Predicting the Winner’s Curse

Phillip J. Lederer
*William E. Simon Graduate School of Business Administration, University of Rochester, Rochester, NY 14627*

**ABSTRACT**

If quantity uncertainty exists in a first price auction that specifies a fixed bid, a participant must answer two questions when evaluating a candidate bid: “What are my chances of winning?” and “What is the effect of the winner’s curse on my quantity estimate?” The winner’s curse is the tendency of the winner of a first-price sealed-bid auction to be the bidder that most overvalues the items being offered. When value uncertainty is due to quantity uncertainty, the winner’s curse implies that the bidder that most overestimates the quantity tends to win. Thus, if there is quantity uncertainty, a participant must adjust its bid for this tendency to overestimate quantities.

This paper presents an empirical method to answer the above questions by estimating a predictive distribution of the highest competing bid and the quantity bias caused by the winner’s curse. The method is developed for timber auctions but is general to auctions where a fixed bid is called for and there is uncertainty in the mix and quantities of items being offered. An example that uses data from timber auctions is used to demonstrate the method.

*Subject Areas: Decision Analysis, Economic Analysis, and Statistical Techniques.*

**INTRODUCTION**

If quantity uncertainty exists in a first price auction that specifies a fixed bid, then the bidder that most overestimates the quantity will tend to win the auction. This effect is known as the “winner’s curse” [1]. The winner’s quantity estimate will be positively biased and the winner will discover that the quantity won is (on average) less than expected. The winner will conclude that he/she bid too high. A rational, profit maximizing bidder must consider the effect of the winner’s curse and reduce its bid. Thus, a participant in auctions of this sort must answer two questions when selecting its bid: “What are my chances of winning?” and “What is the effect of the winner’s curse on my quantity estimate?” This paper helps decision makers by developing an empirical, statistics-based method to answer the two questions. The method finds a predictive distribution of the highest competing bid and estimates the quantity bias caused by the winner’s curse.

This paper was inspired by a real management problem faced by a forest products firm that often bid for timber harvest rights. In every auction, each bidder prepared an independent estimate of the amount and mix of timber. The bidder assumed the risk of quantity uncertainty because the bids were “fixed,” that is, they were lump sum offers. Because there was considerable quantity uncertainty, the bidder risked the winner’s curse. A bidder needed to know if the winner’s curse was significant and if so, how much to adjust its quantity estimates. This paper
indicates how to adjust the bid using a statistical process that uses historical data collected by participants.

The managerial importance of this paper is that it adapts a traditional decision theoretic bidding approach to an environment where the winner’s curse must be considered. The method determines whether the curse is important and advises how to adjust bids for it. The approach is related to decision analysis because historical data is used to estimate a predictive distribution of competing bids. However, this empirical approach also provides information about how quantity uncertainty affects the winner’s curse.

Decision analytic bidding models are developed and described in a number of papers, including Christenson [2], Engelbrecht-Wiggans [3], Friedman [4], and Lavalle [11]. Decision theory bidding models assume the existence of the predictive distribution of the highest competing bid that is either subjectively or analytically assessed. Only a few published papers describe how this distribution is estimated from historical data.

Some examples do exist, though. Hanssman and Rivett [6] use the empirical distribution of the ratio of the highest bid to the bidder’s estimate of the value for past auctions. The probability distribution of the highest competing bid is the product of the empirical distribution and the estimated value of the current bid. Another approach is taken by King and Mercer [9] [10] and Mercer and Russell [13] who first empirically estimated a model of the markup on estimated cost. They estimate the highest competing bid by multiplying the expected markup by the estimated cost. The authors assumed that the prediction’s variance is normally distributed and estimated it empirically without modeling the sources of prediction errors. These papers have two problems that might be serious in practice.

The first problem is that they estimate the predictive distribution on a per “unit” basis (multiplying the distribution by an estimate of the current “size” of the items being offered). In this paper I show that this does not properly account for heteroscedasticity caused by the mix of items in the auctioned bundle and the differences in firms’ size estimates. The distribution of the highest competing bid may be miss-estimated because the predictive distribution’s variance depends upon the mix of items, and increases with larger item sizes and larger variability in bidders’ quantity estimates.

The second problem is that although quantity uncertainty is assumed, the winner’s curse is not. Any auction in which quantity uncertainty exists and bidders have independent quantity estimates potentially suffers from the winner’s curse. As long as bids are monotonic with respect to estimated quantity, quantity overestimation increases the chance of winning and the “curse.”

This paper solves these two problems. The key idea is to statistically estimate a predictive distribution for the highest competing bid conditioned upon actual quantities. Using a predictive distribution of actual quantities conditioned upon a firm’s quantity estimate, the quantity bias introduced by the winner’s curse is calculated for any firm bid.

Although I analyze bidding for timber, the approach applies to many other auction markets with four restrictions. Use is restricted to auctions with a fixed bid, with uncertainty in mix and quantities of items, and where the source of the winner’s curse is quantity uncertainty. This last restriction means that either: each
bidder has its own value for a unit of each of the items auctioned, or bidders value a unit of each item similarly, but the uncertainty in value is small. In auction theoretic terms (see [14]), this implies that bidders' values for a unit of items are either private values or common values with small value uncertainty. The final restriction is that bidders independently sample or collect information to reduce quantity uncertainty and that after sampling they have approximately the same amount of information. Therefore, my method does not apply to bidding situations with information asymmetry.

Many auction markets have these characteristics. Construction contracts are often offered by sealed, first-price, fixed bids and there is uncertainty in the quantities of labor, materials, and equipment needed. Theil [19] empirically studied highway construction contracts in 33 states, and notes that most contracts required more than 30 standard items. Uncertainty in unit item cost is relatively small. Similarly, some maintenance and other service contracts are offered on a fixed bid basis and there is uncertainty in the quantity of services demanded. Natural resource extraction auctions where bidders make a fixed bid (called a bonus) and then pay a prespecified fee per unit (called a royalty payment) for the quantity removed could be modeled with generalization of the distributional assumptions when there is no information asymmetry between bidders. These auctions have large quantity uncertainty, but small per unit value uncertainty because of the existence of commodity futures markets.

The method also applies to bidding situations without quantity uncertainty but with many items. Results derived here compute a predictive bidding distribution that captures heteroscedasticity due to the mix of items. Thus, results are applicable to many bidding situations.

The rest of the paper is organized as follows. A description of timber auctions follows and the next section presents a model that acts as the basis of an econometric specification. Then, a four stage weighted ordinary least squares (OLS) procedure is described to estimate the predictive distribution and estimate the winner's curse. An example using timber auction data is presented. The paper concludes with managerial implications and suggestions for further research.

Bidding for Timber

This method is developed for timber auctions. The first price-fixed bid sealed bid auction is one of the primary means of selling timber cutting rights in the Southeastern United States [12]. Bidders are typically firms that own sawmills or pulpmills in the area. In these auctions, firms submit a fixed bid and the winner is awarded the right to clear timber from a tract for a specified period. All bids are announced at the bid opening.

Each bidder prepares its own estimate of timber quantities by independent sampling, called a cruise estimate. Typically, firms use a "10 percent cruise" meaning that random samples are taken (without replacement) constituting 10 percent of the area of the tract. Even with sampling, uncertainty exists. Bidders know their opponents because sampling is observable.

There are many factors which may affect a firm's valuation for cutting rights. The quantity and mix of trees on the tract affect the value. The distance from tract
to mill affects timber value because transportation costs are borne by the firm. Transportation cost is also important because it limits the geographical area in which a firm can profitably operate. The result is that the same firms often bid against each other. Harvest costs, the costs of building roads, cutting the trees, and loading them onto trucks affect net timber value because these are paid by the firm. Finally, firm needs affect its valuation. A firm with unmet requirements values timber more highly as failure in the auction usually necessitates purchases from a higher cost supplier.

When a firm prepares a bid, it typically has data concerning: the size of the tract; its own quantity estimates and strategic factors, such as which firms may bid; the distances from each potential bidder to the tract; an assessment of others’ need for timber; an estimate of harvest costs; and market data, such as the current spot price for timber.

THE FIRM’S DECISION PROBLEM

The perspective taken is that of a firm that plans to bid in an upcoming auction using historical data of the type described in the preceding paragraph. This firm will be referred to as the “firm”; all other firms will be referred to as the “other bidders.”

The firm’s bidding problem can be understood by the inference diagram of Figure 1. Other bidders bid using their cruise estimates and knowledge of strategic factors. The firm uses its cruise estimate and knowledge of strategic factors when preparing its bid. If the firm’s bid is higher than that of the highest competing bidder, it wins and receives its payoff. The payoff is the firm’s value of the actual timber quantities won minus its bid. The firm understands the structure of the decision problem. In order to estimate the payoff for a proposed bid, the firm calculates its probability of winning and an estimate of actual timber quantities conditioned upon winning.

To simplify, it is assumed that there are only two types of timber at auctions: sawtimber and pulpwood. All results and methods hold for more than two types. It is assumed that private cruise information is the only quantity information any participant has. The framework can be extended to the case where public data exist.

Predicting the Highest Competing Bid

The predictive distribution of the highest competing bid is estimated by regression analysis using historical data. The regression equation’s specification is not derived from a game theory model because there is no game theory model that provides a simple structural form. Timber auctions are viewed as mixed common-private values auctions. Common values are caused by quantity uncertainty and private values are caused by different bidders’ (harvest and transport) costs and use values of the timber obtained. Milgrom and Weber’s [14] work is the most general analysis that considers mixed common-private values auctions. Among their results is a closed form expression for the optimal bidding rule for first price auctions for the general symmetric model. Symmetry requires identical bidders’ costs and valuations, and is unrealistic for timber auctions. For example, bidders have different costs because of transportation costs from tract to mill cost. Therefore, their results do
not apply to the environment studied here. More general models without symmetry
do not provide closed form bidding rules, so it is not possible to estimate bidding
behavior using a structural form derived from a game theory model. Although a
game theory model is not used, the form estimated incorporates important strategic
variables so that competitive effects are empirically estimated.

Used as a bidding tool, my approach is appropriate only to those situations
where other bidders' behavior will continue for an upcoming auction. In particular,
this means that the firm's analysis and the firm's bid amount will not affect the
others' bidding behavior. This is likely to be true for the next auction, but may not
be true for later ones where competitors' react to the firm's success.

The highest competing bid for the rth auction is denoted by $HCB_r$. The following
model is used by the firm to estimate and to predict $HCB_r$, as a function of observable
strategic variables and its cruise estimates:
\[ HCB_t = (\beta_{St} + \varepsilon_{St})(\hat{S}_t + \varepsilon_{St}) + (\beta_{Pt} + \varepsilon_{Pt})(\hat{P}_t + \varepsilon_{Pt}), \]

where

- \( \hat{S}_t(\hat{P}_t) \) = is the firm's cruise estimate of the cords of sawtimber (pulpwood) for the \( t \)th auction,
- \( \varepsilon_{St}(\varepsilon_{Pt}) \) = is a random variable that is the difference between the firm's cruise estimate and the highest competing bidder's estimated quantity for sawtimber (pulpwood) for the \( t \)th auction, as measured in cords,
- \( \beta_{St}(\beta_{Pt}) \) = is the mean of the highest competing bidder's bid-price per cord for sawtimber (pulpwood) for the \( t \)th auction, and
- \( \varepsilon_{St}(\varepsilon_{Pt}) \) = is random variation of the highest competing bidder's bid-price per cord for sawtimber (pulpwood) for the \( t \)th auction.

Note that \( \beta_{St} + \varepsilon_{St} \) can be interpreted as the highest competing bidder's price per cord of sawtimber (similarly for \( \beta_{Pt} + \varepsilon_{Pt} \)). Thus (1) is a decomposition of \( HCB_t \) into four components: bid-price per cord for sawtimber, estimated cords of sawtimber, bid-price per cord for pulpwood, and estimated cords of pulpwood. There are two sources of uncertainty in \( HCB_t \) for each timber-type: uncertainty as to the highest competing bidder's price per unit, and uncertainty as to the highest competing bidder's estimated quantity. It is assumed that \( \varepsilon_{St}, (\varepsilon_{Pt}) \) are independent and identically distributed random variables.

In auction \( t \), if information about whether the firm won or lost the auction is ignored, then the firm's cruise estimates are unbiased estimates of quantities for the various timber-types. Next, it is argued that the highest competing bidder adjusts its cruise estimates for the winner's curse, yielding estimated quantities different from its cruise estimates. The estimated quantities are unbiased estimates for the timber types.

If the winner's curse is important, a rational profit maximizing bidder adjusts its cruise estimates for the winner's curse. This paper assumes that in past and current auctions the other bidders are rational and profit maximizing and by some means adjust their cruise estimates for the winner's curse. This paper presents a method that a bidder can use. The adjusted cruise estimates are referred to as "estimated quantities" (I have used this term in the definition of \( \varepsilon_{St} \) above). The estimated quantities are used by all other bidders when preparing their bids. A competing bidder's cruise estimate for a timber type is not equal to its estimated quantity. The estimated quantity is a lower number than its cruise estimate because of the winner's curse. In summary, assume that:

**Assumption 1**: All other bidders adjust their cruise estimates for the quantity bias due to the winner's curse, yielding an estimated quantity for each timber type.

This is consistent with equilibrium behavior; to assume otherwise would imply non-value maximizing behavior. The assumption implies that the other bidder with the highest competing bid uses unbiased timber estimates. Because of Assumption 1 and the fact that information about whether the firm won or lost auction \( t \) is ignored when estimating (1), the expected values of \( \varepsilon_{St} \) and \( \varepsilon_{Pt} \) are both zero.
The parameters and uncertainties in (1), namely, $\beta_{St}$, $\beta_{Pt}$ and $e_{St}$, $e_{Pt}$, are estimated. The mean values of per cord bid-prices $\beta_{St}$ and $\beta_{Pt}$ are assumed to depend upon independent variables that capture strategic factors. Economic reasoning would imply that the highest competing bid is a function of observable strategic variables, such as the number and identity of opponents, their need for timber and distances from the tract, an estimate of harvest costs, etc. Therefore, per cord bid-prices are a function of these variables. A linear relationship between mean per cord bid-prices and $m$ independent strategic variables, $\{X_{jt}\}$ is assumed:

$$\beta_{Pt} = \sum_{j=1}^{m} \beta_{Pj} X_{jt},$$

and

$$\beta_{St} = \sum_{j=1}^{m} \beta_{Sj} X_{jt}.$$ (2)

Of course, more complicated nonlinear relationships could be adopted. Equation (1) can be rewritten as

$$HCB_t = \sum_{j=1}^{m} \beta_{Pj} X_{jt} \hat{S}_t + \sum_{j=1}^{m} \beta_{Sj} X_{jt} \hat{P}_t + \nu_t,$$

or more simply,

$$HCB_t = \beta_{St} \hat{S}_t + \beta_{Pt} \hat{P}_t + \nu_t.$$ (3)

where

$$\nu_t = \beta_{St} e_{St} + \hat{S}_t e_{St} + e_{St} e_{St} + \beta_{Pt} e_{Pt} + \hat{P}_t e_{Pt} + e_{Pt} e_{Pt}.$$ (4)

When estimating the betas in (3), there is a problem due to heteroscedasticity. Even if $e_{St}$, $e_{Pt}$, $e_{St}$, and $e_{Pt}$ are constant variance for all $t$, $\nu_t$ will generally be heteroscedastic because of the effect of $\hat{S}_t$ and $\hat{P}_t$ on the variance of $\nu_t$. If heteroscedasticity is ignored, bid-price estimates are not efficient. Also, if heteroscedasticity is ignored when estimating the distribution of $HCB$, predictions will miss-estimate the variance of $HCB$. Much of the remainder of this section is devoted to estimating the form of the variance (4) so heteroscedasticity can be considered.

The first two moments of $\nu_t$ are studied. In principle, higher order moments can also be analyzed. For the timber data set studied, assumption of normality of $\nu_t$ is appropriate, which is somewhat expected. The error term $\nu_t$ is the sum of six random variables: two involve quantity variations ($\beta_{St} e_{St}$, $\beta_{Pt} e_{Pt}$) and are mutually independent of the remaining two terms in bid-price variations ($\hat{S}_t e_{St}$, $\hat{P}_t e_{Pt}$). When quantity and bid variations are both important factors in the variance, the central
limit theorem implies that \( v_i \) will possess a distribution that is more "normal" (symmetric and centrally distributed) than the most non-normal of these random variables.

A number of simplifying assumptions about \( e_{S_i}, e_{P_i}, e_{S_t}, \) and \( e_{P_t} \) are made:

**Assumption 2A:**

\[
\text{Cov}(e_{S_i}, e_{S_t}) = \text{Cov}(e_{S_i}, e_{P_t}) = \text{Cov}(e_{P_t}, e_{S_t}) = \text{Cov}(e_{P_t}, e_{P_t}) = 0;
\]

**Assumption 2B:**

\[
\text{Cov}(e_{S_t}, e_{P_t}) = \delta.
\]

Relations in Assumption 2A are made for parsimony. The first and last equality can be justified, at least for the timber bidding data set analyzed later in this paper. If model (3) is estimated with a constant term added \( \text{HCB}_t = \beta_0 + \hat{\beta}_{S_i} \bar{S}_i + \hat{\beta}_{P_t} \bar{P}_t + v_i \), the constant term \( \beta_0 \) is an estimate of the expected value \( e_{S_t} \bar{S}_i + e_{P_t} \bar{P}_t \), which is the sum of the covariances. Using the empirical bidding data, the constant term was not statistically different from zero \( (t=1.0, df=27) \), so that \( \text{Cov}(e_{S_t}, e_{S_t}) = \text{Cov}(e_{P_t}, e_{P_t}) = 0 \) can be justified. Alternately, all of the covariances could be directly estimated, but were not done so because of the small size of the data set. Assumption 2B states that random variations in the highest competing bidder's per cord bid-prices have constant covariance, and this assumption models the unobservable covariance. The assumptions do not affect the ideas behind the method developed here and can be generalized in later research.

Assumption 2 implies that the variance of \( v_i \) is:

\[
\text{Var}[v_i] = \beta_0^2 \text{Var}[e_{S_i}] + \beta_{P_t}^2 \text{Var}[e_{P_t}] + 2\beta_0 \beta_{P_t} \text{Cov}[e_{S_i}, e_{P_t}]
\]

\[
+ \bar{S}_i^2 \text{Var}[e_{S_t}] + \bar{P}_t^2 \text{Var}[e_{P_t}] + 2\delta \bar{S}_i \bar{P}_t
\]

\[
+ \text{Var}[e_{S_i}] \text{Var}[e_{S_t}] + \text{Var}[e_{P_t}] \text{Var}[e_{P_t}].
\]

To proceed further, properties of \( e_{S_t} \) and \( e_{P_t} \) are studied.

**Modeling the Difference In Quantity Estimates**

To derive the form of the variance (and covariance) for \( e_{P_t} \) and \( e_{P_t} \), several assumptions are made. Assume that the firm uses a \( P_{F \times 100} \) percent cruise estimate and all other bidders use a \( P_{F \times 100} \) percent cruise estimate. This is a useful approximation that reflects the fact that most firms use similar sampling plans. Assume that the highest competing bidder's estimated quantity for each type of timber is the highest for all other bidders. This assumption captures the potential effect of the winner's curse on the bidding process. Finally, assume that the square of an item's bid-price is much larger than its variance. This is a realistic assumption for timber auctions. Then, result (A4) in Appendix 1 shows that (5) can be written as:
\[
\text{Var}[v_t] = \text{Var} [\varepsilon_{St}] \hat{S}_{t}^2 + \text{Var} [\varepsilon_{Pr}] \hat{P}_{t}^2 + \frac{c_{St}^2}{N_t} \frac{\text{Var} (NUM_t) P_F + P_H}{P_F P_H} S_{t}^2 \beta_{St}^2 + \frac{c_{Pr}^2}{N_t} \frac{\text{Var} (NUM_t) P_F + P_H}{P_F P_H} \hat{P}_{t}^2 \beta_{Pr}^2 + 2 \alpha_t N_t \frac{\tau P_F + P_H}{P_F P_H} \beta_{St} \beta_{Pr},
\]

(6)

where \(c_{St}^2(c_{Pr}^2)\) is the coefficient of variation of sawtimber (pulpwood) per acre, \(N_t\) is the number of acres, \(NUM_t\) is the number of competing bidders, and \(\alpha_t\) is the covariance of per-acre quantities of sawtimber and pulpwood for auction \(t\). \(\text{Var}(m)\) is the order statistic of the largest of \(m\) independent draws from a standard normal distribution which is tabulated in [16].

Parameters in this equation, such as \(P_H, c_{St}, C_{Pr},\) and \(\alpha_t\) can be observed. Unobservable parameters in (6), such as \(\text{Var} [\varepsilon_{St}], \text{Var} [\varepsilon_{Pr}], \delta,\) and \(\tau\) will be estimated by regressing squared residuals of the regression (3) against the terms on the right hand side. When regressing, a coefficient will weigh each term in (6) to adjust for misspecification when any of the above assumptions do not hold. In this way equation (6) is empirically estimated for the best fit.

Therefore, sources of heteroscedasticity in (3) are the variations in the highest competing bidder’s bid-price (the first three terms) and the differences in the estimated quantities of the firm and the highest competing bidder (the last three terms). The variance increases with timber quantities, and is sensitive to the mix.

Equation (6) has an important implication for auctions involving a single item with quantity uncertainty. The variance of (4) is not proportional to the square of the estimate of the item’s quantity:

\[
\text{Var}[v_t] = \text{Var} [\varepsilon_{St}] \hat{S}_{t}^2 + \frac{c_{St}^2}{N_t} \frac{\text{Var} (NUM_t) P_F + P_H}{P_F P_H} S_{t}^2 \beta_{St}^2,
\]

because \(N_t\) is proportional to \(S_t\). This means that analysis of \(HCB\) on a unit basis will not yield a predictive distribution that accounts properly for the forecast variance. For example, previous papers that estimate predictive distribution on a unit basis [6] [9] [10] [13] are seen to overestimate the variance for large bundles, and underestimate it for small ones.

Equation (6) also has an important implication for auctions without quantity uncertainty but with many items. Heteroscedasticity exists due to the mix (the first three terms of (6)) which must be reflected in the predictive distribution.

Predicting the Winner’s Curse Distribution

This section explains how to estimate the quantity bias introduced by the winner’s curse. In the following, fix all strategic variables in (2) except timber quantities. Suppose that the probability density function of actual timber quantities conditioned upon the firm’s estimate of timber quantities, \(\alpha(S, P)(\hat{S}, \hat{P})\), and the density function of the highest competing bid conditioned upon actual timber quantities, \(h(x|S, P)\), are both known. Then expected timber quantities conditioned upon winning the auction with a bid of \(b\) given cruise estimates \((\hat{S}, \hat{P})\) is denoted by \(E(S', P|b, (\hat{S}, \hat{P}))\) and is:
\[ E(S', P' \mid b_i(\hat{S}, \hat{P})) = \int \int \int (S, P) h(x \mid (S, P)) a((S, P) \mid (\hat{S}, \hat{P})) dS dP dx. \]

This section shows how to estimate the two probability distributions. Later, a simulation experiment is used to estimate \( E(S', P' \mid b_i(\hat{S}, \hat{P})) \).

Suppose that the firm's estimates for auction \( t \) are \( (\hat{S}_t, \hat{P}_t) \). If a diffuse prior on the true timber quantities \( (S_p, P) \) is assumed, then the posterior distribution of \( (S_p, P) \) conditioned upon the sample information is normally distributed with mean \( (\hat{S}_p, \hat{P}) \) and variance-covariance matrix:

\[
\begin{align*}
    &\frac{c_S^2}{P_F N_t} \hat{S}_t^2 \alpha_i P_F \\
    + &\frac{N_t}{P_F} \frac{c_P^2}{P_F} \hat{P}_t^2.
\end{align*}
\]

Therefore, the marginal distribution \( a((S, P) \mid (\hat{S}, \hat{P})) \) is known. A predictive distribution of \( HCB_t \) conditioned upon \( (S_p, P) \) is estimated using ideas similar to that of (1):

\[ HCB_t \mid (S_p, P) = (\hat{S}_t + \varepsilon_{S})(S_p + f_{S}) + (\hat{P}_t + \varepsilon_{P})(P + f_{P}). \]  

(7)

In (7), \( \beta_{St} \) and \( \beta_{Pt} \) are the mean bid-prices defined by (2), and \( f_{S_p} \), \( f_{P} \) is a random variable which is the difference between actual timber quantity and the highest competing bidder's estimated quantity for sawtimber (pulpwood) for the \( t \)th auction. The distribution of \( f_{S_p} \) \( f_{P} \) has mean zero since competitors are assumed to adjust cruise estimates for the effect of the winner’s curse. Expressions for \( \text{Var}[f_{S_p}] \), \( \text{Var}[f_{P}] \) and \( \text{Cov}[f_{S_p}, f_{P}] \) are derived in Appendix 2. Then,

\[ HCB_t \mid (S_p, P) = \beta_{St} S_t + \beta_{Pt} P_t + \nu', \]  

(8)

with

\[ \nu' = \beta_{St} f_{S_p} + S_t \varepsilon_{S} + f_{S_p} \varepsilon_{S} + \beta_{Pt} f_{P_t} + P_t \varepsilon_{P} + f_{P_t} \varepsilon_{P_t}. \]

An expression for the variance of \( \nu' \) is found with similar arguments used to derive equation (6):

\[ \text{Var} [\nu'] = \text{Var} [\varepsilon_{S_p}] S^2_t + \text{Var} [\varepsilon_{P_t}] P^2_t + 2S_t P_t \]

\[ + \frac{c_S^2}{P_H} \frac{\text{Var} [\text{NUM}_t]}{N_t} \hat{S}_t^2 + \frac{c_P^2}{P_H} \frac{\text{Var} [\text{NUM}_t]}{N_t} \hat{P}_t^2 + 2\tau \alpha_i N_t \frac{1}{N_t P_H} \beta_{St} \beta_{Pt}. \]  

(9)
Var [v'] is like (6), and shares many common terms and parameters, except the last three terms are smaller, reflecting that actual quantities are now assumed known. If v_i is normal and Var [v_i] is known, the distribution h(x|S,P) is easily found.

Finding the Predictive Distribution with Weighted Least Squares Regression

Models (3) and (6) determine a predictive distribution for the highest competing bid using weighted ordinary least squares theory.

Dividing all terms in (3) by the weight \( w_i = \sigma_{v_i} \) results in:

\[
\frac{HCB_t}{w_i} = \beta S_i \frac{\hat{S}_i}{w_i} + \beta P_i \frac{\hat{P}_i}{w_i} + \omega_i,
\]

where \( \omega_i \) is the error term in this new model. This model possesses a constant variance error term and Var [\( \omega_i \)] = 1. If \( \beta_{St} \) and \( \beta_{Pt} \) have a linear structure, that is (2) holds, ordinary least squares (OLS) regression of (10) results in efficient and unbiased estimates for the bids and an efficient and unbiased point forecast \( (HCB_t/w_i) \). Substituting (2) into (10) yields

\[
\frac{HCB_t}{w_i} = \bar{\beta}_S \frac{\bar{S}_i}{w_i} + \bar{\beta}_P \frac{\bar{P}_i}{w_i} + \omega_i,
\]

where

- \( \bar{\beta}_S \) = m by 1 column vector: \( (\beta_{S1}, \ldots, \beta_{Sm}) \),
- \( \bar{S}_i \) = m by 1 column vector: \( \hat{S}_i(X_{1i}, \ldots, X_{mi}) \),
- \( \bar{\beta}_P \) = m by 1 column vector: \( (\beta_{P1}, \ldots, \beta_{Pm}) \), and
- \( \bar{P}_i \) = m by 1 column vector: \( \hat{P}_i(X_{1i}, \ldots, X_{Mi}) \).

The point forecast for auction \( t \) is the point forecast for (11) times \( w_i \):

\[
H\hat{CB}_t = w_i \left( \frac{H\hat{CB}_t}{w_i} \right).
\]

The variance of the prediction can be calculated using OLS theory if \( \omega_i \) is normal (which is equivalent to assuming that \( v_i \) is normal):

\[
Var \left[ \frac{H\hat{CB}_t}{w_i} \right] = \left( 1 + \frac{1}{w_i^2} (\bar{\epsilon}_{S_i} \bar{\epsilon}_{P_i})'(T'T)^{-1}(\bar{\epsilon}_{S_i} \bar{\epsilon}_{P_i}) \right),
\]

where \( T \) is the matrix of observations of independent variables in (12) divided by the proper weights; that is, the \( k \)th row of \( T \) is the 2m by 1 vector: \( (1/w_k)(\bar{\epsilon}_{S_k} \bar{\epsilon}_{P_k}) \). Thus,

\[
Var [H\hat{CB}_t] = w_i^2 \left( 1 + \frac{1}{w_i^2} (\bar{\epsilon}_{S_t} \bar{\epsilon}_{P_t})'(T'T)^{-1}(\bar{\epsilon}_{S_t} \bar{\epsilon}_{P_t}) \right).
\]
If $w^2_t$ is unknown then it must be estimated. To obtain an estimate, $\hat{w}_t$, the following four step procedure from Theil [18] can be used:

Step 1. Estimate model (3) by OLS. This model is called the "OLS forecasting model."

Step 2. Given the residuals of Step 1, $r_t$, and estimates, $\hat{\beta}_S$ and $\hat{\beta}_P$, obtained in Step 1, use OLS to estimate $\gamma_1$, $\gamma_2$, $\gamma_3$, $\gamma_4$, $\gamma_5$, and $\gamma_6$ in the model:

$$r_t^2 = \gamma_1\hat{P}_t^2 + \gamma_2\hat{S}_t^2 + \gamma_32\hat{S}_t\hat{P}_t + \gamma_4\frac{c^2_{St}}{N_t}\frac{\text{Var}(NUM_t)P_F + P_H}{P_FP_H} \hat{\beta}_S^2 + \gamma_5\frac{c^2_{Pt}}{N_t}\frac{\text{var}(NUM_t)P_F + P_H}{P_FP_H} \hat{\beta}_P^2 + 2\alpha_n\frac{\gamma_6P_F + P_H}{P_FP_H} \hat{\beta}_S\hat{\beta}_P + \phi_t$$  \hspace{1cm} (14)

Step 3. Use estimates from (14) and compute $\hat{w}_t^2$:

$$\hat{w}_t^2 = \gamma_1\hat{\beta}_t^2 + \gamma_2\hat{S}_t^2 + \gamma_32\hat{S}_t\hat{P}_t + \gamma_4\frac{c^2_{St}}{N_t}\frac{\text{Var}(NUM_t)P_F + P_H}{P_FP_H} \hat{\beta}_S^2 + \gamma_5\frac{c^2_{Pt}}{N_t}\frac{\text{Var}(NUM_t)P_F + P_H}{P_FP_H} \hat{\beta}_P^2 + 2\alpha_n\frac{\gamma_6P_F + P_H}{P_FP_H} \hat{\beta}_S\hat{\beta}_P;$$  \hspace{1cm} (15)

Step 4. Using weights $\tilde{w}_t = \sqrt{\frac{1}{\hat{w}_t^2}}$ obtained in Step 3, more efficient estimates of $\beta_S$ and $\beta_P$ are possible than were obtained in Step 1. This is done by estimating (11) using $w_t = \tilde{w}_t$, that is, 

$$\begin{pmatrix} \hat{HCB}_t \\ \hat{w}_t \end{pmatrix} = \begin{pmatrix} \hat{\beta}_S \\ \hat{\beta}_P \end{pmatrix} \tilde{w}_t - \begin{pmatrix} \hat{z}_S \\ \hat{z}_P \end{pmatrix} = \begin{pmatrix} \beta_S \\ \beta_P \end{pmatrix} \begin{pmatrix} \tilde{w}_t \\ \tilde{w}_t \end{pmatrix} + \omega_t$$  \hspace{1cm} (16)

The resulting model is referred to as the "weighted forecasting model." This four step procedure results in unbiased and asymptotically efficient estimates of $\hat{\beta}_S$ and $\hat{\beta}_P$, and unbiased estimates of $\gamma_1$, $\gamma_2$, $\gamma_3$, $\gamma_4$, $\gamma_5$, and $\gamma_6$ (see [18]). Point forecasts and estimates of the uncertainty in the point forecast for the weighted forecasting model are

$$H\hat{CB}_t = \hat{w}_t \begin{pmatrix} \hat{HCB}_t \\ \hat{w}_t \end{pmatrix},$$  \hspace{1cm} (17)

and

$$\text{Var}[H\hat{CB}_t] = \hat{w}_t^2 \left( 1 + \frac{1}{\hat{w}_t^2}(\hat{\beta}_S\hat{z}_P)'(T' T)^{-1}(\hat{\beta}_S\hat{z}_P) \right).$$  \hspace{1cm} (18)

If $\omega_t$ can be assumed independent and normally distributed, and parameters $\hat{\beta}_S$ and $\hat{\beta}_P$ have jointly uninformative priors; then from a Bayesian perspective, the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
prediction of \( HCB_t \) is asymptotically normally distributed with mean (17) and variance (18) (see [8]).

Next, the coefficient estimates are used to estimate the quantity bias due to the winner’s curse.

**Predicting the Winner’s Curse**

Parameters derived to estimate the mean and variance of \( HCB_t \) can be used to estimate the mean and variance of \( H\hat{CB}_t / (S_t, P_t) \):

\[
H\hat{CB}_t \mid (S_t, P_t) = \hat{\beta}_{S,t} S_t + \hat{\beta}_{P,t} P_t
\]

(19)

and

\[
\text{Var} [H\hat{CB}_t] \mid (S_t, P_t) = \hat{w}_t^2 \left[ 1 + \frac{1}{\hat{w}_t^2} (\bar{S}_t \bar{P}_t)^T (T' T)^{-1} (\bar{S}_t \bar{P}_t) \right] \mid (S_t, P_t).
\]

(20)

In (19), bid-prices are estimated from estimation (16). In (20) the specified conditioning by \((S_t, P_t)\) on the right hand side means that variables \((S_t, P_t)\) are used instead of \((\bar{S}_t, \bar{P}_t)\) in the definition of the column vector \((\bar{S}_t, \bar{P}_t)\). The weight \(\hat{w}_t^2\) which is an estimate of \(\text{Var}[\tilde{v}_t]\) defined by (9) uses parameter estimates found in (15):

\[
\hat{w}_t^2 = \gamma_1 \hat{P}_t^2 + \gamma_2 S_t^2 \hat{S}_t + \gamma_3 2 S_t P_t + \gamma_4 \frac{\text{Var}(\text{NUM}_t) P_F + P_H}{P_F P_H} - \frac{c_3^2}{P_F N_t} S_t^2 \hat{S}_t^2
\]

\[+ \gamma_5 \frac{c_3^2}{P_F N_t} \frac{\text{Var}(\text{NUM}_t) P_F + P_H}{P_F P_H} - \frac{c_3^2}{P_F N_t} P_t^2 \hat{P}_t^2 \]

\[+ \left[ \alpha N_t \gamma_6 P_F + P_H - \frac{N_t}{P_F} \right] 2 \hat{S}_t \hat{P}_t.
\]

(21)

If \( v_t \) can be assumed normal, so can \( \tilde{v}_t \). This is because if \( v_t \) is evaluated at \((S, P)\), then \( v_t = \beta_{S,t} (\bar{S}_t - S_t) + \beta_{P,t} (\bar{P}_t - P_t) \). Assuming large sample sizes, random variables \((\bar{S}_t - S_t)\) and \((\bar{P}_t - P_t)\) are normally distributed by the central limit theorem, and \( v_t \) is by assumption, therefore \( \tilde{v}_t \) is normal.

Next, the simulation experiment for predicting the winner’s curse is described. Suppose that estimates \((\hat{S}_t, \hat{P}_t)\) are found and the firm makes a bid of \( b \). The simulation experiment with \( n \) trials is done as follows:

1. A single draw from the distribution of \((S_t, P_t)\) conditioned upon \((\hat{S}_t, \hat{P}_t)\) is found.
2. Then, an estimate of \( HCB_t (S_t, P_t) \) is obtained by sampling a single observation from the normal distribution with mean (19) and variance (20). If \( b > HCB_t \), the firm wins, and the differences \((\bar{S}_t - S_t)\) and \((\bar{P}_t - P_t)\) are recorded.
3. If Step (1) has been done fewer than \( n \) times, return there, otherwise the means of the observations of \((\bar{S}_t - S_t)\) and \((\bar{P}_t - P_t)\) are calculated, and the procedure ends.
The means of the observations of \( \hat{S}_t - S_t \) and \( \hat{P}_t - P_t \) are denoted respectively by \( \bar{S}_t - S_t \) and \( \bar{P}_t - P_t \), and are predictions of the quantity bias due to the winner’s curse for bid \( b \). Using our previous notation: \( \langle \bar{S}_t - S_t \rangle, \langle \bar{P}_t - P_t \rangle \) estimates \( E(S',P'|b,\hat{S},\hat{P}) = \langle \bar{S},\bar{P} \rangle \).

**USING THE METHOD**

The method presented in the last section is used to derive the predictive distribution and estimates for the winner’s curse for a sample problem. The example uses data from a forest products company operating a single sawmill and pulpmill in the southeastern United States. The data set includes information about one year’s bidding history involving 32 auctions. Table 1 presents definitions of independent variables and the mean and the standard deviation for the highest competing bid. The coefficient of variations, \( (c_S, c_P) \), and the covariance of timber quantities, \( \alpha \), were not in the data set so these parameters were assumed to be constant for all auctions and estimates. As the number of bidders was nearly constant, a constant number of bidders was assumed when estimating Step 3. These assumptions simplified parameter estimation in Step 3 because only a single parameter needs to be estimated for each term on the right hand side of (15).

Table 2 reports results from the four step procedure. Table 2a presents results of Step 1, which derives the OLS forecasting model, using only strategic independent variables that were statistically significant: cruise estimates of sawtimmer and pulpwood. Nonsignificant and colinear variables are dropped for reasons of parsimony. Table 2b presents the results of the Step 2 regression of (13) that finds the weights. Table 2c presents the result for Step 4, the weighted forecasting mode, using only statistically significant strategic variables.

Reviewing the results of the models, the OLS regression of Table 2a has good fit and the statistics of skewness and kurtosis fail to reject the hypothesis that the residuals are normally distributed. Table 2c is the best weighted regression model for HCB. Strategic variables thought a priori to be important were not found to be statistically significant: distance of the closest bidder to the tract and the number of bidders. The statistics of skewness and kurtosis do not reject the hypothesis that the residuals are normally distributed. It is assumed that \( \omega, \xi \) are independent and identically distributed as there is no evidence of serial autocorrelation of the residuals.

A common error is to use the OLS forecasting model of Step 1 rather than the weighted forecasting model of (17) and (18) to make predictions. Although the point predictions of the OLS forecasting model are unbiased, this model’s estimate of the prediction’s variance is incorrect because heteroscedasticity of the data has been ignored.

The variance of prediction for the OLS forecasting model and the weighted forecasting model are different and depend on the “size” of the bundle. Consider the mean value of sawtimmer, pulpwood, and pulpwwood times harvest quality for the 32 auctions: \( \bar{S}, \bar{P}, \text{ and } \bar{P}_Q \). Now consider the prediction errors of the two models with independent variables \( (\lambda \bar{S}, \lambda \bar{P}) \) and \( (\lambda \bar{S}, \lambda \bar{P}, \lambda P_Q) \), respectively. Parameter \( \lambda \) scales the size of the tract up or down, leaving the mix of timber types constant. Figure 2 compares the standard deviation of prediction errors for the two models as a function of the size of the bundle, as parameterized by \( \lambda \). The standard
Table 1: Variables: Definitions and descriptive statistics.

<table>
<thead>
<tr>
<th>Variable (unit)</th>
<th>Definition</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent Variable</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$HCB_t$ (dollars)</td>
<td>Highest competing bid</td>
<td>81,909.0</td>
<td>63,361.0</td>
</tr>
<tr>
<td>Independent Variables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{S}_t$ (cords)</td>
<td>Firm’s cruise estimate of sawtimber</td>
<td>1,190.0</td>
<td>923.1</td>
</tr>
<tr>
<td>$\hat{P}_t$ (cords)</td>
<td>Firm’s cruise estimate of pulpwood</td>
<td>282.5</td>
<td>332.3</td>
</tr>
<tr>
<td>$DIST_t$ (miles)</td>
<td>Distance of nearest competing firm to tract $t$</td>
<td>17.7</td>
<td>8.5</td>
</tr>
<tr>
<td>$NUM_t$</td>
<td>Number of competing bidders</td>
<td>5.2</td>
<td>1.3</td>
</tr>
<tr>
<td>$N_t$ (acres)</td>
<td>Size of tract</td>
<td>153.75</td>
<td>101.4</td>
</tr>
<tr>
<td>$Q_t$ (dummy variable)</td>
<td>1 if low harvest cost</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0 if high harvest cost</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

deviations are equal at $\lambda=1.35$. The OLS forecasting model overestimates prediction error for values of $\lambda$ less than 1.35 and underestimates prediction error for larger values of $\lambda$. Even for the “average tract,” ($\lambda=1$), the simple OLS forecasting model overestimates the standard deviation of prediction by 30.8 percent.

The mix of items also affects the predictive distribution. As an exercise, two very different bundles are considered. The first bundle is 1000 cords of sawtimber, and the second bundle is 2551.18 cords of pulpwood. These bundles have the same $HCB$ for the weighted forecasting model: $60,310$. The estimate of the highest competing bid using the OLS forecasting model is similar for the first bundle, $59,440$, but is very different for the second, $111,920$. Table 3 shows the standard deviation of the forecast for two bundles of timber for the two models. The models have rather different estimates for the predicted standard deviation. The estimate of the standard deviation for the OLS forecasting model is slightly higher than that for the weighted forecasting model for the first bundle, but is much lower for the second bundle.

The winner’s curse can be predicted. Consider a situation where the firm seeks to bid in an auction involving a tract of average size and mix, that is $(\hat{S}, \hat{P})=(1190.0, 282.5)$ and that the firm is considering bidding $81909k$ for various values of $k$, $1.3 \geq k \geq 0.7$. The simulation procedure was used to estimate the effect of the winner’s curse. (Parameters $c_S$ and $c_P$ were estimated from historical data: $c_S=1.8$ and $c_P=3.1$ were estimated by finding the coefficients of variations for all past auctions. The covariance of sawtimber and pulpwood quantities per acre ($\alpha$) was set to zero, and $\beta=-1$, the firm’s actual sampling plan.) Figure 3 presents estimates of the quantity bias, $(\hat{S}-S)$ and $(\hat{P}-P)$, and the probability of winning as a function of $k$. For the simulation trials, pulpwood quantity bias was (surprisingly) negative; however, this bias is never statistically different from zero at the 10 percent level (2-tails). For sawtimber, the quantity bias was positive (as expected) and statistically significant at 10 percent for all $k<1.25$. 

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Table 2: Steps 1, 2, and 4 for the example.

<table>
<thead>
<tr>
<th>a. Step 1 Model</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>OLS Forecasting Model</td>
<td></td>
</tr>
<tr>
<td>( HCB_t = 59.44 \hat{S}_t + 43.87 \hat{P}_t )</td>
<td></td>
</tr>
<tr>
<td>( (34.16)^* \ (7.276) )</td>
<td></td>
</tr>
</tbody>
</table>

\( R^2 \) (adjusted) = .988; \( F = 1339.96 \) (p value = .0001)
Standard deviation of residuals = 11,187.06
Skewness = .46 (.43) Kurtosis = 1.73 (.42)

<table>
<thead>
<tr>
<th>b. Step 2 Model</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Deriving Weight Estimates</td>
<td></td>
</tr>
<tr>
<td>( r_t^2 = 48.57 \hat{S}_t^2 - 258.87 \hat{P}_t^2 + 46.09 2\hat{S}_t\hat{P}<em>t - .21 (1/N) \hat{P}</em>{st}\hat{S}_t )</td>
<td></td>
</tr>
<tr>
<td>( (3.190)^* \ (-.782) \ (42) \ (-.71) )</td>
<td></td>
</tr>
<tr>
<td>+ 15.79(1/N)\hat{P}_{st}\hat{S}_t \hat{P}<em>t + 6.34 2N\hat{P}</em>{st}\hat{S}_t \hat{P}_t )</td>
<td></td>
</tr>
<tr>
<td>( (1.27)^\dagger \ (.06) )</td>
<td></td>
</tr>
</tbody>
</table>

\( R^2 \) (adjusted) = .61; \( F = 1339.96 \) (p value = .001)
Standard deviation of residuals = 150,333,601.
\( ^\dagger \text{Significant at } \alpha = .25 \) (two-tails)

<table>
<thead>
<tr>
<th>c. Step 4 Model</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Weighted Forecasting Model</td>
<td></td>
</tr>
<tr>
<td>( HCB_t / \hat{w}_t = 60.31 (\hat{S}_t / \hat{w}_t) + 23.64 (\hat{P}_t / \hat{w}_t) + 12.76 (\hat{P}_tQ_t / \hat{w}_t) )</td>
<td></td>
</tr>
<tr>
<td>( (39.87)^* \ (3.87) \ (1.63)^{\dagger\dagger} )</td>
<td></td>
</tr>
</tbody>
</table>

\( R^2 \) (adjusted = .9884; \( F = 906.18 \) (p value = .0001)
Standard deviation of residuals = .95
Skewness = .53 (.56) Kurtosis = -.358 (.08)
\( ^{\dagger\dagger} \text{Significant at } \alpha = .15 \) (two-tails)
*Numbers in parentheses are t statistics
All models have 32 observations

If the firm values sawtimber and pulpwood at $60.31 and $23.64 per cord, respectively, the reservation value of the tract is just $81,909 (unadjusted for the winner’s curse). Figure 4 reports the function \( A(b) \), the percent adjustment that should be made in the reservation value to correctly adjust for the winner’s curse as a function of the bid. As expected, the larger the bid, the smaller the effect of the winner’s curse. The degree of the adjustment is small, but potentially important considering the small margins involved in bidding.

To demonstrate this point, the following example shows opportunity losses that occur if the winner’s curse is not considered. Following the above example,
Figure 2: Standard deviation of forecast errors for the two models as a function of the value of auction.

Table 3: Comparison of the standard deviation of forecast errors for two different bundle mixes with the same point forecasts for the weighted forecasting model.

<table>
<thead>
<tr>
<th>Model</th>
<th>$\hat{\sigma}_t=1000.00$</th>
<th>$\hat{\mu}_t=2551.18$</th>
</tr>
</thead>
<tbody>
<tr>
<td>OLS forecasting</td>
<td>13,121</td>
<td>11,512</td>
</tr>
<tr>
<td>Weighted forecasting</td>
<td>9,484</td>
<td>17,947</td>
</tr>
</tbody>
</table>

assume the firm seeks to bid in an auction involving a tract of average size and mix. Suppose the bidder values the tract at $81909r$, for some $r>0$. Parameter $r$ reflects the effect of the bidder’s private valuation of the tract. If the firm sets its bid to maximize its expected winnings, and if $F$ is the cumulative distribution of $HCB$ for this auction, the optimal bid, $b^*$, solves the problem:

$$\text{Maximize}_{b \geq 0} (81909 (r - A(b)) - b) F(b). \tag{22}$$

The optimal bid $b^*$ was calculated as a function of $r$ considering the winner’s curse and ignoring it ($A(b)$=0). Figure 5 reports the opportunity losses that occur if the winner’s curse is ignored, expressed as a percent of the optimal expected value of auction winnings. Opportunity losses are measured as the difference between expected winnings if the firm’s bid considers the winner’s curse and expected winnings if the firm’s bid ignores it, divided by expected winnings if the firm’s bid considers the winner’s curse. The losses are large in percentage terms, more than 200 percent if $r$ is less than .83. If $r$ is close to 1, the losses are virtually zero,
**Figure 3:** Probability of winning and quantity bias due to the winner’s curse for timber types, if the firm’s bid is $81909k.

![Graph showing probability of winning and quantity bias](image)

**Figure 4:** Function A($81909k$): Percent adjustment in the firm’s reservation value due to the winner’s curse if its bid is $81909k.

![Graph showing percent adjustment](image)

suggesting that the winner’s curse may be ignored. This tells us that when the bidder’s reservation value is close to that of the expected winning bid or higher, the effect of the winner’s curse on optimal bids is negligible and can be ignored. However, bargain hunters (firms with low reservation bids and, therefore, low bids) must consider the winner’s curse to avoid opportunity losses.
Figure 5: Opportunity losses if the winner's curse is ignored as a function of the reservation value.

MANAGERIAL IMPLICATIONS AND EXTENSIONS

This paper presented a statistical method to predict the highest competing bid and the winner's curse. As such, it has obvious use as a tool to prepare bids. Beyond this, it provides a manager with a way to determine whether the winner's curse due to quantity uncertainty is important enough to consider when bidding. The pioneering work of Capen, Clapp, and Campbell [1] on oil auctions showed that an auction winner may not be aware of the winner's curse even when it is significant. Therefore, the method has value in alerting managers as to whether they should mind the "curse."

Qualitatively, this paper demonstrates the important effect that bid size has on the curse: all other factors held constant, the higher a bid, the less significant the curse, and vice versa. Thus, a bidder that considers bidding more or less aggressively can evaluate this change on the size of the curse and expected winnings. Importantly, a bidder that previously ignored the curse but changes its strategy by bidding lower can determine how important the curse will be.

The paper also has application for bidding even when the winner's curse is unimportant. When auctions offer bundles of items with differing mixes, the method helps to compute a predictive distribution that fully considers the effect of item-mix on the variance of the distribution.

Future Research

Many improvements to this paper are possible. First, more sophisticated statistical estimation techniques could be used. Full maximum likelihood procedures could be used instead of the four stage OLS method. Then, estimates of the bid coefficients
and the uncertainty of the process could be determined in one stage. However, to do this all the strategic variables used to estimate $HCB$ in Step 4 must be preselected. In practice this may be difficult to do. Second, some of the results rely upon assumption of normality. Further research is required to understand distributional properties of the highest competing bid conditional upon the actual timber volumes when normality is not appropriate.

Further research is required to understand strategic play when the method is used to bid in sequential auctions. For example, if the firm's bidding method becomes known to competitors, they may bid strategically to affect parameter estimates and bids. Some related analysis is provided by Greenberg [5], who studied strategic distortion of a seller's subjective probability distribution in a buyer-seller bargaining model. Also, there may be strategic effects due to common value information obtained about upcoming auctions when a bid is won. (This problem has been analyzed by [7].) These effects are not present in timber auctions because no quantity information about upcoming auctions is obtained when winning. Lastly, as pointed out by Pfeifer and Schmidt [15], there is no guarantee that the increased expected winnings using the method justifies a bidder's time and effort to implement it. [Received: April 20, 1993. Accepted: October 27, 1993.]

REFERENCES

APPENDIX 1

Estimating the Variance and Covariance of $e_{St}$ and $e_{Pt}$

Define:

$\bar{Y}_{SF}$ = Sample mean for sawtimber (measured in cords per acre) observed by the firm, and

$\bar{Y}_{SH}$ = Estimated quantity for sawtimber (measured in cords per acre) of the highest competing bidder.

$\bar{Y}_{SF}$ and $\bar{Y}_{SH}$ are unbiased estimates of sawtimber quantity per acre. Define $Y_{PF}$ and $Y_{PH}$ to be the equivalent variables for pulpwood. If $N_t$ is the size of tract $t$ as measured in acres and the acre is the sampling unit, then $N_tP_F, (N_tP_H)$ is the sample size for the firm (other bidders). Then, $e_{St} = N_t(\bar{Y}_{SF} - \bar{Y}_{SH})$. If $\sigma_{St}^2$ is the variance of sawtimber quantities for randomly selected one acre segments of tract $t$, $NUM_t$ is the number of other bidders and Var$(m)$ is the variance of the order statistic of the largest of $m$ independent draws from a standard normal distribution, then $\text{Var}[\bar{Y}_{SH}] = \text{Var}(NUM_t)\sigma_{St}^2/N_tP_H$. The value of Var$(m)$ is found from tables (for example, [16]).

It follows that

$$\text{Var}[e_{St}] = N_t^2 \left[ \frac{\sigma_{St}^2}{N_tP_F} + \frac{\text{Var}(NUM_t)\sigma_{St}^2}{N_tP_H} \right].$$

If $\mu_t = \bar{Y}_{St}/N_t$ and $c_{St} = \sigma_{St}/\mu_t$, then $\sigma_{St}^2 = (c_{St}^2/N_t)^2$ and

$$\text{Var}[e_{St}] = \frac{c_{St}^2}{N_t} \frac{\text{Var}(NUM_t)P_F + P_H}{P_FP_H} S_t^2.$$

An unbiased estimator of $S_t^2$ is $\hat{S}_t^2 + N_t^2c_{St}^2$. Assume that $\hat{S}_t^2 \gg N_t^2c_{St}^2$, so that

$$\text{Var}[e_{St}] = \frac{c_{St}^2}{N_t} \frac{\text{Var}(NUM_t)P_F + P_H}{P_FP_H} S_t^2. \quad (A1)$$

A similar expression for $\text{Var}[e_{Pt}]$ is readily found. Next note that:

$$\text{Cov}[e_{St}, e_{Pt}] = \text{Cov}[N_t(\bar{Y}_{SF} - \bar{Y}_{SH}), N_t(\bar{Y}_{PF} - \bar{Y}_{PH})]$$

$$= N_t^2 \text{Cov}[(\bar{Y}_{SF}, \bar{Y}_{PF})] + N_t^2 \text{Cov}[(\bar{Y}_{SH}, \bar{Y}_{PH})]$$

$$+ N_t^2 \text{Cov}[(\bar{Y}_{SF}, \bar{Y}_{PH})] + N_t^2 \text{Cov}[(\bar{Y}_{SH}, \bar{Y}_{PF})].$$

Because the firm and all the opponents sample independently, the last two terms are zero. Let $\alpha_t$ be the covariance of per-acre quantities of sawtimber and pulpwood for auction $t$. If the highest competing bidder's estimated quantity for each timber-type
is the highest for all other bidders, then $\bar{Y}_{SH}, \bar{Y}_{PH}$ are asymptotically independent ([17]). Thus

$$\text{Cov}[e_{Si}, e_{Pr}] = N_t^2 \frac{\alpha_t}{P_FN_t^2}. $$

However, it may be desirable to alternately assume that the highest competing bidder’s estimate for each timber type is an independent estimate of timber quantity, in which case $N_t^2 \text{Cov}[\bar{Y}_{SH}, \bar{Y}_{PH}] = N_t^2 (\alpha_t / P_H N_t)$. This is the case when the winner’s curse is unimportant. Assume that

$$\text{Cov}[e_{Si}, e_{Pr}] = N_t^2 \frac{\alpha_t}{P_FN_t^2} + \tau N_t^2 \frac{\alpha_t}{P_H N_t^2}, $$

where $\tau$ is some parameter which must be estimated. Canceling and rearranging yields,

$$\text{Cov}[e_{Si}, e_{Pr}] = \alpha_t N_t \frac{\tau P_F + P_H}{P_F P_H}. $$

(A2)

Using (5), (A1), and (A2),

$$\text{Var}[v_i] = \text{Var}[e_{Si}\hat{S}_i^2] + \text{Var}[e_{Pr}\hat{P}_i^2] + 2\beta \hat{S}_i \hat{P}_i$$

$$+ \frac{c_{Si}^2 \text{Var}(NUM_{Si}) P_F + P_H}{N_t P_F P_H} \beta_{Si}^2 \hat{S}_i^2$$

$$+ \frac{c_{Pr}^2 \text{Var}(NUM_{Pr}) P_F + P_H}{N_t P_F P_H} \beta_{Pr}^2 \hat{P}_i^2$$

$$+ 2\alpha_t N_t \frac{\tau P_F + P_H}{P_F P_H} \beta_{Si} \beta_{Pr}$$

$$+ \frac{1}{N_t} \frac{\text{Var}(NUM_{Si}) P_F + P_H}{P_F P_H} (\text{Var}[e_{Si}] c_{Si}^2 \hat{S}_i^2 + \text{Var}[e_{Pr}] c_{Pr}^2 \hat{P}_i^2). $$

(A3)

If the square of an item’s bid-price is much larger than the bid-price’s variance, in (A3) the seventh term will be small compared with the sum of the fourth and fifth terms and can be ignored:

$$\text{Var}[v_i] = \text{Var}[e_{Si}\hat{S}_i^2] + \text{Var}[e_{Pr}\hat{P}_i^2] + 2\beta \hat{S}_i \hat{P}_i$$

$$+ \frac{c_{Si}^2 \text{Var}(NUM_{Si}) P_F + P_H}{N_t P_F P_H} \beta_{Si}^2 \hat{S}_i^2$$

$$+ \frac{c_{Pr}^2 \text{Var}(NUM_{Pr}) P_F + P_H}{N_t P_F P_H} \beta_{Pr}^2 \hat{P}_i^2$$

$$+ 2\alpha_t N_t \frac{\tau P_F + P_H}{P_F P_H} \beta_{Si} \beta_{Pr}. $$

(A4)
**APPENDIX 2**

*Estimating the Variance and Covariance of $f_{St}$ and $f_{Pt}$*

Use the definitions made in Appendix 1. Then, $f_{St} = N_t \bar{Y}_{SH}$. It follows that

$$\text{Var}[f_{St}] = N_t^2 \left[ \frac{\text{Var}(\text{NUM}_t)}{\text{Var}(\text{NUM}_t) \sigma^2_{St}} \right] N_t P_H,$$

and using the fact that $\sigma^2_{St} = (c^2_{St} / N_t^2) S^2_t$ yields

$$\text{Var}[f_{St}] = \frac{c^2_{St}}{N_t} \frac{\text{Var}(\text{NUM}_t)}{P_H} S^2_t.$$

An unbiased estimator of $S^2_t$ is $\hat{S}^2_t = N_t f_{St}^2$. Again, assume that $\hat{S}^2_t \gg N_t c^2_{St}$ so that

$$\text{Var}[f_{St}] = \frac{c^2_{St}}{N_t} \frac{\text{Var}(\text{NUM}_t)}{P_H} \hat{S}^2_t.$$

A similar expression for $\text{Var}[f_{Pt}]$ can be found.

$$\text{Cov}[f_{St}, f_{Pt}] = \text{Cov}[N_t \bar{Y}_{SH}, N_t \bar{Y}_{PH}] = N_t^2 \text{Cov}[\bar{Y}_{SH}, \bar{Y}_{PH}].$$

Consistent with the derivation of (A2), $\text{Cov}[\bar{Y}_{SH}, \bar{Y}_{PH}]$ will be estimated as $\alpha_t / P_H N_t$, so

$$\text{Cov}[f_{St}, f_{Pt}] = \tau \alpha_t N_t \frac{1}{P_H}.$$

(A5)

Phillip J. Lederer is Associate Professor of Operations Management at the William E. Simon Graduate School of Business Administration, University of Rochester. He received a B.S. (physics) from SUNY at Stony Brook and M.S. and Ph.D. (applied mathematics) degrees from Northwestern University. His research studies are in the economics of operations management. He has published papers in such journals as *Econometrica, Journal of Manufacturing and Operations Management, Operations Research Letters, International Journal of Flexible Manufacturing Systems, Regional Science and Urban Economics*, and *Transportation Science*. 

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.