Estimating Property Value by Replicating One

Tsong-Yue Lai, Kerry Vandell, and Ko Wang *

October 2003

*Tsong-Yue Lai is Professor, Department of Finance, California State University-Fullerton Fullerton, CA 92834 (tylai@fullerton.edu). Kerry Vandell is Professor, Department of Real Estate, University of Wisconsin-Madison. Ko Wang is Professor, Department of Finance, California State University-Fullerton.
Abstract
This paper develops a Replication method for estimating property values. Under this method, an appraiser’s assignment is to estimate an optimal weight vector that can best replicate the property attributes of a subject property using comparable properties. Once the weights are determined, the value of the subject property is simply a linear combination of the values of the comparable properties based on the estimated weight vector. Since there is no need to estimate hedonic prices (or adjustment factors) of property attributes, the Replication method is superior to the traditional Regression method when there are multicollinearity and misspecification problems. The Replication method also outperforms the Grid method when the number of comparables used in the valuation is no less than the number of property attributes identified for the comparable properties. Keywords: Real Estate
1 Introduction

The valuation of real estate assets (properties) differs dramatically from the valuation of financial assets (stocks and bonds). When developing a valuation technique for financial assets (such as the Capital Asset Pricing Model or the Arbitrage Pricing Theory), data availability is rarely the main consideration and the models derived normally have a closed-form solution. Individuals who use the models are only responsible for finding the required parameter values as inputs for the model. As long as the inputs are the same, the results derived from the model should be the same regardless of who performs the valuation assignment. Given this, there is little subjectivity in the valuation process.¹

However, the situation is different for real estate assets. Because of location boundedness, non-homogeneous production, and non-continuous trading, the quantity and quality of data available in the field might not allow individuals to use a model that offer a closed-form solution. For example, the Regression method (or the Hedonic Pricing Model), which has been advocated by academicians as an objective valuation method, has not received a warm welcome from practitioners in the field. Instead, practitioners prefer to use the Grid method and are willing to tolerate the subjectivity inherent in this particular valuation process.

Simply put, the Regression method can be described as a shopping-cart approach. Under this method, each attribute of a property has a price tag. The price of a property is determined by the price and quantity of the attributes of the property. In other words, when shopping for a property, a person selects the attributes she/he wants for the subject property and puts those attributes into a shopping cart. The customer pays a total price (the estimated property value) at the check-out point based on the attributes (and the price of each attribute) the person selected for the property. When implementing this technique, one must estimate the price tag (or β, in statistical terms) for each property attribute. However, it is well-known that the estimation of β requires the use of massive quantity of quality data. It

¹Of course, selections of input still depend upon individual judgments.
might be fair to say that the data requirement has prevented the Regression technique from being widely used in the industry.

The Grid method can best be described as a cut-and-paste approach. Similar to the Regression method, the Grid method also assumes that each property attribute has a price tag. However, to implement the Grid method, one must also find a number of properties that have attributes similar to the subject property. The Grid method assumes that a person bought a comparable property (with a known price) first. The person will then sell (or buy) the attributes she/he dislikes (or likes) for the property. After this cut-and-paste process, the comparable property should be similar to the subject property in terms of property attributes. The price of the subject property (with the attributes the person wants) should be the price of the comparable property adjusted for the attributes deleted (or added). Since there are several comparable properties, this process should be repeated for each of the comparable properties. At the end, a weighted average of the adjusted comparable prices represents the price that one is willing to pay for the subject property.

As one can see quite clearly, the Grid method requires an appraiser to estimate both the weights for the comparables as well as the price of each property attribute. A major drawback of this technique is that there is little guidance on how comparables and weights should be selected. In other words, there is substantial subjectivity in this valuation process. However, when compared to the Regression method, the benefit of this approach is that an appraiser is able to select more carefully the comparables used for the evaluation of the subject properties. Because of the ability to screen the comparable properties in great detail, it is not surprising if the Grid method perform better than the Regression method under certain circumstances. This is especially true if the properties used in estimating $\beta$ are not really comparable to each other (and, therefore, the estimated $\beta$s are not reliable).

The purpose of this paper is to develop a valuation technique for real assets that can remove subjectivity from the valuation process while accommodating the inherent data constraints present in the real estate field. We call this technique the Replication method. Specifically, the Replication method replicates the property attributes of the subject property using comparable properties by finding an optimal weight for each of the comparable properties. The estimated price of the subject property is then calculated based on the weights and prices of the comparable properties. It should be noted that the central theme of both the Regression and the Grid methods is also to
duplicate the subject property by adjusting for the attributes of properties. However, both methods require the estimation of the price of each property attribute. Our proposed Replication method takes the same philosophy of duplicating a subject property, but without the need to estimate the prices of attributes. Clearly, the Replication technique retains the benefits of the Regression and Grid methods, while avoiding the problems associated with the estimation of hedonic prices (and, therefore, the need to have a large quantity of quality data).

Section 2 introduces the Replication method and provides a theoretical foundation for the approach. Section 3 details the estimation issues. Sections 4 and 5 compare the statistical properties of the Replication method with that of the Regression method and the Grid method. In these two sections, we demonstrate that the proposed Replication method performs better or no worse than the Regression method and the Grid method under realistic scenarios. Implementation issues of the Replication method are provided in section 6. Section 7 contains our conclusions.

2 The Replication Method

Our proposed Replication method can be accomplished in two simple steps. The first step is to estimate a 1xn optimal vector \( \omega_{rp} \) such that

\[
\omega_{rp}X_C = X_S.
\]

\( X_C \) is a \( nxk \) matrix representing the number of property attributes and the number of comparable properties used in the valuation assignment, where \( n \) is the number of comparable properties and \( k \) is the number of property attributes. \( X_S \) is a 1xn vector representing the \( k \) property attributes of the subject property. The vector \( \omega_{rp} \) is a 1xn vector representing the optimal vector of the comparable properties to be estimated by an appraiser. Simply put, at this stage the task of an appraiser is to estimate the optimal vector \( \omega_{rp} \) that can best duplicate the property attributes of the subject property \( X_S \) using information (on property attributes) from comparable properties \( X_C \).

\(^2\)For any 1xn vector \( G \) and any kxn matrix \( \Phi \) with rank \( k \), all solutions for equation (1) must be in the form of \( \omega_{rp} = G + [X_S - GX_C](\Phi X_C)^{-1}\Phi \).
Once the optimal weight vector $\omega_{rp}$ is estimated, the second step of the Replication method is to estimate the value of the subject property using the optimal weights and the prices of comparable properties, or

$$\hat{V}_{S(rp)} = \omega_{rp}V_C,$$  \hspace{1cm} (2)

where $V_C$ is a $n \times 1$ vector representing the prices of $n$ comparable properties and $\omega_{rp}$ is the optimal weight derived from equation (1). We define $\hat{V}_{S(rp)}$ as the estimated value of the subject property using the Replication method.

The theoretical foundation of our Replication method is based on one simple assumption, or

$$V = X\beta.$$ \hspace{1cm} (3)

This assumption simply says that property values (of the subject property and all comparable properties) can be estimated by a set of property attributes and the prices of these attributes without errors.

Under this certainty case, the value of a comparable property is determined by

$$V_C = X_C\beta,$$ \hspace{1cm} (4)

while the value of the subject property $V_S$ is

$$V_S = X_S\beta.$$ \hspace{1cm} (5)

If we can obtain a $\omega_{rp}$ that is a unique solution satisfying $\omega_{rp}X_C = X_S$ (see equation (1)), then equation (2) can be re-written as

$$\hat{V}_{S(rp)} = \omega_{rp}V_C = \omega_{rp}X_C\beta = X_S\beta = V_S.$$ \hspace{1cm} (6)

This simple exercise indicates that, as long as the assumption $V = X\beta$ holds and if $\omega_{rp}$ is a unique solution, then the true value of the subject property can always be estimated by our Replication method without errors. Clearly, this is the most desirable condition that anyone can hope for any valuation technique under certainty.
However, in reality, we are working under uncertainty and property values are estimated with errors. When applying the Replication method using real world data, there will often be multiple weight vectors $\omega_{rp}$ that can satisfy equation (1). When this happens, the estimated property value based on our Replication method ($\hat{V}_{S(rp)} = \omega_{rp}V_C$, or see equation (2)) will not be unique. When there are multiple solutions for one problem, then the final decision must subject to errors. However, will this problem make our proposed method un-attractive? The answer is a no because all valuation techniques available in the field are also subject to errors. Consequently, we will have to judge the usefulness of an estimation technique based on established criteria. Given this, a more important question to be addressed is "how does the Replication method perform when compared with the Regression method and the Grid method?".

When compared to the Regression and Grid methods, the advantage of using our Replication technique can be seen clearly from equations(1) and (2). This two-step estimation procedure involves only one estimation: the weights of the comparable properties, or $\omega_{rp}$. Intuitively, we know that the Replication method should be better than the Grid method since the use of the Grid method involves one more estimation (the estimation of adjustment factors $\beta$). Although both the Regression method (involving the estimation of $\beta$) and the Replication method (involving the estimation of $\omega_{rp}$) require one estimation, the use of the Replication should be preferred since the weight is an exact solution while the estimation of $\beta$ in the Regression method is based on a best-fit criterion (which must be accompanied with estimation errors). Clearly, additional errors can be introduced through the estimation of the $\beta$ when the Regression method is used.

With the problems of the Grid method and the Regression method in mind, we are now in a position to discuss the implementation issues of the Replication method. This will be done in the next section.

3 Estimation

Equation (1) is basically a system that solves simultaneous $k$ linear equations of $\omega$ with $n$ unknown variables. When solving this system of equations, we know that $k = n$ (number of property attributes equals number of comparables) and $X_C$ is a full rank matrix are the two required conditions for $\omega_{rp}$ to be a unique solution of equation (1). If we have more property attributes
than comparable properties (or when \( k > n \)), there is no guarantee that we can find a weight vector that satisfies equation (1). If we have more comparable properties than property attributes (or \( k < n \)), there will be an infinite number of weight vectors that satisfies equation (1). In the first case, an appraiser just selects the one and only answer as the optimal weight. In the other two cases (\( k > n \) and \( k < n \)), an appraiser will have to select an optimal weight vector \( \omega_{rp} \) from alternatives. We will start with the most likely scenario where \( k < n \).

3.1 Standard case: when \( k < n \)

From equations (1) and (2), we know that for every \( \omega \) that satisfies equation (1) we can obtain a price for the subject property using \( V_{S(rp)} = \omega V_C \). We also know that when \( k < n \), there is an infinite number of solutions \( \omega \) that satisfies equation (1). Since multiple weight vectors satisfy the condition set by equation (1), an appraiser has to impose additional criteria to select the optimal one among alternatives. We propose to use unbiasedness and minimum variance as the criteria.

To implement these criteria and to make our comparison with the Regression and Grid methods meaningful, we first make a standard assumption that the true property value (of subject property or comparable properties) can be estimated as a linear function of its property attributes, so that

\[
V = X\beta + \epsilon, \quad (7)
\]

which implies

\[
V_C = X_C\beta + \epsilon_C \quad (8)
\]

and

\[\text{Since every set of } k \text{ comparables can produce an exact solution, there will be } n!/[n-k)! \ast k!] \text{ weight vectors that can satisfy equation (1) if there are } n \text{ comparables. (For example, if we have 5 equations and 4 unknowns, we will have 5 possible solutions when solving this system of equations.) Since linear combinations of those } n!/[n-k)! \ast k!] \text{ weight vectors can also satisfy equation (1), the solutions are infinite. To see this, let } \omega_1 \text{ and } \omega_2 \text{ be the two solutions satisfying equation (1). It is clear that any point along the line } a\omega_1 + (1-a)\omega_2 \text{ is also a solution because } [a\omega_1 + (1-a)\omega_2] X_C = aX_S + (1-a)X_S = X_S.\]
\[ V_S = X_S \beta + \epsilon_S, \] (9)

where \( X \) is a \((n + 1) \times k\) matrix representing \( k \) property attributes of the subject property and \( n + 1 \) comparable and subject properties, \( \beta \) is a \( k \times 1 \) vector of unknown parameters, \( \epsilon \) is a \((n + 1) \times k\) vector of error terms, \( \epsilon^T = (\epsilon_S^T, \epsilon) \), \( X_C \) is a \( n \times k \) matrix representing \( k \) property attributes and \( n \) comparable properties, and \( X_S \) is a \( 1 \times k \) vector representing the \( k \) property attributes of the subject property. We also make a standard assumption that the mean of \( \epsilon \) is a zero vector. Note that this is also the standard assumption behind the Regression method.

Since the true property value is generated via equation (7), given any weight vector \( \omega \) (derived from equation (1)), the prediction error should be the difference between the true value \( V_S \) (which is unknown) and the estimated value \( \hat{V}_{S(rp)} = \omega V_C \) (derived from equation (2)), or

\[ \epsilon_{rp} = V_S - \hat{V}_{S(rp)} = V_S - \omega V_C = X_S \beta + \epsilon_S - [\omega(X_C \beta + \epsilon_C)] = \epsilon_S - \omega \epsilon_C. \] (10)

It is important to note that in equation (10) the prediction error \( \epsilon_{rp} \) is independent of the unknown parameters \( \beta \) and relies only on the estimated weight vector \( \omega \). That is, when the Replication method is used, the problems (such as multicollinearity and misspecifications) occurring in the estimation of \( \beta \) in the Regression method can be circumvented.

Given the zero mean assumption on error terms, the expectation of the prediction error \( \epsilon_{rp} \) is given by

\[ E(\epsilon_{rp}) = E(\epsilon_S - \omega \epsilon_C) = 0, \] (11)

where \( E(.) \) is an expectation operator. Equation (11) shows that all Replication estimators have an expected zero error term and are unbiased estimators of the true property value. We can summarize those information in the following proposition.

\footnote{A notation of \(^T\) denotes the transpose vector of a matrix.}
Proposition 1  Under the assumption that the true property value is a linear function of its property attributes with a zero mean error term (as specified in equation (7)), all Replication estimators satisfying equations (1) and (2) are unbiased.

Since all Replication estimators using weight vectors satisfying equation (1) are unbiased, we will have to select the optimal weight vector using the minimum variance criterion. To do this, we minimize the expected square prediction error, or

\[
\text{Min } E(\epsilon_S - \omega \epsilon_C)^2 = \text{Min } [\sigma^2_S - 2\Omega_{CS} \omega^T + \omega \Omega \omega^T]
\]  (12)

subject to

\[
\omega X_C = X_S,
\]

where \(\sigma^2_S\) is the variance of \(\epsilon_S\), \(\Omega\) is \(nxn\) covariance matrix of \(\epsilon_C\), and \(\Omega_{CS}\) is a \(1xn\) covariance vector between \(\epsilon_C\) and \(\epsilon_S\). The optimal solution to equation (12), given the zero mean assumption of error terms, is

\[
\omega_{rp} = \Omega_{CS} \Omega^{-1} + [X_S - \Omega_{CS} \Omega^{-1} X_C](X_C^T \Omega^{-1} X_C)^{-1} X_C^T \Omega^{-1}. \]  (13)

Equation (13) is a necessary and sufficient condition for minimizing equation (12) because \(\Omega\) is a positive definite matrix in quadratic mathematical programming. Given equation (13), we can derive the value of the subject property as

\[
\hat{V}_{S(rp)} = \omega_{rp} V_C = \Omega_{CS} \Omega^{-1} V_C + [X_S - \Omega_{CS} \Omega^{-1} X_C](X_C^T \Omega^{-1} X_C)^{-1} X_C^T \Omega^{-1} V_C. \]  (14)

Defining \(\sigma^2_{(rp)}\) as the variance of the prediction errors by the Replication method and combining equation (12) with equation (13), we obtain

\[
\sigma^2_{(rp)} = \sigma^2_S + X_S(X_C^T \Omega^{-1} X_C)^{-1} X_S^T - 2X_S(X_C^T \Omega^{-1} X_C)^{-1}(X_C^T \Omega^{-1} \Omega_{CS}) X_C^T \Omega^{-1} \Omega_{CS}^T \] 
\[- \Omega_{CS} \Omega^{-1} \Omega_{CS}^T + \Omega_{CS} \Omega^{-1} X_C(X_C^T \Omega^{-1} X_C)^{-1} X_C^T \Omega^{-1} \Omega_{CS}. \]  (15)
Although equation (14) provides a general estimation of the property value, the covariance terms in equation (14) might be difficult to estimate in reality. To simplify the task (and to make the comparison with the Regression and Grid methods easier), we further assume that the distributions of error terms $\varepsilon_C$ and $\varepsilon_S$ are iid with zero mean and standard deviation $\sigma$. Under this assumption (similar to the one used in the Regression method), $\Omega = \sigma^2 I$ and $\Omega_{CS}$ become zero vectors. We can then re-write equation (12) as

$$
\min \sigma^2 + E(\omega\varepsilon_C)^2 = \min \sigma^2 + [\omega\omega^T]\sigma^2
$$

subject to

$$
\omega X_C = X_S,
$$

where $\sigma^2$ is the variance of error terms $\varepsilon_C$ and $\varepsilon_S$. The optimal solution of equation (16) can be obtained from equation (13) and is

$$
\omega_{(rp, iid)} = X_S(X_C^TX_C)^{-1}X_C^T.
$$

Consequently, the value of the subject property is

$$
\hat{V}_{S(rp, iid)} = \omega_{(rp, iid)}V_C = X_S(X_C^TX_C)^{-1}X_C^TV_C.
$$

It is clear that, when $\varepsilon_C$ are iid and un-correlated with $\varepsilon_S$, equation (14) collapses to equation (18). We define $\sigma^2_{(rp, iid)}$ as the variance of the prediction errors by the Replication method under the iid assumption. $\sigma^2_{(rp, iid)}$ can be derived by combining equation (16) with equation (18), or simply from equation (15).

Given that $\Omega = \sigma^2 I$ and $\Omega_{CS}$ is a zero vector,

$$
\sigma^2_{(rp, iid)} = \sigma^2 + [\omega_{(rp, iid)}\omega_{(rp, iid)}^T]\sigma^2
= \sigma^2 + \sigma^2 X_S(X_C^TX_C)^{-1}X_C^TX_C(X_C^TX_C)^{-1}X_S^T
= \sigma^2 + \sigma^2 X_S(X_C^TX_C)^{-1}X_S^T.
$$

Equation (19) is important as we will compare it with the prediction errors derived under the Regression and Grid methods in later sections. The findings of this sub-section can be summarized in the following proposition.
Proposition 2 Under the assumption that the true property values are a linear function of their property attributes and have error terms distributed iid with a zero mean, when the number of comparable properties is more than the number of property attributes used in the evaluation, the optimal weight set of a Replication estimator is

$$\omega_{rp} = X_S(X_C^T X_C)^{-1} X_C^T,$$

the value of the subject property is

$$\hat{V}_{S(rp)} = X_S(X_C^T X_C)^{-1} X_C^T V_C,$$

and the variance of the prediction error of a Replication estimator is

$$\sigma^2_{rp} = \sigma^2 + \sigma^2 X_S(X_C^T X_C)^{-1} X_S^T.$$\[3.2 \text{ Special cases: when } k = n \text{ and } k > n \]

When the number of properties equals the number of property attributes ($n = k$) and $X_C$ is a full rank matrix, there exists only a unique solution for equation (1). Under this circumstance, an appraiser should use this particular weight vector to derive the value of the subject property. When there are more property attributes (equations) than the number of comparable properties (variables), or when $k > n$, there is no solution for equation (1) in general. In other words, there is no guarantee that one is able to replicate a subject property using the limited number of comparable properties. However, when there is no exact solution for equation (1), appraisers can still find an optimal solution of weight vector for estimating property values.

The idea is to create a property (using property attributes of comparable properties) that is most similar to the property attributes of the subject property. In other words, while we cannot completely replicate the property attributes of the subject property, we can mimic the property attributes of the subjective property in terms of the minimum sum of squared errors, or

$$\text{Min } (X_S - \omega_{rp} X_C)(X_S - \omega_{rp} X_C)^T. \quad (20)$$

\[5\text{While it is possible that we can duplicate the subject property when } k > n, \text{ in this case the solution will be the same as that discussed in the scenario where } k < n \text{ or } k = n. \] We, therefore, ignore this possibility.
The optimal solution \( \omega_{rp}^* \) that minimizes the sum of squared errors of attributes between the subject and the mimic property defined by equation (20) is

\[
\omega_{rp}^* = X_S X_C^T (X_C X_C^T)^{-1}.
\]  

(21)

The value of the subject property, under this scenario, can be estimated as

\[
\hat{V}_{S(rp)} = \omega_{rp}^* V_C = X_S X_C^T (X_C X_C^T)^{-1} V_C.
\]  

(22)

Although \( \omega_{rp}^* \) in equation (21) is not an exact solution for equation (1), \( \omega_{rp}^* X_C \) represents the best mimic attributes of the subject property in terms of the least sum of squared errors. In other words, while we do not have an exact answer because of the constraint, the assignment is to use the best answer one can find as the solution to the problem.\(^6\)

### 3.3 Comparison issues

While all the three scenarios \((k > n, k = n, \text{and } k < n)\) are discussed in this section, we believe that, in practice, the scenario \(k < n\) might be the one that should receive the most attention. Clearly, when the number of comparables is less than or equal to the number of property attributes, we will not be able to estimate hedonic price vector \(\hat{\beta}\) using a regression equation. Since \(\hat{\beta}\) cannot be estimated, it is not feasible to use both the Regression method and the Grid method.\(^7\) While we might be able to use the Replication method, we will not be able to ascertain how the estimator performs when compared with the Regression and Grid methods.

Given this, in the following sections we will only concentrate our discussion on the scenario where \(k < n\). We will first show that our Replication approach can provide at least as good an estimator as the Regression

---

\(^6\)When applying the Arbitrage Pricing Model (APT) using real world data, Huberman, Kandel and Stambaugh (1987) demonstrate that it is feasible to use mimicking portfolios to substitute the \(k\) factors in an exact Arbitrage Pricing Model. Ang and Lai (1998) show that the factor loading in APT can be replicated and thus there are \(k+1\) risk premiums under the \(k\)-factor model in a finite economy. In this regard, our approach is similar to their approach.

\(^7\)While it might be feasible to use other subjective methods to derive the adjustment factors for the Grid method, we will not be able to make meaningful comparisons with the Regression or the Replication methods under this circumstance since those subjective methods might not have a closed-form equation.
method (when the regression equation is perfectly specified). However, since the Replication method does not require the estimation of \( \beta \), we will show that it is better than the Regression method if equation (7) suffers from multicollinearity and/or misspecification problems. We will also show that our Replication method is superior to the Grid method, as long as the number of comparable properties is more than the number of property attributes used in the valuation assignment.

4 Comparing with the Regression Method

We will compare the Replication estimator with the estimators derived from the traditional Regression and Grid methods under the same set of assumptions. In other words, we start with the same assumption of a linear pricing model as equation (8), or

\[
V_C = X_C \beta + \epsilon_C,
\]

where \( \beta \) is an unknown \( k \times 1 \) vector that contains the hedonic coefficients of property attributes (which can also be referred to as the adjustment factors of property attributes). Since \( \beta \) is an unknown vector, an appraiser has to estimate it first. It is well-known that, under the Ordinary Least Square (OLS) assumption,

\[
\begin{align*}
\hat{\beta} &= (X_C^T X_C)^{-1} X_C^T V_C = (X_C^T X_C)^{-1} X_C^T (X_C \beta + \epsilon_C) \\
&= \beta + (X_C^T X_C)^{-1} X_C^T \epsilon_C,
\end{align*}
\]

where \( \hat{\beta} \) is an unbiased estimator of the true \( \beta \). The value of the subject property can be estimated using the estimated hedonic coefficients \( \hat{\beta} \) and property attributes of the subject property, or

\[
\hat{V}_{S(ols)} = X_S \hat{\beta} = X_S (X_C^T X_C)^{-1} X_C^T V_C,
\]

where \( \hat{V}_{S(ols)} \) is the estimated value of the subject property using the Regression method. The prediction error of a Regression estimator based upon equations (9) and (24) is
\[ \epsilon_{\text{ols}} = V_S - \hat{V}_{S(\text{ols})} = X_S\beta + \epsilon_S - X_S\hat{\beta} = \epsilon_S - X_S(\hat{\beta} - \beta) \]
\[ = \epsilon_S - X_S(X_C^TX_C)^{-1}X_C^T\epsilon_C. \] (25)

The variance of the prediction errors of a Regression under the OLS estimator is

\[ \sigma^2_{\text{ols}} = \sigma^2[1 + X_S(X_C^TX_C)^{-1}X_S^T]. \] (26)

It is important to note that the estimator provided by equation (24) is exactly the same as that in equation (18). This means that, when the linear pricing model specified in equation (7) and the error terms are \(iid\) with a zero mean, the Replication method provides exactly the same estimator as that provided by the Regression method.

However, it is also well known that equation (7) might not hold in applications. We might face multicollinearity problems. We might also omit relevant variables or include irrelevant variables in actual applications. How well does the Replication method perform when compared with the Regression method under those circumstances?

### 4.1 Multicollinearity

It is well known that, when multicollinearity exists, the variance of the estimated adjustment factors (\(\hat{\beta}\)) and the prediction error will be amplified. Multicollinearity can result in high variance of the estimated parameter \(\beta\) because of the use of the inverse matrix \((X_C^TX_C)^{-1}\), which could run up to infinity when \(X_C\) is not of full rank. With multicollinearity, estimated coefficients might have a wrong sign or an implausible magnitude.\(^8\) The higher the correlation among the regressors, the less the precision of the estimator (and thus the higher the variance of the prediction error). In the extreme case where a perfect collinearity exists, the \(k\times n\) matrix \(X_C\) is not a full rank (and has the rank \(j < k\)). Under this circumstance, a regression equation can only estimate \(j\) parameters (\(\hat{\beta}\)). Since \(k - j\) parameter cannot be estimated in the model, the Regression method will not be able to predict the value of the subject property.\(^9\)

---

\(^8\)See, for example, Johnston (1984, p239) for a detailed discussion of this issue.

\(^9\)See Greene (1998, p419-429) for a detailed discussion of this issue.
On the other hand, a multicollinearity problem does not affect the Replication method. This is true because the Replication method does not rely on the estimation of adjustment factors $\hat{\beta}$. The use of the Replication method depends only on the selection of an optimal weight vector from the available set $\{ \omega \mid \omega X_C = X_S \}$ using linear simultaneous equations. The rank problem (of matrix $X_C$) that exists under the Regression method will not be a problem when one is solving linear simultaneous equations.

To demonstrate this concept with an example, let us assume that there are three property attributes and seven comparables properties with prices $V_i, i = 1, 2, \ldots, 7$. The third property attribute can be duplicated by combining the first two property attributes, or

$$X_C^T = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}. \quad (27)$$

When the Regression method is used, the solution is

$$X_C^T X_C \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix} = \begin{bmatrix} 4 & 0 & 4 \\ 0 & 3 & 3 \\ 4 & 3 & 7 \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix} = X_C^T V = \begin{bmatrix} \Sigma_{i=1}^4 V_i \\ \Sigma_{i=5}^7 V_i \\ \Sigma_{i=1}^7 V_i \end{bmatrix}. \quad (28)$$

Without the perfect collinearity problem in $X_C$, equation $(X_C^T X_C)^{-1} \hat{\beta} = X_C^T V$ should give a unique and optimal solution for the $\hat{\beta}$. However, as can be seen from equation (28), when a perfect multicollinearity problem exists, it is not possible to obtain a unique optimal solution for $\hat{\beta}$.

However, when the Replication method is used, the weight vectors $\{ \omega \mid \omega X_C = X_S \}$ can still be estimated without a problem, or

$$X_C \omega = \begin{bmatrix} \Sigma_{i=1}^4 \omega_i \\ \Sigma_{i=5}^7 \omega_i \\ \Sigma_{i=1}^7 \omega_i \end{bmatrix} = X_S = \begin{bmatrix} X_{S1} \ X_{S2} \ X_{S3} \end{bmatrix}. \quad (29)$$

It is clear that an infinite number of $\omega$ vectors can satisfy equation (29), as long as the difference between the number of comparable properties and the number of property attributes is no less than two. The problem of having $X_{S1} + X_{S2} = X_{S3}$ is not serious for the Replication method because all it does is to reduce the number of simultaneous equations by one. Once the set of $\omega$ vectors is established, an appraiser can find that optimal one by following the method established by equation (16). This example based on
the extreme case tells us that the existence of the multicollinearity problem (regardless of the degree) will not affect the Replication method, as long as there is a sufficient number of comparable properties.

Intuitively, we know that the Regression method minimizes the sum of squared errors between the observed values and the estimated values (by using the best-fit parameters for the equation), which uses \((X_C^TX_C)^{-1}\) to determine parameters values. Multicollinearity affects the estimation of \((X_C^TX_C)^{-1}\), which in turn affects the precision of the estimated parameters and their variances. On the other hand, the implementation of the Replication method involves two steps. The first step is to find the exact solutions \(\{\omega \mid \omega X_C = X_S\}\). Clearly, the solution is based on linear algebra and \((X_C^TX_C)^{-1}\) does not play a role in finding those exact solutions. The second step is an optimization process, which selects the best \(\omega\) from the solution set \(\{\omega \mid \omega X_C = X_S\}\). Since the Replication method selects the weight vector that results in the lowest variance among all alternatives, it is unlikely that the selected minimum variance estimator will be affected by the multicollinearity problem in \((X_C^TX_C)^{-1}\).

### 4.2 Irrelevant variables

It is well known that an inclusion of irrelevant variables in an estimation equation, while still resulting in unbiased estimators, will increase the variance of the estimated parameters (and decrease the precision of a prediction).\(^{10}\) To demonstrate the effect, we specify the true model as defined in equation (8), or

\[
V_C = X_C\beta + \epsilon_C
\]

and the misspecified model (with irrelevant variables) as

\[
V_C = X_CI\beta_CI + \epsilon_CI, \quad (30)
\]

where \(X_CI\) is a \(nx(k + I)\) matrix representing the \(k\) true variables and \(I\) irrelevant variables, \(\beta_CI\) is the \((k + I)x1\) hedonic price vector, and \(\epsilon_CI\) is the \(nx1\) vector of error term. We assume that the product of \(X_C\) and

\(^{10}\)For a detailed discussion of this issue, see, for example, Maddala (1992, p165).
the $I$ irrelevant variables is zero. Together with the iid assumption of the error term, the variance of the prediction error of the misspecified equation (equation (30) is

$$\sigma^2_{\text{ols, irrel}} = \sigma^2 + \sigma^2 S_I (X_{CI}^T X_{CI})^{-1} X_{SI}^T$$  \hspace{1cm} (31)

$$> \sigma^2 + \sigma^2 S (X_C^T X_C)^{-1} X_S^T = \sigma^2_{\text{ols, true}}.$$ \hspace{1cm} (32)

$\sigma^2_{\text{ols, true}}$ (or equation (32) is estimated based on the true model (equation (8)). From the literature, we know that $\sigma^2_{\text{ols, true}}$ is smaller than the variance of the misspecified model, $\sigma^2_{\text{ols, irrel}}$ (or equation (31), which is based upon equation (30)).

When the Replication method is used, from Proposition 2 and under the iid assumption of error terms, we know that the optimal weight and the variance (see equations (16) and (17)) of the prediction error of a Replication estimator are

$$\omega^*_{\text{rp, true}} = X_S (X_C X_C^T)^{-1} X_C^T,$$ \hspace{1cm} (33)

and

$$\sigma^2_{\text{rp, true}} = \sigma^2 + \sigma^2 S (X_C^T X_C)^{-1} X_S^T.$$ \hspace{1cm} (34)

Since $X_{CI}$ includes both the true and irrelevant variables, the optimal weight vector and the variance of the prediction error of a Replication estimator are

$$\omega^*_{\text{rp, irrel}} = X_{SI} (X_{CI} X_{CI}^T)^{-1} X_{CI}^T$$ \hspace{1cm} (35)

and

$$\sigma^2_{\text{rp, irrel}} = \sigma^2 + \sigma^2 S_I (X_{CI}^T X_{CI})^{-1} X_{SI}^T.$$ \hspace{1cm} (36)

Define $A = \{\omega \mid \omega X_C = X_S\}$ and $B = \{\omega \mid \omega X_{CI} = X_{SI}\}$, where $B$ is a subset of $A$ (or $B \subset A$). That is, all the feasible $\omega$ in set $B$ must be also a feasible $\omega$ in set $A$. Given this (and from Proposition 1), we know
that a Replication estimator must be unbiased, even if it includes irrelevant variables in the estimation procedure.

However, Since $B \subset A$, we know that $\sigma_{rp,irrel}^2 \geq \sigma_{rp, true}^2$. In other words, the optimal weight in subset $B$ is also the optimal weight in set $A$, then a Replication estimator must also be the minimum variance estimator, even if it includes irrelevant variables in the estimation procedure. When this happens ($\sigma_{rp,irrel}^2 = \sigma_{rp, true}^2$), we know that a Replication estimator must be preferred when compared with a Regression estimator (since a Regression estimator has a higher variance than a Replication estimator under this circumstance).

However, when the optimal weight in subset $B$ is not the optimal weight in set $A$, then a Replication estimator is not the minimum variance estimator ($\sigma_{rp,irrel}^2 > \sigma_{rp, true}^2$). However, while this is true, it is important to note that under this circumstance, the variance of a Replication estimator (see equation (36)) is identical to that of a Regression estimator (see equation (31)). This means that in terms of the minimum variance criterion, a Replication estimator should perform better than or at least as well as a Regression estimator under all circumstances.

### 4.3 Omitted variables

The statistics literature tells us that an omission of relevant variables in an estimation equation will result in biased estimators. To see this, we specify the true model as

$$V_C = X_{C1} \beta_1 + X_{C2} \beta_2 + \epsilon_C,$$

where $X_{C1}$ is a $nxk$ matrix and $X_{C2}$ is a $nxJ$ matrix. The $kx1$ vector $\beta_1$ and $Jx1$ vector $\beta_2$ are hedonic prices for the attributes. Both $X_{C1}$ and $X_{C2}$ are the true property attributes of comparable properties. We assume that there are $J$ omitted variables. The misspecified model (with omitted variables $X_{C2}$) is

$$V_C = X_{C1} \beta + v_C.$$  

where, $v_C$ is a $nx1$ vector of error terms. We assume that the product of $X_{C1}^T$ and the $J$ omitted variables $X_{C2}$ is zero. Let $\hat{\beta}_O$ be the estimated $\beta$
using the Regression method. The prediction error of a Regression estimator is given by

\[ V_S - X_{S1}\hat{\beta}_O = [X_{S1}\beta_1 + X_{S2}\beta_2 + \epsilon_S - X_{S1}\hat{\beta}_O] \]
\[ = X_{S2}\beta_2 + \epsilon_S - X_{S1}(X_{C1}^TX_{C1})^{-1}X_{C1}^T\epsilon_C \]
\[ -X_{S1}(X_{C1}^TX_{C1})^{-1}X_{C1}^T X_{C2}\beta_2, \]

where \( X_{S1} \) and \( X_{S2} \) are the true property attributes of the subject property. The non-zero expectation of prediction error in equation (39) implies that the Regression method produces a biased estimator because \( \hat{\beta}_O \) is biased.\(^{\text{11}}\)

Define \( X_{C12} = (X_{C1}, X_{C2}) \) and \( X_{S12} = (X_{S1}, X_{S2}) \). That is, the \( nx(k + J) \) matrix \( X_{C12} \) is the true matrix that should be used in the true model (without the omitted variable problem). Under this circumstance and the \( iid \) assumption on error terms, the optimal weight and the variance of the prediction error of a Replication estimator are

\[ \omega_{\text{rp, true}}^* = X_{S12}(X_{C12}^TX_{C12})^{-1}X_{C12}^T, \]  
\[ \sigma_{\text{rp, true}}^2 = \sigma^2 + \sigma^2 X_{S12}(X_{C12}^TX_{C12})^{-1}X_{S12}^T. \]  

With an omitted variable problem, the weight vector and the variance of the prediction error of a Replication estimator are

\[ \omega_{\text{rp, omit}}^* = X_{S1}(X_{C1}^TX_{C1})^{-1}X_{C1}^T, \]  
\[ \sigma_{\text{rp, omit}}^2 = \sigma^2 + \sigma^2 X_{S1}(X_{C1}^TX_{C1})^{-1}X_{S1}^T \]
\[ < \sigma^2 + \sigma^2 X_{S12}(X_{C12}^TX_{C12})^{-1}X_{S12}^T = \sigma_{\text{rp, true}}^2. \]

Define \( C = \{ \omega \mid \omega X_{C12} = X_{S12} \} \) and \( D = \{ \omega \mid \omega X_{C1} = X_{S1} \} \). Since \( X_{C12} \)

\(^{\text{11}}\)See Maddala (1992, p162) for a detailed discussion of this issue.
and $X_{S12} = (X_{S1} X_{S2})$, $C$ must be a subset of $D$ (or $C \subset D$). That is, all the feasible solution $\omega$ for the equation $\omega X_{C12} = X_{S12}$ (or in set $C$) must also be a feasible solution $\omega$ for the equation $\omega X_{C1} = X_{S1}$ (or in set $D$).

When a model has omitted variables, the prediction error based on the Replication method, given $\omega X_{C1} = X_{S1}$, is

$$V_S - \omega V_C = [X_{S1}\beta_1 + X_{S2}\beta_2 + \epsilon_S] - \omega[X_{C1}\beta_1 + X_{C2}\beta_2 + \epsilon_C]$$  (44)

$$= (X_{S2} - \omega X_{C2})\beta_2 + (\epsilon_S - \omega \epsilon_C).$$  (45)

It is clear that if $X_{S2} = \omega X_{C2}$, (i.e., \omega is an element of $C$), then a Replication estimator will be unbiased. Intuitively, we know the true weight vector should be selected from the true set $C$, which is a subset of $D$. However, with an omitted variable problem, the Replication method will falsely select the optimal weight vector from a larger set $D$. Since $C \subset D$, if the optimal weight vector is selected from the subset ($D - C$), then the Replication estimator will be a biased estimator.

However, while both a Regression estimator (see equation (39)) and a Replication estimator (see equation (45)) can be biased when the model omits relevant variables, the problem is more serious in the Regression method than in the Replication method. This is true because of four reasons.

First, when the optimal weight vector is selected from the subset $C$, since $C \subset D$, the Replication estimator will be unbiased. Given this, it is quite likely that the Replication estimator can be unbiased.

Second, since both set $D$ and subset $C \subset D$ contain an infinite number of solutions and the property attributes in both sets differ only in the omitted variables, it is very likely that the estimator selected from subset $C$ does not differ too much from the optimal solution selected from subset ($D - C$).

Third, the probability of having omitted variables is much higher in the Regression method than in the Replication method. When the Regression method is used, it requires a large number of comparable properties in the estimation process. In the most likely case, appraisers will not know the details of each property. Given the large number of unfamiliar properties, the Regression method is highly likely to have omitted variables in the estimation equation. On the other hand, the use of the Replication method requires much fewer comparable properties than the Regression method. Similar to
the Grid method, an appraiser can hand pick and use only comparable properties that are the most similar to the subject property. Given this, the problem of having the omitted variable problem is quite minimal, especially when compared to the Regression method.

Fourth and most importantly, when the Replication estimator is biased (or the optimal weight vector is selected from the subset \((D - C)\)), the Regression method will also select the same predicted value as that of the Replication method. In other words, in the worst scenario, the Replication method will perform the same as the Regression method under the unbiasedness criterion.

Given the fact that \(C\) is a subset of \(D\), from Proposition 1 we know that a Replication estimator must be a minimum variance estimator, even if it omits relevant variables in the estimation procedure. This is true because when the optimal weight vector in the set \(D\) is also the optimal weight in subset \(C\), then \(\sigma_{rp,\text{true}}^2 = \sigma_{rp,\text{omit}}^2\). When the optimal weight in the set \(D\) is not the optimal weight in subset \(C\), then \(\sigma_{rp,\text{true}}^2 > \sigma_{rp,\text{omit}}^2\). Given this, it is safe to say that the omission of relevant variables will not increase the variance of a replication estimator. Given this, we do not need to worry about the minimum variance criterion when the Replication method is used.

To summarize, while the Replication method could result in an unbiased and minimum variance estimator in some circumstances, the estimator can be biased under other circumstances. However, when the Replication estimator is biased, it should be noted that the Regression method will provide the same biased estimator. Given this, the Replication method should be preferred because the probability of having an omitted variable problem is much higher when the Regression method is used.

**Proposition 3** Under the assumption that the true property values are a linear function of their property attributes and have error terms distributed iid with a zero mean, a Replication estimator performs at least as well as a Regression estimator. When a regression equation suffers from multicollinearity and/or misspecification problems, a Replication estimator performs equally well or better than a Regression estimator.

5 Comparing with the Grid Method

The Grid method also uses a weighted average of the observed comparable property prices (after adjusting for differences in property attributes) to esti-
mate the value of a subject property. Defining $V_{S(\text{grid})}$ as the estimated value of the subject property using the Grid method, the estimation equation of the Grid method can be written as

$$\hat{V}_{S(\text{grid})} = \omega_{\text{grid}} [V_C + (X_S - X_C) \hat{\beta}]$$ (46)

and

$$\omega_{\text{grid}} e = 1,$$

where $e$ is an $n \times 1$ vector of one, $\omega_{\text{grid}}$ is the weight vector derived under the Grid method, and $\hat{\beta}$ is the adjustment factor vector (or the hedonic prices of the property attributes). When implementing the Grid method, an appraiser’s assignment is to select the number of comparables $n$ and to estimate both $\omega_{\text{grid}}$ and $\hat{\beta}$. While there is no consensus on how the weight vector $\omega_{\text{grid}}$ should be estimated, it is often agreed that the adjustment factors $\hat{\beta}$ can be the same as that (hedonic coefficients) derived from the Regression method.

Under the same assumption that the true property value is a linear function of its property attributes with a zero mean error term (see equation (7)), the prediction error of a Grid estimator $\epsilon_{\text{grid}}$ can be specified as

$$\epsilon_{\text{grid}} = V_S - \hat{V}_{S(\text{grid})} = V_S - \omega_{\text{grid}} [V_C + (X_S - X_C) \hat{\beta}]$$ (47)

$$= X_S \beta + \epsilon_S - \omega_{\text{grid}} [I_n - X_C (X_C^T X_C)^{-1} X_C^T] (X_C \beta + \epsilon_C)$$ (48)

$$- X_S (X_C^T X_C)^{-1} X_C^T (X_C \beta + \epsilon_C)$$

$$= \epsilon_S - \omega_{\text{grid}} [I_n - X_C (X_C^T X_C)^{-1} X_C^T] \epsilon_C - X_S (X_C^T X_C)^{-1} X_C^T \epsilon_C.$$(49)

12 Other than the constraint that $\omega_{\text{grid}} e = 1$, there is no definite rule on how the weight vector $\omega_{\text{grid}}$ should be selected. Practitioners insist that one should give more weight to the most comparable properties but have a difficult time defining what are the most comparable properties or giving an exact formula for determining the weights. Academics, on the other hand, derive some exact methods for estimating weights. However, few practitioners use them. (See, for example, Colwell, Canaday and Wu (1983), Vandell (1991), Gau, Lai and Wang (1992, 1994), Green (1994), Isakson (1986, 2002), Pace and Gilly (1997, 1998), and Pace (1998) for a detailed discussion on the different methods of estimating $\omega_{\text{grid}}$).
Under the same iid assumption of error terms ($\epsilon_S$ and $\epsilon_C$) as in OLS, the variance of the prediction errors of a Grid estimator can be calculated from equation (49) as

$$
\sigma_{grid}^2 = \sigma^2 + \sigma^2 \{ \omega_{grid} [I_n - X_C (X_C^T X_C)^{-1} X_C^T] [I_n - X_C (X_C^T X_C)^{-1} X_C^T] \omega_{grid}^T \\
+ X_S (X_S^T X_S)^{-1} X_S^T \} \\
+ 2 \sigma^2 X_S (X_S^T X_C)^{-1} X_C^T [I_n - X_C (X_C^T X_C)^{-1} X_C^T] \omega_{grid}^T.
$$ (50)

Since $I_n - X_C (X_C^T X_C)^{-1} X_C^T$ is an idempotent matrix, equation (50) can be re-written as

$$
\sigma_{grid}^2 = \sigma^2 + \sigma^2 \{ \omega_{grid} [I_n - X_C (X_C^T X_C)^{-1} X_C^T] \omega_{grid}^T + X_S (X_S^T X_C)^{-1} X_S^T \} \\
+ 2 \sigma^2 X_S (X_S^T X_C)^{-1} X_C^T [I_n - X_C (X_C^T X_C)^{-1} X_C^T] \omega_{grid}^T.
$$ (51)

Since $X_C^T [I_n - X_C (X_C^T X_C)^{-1} X_C^T] = X_C^T - X_C^T X_C (X_C^T X_C)^{-1} X_C^T = 0$, equation (51) can be rewritten as

$$
\sigma_{grid}^2 \geq \sigma^2 + \sigma^2 \{ X_S (X_S^T X_C)^{-1} X_S^T \}. 
$$ (52)

The last inequality in equation (52) holds because a non-negative definite idempotent matrix $I_n - X_C (X_C^T X_C)^{-1} X_C^T$ implies

$$
\omega_{grid} [I_n - X_C (X_C^T X_C)^{-1} X_C^T] \omega_{grid}^T \\
= \omega_{grid} [I_n - X_C (X_C^T X_C)^{-1} X_C^T] [I_n - X_C (X_C^T X_C)^{-1} X_C^T] \omega_{grid}^T \geq 0. 
$$ (53)

Equation (52) provides an important result of our paper. Note that $\sigma^2 + \sigma^2 X_S (X_S^T X_C)^{-1} X_S^T$ is, in fact, the variance of the prediction error under the Replication method (see equation (19)). Since the variance of the prediction error of the Grid method is higher than or equal to that of the variance of the prediction error of the Replication method, the proposed Replication method should be preferred when compared to the Grid method. This is especially true if we take the unbiasedness criterion into consideration. We know that a Replication estimator is always unbiased. However, from equation (47), it is clear that a Grid estimator is unbiased only if the estimated adjustment
factor $\hat{\beta}$ is an unbiased estimator of the true $\beta$. This condition might not hold all the time and a Grid estimator could be a biased estimator of the true property value. The following proposition summarizes our findings in this section.

**Proposition 4** Under the assumption that the true property values are a linear function of their property attributes and have error terms distributed iid with a zero mean, the variance of the prediction error derived from the Replication method is less than that derived from the Grid method as long as the number of comparables is no less than the number of attributes used in the valuation.

6 When Will They Perform the Best?

All appraisal methods are established under certain assumptions. Some of the assumptions hold better than others under different circumstances. Consequently, the use of appraisal methods should depend on the nature of appraisal assignments. Given this, when is the best time for appraisers to use the Replication method?

Among all the three methods discussed, the Regression method should be the preferred one when an appraiser knows all the comparable properties very well. Under this scenario, the estimation equation used by the appraiser should not suffer from the problems of multicollinearity and misspecification and the Regression method should perform quite well. However, when there is a large number of comparable properties (which the appraiser is not familiar with), then the problems associated with the estimation of hedonic prices (or adjustment factors) can be serious. Under this circumstance, the use of the Replication method or Grid method might not a bad idea.

When compared to the Regression method, the advantage of the Replication and the Grid methods is the ability to pick carefully a few truly comparable properties in the evaluation process. When there are few quality comparables, it might be the best for appraisers to use the Replication or the Grid method. However, the Replication method should be preferred over the Grid method, unless the adjustment factors (hedonic prices) of the Grid method can be derived with high confidence. In other words, the Grid method can be a preferred method if the appraiser can derive reliable estimators for the adjustment factors without using a regression equation.
When the Replication method is used and appraisers are not confident about the true model (or when there is a probability of including irrelevant variables or omitting relevant variables in the estimation equation), it might be best for appraisers to include more (rather than fewer) property attributes in the estimation equation (as long as \( n > k \)). For example, when in doubt, appraisers can add the square of property attributes or the cross product of property attributes to the model. In general, the inclusion of higher orders of attributes and cross products to non-linear models will result in a better prediction than linear models.\(^{13}\) Since there is no multicollinearity problem in the Replication method, an increase in the number of property attributes should not decrease the precision of the estimators. More importantly, the cost of including irrelevant variables seems to be much less than the cost of omitting relevant variables. Given this loss function, it might be preferable that appraisers use all reasonable property attributes they can find in the valuation process.

7 Conclusions

The development of techniques that can remove (or, at least, substantially reduce) the subjectivity in an appraisal process is needed in the real estate field. Without such a method, real estate valuation will always remain an art rather than a science. At the same time, it is also understood that the method for valuing real estate must differ from than used for financial assets because of non-homogeneous products and data limitations.

This paper develops a real estate valuation model that takes objectivity out of the valuation process and performs well under data constraints. We also demonstrate that the proposed Replication method should perform better than the Grid method and the Regression method under most scenarios. In the worst case, the replication method performs the same as the other two methods. However, while this model seems to perform well on theoretical grounds, it is important to see if the model holds in the world applications. Given this, we recommend an empirical study that compares the performance

\(^{13}\)According to Taylor expansion, any function can be approximated by a polynomial function. Hence, if the valuation generating function is a non-linear function of attributes, then using Taylor expansion to include higher degree terms of attribute will improve its accuracy.
among the three methods (Replication, Grid, and Regression) using a large data set.
References


