Optimal Comparable Selection and Weighting in Real Property Valuation: An Extension

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

Vandell (1991) recently developed a rigorous minimum variance technique for selecting and weighting comparables in real estate appraisal. This article extends Vandell's methodology in three areas: (1) an alternative objective function; (2) an approach that explicitly recognizes the non-negativity constraint on comparable weights; and, (3) a more robust comparable inclusion process. Using Vandell's data, we show how our methodology modifies Vandell's results.

Though the sales comparison approach has been used extensively by real estate appraisers, there has been little research on the important issues of how to select comparable sales in a real estate valuation and what weights to assign each comparable in the final estimation of a property's appraised value. Recently, however, Vandell (1991) develops a rigorous minimum variance approach for modelling the selection and weighting of comparables. His technique, together with the grid-adjustment process devised by Colwell et al. (1983) and tested by Kang and Reichert (1991), represent significant progress toward transforming the traditional sales comparison approach into a more scientific method.

While Vandell's methodology is a substantial advancement, there are several areas that can still be improved. First, Vandell's objective function for comparable selection, variance minimization, does not recognize that it may be more appropriate to consider the standard error per dollar of value under certain conditions. Second, Vandell does not explicitly

* University of Texas at Austin, Austin, Texas 78712
** California State University at Fullerton, Fullerton, California 92634

1 Except for Isakson (1986), previous research dealing with these issues is quite qualitative and involves a high degree of subjectivity. See, for example, Cannady, Colwell and Wu (1984) and Grissom, Robinson and Wang (1987).
include a non-negativity constraint on the optimal comparable set in his minimization procedure, but instead proposes an ad hoc procedure for eliminating a comparable with a negative weight—a procedure that is not robust when there is more than one negatively weighted comparable in the optimal set. Third, the comparable inclusion process used by Vandell in his empirical example is based on the magnitude of the variance of each comparable. The potential effect of the covariance terms among comparables is not explicitly considered in the search process. This article extends Vandell’s research in these three areas and formally proves a Vandell assertion that the variance of the optimal comparable set is a non-increasing function of the number of properties in the comparable set.

The methodology proposed by Vandell selects a set of comparables that has the global minimum variance among all possible subsets of available comparables. This approach neglects the magnitude of the final value estimate. It will select the global minimum variance comparable set over a comparable set that may have a slightly higher variance but with a much higher final value estimate which may intuitively seem to be preferred. For example, if the standard errors of comparable set one and comparable set two are $5,000 (with an estimated value of $250,000) and $5,500 (with an estimated value of $300,000), respectively, the minimum variance approach will select set one over set two because of the magnitude of the standard error. However, comparable set two might be preferred because it provides the least standard error per dollar of appraised value.

This concept can be illustrated through an analogy to a concept found in the finance literature, the mean-variance efficient frontier. The efficient frontier represents various combinations of securities (comparables) that produce the minimum variance for a given level of portfolio return (final value estimate). The global minimum variance portfolio on the efficient frontier is the portfolio with not only the minimum variance, but also the lowest expected return among the efficient portfolios. The comparable property set selected using the minimum variance criterion is conceptually equivalent to selecting this global minimum variance portfolio. Since the efficient frontier is a concave function, the minimum variance approach will select the lowest final estimate (return) among all efficient combinations of comparables (securities). This may not be the “best” portfolio, however, depending upon the risk aversion of the investor, just as this may not be the “optimal” weighting of comparables, depending upon the appraiser/analysts’ view of the optimal trade-off between the magnitude of the final value estimate and the magnitude of the variance.
One approach that could address this problem is to redefine the objective function in terms of a measure that would be dimensionless. The coefficient of variation is such a measure and is commonly used in investment analysis and portfolio management to represent relative dispersion. For our purpose the coefficient of variation of a comparable set is defined as the standard error of the final value estimate divided by the final value estimate, and we will derive a comparable selection process with the minimization of this measure as the selection criterion. Graphically, this would be represented as the tangency of the ray from the origin to the efficient frontier, at a slightly higher variance and higher final value estimate than in the minimum variance case.

To impose a non-negativity constraint on the property weights, Vandell proposes a three-step procedure: (1) solve for optimal comparable weights without the non-negativity constraint; (2) if there is a comparable with a negative weight in the optimal feasible solution, eliminate it from the optimal set; and, (3) re-weight the remaining comparables in the set. While this procedure is appropriate when there is only one comparable with a negative weight, it does not provide guidance in applications where the initial optimal set contains more than one negatively weighted property. For instance, if the unconstrained optimal comparable set includes two or more comparables with negative weights, what is the criterion for choosing which comparable(s) to drop? In this article we will present an alternative methodology for comparable and weight selection that formally incorporates the constraint of non-negative weights in the model development. Our approach eliminates the implementation difficulties associated with Vandell’s three-step procedure.

In this article, we will provide a formal proof that the inclusion of an additional comparable never increases the variance of the optimal final value estimate. The intuition of this proof is simple; the weight for the last included comparable can always be zero. Yet, as noted in Vandell (footnote 11), there is a trade-off relationship between the marginal gain in the precision of the estimate and the marginal costs of obtaining more comparables. In addition, there might be regulatory constraints limiting the number of comparables used in an appraisal assignment (e.g., most

---

2 For a discussion of the properties of the coefficient of variation, see Reilly (1989).

3 See Vandell’s (1991) footnote 8 on page 223.
residential appraisals). Under these constraints the comparable inclusion process and the optimal number of comparables are important considerations in practical applications.

For the comparable inclusion process, Vandell points out that it is necessary to test all possible combinations of comparables to decide which comparable should be included in the comparable set. In his empirical example (1991, Table 5), however, Vandell uses a short-cut method that orders the inclusion of comparables according to their variances. This short-cut method is acceptable when the number of comparables in the comparable set is small. Under such circumstances the influence of the diagonal terms (the variance of each comparable) in the variance-covariance matrix, most likely, will dominate the impact of the off-diagonal terms (the covariances among the comparables). On the other hand, as the number of properties in the comparable set increases, the effect of the off-diagonal terms can become more pronounced and dominate the diagonal term.  

Given the practical constraints on the number of comparables, a more robust comparable inclusion process that considers the off-diagonal terms seems warranted.

The second section of this article derives a comparable selection and weighting process that formally incorporates the constraint of non-negative weights while retaining Vandell's minimum variance objective function. The following section develops a comparable inclusion process using both variances and covariances of comparables as the selection criteria. In this section we will also mathematically prove that the variance of the optimal comparable set is a non-increasing function of the number of comparables in the comparable set. The fourth section presents a comparable selection and weighting process using the minimum coefficient of variation as the selection criterion. The last section contains our conclusions.

**Comparable Selection with Non-Negative Weight Constraint**

This section presents a comparable and weight selection methodology that formally recognizes a non-negative weight constraint. Following the notation used by Vandell, let \( V_i \) be the expectation of \( V_i \), the adjusted sales

---

4 See Ross et al. (1990, p. 271) for a good discussion on the composition of a portfolio variance.
price of \( i \)th observed sale, and \( \sigma_i^2 \) be the estimated variance of \( \hat{V}_i \). The weighted average of the individual \( V_i \) selected, \( i = 1, \ldots, n \), is determined by

\[
\omega^T \hat{V} = \omega_1 \hat{V}_1 + \omega_2 \hat{V}_2 + \ldots + \omega_n \hat{V}_n, \tag{1}
\]

where \( \omega_i \) is the weight of \( \hat{V}_i \), for \( i = 1, \ldots, n \), with \( \sum_{i=1}^{n} \omega_i = 1 \), \( \omega^T \) is the transpose of the weighting vector \( \omega \), \( V \) is the \( n \times 1 \) expected adjusted sales price vector. The expectation and variance of \( \omega^T V \) are determined by

\[
\omega^T V = \omega_1 V_1 + \omega_2 V_2 + \ldots + \omega_n V_n, \tag{2}
\]

and

\[
\sigma^2_V = \sum_{i=1}^{n} \sum_{j=1}^{n} \omega_i \omega_j \sigma_{ij} = \omega^T \Omega \omega, \tag{3}
\]

where \( \sigma_{ij} = \text{Cov}(\hat{V}_i, \hat{V}_j) \) is the covariance of adjusted sales prices \( \hat{V}_i \) and \( \hat{V}_j \), and \( \Omega \) is the \( n \times n \) covariance matrix of \( \sigma_{ij} \). Minimizing the variance of the final value estimate through the optimal selection of comparable weights can be specified as the following constrained minimization problem:

\[
\text{Min } \omega^T \Omega \omega / 2, \tag{4}
\]

subject to

\[
\omega^T e = -\sum_{i=1}^{n} \omega_i - 1
\]

\[
\omega \geq 0.
\]

The Lagrangian of (4) is given by

\[
L = \omega^T \Omega \omega / 2 - \lambda_m (\omega^T e - 1) - \omega^T \Gamma_m, \tag{5}
\]

where \( \lambda_m \) and \( \Gamma_m \) are the Lagrangian multipliers of the constraints \( \omega^T e = 1 \) and \( \omega \geq 0 \). The \( e \) column vector is a \( n \times 1 \) vector of one. The first-order conditions of (5) are:

---

\(^5\) A notation with a bold face refers to either a vector or a matrix; "T" denotes the transpose vector or matrix.

\(^6\) It should be noted that minimizing \( \omega^T \Omega \omega / 2 \) is the same as minimizing \( \omega^T \Omega \omega \). We divided \( \omega^T \Omega \omega \) by 2 to simplify the derivation of equation (6).
\[ \Omega \omega - e \lambda_m - \Gamma_m = 0 \] (6)

\[ \omega^T e = 1 \]

\[ \omega_i \Gamma_{m_i} = 0 \quad \text{for all } i \]

\[ \omega \geq 0. \]

Multiplying \( e^T \Omega^{-1} \) on both sides of (6) and applying the constraint of \( \omega^T e - 1 \) to (6) results in \( \lambda_m = [I - e^T \Omega^{-1} \Gamma_m] e^T \Omega^{-1} e \). Substituting \( \lambda_m \) into (6), the minimum variance weight must satisfy

\[ \omega_{mr}^* = \Omega^{-1} e / e^T \Omega^{-1} e + [I - \Omega^{-1} ee^T / e^T \Omega^{-1} e] \Omega^{-1} \Gamma_m \]

\[ = \omega_{mr} + A_m \Gamma_m, \] (7)

where \( \omega_{mr} = \Omega^{-1} e / e^T \Omega^{-1} e \), and \( A_m = [I - \Omega^{-1} ee^T / e^T \Omega^{-1} e] \Omega^{-1} = [I - \omega_{mr} e^T] \Omega^{-1} \). The \( I \) matrix is the \( n \times n \) identity matrix with one on the diagonal and zeros on the off-diagonal. It should be noted that \( e^T A_m = 0 \). The zero weight of the last term, \( A_m \Gamma_m \), in (7) implies that there exists a net zero weight vector to ensure the positive weight of \( \omega_{mr}^* \). Equation (7) shows that the minimum variance weight with a non-negativity constraint is determined by (i) the ordinary minimum variance weight without the constraint on negative weights, and (ii) a net zero weight vector, \( A_m \Gamma_m \), to circumvent the negative weight from the minimum variance weight \( \omega_{mr} \).

It is worthwhile to note that \( A_m \) is determined by the minimum variance weight \( \omega_{mr} \) and the inverse covariance matrix \( \Omega^{-1} \). With the aid of (6), the non-negative \( \Gamma_m \) can be treated as the variables in the linear program that minimize \( e^T M \) subject to (7). In this application, \( M \) could be any vector of “big number” \( M \). In other words, the non-negative minimum variance weights \( \omega_{mr}^* \) can be obtained using the big \( M \) method to solve a linear program with equation (6) as its constraints.

The technique of the big \( M \) method in linear programming is used to penalize artificial variables. Since the dual problem of (4) is to maximize \( e^T \omega \) subject to the variance and since (7) is the reduced form of (6) (\( \Gamma_m \) can be treated as the slack or artificial variable because of the Kuhn–Tucker condition), maximizing \( e^T \omega \) is equivalent to minimizing \( M^T \Gamma_m \) with non-negative constraints on \( \omega \) and \( \Gamma_m \). See Bradley, Hax and Magnanti (1977, p. 57 and B.7) for a detailed discussion on big \( M \) method and artificial variable.
This technique can be illustrated with a simple example. Assume that there are three comparables and let

\[
\Omega = \begin{pmatrix} 0.8 & 0.5 & 0.7 \\ 0.5 & 1.1 & 0.8 \\ 0.7 & 0.8 & 0.9 \end{pmatrix}.
\]

The minimum variance weight without a non-negativity constraint is

\[
\omega_{mr} = \sqrt{\frac{1}{\Omega^{-1}e}} = \frac{\omega_1}{
\begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} = \begin{pmatrix} 0.81818 \\ 0.45455 \\ -0.27273 \end{pmatrix},}
\]

and the net zero weight matrix is

\[
A_m = [I - \omega_{mr}e^T]\Omega^{-1} = \begin{pmatrix} 3.63636 & 0.90909 & -4.54545 \\ 0.90909 & 2.72727 & -3.63636 \\ -4.54545 & -3.63636 & 8.18181 \end{pmatrix}
\]

The optimal weight selection with a non-negativity constraint can then be specified as the following constrained minimization problem:\footnote{The coefficient, 500, used could be any large number. See footnote 7 for a detailed discussion.}

Min $500 \Gamma_1 + 500 \Gamma_2 + 500 \Gamma_3,$

subject to

\[
\begin{align*}
\omega_1 & = -3.63636 \Gamma_1 - 0.90909 \Gamma_2 + 4.54545 \Gamma_3 = 0.81818 \\
\omega_2 & = -0.90909 \Gamma_1 - 2.72727 \Gamma_2 + 3.63636 \Gamma_3 = 0.45455 \\
\omega_3 & = 4.54545 \Gamma_1 + 3.63636 \Gamma_2 - 8.18181 \Gamma_3 = -0.27273
\end{align*}
\]

$\omega_1, \omega_2, \omega_3, \Gamma_1, \Gamma_2, \Gamma_3 \geq 0.$
Under this minimization program, the resulting optimal weights for \( \omega_1 \), \( \omega_2 \), and \( \omega_3 \) are 0.6667, 0.3333, and 0, respectively.\(^9\) It is worthwhile to note that this result is identical to that derived using Vandell's three-stage procedure.

To test the robustness of our proposed methodology, we use revised data provided by Vandell (see his revised Table 4 on page 152) to estimate the optimal weights, final value estimates, standard error of final value estimates, and coefficients of variation for each of the ten comparable sets. Panel A of our Table 1 reports the results. Except for minor rounding differences, the comparables selected and the optimal comparable weights derived from our approach are identical to those reported by Vandell.

**Comparable Inclusion Process and Optimal Number of Comparables**

Vandell uses evidence from simulations to show that the variance of the final value decreases as the number of comparables in the comparable set increases. We formally prove this proposition using a covariance matrix partition method. From the partitioned matrix, we are able to derive a more robust comparable inclusion process. The objective of this inclusion process is to identify one more comparable (from an available pool of \( n \) comparables) to include in an existing comparable set (with \( m \) comparables) so that the final value estimate of the new comparable set (with \( m + 1 \) comparables) will have the least variance.

Given a comparable set with \( m \) comparables, the variance of the final value estimate under optimal condition, \( \omega_{mr}^\top \Omega^{-1} e/(e^\top \Omega^{-1} e) \), is determined by \( \omega_{mr}^\top \Omega \omega_{mr} = e^\top \Omega^{-1} \Omega \Omega^{-1} e/(e^\top \Omega^{-1} e)^2 - 1/(e^\top \Omega^{-1} e) \). That is, the variance of the minimum variance weight, \( \omega_{mr} \), is determined by the reciprocal of the sum of all elements in the inverse matrix \( \Omega^{-1} \). Assume that there is a comparable set with \( m \) comparables (their covariance matrix is denoted by \( \Omega_{11} \)) and there is one more comparable (variance = \( \sigma^2 \)) to be included in that comparable set, the variance of the new comparable set (with \( m + 1 \) comparables) under optimal condition will be determined by the inverse of the matrix\(^{10}\)

---

\(^9\) Using the popular Lotus 1-2-3 and QSB linear programming package, we solved this example in less than 5 minutes. Even if the number of comparables in the example increases significantly, the incremental time to perform the analysis will be marginal.

\(^{10}\) For the derivation of the inverse partitioned matrix, see Johnston (1984, p. 135).
Table 1
Optimal Weight, Final Value Estimate, Estimated Standard Error, and Its Coefficient of Variation for $q = 1, \ldots, 10$ Comparable (using Vandell’s and our methodologies)

<table>
<thead>
<tr>
<th>Number of Comparables</th>
<th>$w_1$</th>
<th>$w_2$</th>
<th>$w_3$</th>
<th>$w_4$</th>
<th>$w_5$</th>
<th>$w_6$</th>
<th>$w_7$</th>
<th>$w_8$</th>
<th>$w_9$</th>
<th>$w_{10}$</th>
<th>Estimated Final Value Estimate ($$)</th>
<th>Standard Error of Final Value Estimate ($$)</th>
<th>Coefficient of Variation (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.0000</td>
<td>0.5073</td>
<td>0.4927</td>
<td>0.3301</td>
<td>0.3615</td>
<td>0.3085</td>
<td>0.2713</td>
<td>0.3198</td>
<td>0.2308</td>
<td>0.1781</td>
<td>256,662</td>
<td>11,136</td>
<td>4.3386</td>
</tr>
<tr>
<td>2</td>
<td>0.5073</td>
<td>0.4927</td>
<td>0.3301</td>
<td>0.3615</td>
<td>0.3085</td>
<td>0.2713</td>
<td>0.3198</td>
<td>0.2308</td>
<td>0.1781</td>
<td>0.1964</td>
<td>250,837</td>
<td>8,088</td>
<td>3.2243</td>
</tr>
<tr>
<td>3</td>
<td>0.3301</td>
<td>0.3615</td>
<td>0.3085</td>
<td>0.2713</td>
<td>0.3198</td>
<td>0.2308</td>
<td>0.1781</td>
<td>0.1964</td>
<td>0.1885</td>
<td>0.2161</td>
<td>244,400</td>
<td>6,765</td>
<td>2.7680</td>
</tr>
<tr>
<td>4</td>
<td>0.2713</td>
<td>0.3198</td>
<td>0.2308</td>
<td>0.1781</td>
<td>0.1964</td>
<td>0.1885</td>
<td>0.2161</td>
<td>0.1523</td>
<td>0.2066</td>
<td>0.2219</td>
<td>241,545</td>
<td>6,353</td>
<td>2.6301</td>
</tr>
<tr>
<td>5</td>
<td>0.1964</td>
<td>0.1885</td>
<td>0.2161</td>
<td>0.1523</td>
<td>0.2066</td>
<td>0.2219</td>
<td>0.2037</td>
<td>0.1233</td>
<td>0.1052</td>
<td>0.1516</td>
<td>241,150</td>
<td>5,132</td>
<td>2.1282</td>
</tr>
<tr>
<td>6</td>
<td>0.1844</td>
<td>0.1514</td>
<td>0.2219</td>
<td>0.2037</td>
<td>0.1233</td>
<td>0.1052</td>
<td>0.1516</td>
<td>0.1587</td>
<td>0.1523</td>
<td>0.1516</td>
<td>239,674</td>
<td>4,876</td>
<td>2.0346</td>
</tr>
<tr>
<td>7</td>
<td>0.1341</td>
<td>0.1349</td>
<td>0.1423</td>
<td>0.1260</td>
<td>0.1516</td>
<td>0.1587</td>
<td>0.1523</td>
<td>0.1069</td>
<td>0.1760</td>
<td>0.1523</td>
<td>237,666</td>
<td>4,224</td>
<td>1.7774</td>
</tr>
<tr>
<td>8</td>
<td>0.1226</td>
<td>0.1166</td>
<td>0.1394</td>
<td>0.1258</td>
<td>0.1105</td>
<td>0.1069</td>
<td>0.1760</td>
<td>0.1760</td>
<td>0.1021</td>
<td>0.1239</td>
<td>234,321</td>
<td>3,991</td>
<td>1.7031</td>
</tr>
<tr>
<td>9</td>
<td>0.1217</td>
<td>0.1139</td>
<td>0.1390</td>
<td>0.1283</td>
<td>0.0918</td>
<td>0.0850</td>
<td>0.1897</td>
<td>0.0778</td>
<td>0.0527</td>
<td>0.1390</td>
<td>234,207</td>
<td>3,922</td>
<td>1.6745</td>
</tr>
<tr>
<td>10</td>
<td>0.1213</td>
<td>0.1124</td>
<td>0.1402</td>
<td>0.1296</td>
<td>0.0866</td>
<td>0.0782</td>
<td>0.1956</td>
<td>0.0697</td>
<td>0.0421</td>
<td>0.0242</td>
<td>233,553</td>
<td>3,909</td>
<td>1.6735</td>
</tr>
</tbody>
</table>

Panel A: with minimum variance objective function and Vandell’s inclusion process.
Table 1 (continued)

Panel B: with minimum variance objective function and our inclusion process.

<p>| | | | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.0000</td>
<td>256,662</td>
<td>11,136</td>
<td>4.3386</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.5073 0.4927</td>
<td>250,837</td>
<td>8,088</td>
<td>3.2243</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.3301 0.3615 0.3085</td>
<td>244,400</td>
<td>6,765</td>
<td>2.7680</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.2623 0.2377 0.3003 0.1998</td>
<td>244,239</td>
<td>5,712</td>
<td>2.3389</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.1935 0.2054 0.1876 0.2694 0.1442</td>
<td>241,781</td>
<td>5,066</td>
<td>2.0953</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.1597 0.1581 0.1696 0.1612 0.1630 0.1885</td>
<td>239,021</td>
<td>4,487</td>
<td>1.8772</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.1341 0.1349 0.1423 0.1260 0.1516 0.1587 0.1523</td>
<td>237,666</td>
<td>4,224</td>
<td>1.7774</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.1226 0.1165 0.1394 0.1258 0.1105 0.1069 0.1760 0.1021</td>
<td>234,321</td>
<td>3,951</td>
<td>1.7031</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.1217 0.1139 0.1390 0.1283 0.0918 0.0850 0.1897 0.0778 0.0527</td>
<td>234,207</td>
<td>3,922</td>
<td>1.6745</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.1213 0.1124 0.1402 0.1296 0.0866 0.0782 0.1956 0.0697 0.0421 0.0242</td>
<td>233,553</td>
<td>3,909</td>
<td>1.6735</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel C: with minimum coefficient of variation function.

<p>| | | | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.0000</td>
<td>256,662</td>
<td>11,136</td>
<td>4.3386</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.5200 0.4800</td>
<td>250,988</td>
<td>8,090</td>
<td>3.2234</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.3501 0.3603 0.2895</td>
<td>244,903</td>
<td>6,772</td>
<td>2.7651</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.2943 0.3225 0.2208 0.1624</td>
<td>242,246</td>
<td>6,362</td>
<td>2.6262</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.2117 0.1906 0.2096 0.1820 0.2060</td>
<td>241,610</td>
<td>5,137</td>
<td>2.1262</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.1990 0.1649 0.2157 0.1937 0.1275 0.0992</td>
<td>240,176</td>
<td>4,881</td>
<td>2.0325</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.1464 0.1382 0.1396 0.1204 0.1540 0.1528 0.1485</td>
<td>238,096</td>
<td>4,228</td>
<td>1.7758</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.1352 0.1217 0.1373 0.1207 0.1176 0.1077 0.1697 0.0900</td>
<td>233,107</td>
<td>3,997</td>
<td>1.7003</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.1338 0.1190 0.1370 0.1234 0.0989 0.0861 0.1835 0.0665 0.0520</td>
<td>234,969</td>
<td>3,928</td>
<td>1.6717</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.1336 0.1181 0.1377 0.1241 0.0958 0.0821 0.1869 0.0618 0.0458 0.0142</td>
<td>234,583</td>
<td>3,917</td>
<td>1.6698</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
\[ e^T \Omega^{-1} e = \begin{bmatrix} e_1^T & \Omega_{11} & \Omega_{12} \end{bmatrix}^{-1} \begin{bmatrix} e_1 \\ \Omega_{21} \\ \sigma_2^2 \end{bmatrix} \begin{bmatrix} e_1 \\ 1 \end{bmatrix} \] (8)

That is,

\[ \omega_{mr}^T \Omega \omega_{mr} = \frac{1}{e_1^T \Omega_{11}^{-1} e_1 + (e_1^T \Omega_{11}^{-1} \Omega_{12} - 1)^2 \beta_{22}}, \] (9)

where \((e_1^T, 1) = e^T\), and \(\beta_{22} = (\sigma_2^2 - \Omega_{21} \Omega_{11}^{-1} \Omega_{12})^{-1}\). Since \(\Omega\) is a positive definite matrix, its inverse matrix \(\Omega^{-1}\) must be positive definite as well. This implies that \(\beta_{22}\) must be positive. Furthermore,

\[ (e_1^T \Omega_{11}^{-1} \Omega_{12} - 1)^2 = (e_1^T \Omega_{11}^{-1} e_1)^2 \left[ \frac{e_1^T \Omega_{11}^{-1} \Omega_{12}}{e_1^T \Omega_{11}^{-1} e_1} - \frac{1}{e_1^T \Omega_{11}^{-1} e_1} \right]^2, \] (10)

\[ = (e_1^T \Omega_{11}^{-1} e_1)^2 (\omega_{mr} \Omega_{12} - \omega_{mr}^T \Omega_{11})^2, \]

\[ = (e_1^T \Omega_{11}^{-1} e_1)^2 [\text{Cov}(1 \hat{V}_{mr}, \hat{V}_2) - \text{Cov}(1 \hat{V}_{mr}, 1 \hat{V}_{mr})]^2, \] (11)

where \(1 \hat{V}_{mr} = \omega_{mr}^T \hat{V}\) and \(\omega_{mr}^T = \Omega_{11}^{-1} e_1 / e_1^T \Omega_{11}^{-1} e_1\). Note that \(\omega_{mr}^T\) is the minimum variance weight based on the covariance matrix \(\Omega_{11}\). From equations (9) and (11), it is clear that the inclusion of an additional comparable, under the minimum variance criterion, will decrease the variance of the comparable set unless \(\text{Cov}(1 \hat{V}_{mr}, \hat{V}_2) = \text{Cov}(1 \hat{V}_{mr}, 1 \hat{V}_{mr})\). The following proposition summarizes equations (9) and (11).

**Proposition**: The necessary and sufficient condition for the proposition—the inclusion of an additional comparable will always reduce the variance of the optimal final value estimate—is that the covariance of the original optimal comparable set and the additional comparable (i.e., \(\text{Cov}(1 \hat{V}_{mr}, \hat{V}_2)\)) is not equal to the variance of the original optimal comparable set (i.e., \(1 \sigma_{mr}^2 = \text{Cov}(1 \hat{V}_{mr}, 1 \hat{V}_{mr})\)). A sufficient condition for \(\text{Cov}(1 \hat{V}_{mr}, \hat{V}_2)\) to be equal to \(1 \sigma_{mr}^2\) is that \(\hat{V}_{mr}\) and \(\hat{V}_2\) are perfectly positive correlated. In this scenario, adding \(\hat{V}_2\) to the weight selection will not reduce the variance of the original comparable set. However, under no condition will the inclusion of an additional comparable increase the variance of the optimal comparable set.
The positive $B_{22}$ and equation (9) establish the rules of comparable inclusion. More specifically, for the inclusion of an additional property into the existing comparable set (with $m$ comparables), the property chosen from the available pool of $n$ comparables would be the one that maximizes $(e_1^T \Omega_{11}^{-1} \Omega_{12} - 1)^2 B_{22}$. This is true because the variance of the original $m$ comparable set under the optimal condition is given by $1/e_1^T \Omega_{11}^{-1} e_1$. Maximizing $(e_1^T \Omega_{11}^{-1} \Omega_{12} - 1)^2 B_{22}$ is equivalent to minimizing $\omega_{nr}^T \Omega \omega_{nr}$, the variance of the new comparable set with $m + 1$ comparables under optimal condition.

Panel B of Table 1 reports the results using our proposed comparable inclusion process. The comparables selected and the optimal comparable weights derived from our approach are different from that reported by Vandell, when the number of comparables in the comparable set are restricted to 4, 5, and 6. For these three comparable sets, our method provides lower standard errors and coefficients of variation than those reported by Vandell.

**Comparable Selection with Minimum Coefficient of Variation Criterion**

In this section, we develop a comparable selection process using the coefficient of variation as the selection criterion. The objective is to select a vector of non-negative weights that provides the least variation, on a per dollar basis, for the final value estimate. Minimizing the coefficient of variation through the optimal selection of non-negative comparable weights can be formulated as the following constrained minimization problem:

$$\text{Min } (\omega^T \Omega \omega)^{1/2}/\omega^T V,$$

subject to

$$\omega^T e - \sum_{t=1}^n \omega_t = 1,$$

and

$\text{11 With an available pool of 10 comparables, this procedure can be implemented in less than 3 minutes using Lotus 1-2-3.}$

$\text{12 Using Lotus 1-2-3, we obtained the results reported in this panel in less than an hour.}$

$\text{13 See Vandell’s (1991) Table 5 on page 233.}$
\[ \omega \geq 0. \quad (14) \]

By applying the Lagrangian multipliers \( \lambda \) and \( \Gamma \) to constraints (13) and (14), the Lagrangian of (15) is given by

\[ L = (\omega^T \Omega \omega)^{1/2} (\omega^T V - \lambda (\omega^T e - 1)) - \omega^T \Gamma. \quad (15) \]

The first-order conditions of (15) are:

\[ \Omega \omega (\omega^T \Omega \omega)^{-1/2} (\omega^T V)^{-1} - V (\omega^T \Omega \omega)^{1/2} (\omega^T V)^{-1} - \lambda e - \Gamma = 0, \quad (16) \]

\[ \omega^T e = 1 \]

\[ \omega_i \Gamma_i = 0 \quad \text{for all } i. \]

Multiplying the optimal weight \( \omega^T \) on both sides of (16) implies that the Lagrangian \( \lambda = 0 \). This is true because the objective function is scale-free and both \( e \) and \( \omega \) are strictly positive vectors. Given \( \lambda = 0 \), equation (16) can be rewritten as

\[ \Omega \omega - V (\omega^T \Omega \omega) (\omega^T V)^{-1} - (\omega^T \Omega \omega)^{1/2} (\omega^T V) \Gamma = 0. \quad (17) \]

Note that \( \omega^T \Omega \omega \) must be positive because \( \Omega \) is a positive definite matrix. The expected adjusted sales price, \( V \), must be positive given the constraint on negative weights. Hence, the last term in (17) is simply a vector \( \Gamma \) multiplied by a positive scalar. To solve the minimum coefficient of variation weight, we multiply \( e^T \Omega^{-1} \) on both sides of (17) and apply (13) to obtain

\[ (\omega^T \Omega \omega)(\omega^T V)^{-1} = [1 - (\omega^T \Omega \omega)^{1/2} (\omega^T V) e^T \Omega^{-1} V] e^T \Omega^{-1} V. \quad (18) \]

Substituting \( (\omega^T \Omega \omega)(\omega^T V)^{-1} \) into (17) and then multiplying by \( \Omega^{-1} \), the minimum coefficient of variation weight \( \omega_{cv}^* \) is determined by

\[ \omega_{cv}^* = \Omega^{-1} V / \rho^T \Omega^{-1} V + [1 - \Omega^{-1} V \rho^T / \rho^T \Omega^{-1} V] \Omega^{-1} \Gamma (\omega^T \Omega \omega)^{1/2} (\omega^T V) \]

\[ = \omega_{mv} + Ay, \quad (19) \]

\[ ^{14} \text{Note that } \Gamma \text{ must be non-negative. Furthermore, (15) must be greater than or equal to 0 in the absence of } \Gamma \text{ because of the Kuhn–Tucker condition for minimizing (12).} \]
where \( \omega_{mv} = \Omega^{-1} V e^T \Omega^{-1} V, \quad A = \Omega^{-1} V e^T \Omega^{-1} V, \quad \Omega^{-1} = \Omega^{-1} \epsilon^T \), and \( \epsilon = (\omega^T \Omega \omega)^{1/2} (\omega^T V) \). The first term \((\omega_{mv})\) on the right-hand side of (19) represents the optimal weight under the mean-variance framework without non-negativity restrictions, which is derived by minimizing the variance subject to an expected adjusted sales price. The last term \((Ay)\) is the vector that prevents the negative weight in \( \omega^*_v \). Note that \( e^T A = 0 \). The zero sum on the vector implies that \( Ay \) is a net zero weight vector.

As shown in equation (19), the mean-variance weight \((\omega_{mv})\) can be obtained easily by the product of the inverse matrix \( \Omega^{-1} \) and the expected adjusted sales price vector \( V \). The matrix \( A \) is determined by the inverse covariance matrix and the mean-variance weight. Given \( A \) and \( \omega_{mv} \), equation (19) is a simultaneous linear equation system with variables \( \omega^*_v \) and \( y \) and with constraints specifying that \( \omega_i y_i = 0 \) for all \( i \). In other words, similar to the previous section, \( \omega^*_v \) in (19) can be derived simply by solving a linear program with variables \( \omega^*_v \) and \( y \) using the big \( M \) method (minimizing \( y^T M \) subject to (19)).

Again, this technique can be illustrated with a simple example. Assume that there are three comparables and let

\[
\Omega = \begin{pmatrix}
0.8 & 0.5 & 0.7 \\
0.5 & 1.1 & 0.8 \\
0.7 & 0.8 & 0.9
\end{pmatrix}
\]

\[
V = \begin{pmatrix}
2.56662 \\
2.44840 \\
2.30762
\end{pmatrix}
\]

The optimal weights without a non-negativity constraint are

\[
\omega_{mv} = \frac{\Omega^{-1} V e^T \Omega^{-1} V}{e^T \Omega^{-1} V} = \begin{pmatrix}
\omega_1 \\
\omega_2 \\
\omega_3
\end{pmatrix} = \begin{pmatrix}
1.10428 \\
0.62019 \\
-0.72447
\end{pmatrix}.
\]
and the net zero weight matrix is

\[
A = (I - \omega_m e^T)\Omega^{-1} = \begin{pmatrix}
3.29757 & 0.77087 & -4.43252 \\
0.71293 & 2.61830 & -3.57098 \\
-4.01050 & -3.33917 & 8.00350 \\
\end{pmatrix}.
\]

The optimal weight selection with a non-negativity constraint can then be specified as the following constrained minimization problem.\(^{15}\)

Min 500 \(y_1 + 500 \cdot y_2 + 500 \cdot y_3\),

subject to

\[
\begin{align*}
\omega_1 &= 3.29757 \cdot y_1 - 0.77087 \cdot y_2 + 4.43252 \cdot y_3 = 1.10428 \\
\omega_2 &= -0.71293 \cdot y_1 - 2.61830 \cdot y_2 + 3.57098 \cdot y_3 = 0.62019 \\
\omega_3 &= 4.01050 \cdot y_1 + 3.33917 \cdot y_2 - 8.00350 \cdot y_3 = -0.72447
\end{align*}
\]

\(\omega_1, \omega_2, \omega_3, y_1, y_2, y_3 \geq 0\).

Under this minimization program, the resulting optimal weights for \(\omega_1\), \(\omega_2\), and \(\omega_3\) are 0.703, 0.297, and 0, respectively. It is worthwhile to note that this result is identical to that derived using Vandell's three-stage procedure.

Using Vandell's revised data, Panel C of our Table 1 reports the optimal weights, final value estimates, standard errors of final value estimates, and coefficients of variation using the minimum coefficient of variation criterion. Although the standard errors of the final value estimates derived using our proposed criterion are higher than those using the minimum variance approach (see Vandell's revised Table 5 on page 153), the estimated final value estimates contain less variation on a per dollar basis. It should also be noted that the final value estimates derived using the minimum coefficient of variation criterion are higher than those derived using the minimum variance criterion.

\(^{15}\)The variable \(y_i\) is similar to the variable \(f_i\) used in the previous example. It should be noted that the coefficient, 500, used could be any large number. See footnote 7 for a detailed discussion.
Conclusions

The comparable and weighting selection approach proposed by Vandell represents a significant contribution to an important but largely ignored area in the real estate literature. The sales comparison approach traditionally has been viewed by the public as an "art" rather than a "science". With the technique developed by Vandell, the subjectivity associated with the comparable and weight selections can be minimized, and the sales comparison approach is one step closer to being viewed as a scientific approach.

In this article, we use a different comparable selection criterion to develop a formal approach for the optimal comparable and weighting selections in the sales comparison approach. Our method provides an easy solution for including the non-negative weight constraint in the comparable search process. It also provides a more robust way for comparable inclusion and results in a final value estimate with the least error on a per dollar basis. These refinements sharpen the Vandell methodology.

Future research in this area should focus on the empirical applications of the proposed technique, with special emphasis on the trade-off relationship between the marginal gain in variance reduction and the marginal cost of including an additional comparable. For financial assets, Statman (1987) indicates that the maximum diversification benefits can be achieved with less than thirty securities in a portfolio when the costs of trading are taken into consideration. Empirical research on the optimal number of comparables, with consideration of search costs and regulatory constraints, would be extremely helpful in advancing our knowledge in this field.

The authors wish to thank Kerry Vandell (co-editor) and an anonymous referee for their helpful comments.

References


