The Magnitude of Random Appraisal Error in Commercial Real Estate Valuation

Richard A. Graff and Michael S. Young***

Abstract. **Analysis of more than seven hundred pairs of simultaneous independent appraisals of institutional-grade commercial properties shows that the standard deviation of the random component of appraisal error is approximately 2%. Random appraisal error appears constant across both time and the institutional-grade investment universe, except during infrequent periods of real estate market gridlock. Most appraisal error is deterministic in nature, even though it usually appears random in routine cross-sectional analysis. Such appraisal error can be constrained and reduced by investment management control systems.**

Introduction

A decade ago, real estate researchers began to consider the possibility that investment statistics derived from appraisal-based commercial real estate valuations might be much less satisfactory measures of real estate investment performance than corresponding statistics derived from transactions in the United States stock and bond markets. The main source for concern was considered to be random appraisal error, since earlier studies of appraisal error in the residential housing market had suggested that typical appraisal error is at least 10% of asset capitalization and that most residential appraisal error is random.[1]

Since researchers did not have access to large numbers of commercial real estate appraisals to aid in formulating and testing hypotheses about appraisal valuation behavior, empirical investigation of commercial real estate appraisal error sources and their effect on real estate investment statistics was not possible. Accordingly, researchers used accepted beliefs about residential real estate appraisal accuracy and aggregate National Council of Real Estate Investment Fiduciaries (NCREIF) statistics to justify assumptions about the magnitude and random behavior of commercial real estate appraisal error.[2]

Unfortunately, in the case of institutional-grade real estate these assumptions neglect two major differences between commercial property and residential housing:

This article is the winner of the Real Estate Valuation category (sponsored by the Appraisal Institute) presented at the 1998 American Real Estate Society Fourteenth Annual Meeting.

*Electrum Partners, Chicago, IL 60611.
**The RREEF Funds, San Francisco, CA 94111 or MYoung@RREEF.com.

institutions purchase property primarily for cash flow and secondarily for capital appreciation, so the hedonic measures institutions apply to value commercial property are more uniform than the hedonic measures by which homeowners value residential housing; and institutional real estate managers usually make real estate investment decisions on behalf of institutional investors, raising the possibility of agency-based (nonrandom) contributions to appraisal error usually absent in the case of residential housing.[3]

This suggests two likely consequences: commercial real estate appraisal error contains both random and nonrandom components, and the average magnitude of random appraisal error in the case of commercial real estate is smaller than the average magnitude of random residential housing appraisal error. It follows that any description of the effect of appraisal error on investment statistics should be more complicated than suggested by previous studies, since random and nonrandom appraisal error components affect sample means, variances and correlations in very different ways.

In the absence of access to a fairly large database of commercial real estate appraisals, it has been impossible for most researchers to determine whether the magnitude of random appraisal error is large enough to have a material impact on sample real estate investment statistics. However, even if such access were available, it would only enable researchers to estimate the magnitude of total appraisal error. Access to appraisal data alone would not enable researchers to compare the relative importance of contributions from random and nonrandom error components.

The magnitude of random appraisal error can be determined empirically if an appraisal database for institutional-grade commercial real estate can be located that includes *at least two simultaneous independent appraisals* whenever an appraisal valuation is updated. One may reasonably conjecture in the case of such data that the deterministic components of appraisal error are virtually identical for each set of simultaneous appraisals of each asset.[4] If this conjecture or something similar can be tested and verified, then it follows that the nonrandom components of the appraisals in each set of simultaneous appraisals are identical. This implies that the sample standard deviation for each set of simultaneous appraisals is a sample standard deviation for the random appraisal error component.

A database of paired simultaneous independent appraisals does exist. Since 1989, The RREEF Funds has assigned internal personnel to conduct a simultaneous independent appraisal of each real estate asset whenever an outside appraiser is retained to conduct a full asset appraisal. Although outside appraisals are no longer conducted as frequently as they were during the era of closed-end and open-end funds, this still provides us with a database of 747 pairs of simultaneous independent asset appraisals.[5]

This study tests a slightly more complex version of the hypothesis that the difference between simultaneous internal and external appraisals is an estimator for random appraisal error. The test confirms the hypothesis in a subsequent section, and the analysis is extended to derive a consistent estimator for the standard deviation of random appraisal error. Finally, the standard deviation for random appraisal error is combined with empirical estimates of total appraisal error from previous studies to measure the average magnitude of nonrandom commercial real estate appraisal error.

Related Research

In the case of residential appraisals, Dotzour (1988a,b) compares appraisals and purchase prices for more than 500 residential properties acquired by corporate relocation companies as part of corporate employee relocations.[6] Dotzour (1988a) shows that the mean absolute difference between appraised value and purchase price is 2.77% for the subsample in which appraisals were performed by professionally designated appraisers, and 5.13% for the subsample in which appraisals were performed by nondesignated appraisers.[7] Since random appraisal error can be viewed as a component of the difference between appraised value and purchase price, this suggests that mean random appraisal error for these subsamples is no larger than the reported mean absolute difference between appraised value and purchase price.

In the case of commercial real estate appraisals, recent studies implicitly cast light on the question of random appraisal error in the course of examining related questions. In other words, random appraisal error is not addressed specifically, but significant conclusions about random appraisal error follow immediately for anyone who wishes to extend the analyses in this direction. Of particular interest in this regard are Diaz (1997) and Diaz and Wolverton (1998). These studies employ a simultaneous appraisal methodology similar to the one in this study, although each study establishes a controlled appraisal environment for a single representative property in which far more simultaneous appraisals are available for each property than are available in the case of the present study.

Diaz (1997) presents the results of a controlled experiment that, among other things, examines thirty simultaneous expert appraisals of a single parcel of vacant prepared industrial-zoned land in the northern Atlanta suburbs. The appraisals are divided into two subsets of fifteen samples according to the following criterion: appraisals in the first subset were conducted by expert local appraisers with knowledge of the previous asset appraisal, whereas appraisals in the second subset were conducted by expert local appraisers who were denied access to previous appraisal results.

Appraisal means, medians and standard deviations for each subset are presented in Diaz (1997, Exhibit 2). Although these appraisal statistics are presented in absolute terms (*i.e.*, dollars per acre), each sample standard deviation can be converted to a relative measurement by dividing the absolute sample standard deviation by the corresponding sample mean. This yields 2.67% for one fifteen-sample subset and 2.61% for the other fifteen-sample subset.

Of particular significance for the present study, these two sample standard deviations virtually coincide, which strongly suggests that the two appraisal subclasses have the same true standard deviation for relative random appraisal error.[8] Accordingly, the sample standard deviations can be combined by root-mean-square summation to produce the following best estimate for relative appraisal error based on the twenty-eight degrees of freedom in the Diaz data set: $s = ((2.67\%)^2 + (2.61\%)^2)^{1/2} \cong 2.64\%$.[9] This estimate is remarkably close to the value obtained in the present study from a data set with approximately twenty times as many degrees of freedom.

The follow-on Diaz and Wolverton (1998) study examines, among other things, two sets of simultaneous expert appraisals of a Phoenix, Arizona apartment complex in which the groups of appraisals were conducted eight months apart. The first set contains sixteen samples, and the second set contains fifteen samples. In each case, appraisers were denied knowledge of previous appraisals of the subject asset. These samples collectively comprise the ''unanchored'' case in the study.

Sample appraisals are listed in absolute terms in Diaz and Wolverton (1998, Table 1). It is straightforward to derive sample standard deviations in relative terms for each set of simultaneous appraisals by computing sample standard deviations for the natural logarithms of the appraisals rather than for the appraisals themselves. This yields 5.32% for the standard deviation of the sixteen-sample set and 5.06% for the standard deviation of the fifteen-sample set.

The corresponding sample means imply with virtual certainty that the true means for the two cases are not identical.[10] By contrast, the proximity of the sample standard deviations suggests once again that the corresponding true standard deviations of relative random appraisal error are identical.[11] Accordingly, the sample standard deviations can be combined by weighted root-mean-square summation to produce the following best estimate for relative appraisal error based on the 29 degrees of freedom in the Diaz and Wolverton unanchored appraisal data set: $s = ((15*(5.32\%)^2 + 14*(5.06\%)^2)/29)^{1/2} \cong 5.20\%$.[12]

The Diaz and Wolverton study also lists updates of fifteen of the sixteen appraisals in the first sample set that were conducted at the same time as the appraisals in the second set. These appraisal updates were conducted by fifteen appraisers involved in the first group of appraisals, and comprise the ''anchored'' case of the study. The study shows that the first appraisals were a statistically significant psychological ''anchor'' for near-term valuations that prevented the *same* appraisers from fully responding to changes in true asset value. This confirms the longstanding hypothesis that appraisal anchoring is a statistically significant problem with respect to near-term reappraisals by the same appraisers, and suggests that institutional investors should strongly consider the establishment of valuation policy controls to minimize any anchoring effects.

Interestingly from the perspective of the present study, the sample standard deviation of relative appraisal error for appraisal updates in the Diaz and Wolverton study is 6.90%. This is substantially larger than the 5.20% sample standard deviation for relative appraisal error in the case of the thirty-one original (*i.e.*, unanchored) appraisals, and is on the edge of being statistically distinguishable from the unanchored value.[13] In fact, the 90% confidence interval for the true standard deviation of relative appraisal error conditioned on the 5.20% sample standard deviation based on 29 degrees of freedom is (4.29%, 6.65%).[14]

Random appraisal error estimated from the Diaz and Wolverton (1998) data appears significantly larger than appraisal error estimated either from the Diaz (1997) data or from data in the present study. There are two apparent explanations for larger appraisal error in the case of the Diaz and Wolverton study. First, the subject asset of the Diaz and Wolverton study is in Phoenix although the appraisers practice in Atlanta, whereas appraisers in the cases of Diaz (1997) and the present study are active in the markets that contain the subject assets. Second, apartment property may be more difficult to appraise accurately than office, industrial and retail property, due to shorter average leases and less creditworthy tenants. This suggests additional experiments of the Diaz and Wolverton type to determine whether appraisal expertise is immediately and fully transferable across geographical regions, and whether major differences in average lease maturity or credit quality have any effect on the magnitude of random appraisal error.[15]

Significantly larger estimates of the magnitude of random appraisal error appear in several other studies, including one conducted in part by the authors of the present study. For example, Geltner, Graff and Young (1994) derives an algebraic model that relates variances for three types of unobservable random noise connected with real estate investment returns, and also hypothesizes a range of reasonable values for the unknown parameters in the model. These unknown parameters include the standard deviation for random appraisal error. Similarly, Geltner and Goetzmann (1998) derive an error estimate for the total magnitude of several types of appraisal error as an incidental result of a repeated-measures regression study of appraisal-based index volatility. Then the study applies a second-stage regression to separate random appraisal error from the other appraisal error components. Both studies apply their respective models to NCREIF appraisal-based returns. Geltner (1998) also addresses random appraisal error, but relies primarily on Geltner, Graff and Young (1994) for quantitative support.

Geltner, Graff and Young (1994) and Geltner and Goetzmann (1998) both require disaggregated NCREIF appraisal-based return series to be serially independent and identically distributed in order for their conclusions about random appraisal error to be supportable. Although analyses in these studies could conceivably be robust with respect to small nonstationarities in return series, major nonstationarities (such as result from real estate investment cycles) in the return series would invalidate the quantitative appraisal error conclusions in both studies.[16] In addition, the analysis in Geltner, Graff and Young (1994) depends upon the restrictive assumption common to early appraisal error studies that all appraisal error is unbiased and random.

Young and Graff (1995) shows empirically that disaggregated NCREIF appraisal-based returns are extremely nonstationary. The investment characteristics of real estate appraisal-based returns observed in Young and Graff are also observed in Graff, Harrington and Young (1997) for appraisal-based returns from the Australian real estate market compiled by the Property Council of Australia, suggesting that investment return nonstationarity is a ubiquitous feature of commercial real estate markets.

This does not imply that the results of Geltner, Graff and Young (1994), Geltner and Goetzmann (1998) and Geltner (1998) are without interest or significance. However, since the results of these studies depend critically on overly restrictive assumptions about stationarity in investment returns and randomness in total appraisal error, they cannot be accorded the same credence as studies such as Diaz (1997), Diaz and Wolverton (1998) and the present study, in which analyses are largely nonparametric and are consequently unaffected by time dependencies in investment returns and nonrandom components in appraisal error.

Finally, three recent studies relate to the present study by showing empirically that nonrandom appraisal error components affect commercial real estate valuations and investment returns, and that agency costs are probable sources for the nonrandom error. First, Hendershott and Kane (1995) examines the office property capital gains component of the Russell-NCREIF Property Index and demonstrates systematic appraisal overvaluation of institutional-grade office properties throughout the second half of the 1980s. The study attributes this nonrandom appraisal error to two sources: an imputed reluctance on the part of appraisers to recognize sudden large valuation changes, and a certain type of agency cost—more precisely, investment manager pressure to maintain high property valuations in order to avoid reductions in percentage-based investment management fees. Next, Graff and Webb (1997) presents evidence that agency costs embedded in transaction prices introduce nonrandom components into appraisal error that generate statistically significant performance persistence in NCREIF annual appraisal-based investment returns. The Graff and Webb study also suggests how excessive agency costs can be detected and eliminated by appropriately structured management control systems. Finally, Wolverton and Gallimore (1998) examines commercial mortgage lending and demonstrates that client feedback exerts statistically significant material effects on appraisal valuation just prior to pending sales, and that the effects can become coercive in extent.[17]

Data Description

Data for this study consist of 747 pairs of simultaneous independent valuations of commercial office, retail, industrial and apartment properties during the 1989–1997 period.

In the case of each set of paired valuations, one is an internal valuation conducted by RREEF portfolio managers and research analysts, and the other is a full appraisal conducted by an external third-party fee appraiser hired either directly by the relevant RREEF client (in the case of properties in designated separate accounts) or else by RREEF.

In every case, internal RREEF staffers and outside appraisers are given the same factual information about property operating expenses, budgets, financial statements and rent rolls. In order to minimize appraisal anchoring effects, outside appraisers are not informed of prior valuations, either internal or external. However, since outside appraisers sometimes value the same property over several years, individual outside appraisers may have knowledge of prior appraisals. The RREEF database does not contain information that would allow us to assess the incidence of occurrences in which outside appraisers do in fact know the previous valuations. Internal appraisers can access prior valuations if they are so inclined.

Since inside and outside valuations of each property take place at the same time, inside valuations are completed before the corresponding outside valuations in approximately one-half of the cases. Thus, about one-half of the time inside appraisers are aware of the current outside valuation before the inside valuation is completed.

In determining property value for reporting or fee assessment purposes, RREEF generally uses the lower of the two appraisals. More precisely, RREEF accepts the lower of the two appraisals unless the RREEF valuation is less than 5% lower than the external appraisal, in which case the outside appraisal is used.

The desire for an independent check on outside appraisals first arose in 1987 when RREEF noticed weakness in the re-leasing market for oil patch properties that RREEF perceived as a warning signal of declining values. RREEF discovered that outside appraisers were unwilling to attribute as much significance to the impact of this development on oil patch property values as RREEF believed to be appropriate. Concern that systematic appraisal overvaluation of oil patch property would result in ultimately unjustifiable portfolio fees led to the creation of portfolio management and research groups and the accompanying system of internal appraisal checks. Also, RREEF changed the cycle of property appraisals in its closed-end real estate funds from fourth-quarter valuations to a more evenly balanced distribution across all four calendar quarters.

Methodology

This objective of this study is to examine the contribution of random appraisal error to uncertainty in *relative* asset value. Accordingly, each appraisal sample in the analysis is the natural logarithm of actual appraised asset value, and *appraisal error* is defined to be the difference between the natural logarithms of appraised asset value and true asset value. Expressed symbolically:

$$\ln(V^*) = \ln(V) + \delta, \tag{1}$$

where V^* is an appraisal value for an asset, V is the true value of the asset as of the appraisal date and δ is the (total relative) appraisal error.[18]

The assumption that the probability distribution for random appraisal error reflects epistemological limitations on the accuracy of appraisal methodology suggests that the probability distribution for relative random appraisal error may be constant across both time and the universe of real estate assets in the case of full appraisals by professionally designated appraisers.[19] Accordingly, our null hypothesis tests this possibility. More precisely, our null hypothesis consists of the hypothesis that (relative) appraisal error may be represented as the sum of two components: a *random* sample from a normal distribution with zero mean that is constant across both time and the institutional-grade real estate universe, and a *deterministic* component that is constant across simultaneous appraisals for individual assets but that can vary across both time and the real estate asset universe.[20]

The null hypothesis is summarized by the following equation:

$$\ln(V^*(p,t,n)) = \ln(V(p,t)) + \varepsilon(p,t,n) + \eta(p,t), \tag{2}$$

where $V(p,t)$ is the true value of property p at time t, $V^*(p,t,n)$ is the appraisal value for property p at time t derived by appraiser n, $\varepsilon(p,t,n)$ is the random appraisal error component and $\eta(p,t)$ is the deterministic/nonrandom appraisal error component. Note that Equation (2) is essentially the same as Equation (1), except that total appraisal error $\delta(p,t,n)$ has been decomposed into $\varepsilon(p,t,n) + \eta(p,t)$ and restrictions have been imposed on the form of total appraisal error via restrictions on each of the components.[21]

The deterministic component $\eta(p,t)$ reflects a panoply of asset-specific and time-varying effects relating to agency costs. Consequently, the distribution of the function values of the deterministic component appears consistent with the distribution of a random variable when subjected to cross-sectional analysis.[22]

The assumption that $\varepsilon(p,t,n)$ is a random sample from a probability distribution that does not vary with time, property or appraiser implies that $\sigma(\varepsilon(p,t,n)) \equiv \sigma(\varepsilon)$ is the same for all appraisals. Together with the assumption that simultaneous appraisals of the same asset have the same deterministic component, this implies that a set of sample standard deviations for simultaneous appraisals of individual assets constitutes a set of sample standard deviations for random appraisal error.[23]

Although each sample standard deviation is based only on the information available from simultaneous appraisals of a single asset, an estimate for the standard deviation of random appraisal error that incorporates information available from all appraisals in the data set can be obtained from the distribution of the sample standard deviations.[24]

If the null hypothesis is not correct, the essential aspects of this analysis may remain valid with appropriate modifications. For example, if the deterministic components of simultaneous appraisals of individual assets are not identical, then each expected sample variance for simultaneous appraisals of the same asset is an upper bound for the variance of random appraisal error. With a corresponding modification in the interpretation of 'sample standard deviation' for each appraisal pair, the analysis in this study can be shown to generate a consistent estimator for an upper bound for the standard deviation of random appraisal error.

The null hypothesis can be tested in two ways. First, if the null hypothesis is correct then the true probability is exactly 50% that each external appraisal of RREEF-managed assets is greater than the corresponding simultaneous internal RREEF appraisal. Furthermore, the null hypothesis implies that the set of differences between simultaneous paired appraisals constitutes a set of independent samples. Consequently, if the null hypothesis is correct then it follows that the incidence of samples in which the external appraisal is greater than the corresponding internal appraisal is binomially distributed around 50%.[25]

Second, if the null hypothesis is correct in assuming the magnitude of random appraisal error to be independent of time, asset and appraiser, then the distribution of standard deviations for pairs of simultaneous asset appraisals should be substantially unchanged for selected subsets of the total sample set. Accordingly, the distribution of sample standard deviations for appraisal pairs from each individual year in the test interval is examined separately for evidence of time-varying behavior.[26]

It is also useful to recall that some investigators assert that there are differences between investment return behavior for large-capitalization and small-capitalization real estate assets. Since the midpoint of the appraisal set is $10.67 million, the assertions about differences between large-capitalization and small-capitalization data and the independence of random appraisal error can be tested jointly by dividing the data into above-median-capitalization and below-median-capitalization subsets (*i.e.*, large-capitalization and small-capitalization subsets) and performing the tests on each subset.

Empirical Results

Selected descriptive characteristics of the sample distribution for the entire sample set and several subsets are presented in Exhibit 1. The incidence of paired appraisal samples in which external appraisals exceed internal appraisals is presented in Exhibit 2.

Exhibit 1 reveals year-by-year variations in sample standard deviation means and percentile levels that do not support the null hypothesis. The ratio of maximum-to-minimum sample standard deviations for fixed year-by-year percentile levels varies from more than three-to-one for median standard deviations to nearly six-to-one for 90th percentile standard deviations, which does not appear particularly time-independent. Even worse, the aggregated data have a median sample standard deviation of only 1.69% plus at least four sample standard deviations above 22%, which cannot be considered consistent with time-independent sample standard deviations for any reasonable distribution.

When data for 1991–1992 are omitted, the ratio of maximum-to-minimum sample standard deviations for fixed annual percentile levels declines in every case to a value of approximately two-to-one, which is more nearly consistent with the assumption that the true standard deviation for random appraisal error is time-independent.[27]

Exhibit 1
Percentile Levels, Minima, Maxima and Means for Distributions of Sample Standard Deviations for Simultaneous Asset Appraisals

Year(s) and/or Cap Group	Min. (%)	Median (%)	Mean (%)	75th Percentile (%)	90th Percentile (%)
1989	0.23	1.02	1.88	2.25	5.22
1990	0.00	2.08	3.29	4.48	6.97
1991	0.00	3.63	5.98	7.22	13.07
1992	0.00	3.71	7.33	9.61	20.34
1993	0.00	1.69	2.96	3.60	8.06
1994	0.00	1.55	2.60	2.91	4.46
1995	0.00	1.03	1.42	2.12	3.52
1996	0.00	1.08	2.08	3.36	5.45
1997	0.00	1.30	1.96	2.68	5.15
All Years, 1989–1997	0.00	1.69	3.26	3.63	7.17
All Years Except 1991–1992	0.00	1.35	2.27	2.99	5.27
Capitalization $>$ Median, All Years, 1989–1997	0.00	1.35	2.99	3.37	6.60
Capitalization $>$ Median, All Years Except 1991–1992	0.00	0.99	2.13	2.69	5.02
Capitalization $\leq$ Median, All Years, 1989–1997	0.00	1.99	3.53	3.94	7.57
Capitalization $\leq$ Median, All Years Except 1991–1992	0.00	1.59	2.41	3.22	5.26

Exhibit 2
Incidence of Internal Appraisals That Are Less Than External Appraisals

Year(s) and/or Cap Group	Number of Samples	Internal < External for All Samples (%)	Internal < External for $s \leq$ Median (%)	Inte for (%)
1989	16	50.0	62.5	37.5
1990	53	64.2	44.4	84.6
1991	84	79.8****	69.0*	90.5
1992	85	80.0****	67.4*	92.9
1993	81	74.1***	68.3*	80.0
1994	114	66.7***	64.9*	68.4
1995	103	61.2*	51.9	70.6
1996	87	50.6	25.0**	76.7
1997	124	47.6	38.7	56.5
All Years, 1989–1997	747	64.1*****	54.0	74.3
All Years Except 1991–1992	578	59.5***	49.8	69.2
Capitalization > Median, All Years, 1989–1997	373	61.1***	46.2	75.9
Capitalization > Median, All Years Except 1991–1992	289	56.4*	47.2	65.5
Capitalization ≤ Median, All Years, 1989–1997	374	67.1*****	56.1	78.1
Capitalization ≤ Median, All Years Except 1991–1992	289	62.6***	55.6	69.7

*Null hypothesis rejected at the 5% significance level.
**Null hypothesis rejected at the 1% significance level.
***Null hypothesis rejected at the 0.1% significance level.
****Null hypothesis rejected at the 0.00001% significance level.
*****Null hypothesis rejected at the 0.0000001% significance level.

However, the aggregated data again have a median sample standard deviation of only 1.35% plus at least two sample standard deviations above 22%, which still is not consistent with time-independent sample standard deviations.

Exhibit 2 is equally discouraging at first glance for the null hypothesis. The null hypothesis implies a probability of 50% for the proposition that the difference between external and internal appraisals is positive. Unfortunately, the exhibit shows that the test rejects the proposition for five of the nine years in the test interval, rejects the proposition in all six cases of aggregated annual data, and rejects the proposition definitively at the 0.1% level or lower in nine of the fifteen cases examined.

The evidence is more encouraging when the appraisal pairs are separated into subsets of above-median and below-median sample standard deviations. Exhibit 2 shows that test rejections of the proposition can be explained completely by the contribution to the test statistic from paired appraisals with sample standard deviations above the median value. In the case of appraisal pairs with sample standard deviations at or below the median sample standard deviation, the test statistic is consistent with the null hypothesis in every case involving aggregated data.[28]

This suggests that the null hypothesis is consistent with the data for properties that are relatively uncomplicated to appraise. This suggests in turn that the null hypothesis could well be correct for most institutional-grade properties under most circumstances, but that there is some exceptional agency effect that is clouding the RREEF data for properties that have greater-than-average investment risk.[29]

An explanation consistent with this suggestion is apparent when we recall the original RREEF concern that motivated the introduction of internal appraisals. As discussed earlier, RREEF management was concerned that downside risk was not being fully factored into independent asset appraisals, and that this could generate valuation-based management fees that might ultimately appear unjustified by clients if projected worst-case investment scenarios turned out to be correct. RREEF staffers who perform the internal appraisals are certainly aware of the motivation for internal appraisals. Although encouraged toward objectivity in their valuations, they can also be hypothesized to display a tendency toward conservatism in cases that represent unusually large investment risk, a tendency that would not be expected from corresponding external appraisers.[30]

A related source for upward bias in external appraisals relative to internal appraisals under some economic scenarios is reluctance on the part of external appraisers to incorporate expected sources of declining value into appraisals before the declines actually materialize.[31]

The results of the data analysis are consistent with these explanations. Exhibit 2 shows the incidence of paired appraisal samples in which external appraisals exceed internal appraisals to be greater than 50% for all fifteen tests involving sets of sample standard deviations that exceed the median standard deviation (*i.e.*, in which the appraised assets involve greater-than-average investment risk). Furthermore, the test values are statistically significant in thirteen of the fifteen cases at the 1% level or lower, and are significant in all six cases of aggregated data at the 0.1% level or lower.

This suggests that, although the null hypothesis appears correct in the case of the RREEF data for sets of sample standard deviations that do not exceed the median standard deviation, in general the deterministic component of appraisal error can be expected to vary across the appraiser universe as well as across time and the real estate asset universe.

Nonetheless, it is possible to estimate the standard deviation of random appraisal error empirically from the data. This estimate is not completely straightforward, because our test has rejected the null hypothesis for the upper half of the distribution of sample standard deviations. More precisely, the analysis thus far has shown that the upper half of the distribution of sample standard deviations for simultaneous appraisal pairs is biased upward relative to the corresponding distribution of sample standard deviations for random appraisal error. It follows that the expected mean value of sample standard deviations for simultaneous appraisal pairs is strictly larger than the expected mean value of the corresponding sample standard deviations for random appraisal error.

However, the test results have confirmed that the data are consistent with the assumption that the lower half of the distribution of sample standard deviations for simultaneous appraisal pairs is identical to the lower half of the corresponding distribution for sample standard deviations of random appraisal error. This conclusion, together with the assumption of normality for random appraisal error, can be applied to circumvent the problem that the upper half of the distribution of sample standard deviations for random appraisal error is not observable.

In particular, a consistent estimator for the true standard deviation of random appraisal error can be constructed from the observation that the sample median for the distribution of sample standard deviations for simultaneous appraisal pairs is a consistent estimator for the true median of the distribution. Hence, the sample median for this distribution is also a consistent estimator for the true median of the distribution of sample standard deviations for random appraisal error. For a normal distribution, the ratio of the true standard deviation to the true median for the distribution of sample standard deviations is 1.4826027 to eight significant figures.[32] It follows that, in the case of normally distributed random appraisal error, the sample median for the distribution of sample standard deviations of simultaneous appraisal pairs multiplied by 1.4826027 is a consistent estimator for the true standard deviation of random appraisal error.

We select the aggregate set of appraisal pairs that excludes the 1991–1992 data for the median standard deviation estimate, since as discussed earlier the year-by-year cross-sectional distributions for the entire data set appear time-dependent due to the inclusion of exceptionally noisy 1991–1992 data.[33] Thus, Exhibit 1 implies the best estimate for the standard deviation of random appraisal error based on the available data to be $1.4826027 * 1.35\% \cong 2.00\%$. To this determination must be added the qualification that the standard deviation will be slightly larger and possibly also asset-dependent during occasional periods of extreme transactional market gridlock, such as occurred during 1991–1992 and at the beginning of the 1970s.[34]

Although the differences between the medians in Exhibit 1 are not large, sample medians for the six aggregated data sets suggest that random appraisal error for large-capitalization properties is likely to be marginally smaller than random appraisal error for small-capitalization properties, if the standard deviation of random appraisal error is not asset-independent. Thus, the 2.00% estimate for the standard deviation of random appraisal error may be upwardly biased in the case of large-capitalization assets.

Finally, any point estimate of a parameter should be accompanied by some information about the accuracy of the estimate. Since the standard deviation of random appraisal error is estimated by multiplying the point estimate for the median of the sample standard deviations by 1.4826027, it follows that standard error for the former parameter equals the standard error for the latter parameter multiplied by 1.4826027. Standard error for the sample median of the sample standard deviations can in turn be estimated by a general result on the asymptotic standard error of quantile estimators for large sample sizes.

The general result shows that the variance of a sample quantile ξ_p of order p for a cumulative distribution with a probability density function f is given asymptotically by $\mathrm{var}(\xi_p) \cong pq/(n(f(\xi_p)^2)$ for large sample sizes n, where p is any proper fraction between 0 and 1, $q = 1 - p$, and f is continuous and positive at ξ_p.[35]

In the case of the sample median of sample standard deviations with one degree of freedom, $p = q = 0.5$, and $f(\xi_{0.5}) \cong 47.078018/n$. Exhibit 2 shows that the median of the distribution of sample standard deviations is estimated from 578 samples. It follows that $\mathrm{var}(\xi_{0.5}) \cong 0.00011280/n \cong 0.0000001952$, and thus that $\sigma(\xi_{0.5}) \cong 0.000442 = 0.0442\%$.

This estimate can be considered a lower bound for the standard error of the sample median in the case of large sample sizes. However, the estimate does not account for all random error sources that contribute to the sample median in the case of the data for this study. Derivation of the above asymptotic formula for the variance of a quantile depends on the hypothesis that the sample distribution contains n independent samples from the probability distribution, a hypothesis that the binomial test rejected in the case of above-median sample standard deviations for the appraisal pair data. Thus, an adjustment to the standard error is almost certainly necessary to account for uncertainty in the sample median due to upward bias in some above-median samples. Additional research remains to be done on this problem.

Conclusion

Pairs of simultaneous independent appraisals for the same commercial real estate assets can be used to isolate the random component of appraisal error for institutional-grade real estate. Data of this type collected since 1989 shows that the standard deviation of random appraisal error is approximately 2%, except during infrequent periods of transactional market gridlock that occur every couple of decades. This suggests that appraisal methodology applied by professional appraisers to institutional-grade real estate is potentially more accurate than previously supposed by investment industry observers.[36]

Investor and investment manager views of the impact of random appraisal error on asset valuation can have a material impact upon the structure of institutional real estate portfolios. Widely accepted perception that the contribution of random appraisal error to portfolio valuation is relatively significant has encouraged the belief that only large diverse portfolios can be valued accurately by appraisal methodology. This has provided both institutional investors and investment managers with incentives to assemble large diverse real estate portfolios, even though incremental agency costs associated with creating and holding an array of properties diversified by property type and geographical region could well be substantial.[37]

The standard deviation of random appraisal error obtained in this study becomes particularly significant when combined with the results of studies such as Cole, Guilkey and Miles (1986) and Miles, Guilkey, Webb and Hunter (1991) that determine the typical magnitude of total appraisal error to be approximately 10% of appraised value. The joint implication of these studies is that most appraisal error is deterministic. However, deterministic appraisal error is highly nonlinear and appears random when subjected to routine cross-sectional analysis. Since the deterministic component of appraisal error appears randomly distributed for purposes of cross-sectional analysis, the quadratic formula for subtraction of variances shows that virtually all cross-sectional variance in appraisal error observed in previous studies is due to the contribution of the deterministic component.[38]

This conclusion has material consequences for institutional investors. Random error represents the combined effect of a virtually infinite number of largely independent inputs, each of which has such an insignificant impact by itself that individual effects are impossible to detect. By contrast, deterministic error represents the impact of a finite number of inputs, each of which can in principle be identified, constrained and controlled; and once controlled they can be reduced, and possibly even eliminated.[39] Thus, this study provides evidence that previous appraisal research has inadvertently encouraged investor tolerance of excessive agency costs by misinterpreting the aggregate statistical effect of these costs as random error.

This suggests a basic shortcoming in top-down real estate portfolio strategies: informationally efficient asset pricing is not automatically available to real estate investors through routine operation of market-based mechanisms. Inattention to asset pricing can impose a significant penalty on subsequent portfolio returns. In order to avoid excessive agency costs, significant investor attention must necessarily be devoted to asset valuation during investment acquisitions and disposals, or at least to the imposition and implementation of investor-oriented controls on the asset valuation process.

Asset pricing takes care of itself in liquid markets, where well-informed investors supported by a continuously flowing stream of economic information about market assets that is protected by authorities such as the Securities and Exchange Commission and the Financial Accounting Standards Board generate informationally efficient market prices. In such cases, investors have the luxury of focusing on secondary strategic investment issues such as efficient portfolio diversification. However, in an illiquid market without any overseeing authority, a bottom-up approach to asset selection and portfolio construction is a practical consequence of the need for investors to minimize the impact of deterministic influences on the asset valuation process.

Notes

[1] Data for pre-1980 empirical studies in housing economics usually were based wholly or in part on owner estimates of market value. For example, Kish and Lansing (1954) and Kain and Quigley (1972) compare residential valuations by appraisers and owner-occupants. The studies suggest that the nonrandom component of valuation error is much smaller than the random component and of marginal statistical significance. Data from the studies can be used easily to infer that the standard deviation of appraisal error is between 15% and 20% if valuations by appraisers and owners are equally accurate, or somewhat smaller if appraiser valuations are more accurate than owner valuations. In addition, the studies rely on simplified appraisals rather than full appraisals, so random appraisal error in this data may be larger than in the case of full appraisals by professionally designated appraisers. In a related study, Robins and West (1977) apply a hedonic model to similar data and suggest that the standard deviation of appraisal error is about 12%. However, the definition of appraisal error in this study includes any appraisal components explainable by variables not included in the model.

[2] Initial research studies supported the hypothesis that housing and commercial property appraisal error are similar. For example, Cole, Guilkey and Miles (1986) report the mean absolute difference between transaction prices and immediately preceding external appraisals for assets in the NCREIF database to be 9.5% of appraised value, and Miles, Guilkey, Webb and Hunter (1991) update the estimate to 10.7% for the mean absolute difference between transaction prices for NCREIF assets and reported appraisal values for the calendar quarter immediately preceding the quarter of each corresponding transaction. However, the reported values include the combined effects of instantaneous transaction error, instantaneous appraisal error and temporal aggregation error (*e.g.*, see Geltner, Graff and Young, 1994; and Geltner, 1993).

[3] This is not to suggest that residential appraisals are immune from pressures to value agency costs, only that there are fewer sources for such pressures than in the case of commercial real estate appraisals. For example, Ferguson (1988) presents statistically significant empirical evidence that appraisers are susceptible to agency cost pressure to inject upward bias into residential appraisals conducted for mortgage lenders as part of the loan application evaluation process in order to justify previously negotiated prices in purchase contracts. The study also shows that upward bias is statistically more likely in appraisals conducted for lenders by independent appraisers than in appraisals conducted by in-house staff.

[4] If the conjecture is not correct, then it still follows that the sample standard deviation for the pair of simultaneous appraisals is an upper bound for the magnitude of the random component of appraisal error. However, in this case it is unlikely that the distribution of sample standard deviations for pairs of simultaneous appraisals would appear to be stationary, since it is probable that the average magnitude of nonrandom appraisal error varies with time.

[5] In recent years, the focus of pension plan real estate investments has shifted away from commingled funds and toward dedicated separate investment manager accounts, each wholly owned by a single plan sponsor. To reduce expenses, many public plan sponsors require outside appraisal of assets in dedicated separate accounts as infrequently as once every three years. Nevertheless, some investment managers of these accounts conduct annual internal appraisals of all assets for reporting and decision-making purposes.

[6] Appraisal in these situations precedes negotiation of a sale contract, so agency pressure for upward bias to justify a previously negotiated contract price is absent, cf. Note 3. Dotzour (1988b) continues examination of the data set and shows that appraisals in the sample are on average unbiased, although several subsamples (defined by economic geographical region, or calendar quarter in which properties were sold) display evidence of statistically significant appraisal bias.

[7] The Dotzour (1988a) study shows that the difference between the accuracies suggested by the two sample values is statistically significant. This suggests that, in the absence of agency cost pressures, full residential appraisals by professionally designated appraisers are usually more accurate than full residential appraisals by nondesignated appraisers.

[8] The description of the appraisal environment implies that sample appraisal variances in the subclasses are independent. Under the assumption that each subclass is normally distributed, the F distribution can be applied to test the hypothesis that the two subclasses have the same true standard deviation. The ratio of the sample variances is $(2.67/2.61)^2 \cong 1.045 < 1.073 \cong F_{.55}(14,14)$, where 14 is the number of degrees of freedom in both the numerator and the denominator. This implies acceptance of the null hypothesis at the 90% significance level. In other words, even under the null hypothesis there is a 90% probability that randomly generated sample variances for two subclasses with 14 degrees of freedom in each subclass will differ by more than the Diaz sample variances.

[9] A relative appraisal error estimate with 29 degrees of freedom can be derived with respect to a single combined sample mean if the true distributional means for the two appraisal subclasses are identical, a hypothesis tested and accepted in the Diaz study. First, the combined sample mean m_0 is computed as the average of the two sample means m_1 and m_2 for the individual subclasses. Next, the absolute sample variance s_0^2 for the combined appraisal sample is decomposed by two-way analysis of variance into the sum of within-subclass and between-subclass variances by the following equation: $(2n - 1)s_0^2 = (2n - 2)s_w^2 + (n)s_b^2 = (n - 1)s_1^2 + (n - 1)s_2^2 + (2n)(m_i - m_0)^2$, where n is the number of samples in each subclass, s_1^2 and s_2^2 are the individual subclass variances, the within-subclass variance s_w^2 is the average $(s_1^2 + s_2^2)/2$ of the individual subclass variances, m_i may be either of the two subclass sample means, and the between-subclass variance s_b^2 is equal to $2(m_i - m_0)^2$. By numerically evaluating the expression for between-subclass variance, it is easily seen that $s_b^2 << \min(s_1^2, s_2^2)$. It follows from the analysis of variance equation that $s_0^2 < \min(s_1^2, s_2^2)$, and similarly that relative appraisal error for the combined sample is less than the minimum subclass relative appraisal error. Finally, the value of relative appraisal error with 29 degrees of freedom is estimated to be $s_0/m_0 \cong$ 2.60%, which is confirmed by direct comparison to be less than the relative appraisal error estimate for each subclass.

[10] The sample means of the natural logarithms of the appraisals are 8.3673 and 8.2904, so the difference between the sample means is 0.0769. The sample standard deviation for the difference between sample means is $s \cong ((0.0532)^2/15 + (0.0504)^2/14)^{1/2} \cong 0.0193$. Thus the t-Statistic for the difference between the sample means is $0.0769/0.0193 \cong 3.9845$, which implies that the equality hypothesis for the true means of the natural logarithms of the appraisals is rejected at the 0.01% significance level.

[11] Diaz and Wolverton (1998, Figure 1) plots cumulative distributions for the two unanchored appraisal sets. The study observes that the shapes of the two distributions are virtually identical.

[12] The description of the appraisal environment implies that the sample appraisal variances for the two sets are independent. The W' approximation to the Shapiro-Wilk W test (see Shapiro and Francia, 1972) implies that logarithmic normality for the appraisal distributions is accepted at the 10% significance level in the case of the sixteen-sample set and at the 3% significance level in the case of the fifteen-sample set. Thus, the F distribution can be applied to test the hypothesis that the two sets have the same standard deviation. The sample variance ratio is $(5.32/5.04)^2 \cong 1.105 < 1.110 \cong F_{.575}(15,14)$. This implies acceptance of the null hypothesis at the 85% significance level.

[13] The observation that relative appraisal error for anchored appraisals in the Diaz and Wolverton (1998) data set is *larger* than relative appraisal error for the unanchored appraisals contradicts the predictions of most appraisal-smoothing studies. However, this empirical result is consistent with recent theoretical results in Lai and Wang (1998), which implies that appraisal anchoring *exaggerates* rather than diminishes many quadratic investment statistics.

[14] Derivation of the confidence interval for the true standard deviation then follows from the observation that sample variance for a normal distribution has a chi-square distribution, *e.g.*, see Snedecor and Cochran (1989). The Shapiro-Francia W' test implies that logarithmic normality for the anchored appraisals is accepted at the 10% significance level. The ratio of the sample variances is $(6.90/5.20)^2 \cong 1.762 > 1.745 \cong F_{.90}(14,29)$, which implies rejection of the equality hypothesis for the true standard deviations of the anchored and unanchored appraisals at the 80% significance level cf. Notes 8 and 12.

[15] See Graff and Cashdan (1990), Graff (1992) and Graff and Webb (1997) for discussions of the impact of average lease maturity and tenant creditworthiness on real estate investment characteristics.

[16] the case of Geltner and Goetzmann (1998), stationarity in returns is essential for the second stage regression designed to separate random appraisal error from other appraisal error components. More precisely, this regression requires among other things that ex ante (*i.e.*, expected) variance of investment return between appraisals be linearly proportional to the length of time between appraisals for each real estate asset. This can only be the case if the ex ante annual return variance for each real estate asset does not change with time.

[17] Cf. Note 3.

[18] A simple formula relates relative appraisal error to absolute appraisal error provided the magnitude of relative appraisal error is reasonably small: $\delta = \ln(V^*) - \ln(V) = \ln(V^*/V) = \ln((V + V^* - V)/V) = \ln(1 + (V^* - V)/V) \approx (V^* - V)/V$, provided that $(V^* - V)/V$ is reasonably close to zero, *e.g.*, less than 15% in magnitude.

[19] It would be unrealistic to test the hypothesis that the probability distribution for absolute appraisal error is constant across the universe of real estate assets, since this hypothesis would imply that relative appraisal accuracy varies inversely with asset value. For example, if the standard deviation of absolute appraisal error were \$100,000, then the hypothesis would imply that the standard deviation of relative appraisal error is approximately 10% in the case of a \$1 million property but only 0.01% in the case of a \$1 billion property. Similarly, if the standard deviation of absolute appraisal error were \$4,000,000, then the hypothesis would imply that the standard deviation of relative appraisal error is approximately 2% in the case of \$200 million property but 200% in the case of a \$2 million property. Such extreme variations in relative appraisal error are impossible to reconcile with vast investment industry experience with appraisal accuracy, and also are contradicted by the data in this study.

[20] A growing body of research suggests that stable fat-tailed distributions are more appropriate models for investment returns than normal distributions (*e.g*., Young and Graff, 1995; and Graff, Harrington and Young, 1997) in the case of real estate returns. However, this research does not conflict with a vast body of experience in both the natural and social sciences suggesting that normal and log-normal distributions are the appropriate models in general for measurement error.

[21] For example, the assumption that the probability distribution of $\varepsilon(p,t,n)$ has mean value zero implies that any appraisal bias is contained in $\eta(p,t)$. Thus, the assumption that $\eta(p,t)$ is independent of appraiser for simultaneous appraisals of the same asset implies that any bias in the appraisals of property p at time t is the same for each appraisal. This implication of the null hypothesis is testable, and is examined in the next section.

[22] Graff and Webb (1997) implies that the cross-sectional distribution of the deterministic component is fat-tailed and gives rise to the fat-tailed asset returns observed in Young and Graff (1995) and Graff, Harrington and Young (1997). The Graff and Webb study concludes that the nonrandom nature of this component only becomes apparent at the individual asset level when all investment information about each asset and agency information about asset management is available for detailed investment analysis, although information about nonrandom appraisal error at the individual asset level can be inferred from statistical tests of individual appraisal-based return series for nonrandom performance persistence.

[23] Since true asset value and the deterministic component of appraisal error are functional values rather than random variables, it follows that $\sigma(\ln(V^*(p,t,\cdot))) \equiv \sigma(\varepsilon)$. In other words, the only random variable on the right side of Equation (2) is the random appraisal error component, so the standard deviation of total appraisal error equals the standard deviation of the random appraisal error component.

[24] By contrast, since the deterministic component appears random in aggregate analysis of appraisal-based returns, routine data analysis substantially overestimates the magnitude of random appraisal error by incorporating variations in the deterministic component into the estimate. This mistake has material consequences for institutional investor returns. Excessive agency costs can be eliminated by appropriately structured controls on portfolio managers [for example, see Graff and Webb (1997) for suggestions on performance measures designed to detect the effect of excessive agency costs on investment return series]. Thus, our results imply that previous real estate investment research has inadvertently encouraged passive investor reactions to excessive agency costs by misinterpreting the aggregate statistical effect of these costs as random appraisal error.

[25] Since this is a nonparametric test, the conclusion follows even if the standard deviation for random appraisal error varies with respect to time and/or across the asset universe.

[26] If V_1^* and V_2^* are simultaneous appraisals of the same asset, then $s(V_1^*, V_2^*) = |V_1^* - V_2^*|/\sqrt{2}$, where $s(V_1^*, V_2^*)$ is the sample standard deviation for the appraisal pair. In other words, the sample standard deviation in the case of two appraisals is simply the absolute value of the difference between the appraisals divided by $\sqrt{2}$. It follows that the distribution of sample standard deviations for appraisal pairs is equivalent to the distribution of unsigned differences between appraisal pairs. It will be shown in the next section that an examination of the unsigned differences between appraisal pairs is more productive in the case of our data set than an examination of the corresponding signed differences.

[27] This is consistent with the observation in Young and Graff (1995) that 1991 appraisal-based returns display cross-sectional distributional characteristics inconsistent with characteristics displayed by appraisal-based returns in other years. That study attributes exceptional distributional behavior of 1991 returns to transactional market gridlock, which made it difficult

for appraisers to extrapolate from fewer-than-usual transactions to prices for properties that were not sold. That explanation can also account for the exceptional uncertainties observed in the present study for a significant fraction of 1991 and 1992 appraisals.

[28] The null hypothesis is rejected at the 5% level for below-median sample standard deviations in five of the nine cases of annual data. These cases all involve much smaller sample sets than the aggregated data cases, and the rejections are borderline. Since these rejections include the years (1991–1992) of greatest appraisal uncertainty, this could signal the existence of a small amount of time-dependence in the data that vanishes when data are aggregated across a real estate cycle.

[29] This implies that it is inappropriate to attempt to estimate standard appraisal error by computing the sample standard deviation of the set of differences between external and internal appraisals and dividing the result by $\sqrt{2}$, cf. Note 26. The problem with this approach is that appraisal differences that are closer to zero cannot be samples from the same probability distribution as appraisal differences that are farther away from zero, since smaller-magnitude differences are as likely to be negative as positive, whereas larger-magnitude differences are biased in the positive direction. In fact, it is straightforward to show that the expected sample variance of the set of differences between external and internal appraisals equals the variance of random appraisal error multiplied by 2 plus one-half the average squared difference between external and internal deterministic appraisal components.

[30] Individual internal appraisals do show conservative bias in cases that present identifiable exceptional risk factors, such as large single-tenant properties approaching lease expiration in soft rental markets with tenants whose prospects for lease renewal appear uncertain, and assets in areas that are in a regional economic decline (*e.g.*, oil patch properties that provided the impetus for internal appraisals).

[31] Observation of this phenomenon provided the impetus for RREEF to institute the simultaneous appraisal system, cf. the data description section. RREEF has observed that fee appraisers seem unwilling to assume negative market rental growth rates. This results in upward appraisal bias in specific economic situations such as the oil patch properties in the 1980s, although not in most economic circumstances. As noted, Hendershott and Kane (1995) observed this type of systematic upward bias in office property appraisals from the second half of the 1980s.

[32] Under the assumption that random appraisal error is normally distributed, sample variance in the case of two error samples has a chi-square distribution with one degree of freedom, and the cumulative probability of the distribution has a simple integral representation (*e.g.*, see Hogg and Craig, 1978). Since sample standard deviation is the square root of sample variance, this immediately yields a corresponding integral for the cumulative probability of the sample standard deviation with one degree of freedom in terms of the following piecewise continuously differentiable probability density function f: $f(t) = 0$ for $t < 0$, and $f(t) = (\sqrt{2}/(\sigma\sqrt{\pi}))\exp(-t^2/(2\sigma^2))$ for $t \geq 0$, where σ is the true standard deviation. It follows that the median value for the distribution of sample standard deviations is the solution c to the integral equation $0.5 = (\sqrt{2}/(\sigma\sqrt{\pi}))\int_0^c \exp(-t^2/(2\sigma^2))\mathrm{d}t$, *i.e.*, $c = 0.67448952\sigma$ to eight significant figures.

[33] Cf. Note 27.

[34] For these periods, the median sample standard deviation for our data from the years 1991–1992 can be used to estimate the standard deviation of random appraisal error, *i.e.*, 1.4826027*3.67% = 5.42%.

[35] For example, see Kendall and Stuart (1963). Minor technical constraints must also be imposed on the probability density function for this result to be valid. The function f must be square integrable, piecewise continuously differentiable and continuously differentiable at ξ_p. Note 32

implies that the density function for the sample standard deviation with one degree of freedom satisfies these constraints provided that $\xi_p \neq 0$.

[36] This assertion comes with caveats. For example, as noted, the results of Diaz (1997) and Diaz and Wolverton (1998) together suggest that valuation expertise on the part of appraisal professionals is limited geographically.

[37] It is currently popular among real estate professionals and academics to argue that consolidation of real estate investment management functions reduces investor costs through economies of scale with no loss of efficiency at the portfolio management or asset management level. However, these assertions bear a striking resemblance to the concepts promoted by 1960s advocates of conglomeration to justify the assemblage of corporate hodgepodges. Vogel (1997) and Campbell, Ghosh and Sirmans (1998) discuss some shortcomings of recent real estate consolidations also cf. Note 35.

[38] If total appraisal error appears random when viewed cross-sectionally, then both total and deterministic appraisal error have apparent standard deviations, and the apparent magnitude of deterministic appraisal error is given by $s_{nr} = (s_{tot}^2 - s_r^2)^{1/2}$, where s_{tot} is the apparent sample standard deviation of appraisal error, s_r is the sample standard deviation of random appraisal error and s_{nr} is the apparent sample standard deviation of nonrandom albeit nonlinear appraisal error. Based on the Miles, Guilkey, Webb and Hunter (1991) estimate $s_{tot} \approx 10.00\%$ and the estimate $s_r \approx 2.00\%$ from the present study, it follows from the above formula that $s_{nr} \approx 9.80\%$.

[39] For example, Graff and Webb (1997) presents evidence that a significant component of nonrandom appraisal error is due to the impact of excessive transaction-based agency costs on subsequent appraisal valuations. The study also suggests how excessive agency costs can be detected and eliminated by appropriately structured management control systems, and designs a statistically-based mechanism to detect unusually large agency costs in appraisal-based returns.

References

Campbell, R. D., C. Ghosh and C. F. Sirmans, The Great REIT Consolidation: Fact or Fancy?, *Real Estate Finance*, 1998, 15:2, 45–54.

Cole, R., D. Guilkey and M. Miles, Toward An Assessment of the Reliability of Commercial Appraisals, *The Appraisal Journal*, 1986, 54:3, 422–32.

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.

Dotzour, M. G., The Impact of Professional Designation on Appraisal Variation in Corporate-Relocation Appraisals, *The Appraisal Journal*, 1988a, 56:4, 522–30.

——, Quantifying Estimation Bias in Residential Appraisal, *Journal of Real Estate Research*, 1988b, 3:3, 1–11.

Ferguson, J. T., After-Sale Evaluations: Appraisals or Justifications, *Journal of Real Estate Research*, 1988, 3:1, 19–26.

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

——, How Accurate is the NCREIF Index as a Benchmark, and Who Cares?, *Real Estate Finance*, 1998, 14:4, 25–37.

Geltner, D. M. and W. N. Goetzmann, Two Decades of Commercial Property Returns: A NCREIF Index Using Independent Appraisals, Working Paper, Cincinnati, OH: University of Cincinnati, January 1998.

Geltner, D. M., R. A. Graff and M. S. Young, Random Disaggregate Appraisal Error in Commercial Property: Evidence from the Russell-NCREIF Database, *Journal of Real Estate Research*, 1994, 9:4, 403–19.

Graff, R. A., Perspectives on Debt-and-Equity Decomposition for Investors and Issuers of Real Estate Securities, *Journal of Real Estate Research,* 1992, 7:4, 449–67.

Graff, R. A. and D. M. Cashdan, Jr., Some New Ideas in Real Estate Finance, *Journal of Applied Corporate Finance,* 1990, 3:1, 77–89.

Graff, R. A., A. Harrington and M. S. Young, The Shape of Australian Real Estate Return Distributions and Comparisons to the United States, *Journal of Real Estate Research,* 1997, 14: 3, 291–308.

Graff, R. A. and J. R. Webb, Agency Costs and Inefficiency in Commercial Real Estate, *Journal of Real Estate Portfolio Management,* 1997, 3:1, 19–36.

Hendershott, P. H. and E. J. Kane, U.S. Office Market Values During the Past Decade: How Distorted Have Appraisals Been?, *Real Estate Economics,* 1995, 23:2, 101–16.

Hogg, R. V. and A. T. Craig, *Introduction to Mathematical Statistics,* Third Edition, New York: Macmillan, 1978.

Kain, J. F. and J. M. Quigley, Note on Owner's Estimate of Housing Value, *Journal of the American Statistical Association,* 1972, 67, 803–806.

Kendall, M. G. and A. Stuart, *The Advanced Theory of Statistics,* Second Edition, Volume 1, New York: Hafner Publishing Company, 1963.

Kish, L. and J. B. Lansing, Response Errors in Estimating the Value of Homes, *Journal of the American Statistical Association*, 1954, 49, 520–38.

Lai, T. Y. and K. Wang, Appraisal Smoothing: The Other Side of the Story, *Real Estate Economics,* 1998, 26:3, 511–35.

Miles, M., D. Guilkey, B. Webb and K. Hunter, An Empirical Evaluation of the Reliability of Commercial Appraisals, NCREIF Research Paper, Chicago: National Council of Real Estate Investment Fiduciaries, 1991.

Robins, P. K. and R. W. West, Measurement Errors in the Estimation of Home Value, *Journal of the American Statistical Association,* 1977, 72, 290–94.

Shapiro, S. S. and R .S. Francia, An Approximate Analysis of Variance Test for Normality, *Journal of the American Statistical Association,* 1972, 67:3, 215–16.

Snedecor, G. W. and W. G. Cochran, *Statistical Methods,* Eighth Edition, Ames, IA: Iowa State University Press, 1989.

Vogel, J. H., Why the New Conventional Wisdom About REITs is Wrong, *Real Estate Finance,* 1997, 14:2, 7–12.

Wolverton, M. and P. Gallimore, Client Feedback and Perception of the Role of the Appraiser: An International Study of Real Estate Appraisers, Paper presented at the American Real Estate Society Annual Meeting, Monterey, CA, April 1998.

Young, M. S. and R. A. Graff, Real Estate is Not Normal: A Fresh Look at Real Estate Return Distributions, *Journal of Real Estate Finance and Economics,* 1995, 10:3, 225–59.