituting this value into equation (8), we obtain the implied standard deviations of the random component of the disaggregate appraisal error, under the three assumptions of our model. The "most likely" (equation 8) estimate is that the sample standard deviation was 11.07% of the property value in the 1980–92 Russell-NCREIF database. The lower (equation 8a) and upper (equation 8b) bound estimates are between 9.0% and 12.5%.

Implications Regarding Real Estate Heterogeneity

As noted previously, the analytical approach taken in this paper allows some exploration of the question of real estate "heterogeneity," or the extent to which investors can reduce risk exposure by diversifying across properties. Heterogeneity is defined to be the ratio of the variance of the average disaggregate return divided by the variance of the aggregate return.[28] As noted previously, by this measure the stock market has a heterogeneity between 2 and 3. In contrast, previous literature has suggested considerably greater heterogeneity within the institutional grade commercial property asset class.

The most detailed previous analysis of commercial property heterogeneity is that of Hartzell–Hekman–Miles (1986).[29] Using appraisal-based returns, they find a heterogeneity ratio of 11.00 in the quarterly returns of a 220-property sample portfolio diversified both geographically and by property type. However, quarterly returns tend to exaggerate the true heterogeneity in a database where many properties are effectively reappraised only once per year, staggered at different times throughout the year. Hartzell–Hekman–Miles report heterogeneity of only 5.49 in the *annual* returns to the same 220-property portfolio. Within property-type market sectors, they find average annual return heterogeneity of only 4.04.[30]

It is this annual, within-property-type sector heterogeneity that is most comparable to the heterogeneity we are dealing with in the present study. However, Hartzell–Hekman–Miles did not adjust their heterogeneity estimates for random disaggregate appraisal error. As this type of error acts to increase the volatility of individual property returns, appraisal-based heterogeneity will generally overstate the true heterogeneity.

We may gain some insight on the likely magnitude of both appraisal error and true heterogeneity by examining the implications of the relationship between heterogeneity, disaggregate error, aggregate volatility, smoothing, and the observed residual dispersion in the Russell-NCREIF disaggregate return database.

Given an assumption about the magnitude of true aggregate volatility, assumptions about heterogeneity can be translated into implied assumptions about the volatility of the disaggregate idiosyncratic return component. This can then be substituted into equation (7) and applied to the Russell-NCREIF database to produce an implied magnitude of purely random disaggregate appraisal error. Thus, the set of mutually consistent heterogeneity and appraisal error values can be explicitly identified for the Russell-NCREIF database, which may shed some light on the plausible range of values for both of these variables.

To clarify this procedure, we note that, from equation (2):

$$\mathrm{VAR}[e_t(p)] = \mathrm{VAR}[r_t(p)] - \mathrm{VAR}[m_t]\,. \tag{9}$$

Now substituting (9) into (7), taking cross-sectional sample averages, and invoking the definition of "heterogeneity," we obtain:

$$\mathrm{VAR}[\eta_t(p)] = (1-\alpha/2)\mathrm{VAR}[\varepsilon_t(p)] - (\alpha/2)(\mathrm{VAR}[r_t(p)] - \mathrm{VAR}[m_t])\,.$$

Therefore:

$$\begin{aligned} E_p[\mathrm{VAR}[\eta_t(p)]] &= (1-\alpha/2)E_p[\mathrm{VAR}[\varepsilon_t(p)]] - (\alpha/2)\;(E_p[\mathrm{VAR}[r_t(p)]] - \mathrm{VAR}[m_t]) \\ &= (1-\alpha/2)E_p[\mathrm{VAR}[\varepsilon_t(p)]] - (\alpha/2)\;(H-1)\mathrm{VAR}[m_t]\,, \end{aligned} \tag{10}$$

where: $E_p[\ldots]$ is the cross-sectional sample mean operator (i.e., the mean taken across individual properties); H is the true heterogeneity factor as we have defined it, $H = E_p\,[\mathrm{VAR}[r_t(p)]]/\mathrm{VAR}[m_t]$; and $\mathrm{VAR}[m_t]$ is the true aggregate return variance.[31] Equation (10) therefore establishes a relation between the random disaggregate appraisal error dispersion and the true heterogeneity of the sample, given the values of the other parameters in the equation.

Based on equation (10) and the level of disaggregate residual variance observed in the Russell-NCREIF database, Exhibit 1 shows the mutually consistent values of disaggregate appraisal error and true within-sector heterogeneity. The exhibit encompasses three alternative assumptions about true market-sector volatility and five alternative assumptions about disaggregatel-evel appraisal smoothing. Given the empirically observed volatility in the appraisal-based aggregate index, low true volatility is not consistent with high levels of appraisal smoothing, and vice versa. Thus, some volatility/smoothing combinations are omitted from Exhibit 1.[32] (Note that smaller values of the smoothing factor, α, imply more smoothing.)

Exhibit 1 reveals that, ceteris paribus, smaller appraisal error is associated with greater heterogeneity and greater true volatility at the individual property level. Also, other things being equal, smaller random error dispersion would imply less smoothing. It is interesting to note that some heterogeneity and smoothing combinations are mathematically impossible, as they would imply logically impossible negative error variance.

Exhibit 1
Mutually Consistent Values of True Heterogeneity and Disaggregate Appraisal Error Dispersion

Panel A: Random Disaggregate Appraisal Standard Error Assuming True Market Sector Volatility Equals 10%

Market Sector Heterogeneity Factor	Implied: True Property Volatility %	True Idiosyncratic Variance	Smoothing Factor Assumption (*alpha*): 1.00 — Implied S.D. of Random Disaggregate Appraisal Error: %	.67 %	.50 %
1.00	10.00	.0000	9.92	11.44	12.14
1.25	11.18	.0025	9.26	11.06	11.88
1.50	12.25	.0050	8.56	10.68	11.62
1.75	13.23	.0075	7.80	10.28	11.35
2.00	14.14	.0100	6.95	9.86	11.07
2.25	15.00	.0125	5.99	9.43	10.78
2.50	15.81	.0150	4.83	8.97	10.49
3.00	17.32	.0200	NA	7.99	9.87
3.50	18.71	.0250	NA	6.86	9.22
4.00	20.00	.0300	NA	5.50	8.51
4.50	21.21	.0350	NA	3.68	7.75
5.00	22.36	.0400	NA	NA	6.89
6.00	24.49	.0500	NA	NA	4.74

Panel B: Random Disaggregate Appraisal Standard Error Assuming True Market Sector Volatility Equals 12%

Market Sector Heterogeneity Factor	Implied: True Property Volatility %	True Idiosyncratic Variance	Smoothing Factor Assumption (*alpha*): .67 — Implied S.D. of Random Disaggregate Appraisal Error: %	.50 %	.33 %
1.00	12.00	.0000	11.44	12.14	12.81
1.25	13.42	.0036	10.90	11.77	12.58
1.50	14.70	.0072	10.33	11.38	12.34
1.75	15.87	.0108	9.73	10.98	12.10
2.00	16.97	.0144	9.08	10.56	11.85
2.25	18.00	.0180	8.39	10.12	11.60
2.50	18.97	.0216	7.64	9.67	11.34
3.00	20.78	.0288	5.86	8.69	10.80
3.50	22.45	.0360	3.19	7.58	10.24
4.00	24.00	.0432	NA	6.28	9.64
4.50	25.46	.0504	NA	4.64	9.00
5.00	26.83	.0576	NA	1.87	8.32
6.00	29.39	.0720	NA	NA	6.74

Panel C: Random Disaggregate Appraisal Standard Error Assuming True Market Sector Volatility Equals 14%

Market Sector Heterogeneity Factor	Implied: True Property Volatility %	True Idiosyncratic Variance	Smoothing Factor Assumption (*alpha*): .50 — Implied S.D. of Random Disaggregate Appraisal Error: %	.33 %	.25 %
1.00	14.00	.0000	12.14	12.81	13.12
1.25	15.65	.0049	11.63	12.49	12.88
1.50	17.15	.0098	11.09	12.17	12.64
1.75	18.52	.0147	10.52	11.83	12.40
2.00	19.80	.0196	9.92	11.48	12.15
2.25	21.00	.0245	9.29	11.13	11.89
2.50	22.14	.0294	8.60	10.76	11.63
3.00	24.25	.0392	7.03	9.98	11.09
3.50	26.19	.0490	5.00	9.13	10.53
4.00	28.00	.0588	0.70	8.20	9.93
4.50	29.70	.0686	NA	7.14	9.29
5.00	31.30	.0784	NA	5.90	8.61
6.00	34.29	.0980	NA	1.58	7.04

Notes: All figures are based on the Russell-NCREIF All-Properties 1980–92 Disaggregate Residual Variance of .019665. NA indicates impossible value – negative variance.
Smaller *alpha* implies more smoothing.
Source: Authors

The overall implication of Exhibit 1 is that very low values of disaggregate appraisal error dispersion (say, less than about 5% of property value) require rather special assumptions about the other parameter values. An assumption that appraisal error magnitudes (as represented by the root mean square error) have been on the order of 10% of property value (say, in the range of 6% to 13%), would appear to be quite robust and supported by a wide range of plausible values of the other relevant parameters.

Summary and Conclusions

The preceding analysis has attempted to estimate the average magnitude of the purely random component of disaggregate appraisal error among institutionally held commercial properties, using a unique database consisting of thousands of individual property appraisal-based returns. The analytical methodology, the database, and the definition of the type of error being measured, all differ to some extent from previous empirical studies that attempted to quantify disaggregate-level appraisal error, namely, the transaction-based studies of Miles, Webb, and others. Our analysis has required a number of assumptions and simplifications.[33] Nevertheless, we feel our results cast some light on the nature and magnitude of disaggregate appraisal error.

Our point estimate is that random disaggregate appraisal error in the Russell-NCREIF database during the 1980–92 period exhibited a root mean square value of 11.07% of the property value. This is similar to the 10% average absolute error found in the transaction-based studies of the same class of properties by Miles, Webb, and others. Thus, the "10% rule" used in REIT appraisal audits and ERISA would conform approximately to a single standard error rule, while the "20% rule" used by the FDIC would correspond to two standard errors.

We have also examined the implications that our findings hold regarding the degree of heterogeneity within the real estate asset class. By analyzing the relations among true volatility, appraisal smoothing, magnitude of random appraisal error, true real estate heterogeneity, and observed disaggregate return dispersion, we are able to arrive at mutually consistent values for all of these variables. This analysis supports a general conclusion that random disaggregate error in the Russell-NCREIF database has been in the range of 6% to 13% of property value.

Appendix
Algebraic Derivation of the Residual Dispersion Expansion Equation for the Russell-NCREIF Database

We begin with equation (5) from the main body of the text, the observed disaggregate appraisal-based "residuals," labelled $\varepsilon_t(p)$:

$$\varepsilon_t(p) = \alpha \sum_{i=0}^{\infty} (1-\alpha)^{i} e_{t-i}(p) + \sum_{i=0}^{\infty} (1-\alpha)^{i} \omega_{t-i}(p), \qquad (A.1)$$

where: $\omega_t(p) = \eta_t(p) - \eta_{t-1}(p)$.[a]

From these residual observations we can obtain an empirical estimate of the magnitude of the dispersion in random disaggregate appraisal error in the Russell-NCREIF database, as measured by $\text{VAR}[\eta_t(p)]$.[b] From (A.1), and assuming that the idiosyncratic return, $e_t(p)$, and the random component of appraisal error, $\eta_t(p)$, are uncorrelated across time, we see that the dispersion in the observable disaggregate appraisal-based residuals is given by:

$$\text{VAR}[\varepsilon_t(p)] = \alpha^2 \left(\sum_{i=0}^{\infty} (1-\alpha)^{2i} \text{VAR}[e_{t-i}(p)] \right) \qquad (A.2a)$$

$$+ \sum_{i=0}^{\infty} (1-\alpha)^{2i} \text{VAR}[\omega_{t-i}(p)]$$

$$+ 2 \sum_{i=0}^{\infty} (1-\alpha)^{2i+1} \text{COV}[\omega_{t-i}(p), \omega_{t-i-1}(p)],$$

which, after some algebra (and assuming stationarity, i.e., that the moments are constant across time), is seen to be equivalent to:[c]

$$\text{VAR}[\varepsilon_t(p)] = \left(\frac{\alpha}{2-\alpha}\right) \text{VAR}[e_t(p)] + \left(\frac{2}{2-\alpha}\right) \text{VAR}[\eta_t(p)], \qquad (A.2b)$$

which is equivalent to equation (6) in the text.

---

[a]Note that, by definition:

$$\text{VAR}[\omega_t(p)] = \text{VAR}[\eta_t(p)] + \text{VAR}[\eta_{t-1}(p)],$$

$$\text{COV}[\omega_t(p), \omega_{t-1}(p)] = -\text{VAR}[\eta_{t-1}(p)],$$

and $$\text{COV}[\omega_t(p), \omega_{t-i}(p)] = 0, \text{ for } i \geq 2.$$

[b]Our database is pooled, so that the sample second moment, $VAR[\eta_t(p)]$, is necessarily taken across both time and properties. This should not matter under the assumption that $e_t(p)$ and $\eta_t(p)$ are independent both across properties and across time. Data masking by NCREIF prevents the computation of pure time-series moments in our sample.

[c]Stationarity implies that:

$$VAR[\eta_{t-i}(p)] = VAR[\eta_{t-j}(p)], \text{ all } i \,\&\, j,$$

and similarly for the other moments. Note also that in this simplification we make use of the relationship between the moments of ω and the moments of η noted previously, and we use the formula for the sum of a geometric series, as follows:

$$VAR[\varepsilon_t(p)] = (\alpha^2 \sum_{i=0}^{\infty} (1-\alpha)^{2i})\, VAR[e_t(p)] + (2 \sum_{i=0}^{\infty} (1-\alpha)^{2i})\, VAR[\eta_t(p)]$$

$$-2 \sum_{i=0}^{\infty} (1-\alpha)^{2i+1}\, VAR[\eta_t(p)]$$

$$= (\alpha^2 \sum_{i=0}^{\infty} (1-\alpha)^{2i})\, VAR[e_t(p)] + 2(\sum_{i=0}^{\infty} (-1)^i (1-\alpha)^i)\, VAR[\eta_t(p)]$$

$$= \frac{\alpha^2}{1-(1-\alpha)^2}\, VAR[e_t(p)] + 2(\frac{1}{1+(1-\alpha)})\, VAR[\eta_t(p)].$$

Notes

[1]See, in particular, Blundell–Ward (1987), Fisher–Geltner–Webb (1994), Geltner (1989, 1991, 1993a), Miles–Cole–Guilkey (1990), Ross–Zisler (1991), Webb–Miles–Guilkey (1992).

[2]See, for example, Giaccotto–Clapp (1992), Quan–Quigley (1989, 1991), Vandell (1991).

[3]Giaccotto–Clapp (1992) and Geltner (1993a) have pointed out that the difference in the nature of the error, and in the purpose for which appraisals are used, suggests that different appraisal techniques would be optimal, depending on whether the appraisal is to be used in an aggregate index or for disaggregate, property-specific purposes.

[4]In gauging the reliability of individual appraisals both the smoothing component and the purely random error component are important. We focus on the purely random component in this paper because the smoothing component has already been studied extensively.

[5]Young and Graff (1994) observe that real estate investment risk data displays nonnormal distributional characteristics that may limit the effectiveness of standard portfolio diversification techniques in reducing idiosyncratic risk. In the present paper we avoid this complication by assuming that real estate investment is normally distributed. This implies the feasibility of assembling real estate portfolios with sufficiently many assets to diversify the idiosyncratic component of portfolio risk to negligible levels.

[6]NCREIF requires that all properties in the database be appraised at least once per year.

[7]These studies do not report the standard error, i.e., the standard deviation of the distribution of price differences. For normally distributed errors the mean absolute error is somewhat smaller than the standard error (square root of the average squared error). If the errors are normal, a mean absolute error of 10% would correspond to a standard error of about 12.5%.

[8]In highly liquid markets such as the stock and bond markets, the standard deviation of the instantaneous distribution of possible transaction prices for each asset is so small that any random sample from the distribution almost always differs negligibly from the distribution mean. Thus, as a practical matter, the notions of "transaction price" and "market value" may be assumed to coincide in the case of securities, and economists typically make this identification without comment.

[9]A similar rule is included in the Employee Retirement Income Security Act (ERISA). In addition, Johnson (1992) reports that the FDIC also has a similar policy. All properties acquired by the FDIC are appraised by an independent fee appraiser as a matter of policy at the time of acquisition. The agency's policy is to hire two appraisers to independently appraise properties which are estimated by the FDIC to be worth more than $1 million. If the two appraisals vary by more than 20%, the policy is to obtain a third appraisal (Johnson, 1992, 36–37).

[10]Webb (1994) reports that out of a total usable transaction sample of 569 properties, 28 had appraised values *exactly* equal to their subsequent transaction prices, with a further 71 properties appraised at within 1% of the subsequent transaction price.

[11]In the extreme case, suppose both the appraisal error and the transaction noise have a standard error of 10% around the true market value, but the appraisal error and the transaction noise are perfectly correlated. Then the observed difference between the appraised value and the transaction price will always be zero, with no dispersion (as in the case where the appraiser always knows, at the time of the appraisal, the exact price at which the property will subsequently transact, and the appraiser uses that price as the appraised value).

[12]The Russell-NCREIF disaggregate returns are made available to the public for research purposes in a masked form, so that it is impossible to identify individual properties, and hence to trace out individual property returns across time. While the authors have access to the individual property returns each year, we have no way to know which returns in one year go with which individual returns in another year. Thus, the disaggregate database is necessarily pooled, so that cross-sectional and time-series elements cannot be separated.

[13]$r^*_t(p) = V^*_t(p) - V^*_{t-1}(p)$; $r_t(p) = V_t(p) - V_{t-1}(p)$.

[14]See, e.g., Geltner (1989, 1993a), Giaccotto-Clapp (1992), Quan-Quigley (1989, 1991). The appraisal error and smoothing model in equation (1) has been used and discussed in all of these previous articles. Note that smoothing observed in aggregate-level appraisal-based returns comes from temporal aggregation in the construction of the aggregate index as well as from disaggregate-level appraisal smoothing (Geltner, 1993b). Thus, the smoothing factor referred to here, α, would not generally be the same as what has been found or applied in analyses of aggregate-level appraisal smoothing, such as Ross-Zisler (1991) or Fisher–Geltner–Webb (1994).

[15]In our case, data availability limits the market sectors to Office, Retail, Warehouse, and R&D properties.

[16]Note that by the absence of a "*beta*" parameter multiplying m_t in equation (2), we effectively assume that all properties in the market sector have the same expected return and are of the same systematic risk class (i.e., respond with the same sensitivity to macroeconomic and market sector news). This is a simplification of reality made for analytical tractability. To the extent that properties have heterogeneous "betas" within property-type sectors, some of the disaggregate return dispersion we observe and attribute to appraisal error would in fact be due to "*beta*-heterogeneity" rather than to appraisal error, leading our procedure to overestimate the magnitude of random appraisal error. This problem could be avoided if the Russell-NCREIF database were sufficiently "unmasked" to allow individual property returns to be identified across time.

[17]The aggregate return will also contain some temporal aggregation smoothing, due to appraisals being staggered throughout the year. Some of the dispersion in our database will therefore result from the staggered timing of the appraisals rather than from appraisal error,

causing us to overestimate the magnitude of typical appraisal error. However, the bias introduced by this effect should be minor. Given the likely order of magnitude of aggregate volatility (on the order of 10% per year), and assuming that random appraisal error is independent of appraisal timing, the upward bias in our appraisal error dispersion estimate should be less than one percentage point.

[18]See the Appendix for the algebraic derivation of this result.

[19]In theory this is a cross-sectional moment, though in our empirical analysis we are using a pooled database so the dispersion is observed across both time and properties.

[20]Geltner (1993a) uses the value of $\alpha = 1/2$ in an unsmoothing model which, after also accounting for aggregation effects in the index, produces a simulated historical value index of commercial property which agrees broadly with evidence from REIT prices and the perceptions of market participants. The index has annual volatility just under 9%, about half that of the S&P500.

[21]In fact, we explore a broader range of possible smoothing values in our analysis of heterogeneity in the next section.

[22]See, for example, Ross–Zisler (1991), Geltner (1993a), Fisher–Geltner–Webb (1994).

[23]The rationality of this estimate of disaggregate property true volatility is suggested by Ciochetti's (1994) empirical analysis of commercial mortgage pricing using an option-based valuation model. Ciochetti's technique allows one to "back out" the disaggregate volatility implied by the option model of mortgage value (yield spread). He finds implied volatilities in the range of 17%–19%. However, his sample of mortgages includes a few loans backed by raw land, hotel properties and some projects still in the lease-up phase, all of which would be expected to be more volatile than the Russell-NCREIF properties. Also, Ciochetti's mortgage valuation model does not correct for the effect of illiquidity in commercial loans, which would cause his implied volatilities to be biased on the high side.

[24]In other words, we are saying, in the context of equation (2), that a reasonable estimate would be as follows:

$$\mathrm{VAR}[r_t(p)] = \mathrm{VAR}[m_t] + \mathrm{VAR}[e_t(p)] = .015 + .01 = .025\,,$$

implying an average true disaggregate volatility of $\sqrt{.025} = 16\%$.

[25]Again, an even broader range is effectively explored in the next section.

[26]The Russell-NCREIF database is managed by the Frank Russell Company, Tacoma, Washington, for the National Council of Real Estate Investment Fiduciaries (NCREIF). There are currently nearly 2000 properties in the database, with an aggregate equity value on the order of $40 billion. We have used the (deleveraged) leveraged property database, as well as the unleveraged properties of the traditional Russell-NCREIF Property Index. Note that our database has been purged of properties that were either acquired or sold during the calendar year, so it represents a pure appraisal value sample. The masked, disaggregate returns data used in this paper are available to the public from Frank Russell Company for a nominal charge of $100.

[27]The data is available quarterly. However, most of the properties are not reappraised every quarter, so we are working with annual returns.

[28]In this context, the term "disaggregate" refers to individual asset (e.g., individual stock or individual property), and the term "aggregate" refers to an equally weighted portfolio or index consisting of all the assets in the asset class or index. The variance referred to is the time-series sample second moment, that is, taken across time. Thus, the "heterogeneity ratio" is defined on historical sample moments. As aggregate variance equals the average covariance among the individual assets (including each asset with itself), the inverse of the heterogeneity ratio approximately equals the average historical cross-correlation coefficient among pairs of assets in the asset class. Thus, for example, a heterogeneity ratio of 2.5 corresponds to an average correlation coefficient among pairs of individual assets equal to 40%.

[29]Other studies are by Miles-McCue (1984) and Hartzell–Shulman–Wurtzebach (1987). All these studies use quarterly appraisal-based returns uncorrected for error.

[30]This is the average of their findings of 5.05 for industrial properties, 3.99 for office, and 3.08 for retail property. Note that there are typographical errors in the Hartzell–Hekman–Miles article. Their Table 12 on page 250, labelled "Quarterly" is actually "Annual," and their footnote to Table 11 on page 249 which says variances are multiplied by 10^3, actually means variances are multiplied by 10^4.

[31]$E_p[\mathrm{VAR}[\mathcal{E}_t(p)]]$ is thus the average disaggregate residual dispersion observed in the Russell-NCREIF database, and α is the disaggregate appraisal smoothing factor. Since our database is pooled, the sample moments we observe are already averaged across properties, so that $E_p[\mathrm{VAR}[...]] = \mathrm{VAR}[...]$ in our sample. Thus, equation (10) applies directly to our observed residual variance.

[32]Without correcting for smoothing, the Russell-NCREIF property-type subindices had an historical annual volatility of 7.46% during the 1980–92 period. Thus, the assumption of true aggregate volatility equal to 10% in Panel A effectively assumes very little smoothing.

[33]Some of these could be avoided if the masking of the disaggregate Russell-NCREIF database were relaxed enough so that individual property returns could be traced across time.

References

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

——, The Accuracy of Valuations – Expectation and Reality, Working paper, Department of Land Management and Development, University of Reading, April 1993.

Ciochetti, B., Credit Risk Diversification in Commercial Mortgages, Ph.D. dissertation, Department of Real Estate, University of Wisconsin, 1994.

Damodaran, A. and C. Liu, Insider Trading as a Signal of Private Information, *Review of Financial Studies*, Spring 1993, 6:1, 79–19.

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 & Economics*, 1994, 9:2, 137–64.

Geltner, D., Estimating Real Estate's Systematic Risk from Aggregate Level Appraisal-Based Returns, *AREUEA Journal*, Winter 1989, 17:4, 463–81.

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

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

——, Temporal Aggregation in Real Estate Return Indices, *Journal of the American Real Estate and Urban Economics Association*, Summer 1993b, 21:2, 141–66.

Giaccotto, C. and J. Clapp, Appraisal-Based Real Estate Returns Under Alternative Market Regimes, *Journal of the American Real Estate and Urban Economics Association*, Spring 1992, 20:1, 1–24.

Hartzell, D., J. Hekman and M. Miles, Diversification Categories in Investment Real Estate, *AREUEA Journal*, Summer 1986, 14:2, 230–54.

Hartzell, D., D. Shulman and C. Wurtzebach, Refining the Analysis of Regional Diversification for Income-Producing Real Estate, *Journal of Real Estate Research*, Winter 1987, 2:2, 85–96.

Johnson, D., Accuracy in Commercial & Industrial Real Estate Appraisals, Ph.D. dissertation, Department of Real Estate and Business Law, University of Georgia, Athens, Georgia, 1992.

McEnally, R. and C. Boardman, Aspects of Corporate Bond Portfolio Diversification, *Journal of Financial Research*, Spring 1979, 27–36.

Miles, M. and T. McCue, Diversification Within the Real Estate Portfolio, *Journal of Financial Research*, Spring 1984, 7:1, 57–67.

Miles, M., R. Cole and D. Guilkey, A Different Look at Commercial Real Estate Returns, *AREUEA Journal*, Winter 1990, 18:4, 403–30.

Miles, M., D. Guilkey, B. Webb and K. Hunter, An Empirical Evaluation of the Reliability of Commercial Appraisals, NCREIF Research Paper, 1991.

Quan, D. and J. Quigley, Inferring an Investment Return Series for Real Estate from Observations on Sales, *AREUEA Journal*, Summer 1989, 17:2, 218–30.

——, Price Formation and the Appraisal Function in Real Estate Markets, *Journal of Real Estate Finance & Economics*, 1991, 4:2, 127–46.

Ross, S. and R. Zisler, Risk and Return in Real Estate, *Journal of Real Estate Finance & Economics*, 1991, 4:2, 175–90.

Vandell, K., Optimal Comparable Selection and Weighting in Real Property Valuation, *AREUEA Journal*, Summer 1991, 19:2, 213–39.

Webb, R. B., M. Miles and D. Guilkey, Transactions-Driven Commercial Real Estate Returns, *Journal of the American Real Estate and Urban Economics Association*, Summer 1992, 20:2, 325–57.

Webb, R. B., On the Reliability of Commercial Appraisals: An Analysis of Properties Sold from the Russell-NCREIF Index (1978–1992), *Real Estate Finance*, Spring 1994, 11:1, 62–65.

Young, M. and R. Graff, Real Estate Is Not Normal: A Fresh Look at Real Estate Return Distributions, *Journal of Real Estate Finance & Economics*, forthcoming 1994.