A Further Discussion of Optimal Comparable Selection and Weighting, and A Response to Green

George W. Gau,* Tsong-Yue Lai** and Ko Wang**

Using a revised model framework that views expected adjusted prices of comparables as random variables, Green (1994) demonstrates that Vandell’s (1992) minimum variance estimator is preferred under the classical ordinary least squares (OLS) assumptions. As a result, the minimum coefficient of variation estimator proposed by Gau, et al. (1993) is preferred only when the classical OLS assumptions are relaxed. We demonstrate that, even under Green’s revised framework, there is not sufficient evidence in the paper to justify this claim.

Vandell (1991) contributes to the appraisal and valuation literature by formalizing the market comparison approach using a minimum variance selection criterion. Gau, et al. (1992) extend Vandell’s work by introducing a minimum coefficient of variation estimator as an alternative selection criterion. In the preceding article, Green (1994) shows that, under the classical ordinary least squares (OLS) assumptions, Vandell’s variance minimization approach is preferred because of two important statistical properties: unbiasedness and minimum variance. Gau, et al.’s minimum coefficient of variation approach is preferred only when the classical OLS assumptions (namely, omitted variables) are relaxed.

While we agree with Green (1994) that a formal study of the statistical properties of comparable selection processes is important, we believe that Green’s proposition is not supported by sufficient evidence in the paper. Specifically, we believe that there are at least three points deserving further investigation.

First, the model framework developed by Vandell and adopted by Gau, et al. views expected adjusted prices as parameters. Green revises that

*University of Texas, Austin, TX 78712
**California State University, Fullerton, CA 92634
framework by assuming expected adjusted prices are random variables. Whether these prices should be viewed as random variables or parameters when they are used as inputs in an optimization process is debatable. More importantly, under Green’s framework, the minimum variance estimator will also be biased. We will discuss this point in detail.

Second, Green takes the position that the coefficient of variation estimator is biased because it truncates the distribution of the adjusted prices of comparables. We disagree with this position. We argue that, if the “bias” exists, it is because the ratio of two unbiased estimators is (by definition) biased. Under any circumstances, the bias does not result from the sampling problem described by Green. We will clarify this issue.

Third, Green suggests that the magnitude of the bias (if it exists) is so significant that the coefficient of variation of Gau, et al.’s estimator is, in fact, higher than the estimator derived using Vandell’s methodology. Because of this, Green argues that Vandell’s methodology is preferred even if the minimum coefficient of variation is used as the comparables’ selection criterion. The evidence presented by Green, however, is not sufficient to justify this claim. We will address this issue.

Indeed, it is unlikely that a weighted average of a set of unbiased estimators can be biased, while a different set of weights applied to the same set of unbiased estimators is unbiased. It should be noted that, in practical applications, appraisers always take the adjusted prices of comparables into consideration when they reconcile the final value estimates. The minimum coefficient of variation selection criterion uses the same methodology and same information as the minimum variance selection criterion, except that the expected adjusted prices of comparables are also taken into consideration.¹ The claim that the minimum coefficient of variation estimator is a biased estimator (just because it examines more information) seems counter-intuitive.

The next section discusses the bias issue. We show that, if the minimum coefficient of variation estimator is biased, the minimum variance estimator could also be biased. The third section discusses the source of the bias. Our view on the source of bias is significantly different from the

¹ Indeed, especially for commercial properties that have diversified property attributes (and, therefore, a wide dispersion of adjusted prices) and have a relatively low number of transactions (and, therefore, less confidence in the quality of the comparable), it is unlikely that appraisers will reconcile a final value estimate without taking the adjusted price of each comparable into consideration.
one offered by Green. The fourth section addresses the impact and magnitude of the bias. We believe that there is not sufficient evidence to support Green’s claim on this issue. The last section contains our conclusions.

Is There a Bias?

The optimal comparable selection and weighting model developed by Vandell and adopted by Gau, et al. consists of two stages. The first stage estimates the adjusted sales prices (or the adjustment factors) and the variance of the adjusted sales price (or the variance/covariance matrix of the adjustment factors). In Vandell’s framework, the expected value and variance of adjustment factors can be derived from any acceptable method (1991, page 218); OLS is only one of the alternative choices. When the OLS technique is used, Vandell’s equations (5) and (6) define the expected adjusted sales price and the variance of the adjusted sale price as

\[
\begin{align*}
\mu_{V,\alpha} &= E(V_{\alpha}) + \sum_{j=1}^{m} E(\hat{a}_j)(X_{\alpha j} - X_{\alpha j}) \\
&= V_{\alpha 0} + \sum_{j=1}^{m} a_j(X_{\alpha j} - X_{\alpha j}).
\end{align*}
\]

(1)

\[
\sigma_{V,\alpha}^2 = \sum_{k=1}^{m} \sum_{j=1}^{m} (X_{\alpha k} - X_{\alpha k})(X_{\alpha j} - X_{\alpha j})\sigma_{a_k a_j}
+ 2\sum_{j=1}^{m} (X_{\alpha j} - X_{\alpha j})\sigma_{V,\alpha j} + \sigma_{V,\alpha}^2.
\]

(2)

It should be noted that, in Vandell’s equations (5) and (6), both the expected adjusted prices (\(\mu_{V,\alpha}\)) and the variances (\(\sigma_{V,\alpha}^2\)) are viewed as parameters and not random variables. They are considered to be parameters with the understanding that the \(\mu_{V,\alpha}\) is derived from point estimates \(a_j\)'s (adjustment factors), and \(\sigma_{V,\alpha}^2\) is a function of point estimators, \(\sigma_{a_k a_j}\) (the variance/covariance matrix of the adjustment factors).\(^2\)

The second stage of Vandell’s approach uses a quadratic programming technique to estimate the weights of the comparables. The objective is to select the weights that minimize the variance of the final value estimate. It should be noted that the “optimal” weights derived from an optimization approach should be parameters. In other words, an “optimal” solution should not be point estimators that involve uncertainty and have error terms. Vandell’s equations (11) and (12) define the final value estimate (\(\hat{V}_V\)) and its variance (\(\sigma_{V}^2\)) as

\[
E(\hat{V}) = \mu_V = \sum_{i=1}^{n} \omega_i E(V_i) = \sum_{i=1}^{n} \omega_i \mu_{V_i} = \sum_{i=1}^{n} \omega_i V_i,
\]

(3)

\(^2\) In Vandell’s model, \(\sigma_{V,\alpha j}\) is set to zero and \(\sigma_{V,\alpha}^2\) is assumed to be a constant.
and

\[
\text{VAR}(V) = \sigma^2_V = \sum_{i=1}^{n} \sum_{j=1}^{m} \omega_k \omega_j \sigma_{kj}.
\]  

(4)

Again, both the \( \mu_{vi} \) and \( \sigma_{kj} \) are parameters in this constrained minimization problem. Furthermore, the weight \( (\omega_k) \) of each comparable in equations (3) and (4) is not specified as a random variable.

Green, however, views \( \mu_{vi} \)'s as point estimators of random variables. When \( \mu_{vi} \)'s are treated as stochastic, Green shows that the minimum coefficient of the variation estimator is biased. The estimator is biased because the weights \( (\omega's) \) of comparables are estimated using \( \mu_{vi} \)'s. and \( \mu_{vi} \)'s are located in both the numerator and denominator of the equations (see Green's equations (6) and (7)).

We agree that the minimum coefficient of the variation estimator is biased if \( \mu_{vi} \)'s are not parameters. It is well known that the ratio (or product) of two unbiased estimators is a biased estimator.\(^3\) We, however, disagree with Green on two fundamental issues. First, under Green's framework, the minimum variance estimator is also a biased estimator. Second, similar to Vandell, we also believe that, although \( \mu_{vi} \)'s are estimated using sample data, they could be viewed as parameters when used in the comparable selection and weighting process. When \( \mu_{vi} \)'s are considered to be parameters, both the minimum variance and minimum coefficient of the variation estimators are unbiased.

Based on Vandell's equation (5), Green argues that \( \mu_{vi} \) should be a point estimate of a random variable because it is estimated from adjustment factors \( (a_j)'s \). In other words, because adjustment factors are point estimators that contain stochastic terms, Green proposes that \( \mu_{vi} \) should be stochastic. Even if we adopt Green's argument, it is important to note that the variance \( \sigma^2_{V_i} \) specified in Vandell's equation (6) is not a population variance, but rather a sample, or estimated, variance. It is even more important to note that both the \( \mu_{vi} \) and \( \sigma^2_{V_i} \) are estimated from the same set of data.

Vandell's equation (6) clearly shows that both \( \mu_{vi} \) and \( \sigma^2_{V_i} \) are functions of the adjustment factors \( (a_j \text{ and } \sigma_{aaj}) \). If it is assumed that \( \mu_{vi} \) is a point estimate, then \( \sigma^2_{V_i} \) should also be treated as a point estimator. If \( \sigma^2_{V_i} \) is a

\(^3\) See Raj (1968, pp. 86–87) for a discussion on the bias of a ratio estimator.
point estimator, Green’s equations (4) and (5) indicate that the minimum variation estimator must also be a biased estimator since both the numerator and denominator of the equations contain point estimators.

We believe that the framework established by Vandell and adopted by Gau, et al. with \( \mu_v \) and \( \sigma_v^2 \) being parameters, is preferred for two reasons. First, if we adopt Green’s position, the weights derived from a constrained minimization process are point estimators of random variables. If the weights derived from an “optimal weighting scheme” are point estimates that involve uncertainty and have error terms, then the “optimal weighting” issue would become moot.

Second, and more importantly, the framework set up by Vandell follows the literature. The comparable selection technique developed by Vandell is, in spirit, similar to the security selection technique (such as the selection of the minimum variance efficient portfolio) developed in the finance field. It should be noted that, empirically, the mean-variance efficient frontier also depends upon estimated expected returns (they are sample, not population means) and estimated variances/covariances (they are sample, not population variances/covariances) of individual securities. Under Green’s model framework, the mean-variance efficient frontier should also be biased because both the numerator and denominator of the estimators (weights) also contain point estimators.

To summarize, our position is that both the minimum variance and the minimum coefficient of variation estimators are unbiased because expected adjusted prices and variance/covariance of the adjustment factors are parameters when they are used as inputs in an optimization process.

---

4 Given one set of securities, the weights of securities will change if we increase (or decrease) the estimation period used for estimating the expected returns and the variance/covariance matrix. As an example, if the sample-period increases from 180 days to 181 days.

5 Green argues that the expectations of the adjusted prices of comparables should be the same for all comparables. It should be noted that, in Vandell’s framework, the adjustment factors \( \alpha_i \) are estimated independently of the comparables’ observed selling prices (see page 218). Under this framework, the expectation of the adjusted price of each comparable could be different from each other if individual comparables have their own unique property attributes that are not captured in the estimation equation. (Because comparables are not included in the regression sample, this argument does not imply that the OLS equation used to estimate the adjustment factors has the omitted variable problem.) Furthermore, in Vandell’s framework (page 218), the OLS technique is only one of the alternative methods for estimating the adjustment factors. When adjustment factors are derived from methods other than the OLS technique, are there reasons to assume that the expectations of the adjusted prices of comparables must be the same? Empirical validation on the issue raised by Green seems in order.
In other words, we have argued that the weights derived from an optimization process should not be point estimators that involve uncertainty and have error terms. Furthermore, under Green's framework where $\mu_v$ and $\sigma_v^2$ are viewed as point estimators, both minimum coefficient of variation and minimum variance estimators would be biased estimators. Regardless of whether $\mu_v$ and $\sigma_v^2$ are treated as parameters or as point estimators, the minimum variance estimator would not be preferred under the "unbiasedness" statistical property.

**Source of the Bias**

Green takes the position that the coefficient of the variation estimator is biased because it uses a truncated sampling process for the comparables selection. He argues that the coefficient of the variation estimator tends to "reward" comparables that have higher adjusted values. Because of this tendency, the comparable selection process cuts off the bottom portion of the distribution of adjusted prices. Using a random sampling concept, Green proposes that the minimum coefficient of the variation estimator must be biased. Even if it is assumed that the bias exists, we cannot accept Green's argument.

It is important to note that an optimization (minimization) selection process is, by definition, different from a random sample selection process. An optimization process selects observations based on a pre-determined criterion (the objective function). If the optimization process selects observations based on a random sampling procedure, the optimization issue would become moot. It is clear that any bias could not be caused by the random sampling process described by Green.

The bias, if it exists, comes from the ratio (or product) of two unbiased estimators, which could be a biased estimator. If the coefficient of variation estimator is biased, it is because the $\mu_v$'s (adjusted prices) are viewed as point estimators of random variables. It should be noted that, using Gau, et al.'s methodology and a given set of comparables, we will obtain one (and only one) set of weights regardless of whether the adjusted prices are treated as random variables or parameters. However, the set of weights

---

6 If it is true that the minimum coefficient of variation estimator systematically "rewards" comparables with higher adjusted estimates, then it can be argued that the minimum variance estimator also systematically "rewards" comparables with lower variances. Consequently, according to Green's argument, it seems that the minimum variance estimator also uses a truncated sampling process for the comparables' selection.
would be biased if the adjusted prices are random variables, while the same set of weights would be unbiased if the adjusted prices are parameters. Based on this example, it is clear that the magnitude (or the distribution) of adjusted prices has nothing to do with the bias proposed by Green.

**Magnitude of the Bias**

Assuming the minimum coefficient of the variation estimator is biased, Green uses Taylor's approximation technique, together with a numerical example, to demonstrate that the unbiased final value estimate produced by Gau et al.'s technique is actually lower than the estimate produced by Vandell's minimum variance estimator. In other words, he suggests that Vandell's minimum variance estimator should, in fact, have a smaller coefficient of variation than Gau et al.'s coefficient of the variation estimator.

To substantiate Green's arguments, at least three conditions must hold. First, the bias specified in Green's equation (12) must be a good approximation of the true bias. Second, the bias must always be positive. Otherwise, it will increase rather than decrease the final value estimate and, therefore, decrease rather than increase the coefficient of variation of Gau et al.'s estimator. As a result, a negative bias will work against Green's argument. Third, the average magnitude of the bias should be similar to the bias estimated from the single numerical example reported by Green.

Green's equation (11) may not be a good approximation of the bias. It is true that, when applying Taylor's approximation technique, researchers often omit the terms after the second moment (variance) if the expectation is applied to a quadratic function. However, in this particular application, there is no reason to believe that Green's equation (11) should take a quadratic form. As a result, the more complete form of the equation would be

\[
E \left( \frac{X}{Y} \right) = \frac{E(X)}{E(Y)} - \frac{COV(X,Y)}{[E(Y)]^2} + E(X) \frac{VAR(Y)}{E(Y)^3} + \sum_{i=2} \left( -1 \right)^i \frac{E[(Y - E(Y))^i (X - E(X))]}{[E(Y)]^{i+1}} + E(X) \sum_{i=2} \frac{E[(Y - E(Y))^i]}{[E(Y)]^{2i+1}}
\]

(5)

As can be seen from this equation, the magnitudes of the omitted terms (the terms after the second moment) might not be trivial when compared
to the magnitude of the terms used by Green. Given the magnitude of the omitted terms, we are not sure whether Green's equation provides a good approximation of the bias.

An examination of Green's equation (12) indicates that the bias is not always positive. In fact, based on his equation, it is possible to construct a numerical example where the bias is negative. To see this, we rewrite Green's equation (10) as

$$V_s = x_0 + \nu_1 + \frac{\sigma_{11}V_2 - \sigma_{12}V_1}{(\sigma_{22} - \sigma_{12}) V_1 + (\sigma_{11} - \sigma_{12})V_2} (\nu_2 - \nu_1) \quad (6)$$

If the $V_i$'s are random variables, the bias is determined by

$$COV\left[\nu_2 - \nu_1, \frac{\sigma_{11}V_2 - \sigma_{12}V_1}{(\sigma_{22} - \sigma_{12}) V_1 + (\sigma_{11} - \sigma_{12})V_2}\right] \quad (7)$$

The second variable in the covariance term is the weight of the second comparable in Green's equation (10). It is obvious that the sign of the bias is determined by the correlation between the $\nu_2 - \nu_1$ and the weight (which takes a value between 0 and 1). Since the distribution of $\nu_2 - \nu_1$ is the same as $\nu_1 - \nu_2$ (under the assumption that $\nu_i$ is symmetrically distributed), the sign of the covariance could be positive or negative. If the bias specified by Green's equation (11) can be positive or negative, and there is not sufficient evidence to indicate that the magnitude of the bias should be as large as the one proposed by Green, more in-depth analyses are required before the impact and the magnitude of the bias can be judged.

Conclusions

The appraisal methodology is an area that has received little attention in the real estate literature. Considering the importance of unbiased and accurate property value estimates to real estate lending and investment practices, and to the real estate investment literature (for estimating real estate return series), a scientific investigation into the appraisal methodology seems long overdue.

Vandell (1991) is the first to develop a scientific technique (minimum variance estimator) for selecting and weighting comparables in real estate appraisals. Gau, et al. (1992) extend Vandell's pioneering work by proposing the use of an alternative objective function (coefficient of varia-
tion) as the selection criterion. Green (1994) develops arguments as to which objective function (minimum variance or minimum coefficient of variation) should be the preferred approach. We believe that, without empirical evidence, it is difficult to judge the desirability of the two objective functions.

Future research on the optimal comparable selection and weighting should concentrate on empirical applications of the proposed (or other alternative) techniques. Judging from the differing market characteristics, we believe that the preferred technique for appraising commercial properties may differ from the approach used for residential properties. When performing residential property appraisals, comparable properties are usually quite homogeneous and easy to find. Furthermore, because the adjusted prices of those comparables might be quite similar to each other, the minimum variance approach might perform similarly to the minimum coefficient of variation approach. However, for commercial properties with more heterogeneous property attributes (and, hence, a wider dispersion of transaction prices) and a limited number of transactions (and, hence, a low level of confidence in the quality of the comparables), appraisers will not be able to reconcile a final value estimate without taking the adjusted price of each comparable into consideration.

References


7 This is especially true because appraisers often are required to use only the three best comparables. Under this restriction, the comparables used for the appraisal should be quite similar to each other.