An Application of Bootstrapping for Determining a Decision Rule for Site Location

Paul P. Biemer

New Mexico State University

Sheryl E. Kimes

Cornell University

This article provides a general methodology for determining and evaluating a decision rule for hotel site location. Given (a) an indicator of hotel success, (b) an ideal decision rule based on this indicator if it were known without error, and (c) a model for predicting the value of the success indicator at a proposed site, we propose a procedure for finding the optimal model based decision rule for any specified optimality criterion and for evaluating the worth of the rule. The methodology is based on the statistical technique called bootstrapping. This method reduces the bias of conventional methods of estimation that have been applied in the context of sitelocation modeling. The methodology is illustrated using data from a large hotel chain in the United States and evaluated using an independent evaluation sample.

Keywords: Cross-validation; Discriminant analysis; Prediction error; Service management.

The motivation for this article arose from our work with a national hotel chain to develop an objective decision process for identifying sites for their new hotels. The firm's approach to hotel site location had previously not taken advantage of the extensive bank of information on their existing hotels but rather had been based more on the subjective judgment of their location experts. If a statistical model could be

developed that could reasonably predict success or failure of a potential site, this prediction could provide the experts with additional information on which to formulate a decision to build or not to build a hotel. This article describes and illustrates the modeling approach taken to help achieve their objectives.

The service-location literature indicates an extensive use of modeling for forecasting consumer demand of some services, particularly grocery stores, malls, and other shopping locations. For example, consider the class of spatial-interaction or gravity models that use the probabilities p_{ij} to estimate how many of the ith type of customer will select the jth facility to service their demands (Achabal, Gorr, and Mahajan 1984; Huff 1962, 1979). For these models, demand is defined as the product of the number of potential customers in the market area times p_{ij}. The models range from simple to complex but generally assume that customers will be drawn from the population proximate to a site. Since hotels draw their business from areas far removed from their physical site, this class of models is not suitable for hotel site selection. Other service-location models [e.g., Applebaum's (1966) analog model] that were not specifically developed for the lodging industry suffer from similar limitations.

The problem of hotel site location can also be viewed in a general regression-model setting in a way that can accommodate additional variables. Furthermore, rather than predicting the probability that a consumer will select a service at a particular site, other researchers have tried to predict total sales or some related variable at a new location (e.g., see Clawson 1974; Cottrell 1973; Hise, Kelly, Gable, and McDonald 1983; Martin 1967; Olsen and Lord 1979). Regression models are attractive for hotel site location because they allow us to consider alternative dependent variables, as well as numerous predictor variables such as site-specific socioeconomic factors, price, convenience, and degree of competition in forming a characterization of a desirable new site.

Unfortunately, many published location models suffer from methodological problems such as over specificity, multicollinearity, and inadequate validation. The models also tend to be firm specific and are usually not informative about the industrywide generators of demand for a service [cf. Craig, Ghosh, and

McLafferty (1984) for a review of such models]. These problems, combined with the problems of defining the hotel "market area" and the difficulty in measuring many location characteristics, have raised serious questions about the effectiveness of modeling for general site-location problems.

The remainder of the article provides the details of our modeling approach, which gave special emphasis to dealing with the difficulties apparent in published site-location models. Of central importance to our approach is the method of bootstrapping for evaluating the validity of alternative models. The bootstrap-validation step is an integral part of our model-building procedure and, as we shall see, provides information beyond validation statistics that may be used in the decision process.

1. THE PROBLEM

The hotel-location problem may be stated as follows: Given the information on existing hotels and a proposed site, we wish to predict y, some specified measure of performance of a hotel built at the site. For example, y may be occupancy rate, total revenue, profit margin, the probability that a traveler chooses the hotel, a (0, 1) variable indicating a "good" site $(y = 1)$ or a "bad" site $(y = 0)$, or any other suitable indicator of hotel performance.

Let n denote the number of existing sites (or a representative sample thereof), and suppose that at each site we have the data $\mathbf{x}_i = (\mathbf{t}_i, y_i)$ consisting of a k dimensional vector of predictor variables $\mathbf{t}_i$ and the performance indicator y_i. Given a new vector of predictors, $\mathbf{t}_0$, from a proposed site, we wish to predict the corresponding performance indicator y_0. We assume that $\mathbf{x}_1 \ldots, \mathbf{x}_n$ are independent realizations of $X =$ (T, Y), a random vector having some distribution F on $k + 1$ space, and we assume the same for the new case $(\mathbf{t}_0, y_0)$.

The assumption regarding the distribution of a new case $(\mathbf{t}_0, y_0)$ has important consequences for the way by which future sites are identified. Given that the n existing sites were selected using the traditional procedures of the site-evaluation experts, any new sites must be selected in the same way if the subsequent regression model is to be applied for predicting y. This requirement is consistent with our goal

of providing *supplementary* information to the decision makers and not eliminating the role of the evaluation specialists in selecting good candidate sites.

If y were known for the proposed site, a decision D(y, C) could be made in which

$$D(y, C) = D_1 \quad \text{if } y \geq C$$
$$= D_2 \quad \text{if } y < C \qquad (1.1)$$

for some specified constant C. For example, D_1 may be "build a hotel" or "investigate the site further," and D_2 is "not D_1" We refer to $D(y, C)$ as the *ideal* decision rule. (Generalizations to polytomous decision rules are straightforward.)

Given the predictor t_0 at a proposed site, denote by $\pi(t_0; \mathbf{X})$, where $\mathbf{X} = (\mathbf{x}_1 \ldots , \mathbf{x}_n)$, the model for predicting y_0 obtained by some model-fitting technique, perhaps least squares. Thus $\pi(t_0; \mathbf{X})$ is the model prediction of y_0. Consider the consequences of simply substituting the prediction $\pi(t_0; \mathbf{X})$ for y_0 in the decision rule $D(y, C)$. Of course, unless the model perfectly predicts y_0, we would expect that, due to the uncertainty in $\hat{y}_0 = \pi(t_0, \mathbf{X})$, $D(y_0, C)$ and D($\hat{y}_0$, C) might disagree; that is, $D(\hat{y}_0, C)$ might be the wrong decision. We may consider a decision rule $D(\hat{y}_0, C_\pi)$, however, in which C_π is chosen so that the probability of agreement between $D(\hat{y}_0, C_\pi)$ and $D(y_0, C)$ is maximized. Other criteria may also be used for choosing the "best" C_π.

In general, let

$$\mu(C_\pi) = \text{prob}(D(\hat{y}, C_\pi) = D_1 \mid D(y, C) = D_2)$$
$$= \text{prob}(\pi(t; \mathbf{X}) \geq C_\pi \mid y < C) \qquad (1.2)$$

and

$$\lambda(C_\pi) = \text{prob}(D(\hat{y}, C_\pi) = D_2 \mid D(y, C) = D_1)$$
$$= \text{prob}(\pi(t; \mathbf{X}) < C_\pi \mid y \geq C), \qquad (1.3)$$

referred to as the *error characteristic functions* of the decision rule $D(\hat{y}, C_\pi)$. Thus the risk of erroneously

deciding D_1 with the rule $D(\hat{y}, C_\pi)$ is $\mu(C_\pi)$ and the risk of erroneously deciding D_2 is $\lambda(C_\pi)$. Any criteria for determining the best C_π will typically be functions of the error characteristic functions (1.2) and (1.3).

As an example, let D_x denote "build a hotel" and let D_2 denote "do not build a hotel." Then $\mu(C_\pi)$ is the risk of making a poor "build" decision (referred to as a false positive decision) using decision rule $D(\hat{y}, C_\pi)$. Likewise, $\lambda(C_\pi)$ is the risk of making a poor "do not build" decision (false negative decision). Let the expected costs associated with the false positive and false negative decisions be k_1 and k_2, respectively; then a criterion for choosing the C_π that minimizes cost is

$$C^*: \min_{C_\pi} [k_1\mu(C_\pi) + k_2\lambda(C_\pi)]. \qquad (1.4)$$

Unfortunately, it is usually difficult to obtain adequate estimates of k_1 and k_2. Alternatively, one may choose to minimize the probability of making either a false negative or false positive decision. This may be accomplished by assuming that $k_1 = P(y < C)$ and $k_2 = P(y \geq C)$ in (1.4).

Given the potentially high cost of an erroneous "build" decision, the following criterion may be preferred:

$$C^*: \min_{C_\pi} (\lambda(C_\pi) \mid \mu(C_\pi) < \mu_0) \qquad (1.5)$$

for some constant μ_0; that is, choose C_π to minimize the risk of a false negative error, while maintaining the risk of a false positive error at a predetermined level. Other criteria are also possible.

In Section 2, we consider the following problems:

1. Given the prediction rule $\pi(y; \mathbf{X})$ and the ideal decision rule $D(y, C)$, find C_π to satisfy certain optimality criteria for the "model-based" rule $D(\hat{y}, C_\pi)$, where $\hat{y} = \pi(\mathbf{t}, \mathbf{X})$.

2. Let C^* denote the optimal value of C_π just obtained. Estimate from the data $\mathbf{X}$ the error rates $\mu(C^*)$ and $\lambda(C^*)$ associated with the model-based decision rule $D(\hat{y}, C^*)$.

Thus from problem 1 we obtain the "best" decision rule for applying the model $\pi(t, \mathbf{X})$. And from Problem 2

we obtain estimates of the probabilities that our model- based decision rule will lead to false positive and false negative decisions.

2. A SOLUTION

From the previous discussion, we see that any practical criteria for optimizing the rule $D(\hat{y}, C_\pi)$ will typically involve the minimization of some function of the error characteristic functions $\lambda(C_\pi$ and $\mu(C_\pi)$. Thus our primary goal is to estimate, on the basis of the data X, $\mu(C_\pi)$ and $\lambda(C_\pi)$ for C_π in, say, $[C_L, C_U]$ for constants C_L and C_U. We can then apply the optimality criteria to determine C^* using the estimated error functions. Furthermore, the error rates $\mu(C^*)$ and $\lambda(C^*)$ associated with $D(\hat{y}, C^*)$ are by-products of the optimization step.

So far, little has been said about the model $\pi(t, X)$ for predicting y. Of course, from a *practical* perspective it is important that the assumptions of the model-fitting technique hold and that $\pi(t, X)$ provide the best predictor of y available from X. These are *not* necessary conditions of the methodology for estimating the error rates associated with the decision rule $D(\hat{y}, C_\pi)$, however. All that is assumed is that the model-fitting process is replicable.

As before, let t_i denote the observation for the ith site and $\hat{y}_i$ the prediction of y_i using $\pi(t_i, X)$. Let $I(\cdot \mid \cdot)$ indicate the error of the decision $D(\hat{y}_I \mid C_\pi)$; that is,

$$I[D(\hat{y}_i, C_\pi), D(y_i, C)] = 0 \quad \text{if } D(\hat{y}_i, C_\pi) = D(y_i, C)$$
$$= 1 \quad \text{if } D(\hat{y}_i, C_\pi) \neq D(y_i, C).$$

Thus

$$\mu(C_\pi) = E\{I[D(\hat{y}_i, C_\pi), D(y, C)] \mid D(y, C) = D_1\}$$

$$(2.1)$$

And

$$\lambda(C_\pi) = E\{I[D(\hat{y}_i, C_\pi), D(y, C)] \mid D(y, C) = D_2\}$$

$$(2.2)$$

where expectation is taken over all $X = (T, Y) \sim F$ for F defined in Section 1.

The obvious estimators of $\mu(C_\pi)$ and $\lambda(C_\pi)$ are the *apparent error rates*

$$\hat{\mu}_a(C_\pi) = \sum_{i=1}^{n_1}{}' I[D(\hat{y}_i, C_\pi), D(y_i, C)]/n_1 \quad (2.3)$$

And

$$\hat{\lambda}_a(C_\pi) = \sum_{i=1}^{n_2}{}'' I[D(\hat{y}_i, C_\pi), D(y_i, C)]/n_2, \quad (2.4)$$

where $\sum'$ ($\sum''$) denotes summation over the n_1 (n_2) sites from the data base such that $D(y_i, C) = D_1$ (D_2) and $n_1 + n_2 = n$ (Efron 1983). Thus the apparent error rates are the proportions of observed errors made by D($\hat{y}$, C_π) on the data from which $n(t, X)$ was constructed. As Efron (1983) showed, both $\hat{\mu}_a(C_\pi)$ and $\hat{\lambda}_a(C_\pi)$ tend to be smaller than $\mu(C_\pi)$ and $\lambda(C_\pi)$ because the same data have been used both to construct and to evaluate D($\hat{y}$, C_π). Stone (1974), Geisser (1975), and Efron (1983) provided methods for estimating the mis- classification error rates associated with dichotomous rules such as D($\hat{y}$, C_π). Efron compared the performance of the commonly used *cross-validation* or "leave- one-out" approach as in "jackknife procedures" (Lachenbruch and Mickey 1968) with an approach based on the *bootstrap* (Efron 1979). Efron concluded that *cross-validation* "gives a nearly unbiased estimate of [the error rate], but often with unacceptably high variability, particularly if n is small" (Efron 1983, p. 328). Bootstrap procedures provide estimates with low variability and low bias, the price being paid by the much heavier computational load. In this age of inexpensive and fast computers, however, the price is easily affordable for many situations and especially for those like hotel site selection in which large construction or lost opportunity costs may be involved.

Efron and Gong (1983) described a procedure for estimating model-misclassification error in a dichotomous prediction model. By bootstrapping the model building and prediction process, the model-misspecification error, as well as the sampling error, is included in the estimated error rates. Our approach retains this feature of their procedure but extends their work in the following ways: (a) Both false positive and false negative error rates are estimated, (b) the error curves $\mu(C_\pi)$ and $\lambda(C_\pi)$ are estimated by systematically varying C_π, and (c) rather than a priori choosing a cutoff for classifying an observation as either 0 or 1, the bootstrapping procedure is employed for choosing the cutoff, C_π, that minimizes the model-misclassification error. The remainder of this section provides the general methodology. In Section 3, the methodology is applied to the data of our client firm.

An alternative approach to bootstrapping regression models was described by Freedman and Peters (1984). Their approach was based on a resampling of residuals for a single specification of the model. Because of the importance of model misspecification error in site location models, however, the Efron and Gong (1983) approach was taken so that misspecification error would be accounted for in the model-validation stage.

Let $\hat{F}$ denote the empirical probability distribution of the data $\mathbf{x}_1, \ldots, \mathbf{x}_n$, putting probability mass $1/n$ on each X/. Let $\hat{E}(\cdot)$ denote expectation with respect to random sampling from $\hat{F}$; that is, $\hat{E}(\cdot)$ is expectation over all n^n possible samples of size n drawn with replacement from $\hat{F}$,

$$\mathbf{x}_1^*, \mathbf{x}_2^*, \ldots, \mathbf{x}_n^* \overset{\text{iid}}{\sim} \hat{F}.$$

Then the bootstrap estimates of $\mu(C_\pi)$ and $\lambda(C_\pi)$ are given by (2.1) and (2.2), respectively, replacing $E(\cdot)$ with $\hat{E}(\cdot)$.

Denote by M the model-building process that was applied to the data $\mathbf{X}$ to produce the model $\pi(\mathbf{t}; \mathbf{X})$. To compute bootstrap estimates of the functions $\mu(C_\pi)$ and $\lambda(C_\pi)$ at some value $C_\pi = C_0$, the following steps are performed:

1. From the population of n sites in the data base, select a sample of n sites using simple random sampling *with replacement;* that is, repetitions are allowed. This sample will be referred to as the bootstrap sample (BSS) for iteration 1 and denoted by s_1.

2. Let $\mathbf{X}^* = [x_1^*, x_2^*, \ldots, x_n^*]$ denote the BSS data base corresponding to the n sites in s_1. This data base contains information on the explanatory variables as well as the dependent variable for all sites in s_1 repeated for each occurrence of each site in s_1. Apply the model building process M to the data base $\mathbf{X}^*$ to obtain the model $\pi_1(\mathbf{t}^*; \mathbf{X}^*)$.

3. Use the model $\pi_1(\mathbf{t}^*; \mathbf{X}^*)$ to predict $y_1 \ldots, y_n$, the performance indicators associated with all sites in the *original* population of n sites, and denote these predictions by $\hat{y}_{1i}$—that is, $\hat{y}_{1i} = \pi_1(t_i; \mathbf{X}^*)$—for $i = 1, \ldots, n$.

Finally, compute the BSS error rates given by

$$\hat{\mu}_1(C_0) = \sum_{i=1}^{n_1}{}' I[D(\hat{y}_{1i}, C_0), D(y_i, C)]/n_1 \quad (2.5)$$

and

$$\hat{\lambda}_1(C_0) = \sum_{i=1}^{n_2}{}'' I[D(\hat{y}_{1i}, C_0), D(y_i, C)]/n_2, \quad (2.6)$$

where $\sum'$, $\sum''$, n_1, and n_2 are as defined for (2.3) and (2.4). Thus $\hat{\mu}_1(C_0)$ is the proportion of n_1 sites having correct decision D_1 that are misclassified as D_2, and $\hat{\lambda}_1(C_0)$ is the proportion of n_2 sites having correct decision D_2 that are misclassified as D_1.

Steps 1-3 complete one bootstrap iteration. Let B denote the number of iterations required to obtain a desired precision in the estimates of $\mu(C_0)$ and $\lambda(C_0)$ [see Efron (1983) for a discussion on the choice of B]. Then repeat steps 1-3 for a total of B times for independent bootstrap samples $s_1, \ldots, s_B$. The

bootstrap estimators of $\mu(C_0)$ and $\lambda(C_0)$ are

$$\hat{\mu}(C_0) = \sum_{k=1}^{B} \hat{\mu}_k(C_0)/B$$

And

$$\hat{\lambda}(C_0) = \sum_{k=1}^{B} \hat{\lambda}_k(C_0)/B.$$

Finally, instead of computing $\hat{\mu}_k(\cdot)$ and $\hat{\lambda}_k(\cdot)$ for only one value of C_π at bootstrap iteration k, the procedure can be extended by computing error rates at each iteration for multiple C_π values, say $C_L = C_1 < C_2 < \ldots < C_m = C_u$, so that the functions $\mu(C_\pi)$ and $\lambda(C_\pi)$ can be approximated over the range $C_L \leq C_\pi \leq C_U$. Thus, for BSS k $(k = 1, \ldots, B)$, $2m$ error rates $\hat{\mu}_k(C_j)$ and $\hat{\lambda}_k(C_j)$ $(j = 1, \ldots, m)$ are computed. The average over B BSS's for each cutoff C_j yields the $2m$ bootstrap estimates $\hat{\mu}(C_j)$ and $\hat{\lambda}(C_j)$ $(j = 1, \ldots, m)$. With the estimates of $\mu(C_\pi)$ and $\lambda(C_\pi)$, any of the optimality criteria discussed earlier can be applied to the estimated functions to obtain an estimate of the optimal C_π—that is, C*.

3. AN EXAMPLE

As an illustration of the general methodology, we consider the specific situation of our client firm. [For a full discussion, see Kimes (1987).] The firm operates approximately 200 inns aimed primarily at frequent business travelers. The inns are all in the range of 100150 rooms, are very homogeneous in terms of quality of construction, and are nearly all located on major highways.

3.1 Data Analysis

Because new inns typically experience initial instability in occupancy, operating costs, and so forth, the analysis was confined to mature inns operated by the firm. (A mature inn was defined as one that had been in operation for at least three years and that had experienced a leveling off of occupancy rates.) Data

10

for a sample of 57 mature inns were collected for numerous response and independent variables for the years 1983 and 1986. For the purpose of collecting data on variables relating to potential attractiveness of a location, we defined an inn's market area as the four-mile radius surrounding the inn (Tallis 1983). For each inn's market area, we collected data on demographic variables (e.g., population, income, unemployment), physical variables (e.g., accessibility, sign visibility, traffic count), competitive variables (e.g., amount of competition, competitive room rate), and demand-generator variables (e.g., hospitals, office space, colleges, military bases). We experimented with alternative dependent variables including total occupancy, occupancy rate, total revenue, operating income, and operating margin. Operating margin [defined as operating revenue minus operating costs (not including depreciation and interest expense), all divided by operating revenue] exhibited the highest correlations with the independent variables and appeared the most stable over time. Furthermore, this was a variable of critical importance for which management was readily able to formulate decision rules based on the expected profitability of a proposed site. Thus operating margin was chosen for the dependent variable y with ideal decision rule

$$D(y, 35) = \{\text{build}\} \quad \text{if } y \geq 35\%$$
$$= \{\text{do not build}\} \quad \text{if } y < 35\%. \quad (3.1)$$

In words, if it were known that a proposed site would achieve at least a 35% operating margin, the decision would be made to build at the site. Otherwise, the site would be rejected. In actuality, our client firm was interested in observing the model decisions for several cutoff values, C, not just 35%. We chose 35% here, however, to illustrate the methodology.

Prior to the model-building phase, exploratory data analysis techniques were used to obtain suitable transformations of some of the independent variables (Tukey 1977). This resulted in 20 variables, which entered into the data base $\mathbf{X}$ for the analysis. To obtain the best predictor of operating margin y, a model-building procedure (Efron and Gong 1983) equivalent to the following was used:

1. To reduce the number of variables to the more important subset, single variable regressions were run for each independent variable. Only those significant at the .05 level were retained.

2. A forward stepwise regression was run for the retained variables. Only those significant at the .10 level were allowed to enter the model.

3. To obtain a more parsimonious model, a second stepwise regression was run for the remaining retained variables. All of those significant at the .05 level were retained and constituted the final model.

Multicollinearity was checked using condition numbers (Belsley, Kuh, and Welsch 1980), and appropriate regression diagnostic tests were conducted. The final data matrices were found to be well conditioned with no serious outlier or leverage problems. This procedure then constituted the model-building process, M, to be replicated for each bootstrap sample, which explains its mechanical nature. In actual practice, the process M may only approximate the model-building process actually used for constructing the original model.

3.2 Results

The outcome of the model-building process for the year 1986 was the model in Table 1 for predicting profit margin (R^2 = .46). The model condition number was 2.5, indicating a well-conditioned data matrix—that is, no problem with multicollinearity. Since this is a predictive model, no cause-effect relationships should be attributed to variables in the model. The model reflects, however, the importance of predicting the operating margin of a site of market penetration, room rate (set in relation to the competition), and market-area income.

The negative coefficient associated with income indicates that this chain does better when located in moderate income areas.

Table 1. Final Model for Predicting Profitability

Variable	Coefficient	Standard error
Intercept	39.05	1.06
State population per inn (1,000s)	−2.34	1.09
Inn room rate (dollars)	5.82	1.09
Root median income in market area (1,000s)	−2.87	1.07

Although this may seem odd, moderate income areas are usually associated with more commercial demand generators than higher income areas. The negative coefficient associated with market penetration is a function of the way in which the variable was measured. Higher state population per inn (poor market penetration) leads to poorer performance.

The next step was to obtain estimates of the error functions $\mu(C_\pi)$ and $\lambda(C_\pi)$ associated with the model in Table 1 and the model-based decision rule $D(\hat{y}, C_\pi)$ using the three-step bootstrap procedure described in Section 1. The model-building process M for the bootstrap procedure was described in Section 3.1. The parameter B was set at 200.

A graph of the resulting estimates of $\mu(C_\pi)$ and $\lambda(C_\pi)$ for $0 \leq C_\pi \leq 60$ appears in Figure 1 (see the dashed curves). The corresponding apparent error rate curves for the same range of C_π are also shown in Figure 1 (see the solid curves). The discrepancy between the curves is an estimate of the bias in the apparent error rates. Note that the estimated bias may be either positive or negative but tends to be negative in the range of interest—that is, $.01 \leq \mu(C_\pi) \leq .1$. Thus in this range the apparent error rates are "too optimistic." Hence decision makers would be taking a somewhat greater than desired risk in using the apparent error rates to choose C*. This point is expounded in Section 3.3.

Of primary importance to management was the ability to be able to evaluate the risk of building a hotel at an unprofitable site. Therefore, Criterion (1.5) was used for selecting the optimal value of C_π. For any given μ_0 (i.e., risk of building at an unprofitable site), the corresponding value of C_π can be obtained directly from Figure 1; C* then is the smallest C_π satisfying $\mu(C_\pi) \leq \mu_0$. Table 2 provides values of C* for

alternative levels of risk. Thus, for example, to limit the risk of a false positive decision to 5%, the following

model-based decision rule would be used with the model for $\hat{y}$ in Table 1:

$$D(\hat{y},\ 43) = \{build\} \quad if\ \hat{y} \geq 43$$
$$= \{do\ not\ build\} \quad if\ \hat{y} < 43. \quad (3.2)$$

From Table 2 we see that the corresponding risk of *not* building at a profitable site is .62 for this decision

rule.

Table 2.   Estimates of C^* and $\lambda(C^*)$ for Alternative Values of $\mu(C^*)$

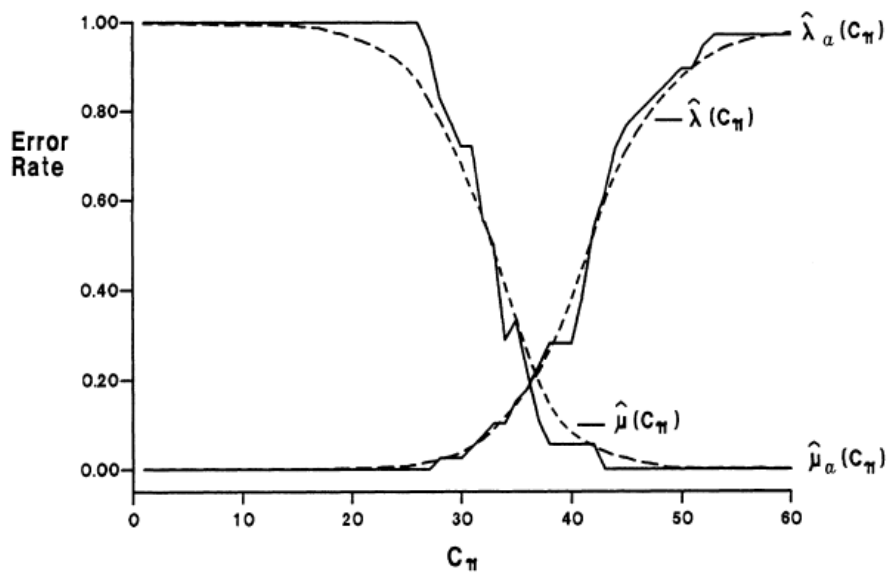| Probability of a bad build decision, $\hat{\mu}(C^*)$	Optimal C^* for $D(\hat{y},\ C^*)$	Probability of a bad no-build decision, $\hat{\lambda}(C^*)$
.02	43%	.62
.05	43%	.62
.10	38%	.28

Figure 1. The Functions $\hat{\lambda}(C_\pi)$, $\hat{\mu}(C_\pi)$, $\hat{\lambda}_a(C_\pi)$, and $\hat{\mu}_a(C_\pi)$ for the Model $\pi(t,X)$.

14

It should be emphasized that the error rates associated with the model decision rule (3.2) applies to agreement or disagreement with the ideal decision rule (3.1). If some criterion for determining a build/no-build decision other than (3.1) is formulated—for example, $D(y, 10)$—the bootstrap procedure would necessarily have to be repeated to obtain the corresponding error characteristic functions.

3.3   External Validation

In 1987, the firm operated 151 mature inns. Current data were collected on all inns, including the 57 inns selected for the original analyses and model development, to test the approach suggested here. Eventually, the bootstrap-based decision methodology will be replicated for all 151 inns. For now, however, the 1987 data affords an opportunity to evaluate the 1986 decision model on an independent sample of 94 inns. In particular, we wish to determine the number of false positive and false negative decisions that would be made if the decision model were used to classify the 94 inns as having operating margins above 35% (corresponding to a build decision at a proposed site) or less than 35% (corresponding to a no-build decision). These results are summarized in Table 3.

The "Expected" columns are the error rates expected from Figure 1 using a profitability critical value of 43, which corresponds to a false positive risk of .05. The "Actual" columns report the observed rate of false positive (μ) and false negative (λ) error in applying the decision rule $D(\hat{y}, 43)$ to the data.

Several factors need to be considered when interpreting the results in Table 3. First, since the model was fit to 1986 data and is being evaluated one year later on 1987 data, the effect of "aging" of the model is also seen in the error rates. For example, if the firm's economic condition in 1987 was such that operating margins for hotels generally were less than in 1986, this could increase the probability of a false positive error. Although the full 1986 data is not available, the available evidence indicates that operating margins for the firm tended to be less in 1987 than in 1986. This effect could be alleviated, however, if the

decision rule were updated annually using current data. Second, a substantial number of inns "matured" in the intervening year. Thus the independent sample has a lower average age than the data-base sample. This would also tend to affect the model's predictive success, since the longevity of an inn is strongly correlated with operating margin. In fact, the variable number of years in operation was deliberately excluded from the prediction model, since it was uninformative for new-site selection and tended to "overwhelm" the other variables in the data base. Again, if the model-building procedure were updated perhaps annually, using all currently mature inns, this effect could also be eliminated. Finally, although the bootstrap procedure eliminates much of the bias in the estimated error rates, it does not eliminate all of it. Further, the estimates of μ and λ are also subject to sampling variance.

Despite these limitations, the model's overall performance was quite acceptable even though the risk of a false positive error was somewhat higher than predicted for the total population. The results of this analysis serve to emphasize the importance of using the model-based decisions as an objective *guide* to hotel site location, not a substitute for human judgment.

Table 3. Comparison of Predicted and Actual Error Rates Using the 1986 Decision Rule on 1987 Data

Sample	Number of inns	Expected		Actual	
		μ	λ	μ	λ
Data-base sample	57	.05	.60	.042	.545
Independent sample	94	.05	.60	.132	.707
Total population	151	.05	.60	.100	.648

Finally, consider the decision rule obtained based on apparent error rates; that is, suppose that C_π was chosen with the same objective as before—that is, $\mu(C_\pi) \leq .05$ and $\lambda(C_\pi)$ is minimized—but now using the apparent error-rate curve in Figure 1. From the figure, we see that this corresponds to a profitability-critical value of 38%. Applying the rule $D(\hat{y}, 38)$ to the 151 inns resulted in 22 false positive

decisions (37% error rate) and 16 false negative decisions (18% error rate) as opposed to only six false positive decisions (10% error rate) and 43 false negative decisions (47% error rate) using the bootstrap method.

A question that might be raised regarding our procedure is this: What happens in the long run as fewer unprofitable inns are built and the range of dependent variables is truncated? Will not the model then become unstable and biased due to "censoring" of the sample? First, it is unlikely that the client firm will have a perfect record in always choosing new sites with profit margins that exceed their desired cutoff value C. Second, it is conceivable that as the average profit margin for existing sites increases, so too will the firm's desired cutoff value C increase. Therefore, there should always be sites with values of the dependent variable that adequately span the range of interest and thus provide for model predictability.

Finally, an advantage of our bootstrap method is that model-misspecification error and model instability are reflected in the estimates of μ and λ, the misclassification error rates. Our recommendation to the firm is to respecify the model and estimate λ and μ each year when new data become available, thereby monitoring the validity of the model. As long as μ and k are acceptable, the model can provide important information to aid the firm in the selection of new hotel sites.

4. SUMMARY

A general methodology was developed for obtaining an "optimum" decision rule for building a new hotel at a proposed site based on a model prediction of the success of a hotel at the site. The methodology incorporates bootstrapping to avoid the short falls of other methodologies for evaluating model prediction error such as cross-validation. By bootstrapping the decision rule itself, an improved decision rule is obtained that is better calibrated to the maximum acceptable levels of false positive and false negative error than decisions relying on apparent misclassification-error rates. The methodology is applicable for any ideal decision rule D(y, C) and any criterion for optimizing the model- based decision rule D($\hat{y}$, C_π), which is a function of the error characteristic functions (C_π) and (C_π).

The methodology was illustrated for a sample of 57 mature inns owned by a national hotel chain. For this example, the bias in the use of apparent error rates for optimizing the choice of C* in the model-based rule D(, C*) was demonstrated.

Finally, the bootstrap decision rule was evaluated on an independent sample of 94 mature inns owned by the firm. Furthermore, the bootstrap decision rule was compared with the standard regression-forecasting approach (based on apparent error rates). Consistent with the theory, our bootstrap procedure outperformed the standard regression procedures.

In general, applications of this methodology are limited to situations in which the dependent variable (i.e., profitability) has not been censored by management decisions to close unprofitable sites. For our client's hotel chain, inns operating below an acceptable profit margin were allowed to continue in operation. Thus the parameter estimates for our model are consistent estimates. This is not atypical because hotel chains are usually more concerned with long-term profitability than with short-term fluctuations in the profitability of a particular inn. In addition, hotel chains may want to provide national or regional coverage and may keep some unprofitable properties open in an attempt to maintain this coverage. In situations in which truncation of the dependent variable has occurred, however, the estimates of the model parameters will not, in general, be consistent, and predictive ability of the model could be diminished.

As explained in Section 3.3, the number of years that an inn is in operation was deliberately excluded from our model, since inclusion of this variable eliminated other variables that were more informative and useful for selecting *new* sites. Exclusion of this variable comes with a cost that is reflected in the model error rates. In other situations, however, it may be improper to exclude the age of the inn. For example, managers may wish to know *when,* if ever, a new site will be profitable. Therefore, a general recommendation for treating the age variable is not possible and will depend on the data set and the objectives of the model-building process.

References

Achabal, Dale, Gorr, William L., and Mahajan, Vijay (1984), "MULTILOC: A Multiple Store Location Decision Model," *Journal of Retailing,* 58, No. 2, 5-25.

Applebaum, William (1966), "Methods for Determining Store Trade Areas, Market Penetration and Potential Sales," *Journal of Marketing Research,* 3, 127-141.

Belsley, David A., Kuh, Edwin, and Welsch, Ray E. (1980), *Regression Diagnostics,* New York: John Wiley.

Clawson, C. Joseph (1974), "Fitting Branch Locations, Performance Standards, and Marketing Strategies to Local Conditions," *Journal of Marketing,* 38, *12-1 A.*

Cottrell, James L. (1973), "An Environmental Model for Performance Measurements in a Chain of Supermarkets," *Journal of Retailing,* 48, No. 3, 51-63.

Craig, C. Samuel, Ghosh, Avijit, and McLafferty, Sara (1984), "Models of the Retail Location Process: A Review," *Journal of Retailing,* 60, No. 1, 5-36.

Efron, Bradley (1979), "Bootstrap Methods: Another Look at the Jackknife," *The Annals of Statistics. 1,* 1-26.

--- (1983), "Estimating the Error Rate of a Prediction Rule: *Improvement on Cross-Validation,"* Journal of the American Statistical Association, *78, 316-331.*

Efron, Bradley, and Gong, Gail (1983), "A Leisurely Look at the Bootstrap, the Jackknife and Cross-Validation," *The American Statistician,* 37, 36-48.

Freedman, D. A., and Peters, S. C. (1984), "Bootstrapping a Regression Equation: Some Empirical Results," *Journal of the American Statistical Association,* 79, 97-106.

Geisser, S. (1975), "The Predictive Sample Reuse Method With Applications," *Journal of the American Statistical Association,* 37, 320-328.

Hise, Richard T., Kelly, J. Patrick, Gable, Myron, and McDonald, James B. (1983), "Factors Affecting the Performance of Individual Chain Store Units: An Empirical Analysis," *Journal of Retailing,* 59, No. 2, 22-39.

Huff, David L. (1962), "A Probabilistic Analysis of Consumer Behavior," in *Emerging Concepts in Marketing,* ed. W. D. Decker, Chicago: American Marketing Association.

---(1979), "Retail Location Theory," working paper, University of Texas, Dept, of Management Science and

Information Systems.

Kimes, Sheryl E. (1987), *Location Analysis in the Lodging Industry,* unpublished Ph.D. dissertation, University of Texas at Austin, Dept, of Management Science and Information Systems.

Lachenbruch, P., and Mickey, R. (1968), "Estimation of Error Rates in Discriminant Analysis," *Technometrics,* 10, 1-11.

Lord, J. Dennis, and Lynds, Charles D. (1981), "The Use of Regression Models in Store Location Research: A Review and Case Study," *Akron Business and Economic Review,* 12, 2.

Martin, Preston (1967), "Savings and Loans in New Submarkets: Search for Predictive Factors," *Journal of Marketing Research,* 4, 163-166.

Olsen, Lola M., and Lord, J. Dennis (1979), "Market Area Characteristics and Branch Bank Performance," *Journal of Bank Research,* 10, 102-110.

Stone, M. (1974), "Cross-Validatory Choice and Assessment of Statistical Predictions," *Journal of the Royal Statistical Society,* Ser. B, 36, 111-113.

Tallis, Alan (1983), "Trade Area Analysis," professional report, University of Texas at Austin, Graduate School of Business.

Tukey, John W. (1977), *Exploratory Data Analysis,* Reading, MA: Addison-Wesley.