Journal of Accounting and Economics 8 (1986) 3–35. North-Holland

PREDICTING TAKEOVER TARGETS

A Methodological and Empirical Analysis

Krishna G. PALEPU*

Harvard University, Boston, MA 02163, USA

Received January 1985, final version received August 1985

Several published studies claim that acquisition targets can be accurately predicted by models using public data. This paper points out a number of methodological flaws which bias the results of these studies. A fresh empirical study is carried out after correcting these methodological flaws. The results show that it is difficult to predict targets, indicating that the prediction accuracies reported by the earlier studies are overstated. The methodological issues addressed in this paper are also relevant to other research settings that involve binary state prediction models with skewed distribution of the two states of interest.

1. Introduction

A number of empirical studies have attempted to construct statistical models using publicly available financial information to predict acquisition targets. These include Simkowitz and Monroe (1971), Stevens (1973), Castagna and Matolcsy (1976), Belkoui (1978), and Dietrich and Sorensen (1984). The results reported by these studies indicate that such models have impressive ability to predict acquisition targets six to twelve months before the announcement of takeovers. For example, Simkowitz and Monroe report that their multiple discriminant model correctly predicts 83% of the targets and 72% of the non-targets in the sample used in estimating the model, and 64% of the targets and 61% of the non-targets in a holdout sample. The other studies report prediction accuracies ranging from 70% to 90%.

In contrast to the predictive ability claimed by the above studies, however, the stock market does not seem to predict acquisition targets with a high degree of accuracy even three months prior to the announcement of takeover bids. The pre-takeover stock price movement of target firms reported by Dodd

*I would like to thank Michael Jensen and Robert Kaplan for suggesting the idea for this paper and providing valuable comments on an earlier version. Parts of this paper are based on my doctoral dissertation research at MIT and I am indebted to my advisor Stewart Myers for his helpful guidance throughout that project. I also benefited from comments of Ken Gaver, Daniel McFadden, Morris McInnes, Jim Ohlson, Ram Ramkrishnan, Ross Watts, Jerold Zimmerman, and especially Richard Ruback. Partial funding for this research has been provided by the Division of Research, Harvard Business School.

and Ruback (1977), Asquith (1983) and others indicates that the market receives most of the signals that a firm is a probable target during a very short period around the announcement of a takeover bid. In a recent review of this evidence, Jensen and Ruback (1983, p. 29) argue that 'it is difficult, if not impossible, for the market to predict future targets'.

In light of the above, the results reported by the earlier acquisition studies imply that the models developed by them are better able than the stock market to identify future takeover targets. Stated differently, if the claims of the earlier prediction studies are valid, it is possible to earn abnormal returns using the prediction models. To probe this further, this paper undertakes a methodological and empirical analysis of takeover prediction with two related objectives. The first objective is to analyze the methodological problems associated with the development of binary state prediction models when the distribution of the two states of interest is skewed and to illustrate ways to avoid these problems in an applied context. The second objective is to examine whether it is possible to predict targets with a high degree of accuracy after correcting the methodological flaws of the earlier studies.

A critical examination of the methodology used by the earlier acquisition studies shows that there are three principal methodological flaws which make their reported prediction accuracies unreliable. First, the use of non-random, equal-share samples in the model estimation, without appropriate modification to the estimators, leads to inconsistent and biased estimates of the model parameters and the acquisition probabilities. This results in overstating the model's ability to predict targets. Second, the use of equal-share samples in prediction tests leads to error rate estimates that fail to represent the model's predictive ability in the population. Third, the use of arbitrary cutoff probabilities in prediction tests without specifying a decision context, the relevant state-payoff matrix, and the prior state probabilities, makes the reported prediction accuracies difficult to interpret.

The empirical study described in this paper attempts to correct the above methodological problems. The estimation procedure explicitly considers the sampling scheme employed. The prediction tests are conducted on a group of firms that approximates the population over which the model would be used in a realistic forecasting application. The predictive usefulness of the model is tested in the context of a specific forecasting application, and the optimal cutoff probability used in the tests is derived by explicitly considering the relevant payoff function and prior probabilities. The study also improves upon the earlier ones by employing an acquisition probability model that is developed from the economics of the acquisition process and by specifying variables based on hypotheses suggested by the literature.

The empirical results reported in this study differ markedly from those of the earlier acquisition prediction studies. A group of 163 targets and 256 non-targets listed on the New York and American stock exchanges is used to estimate a logit probability model with nine independent variables. While the estimated model is found to be statistically significant, its explanatory power is small. The magnitudes of the estimated acquisition probabilities are in general very small. The predictive ability of the model is tested on a group of firms consisting of 30 targets and 1087 non-targets. The results indicate that, while the model correctly identifies a high percentage of actual targets, it erroneously predicts a large number of non-targets as targets. Hence, it is not possible to earn significant abnormal returns by investing in firms that are predicted by the model to be potential acquisition targets.

While the methodological issues raised in this paper are analyzed in the context of acquisition prediction, they are relevant to any research problem that involves the development of a binary state prediction model, especially when the two states of interest are present in the population with unequal frequencies. Prominent areas of accounting research where the issues addressed in this paper are relevant include the prediction of corporate bankruptcy and the explanation/prediction of accounting policy choices of firms. A number of bankruptcy prediction studies employ methodologies similar to those of the previous acquisition prediction studies and hence suffer from the methodological flaws discussed in this paper.[1,2] Similarly, the accounting policy choice studies that analyze policy alternatives which are chosen by firms in the population with vastly different frequencies face the sampling and prediction testing problems discussed in this paper.[3,4]

The rest of the paper is organized as follows. In section 2, the methodological issues in the acquisition prediction literature are analyzed. This includes the issues related to sampling and the optimal cutoff probability. Section 3 describes the data and methods used in the present study. The empirical results

[1] In a recent paper, Zmijewski (1984) provides a methodological critique of the bankruptcy prediction literature focusing specifically on three issues: the choice between various statistical specifications of the bankruptcy probability model, the problem of incomplete data availability for some members of the population, and the effects of using non-random state-based samples for model estimation. Zmijewski's analysis overlaps the analysis of this paper in examining the biases from using state-based samples for model estimation. His study, however, does not touch upon two other issues addressed by this paper, namely, the bias introduced by the use of state-based samples in prediction tests, and the problems of using an arbitrary cutoff probability in prediction tests. Also, in a more general critique of the use of discriminant analysis, Eisenbeis (1981) points to some of the problems discussed here.

[2] As pointed out later in this paper, the bankruptcy prediction studies by Ohlson (1980) and Altman, Haldeman and Narayanan (1977) avoid some of the problems pointed out in this study.

[3] See Holthausen and Leftwich (1983) for a recent review of this literature.

[4] Consider, for example, a study that investigates the determinants of depreciation accounting policy choice of firms. Since a very large proportion of firms in the population chooses straight-line depreciation policy, a non-random sample is usually employed to obtain adequate representation of other depreciation policy choices. As shown in this paper, in the absence of appropriate modification to the estimation procedure, this results in biased coefficient estimates and impairs the validity of hypotheses tests.

are presented in section 4. The paper ends with a summary and discussion of conclusions in section 5.

2. Methodological issues in acquisition prediction

Several methodological flaws in the acquisition prediction literature lead to erroneous conclusions on the predictive ability of the estimated models. These are discussed in this section and modifications to the methodology to avoid these problems are proposed.

2.1. Sampling

The typical procedure used in the acquisition studies is to draw a sample with an approximately equal number of targets and non-targets.[5] This type of sample, referred to as a state-based sample in this paper, is not a pure random sample because, unlike in random sampling, a firm's probability of being selected into a state-based sample is a function of its acquisition status, i.e., whether a firm is a target or not. The practice in the acquisition prediction literature has been to employ state-based samples in conjunction with inference procedures which assume random sampling. This leads to biased and incorrect inferences as shown below.

2.1.1. State-based sample for model estimation

Consider a population of N firms consisting of N_1 targets and N_2 non-targets. Suppose the desired sample size is n. In the case of random sampling, n firms are drawn randomly from the entire population. Under state-based sampling, n_1 firms are drawn randomly from the target subpopulation and n_2 firms are drawn from the non-target subpopulation, n_1 and n_2 totaling up to n. Typically, n_1 and n_2 are set to be approximately equal.

There is a valid econometric justification for preferring a state-based sample over a random sample in the estimation of an acquisition prediction model because the number of targets is very small compared to the number of non-targets in the population. If a random sample were to be drawn from such a population, the sample would be likely to consist of an overwhelming majority of non-targets and very few targets. The 'information content' of such a sample for model estimation is quite small, leading to relatively imprecise parameter estimates. The sample can be enriched informationally by making the sample proportions of targets and non-targets more evenly balanced. A state-based sample accomplishes just this.

[5] For example, Stevens (1973) uses a sample consisting of 40 targets and 40 non-targets for estimating the model, and another sample of 20 targets and 20 non-targets for testing the model's predictive ability. Similar samples are used by the other studies.

Manski and Lerman (1977) and Manski and McFadden (1981) show that in a population like the one described above, an appropriate state-based sample provides more efficient estimates compared to a random sample of the same size. Alternatively, for a given level of precision, state-based sampling reduces the required sample size. Based on an extensive simulation analysis, Cosslett (1981) reports that a state-based sample of equal proportions, like the type used by the acquisition prediction literature, is usually a close-to-optimum design.

While the use of a state-based sample is justified on the grounds of efficiency, realization of this efficiency gain is predicated on the use of a suitable estimation procedure that recognizes the nature of the state-based sampling process. Unfortunately, this is where the earlier studies fail; they employ estimators which assume random sampling. As Manski and Lerman (1977) show, this leads to inconsistent and asymptotically biased estimates of the model parameters and hence biased estimates of the acquisition probability. Cosslett (1981) reports possible biases of 30% or more, indicating the seriousness of this problem.

To see the nature of the bias, consider a firm *i in the population* with a probability p of being a target. Let p' be the probability that the firm *i in the sample* is a target. Using Bayes' formula for conditional probability,

$$p' = \text{probability } (i \text{ is a target} | i \text{ is sampled})$$

$$= \frac{\begin{array}{c}\text{probability } (i \text{ is a target})\\ \times \text{probability } (i \text{ is sampled} | i \text{ is a target})\end{array}}{\begin{array}{c}[\text{probability } (i \text{ is a target})\\ \times \text{probability } (i \text{ is sampled} | i \text{ is a target})\\ + \text{probability } (i \text{ is a non-target})\\ \times \text{probability } (i \text{ is sampled} | i \text{ is a non-target})]\end{array}}.$$

In the case of random sampling, the probability of firm i being sampled is the same whether it is a target or not. Hence the above expression simplifies to p. However, under state-based sampling, this is not so. If N_1 and N_2 are the number of targets and non-targets in the population and n_1 and n_2 are the corresponding numbers in the sample, then

$$p' = \frac{p(n_1/N_1)}{p(n_1/N_1) + (1-p)(n_2/N_2)} \neq p.$$

The practice in the acquisition prediction literature is to use the simple maximum likelihood (MLE) procedure to estimate the model parameters and hence the state probabilities. Note that the maximum likelihood procedure consists of maximizing the sample likelihood function. When a state-based sample is used, the sample likelihood is formed using p'. The maximization of the sample likelihood, thus, yields an unbiased estimate of p'. However, since p' does not equal p, the procedure does not yield an unbiased estimate of p, the population acquisition probability. Hence, the simple maximum likelihood estimation procedure, when used on state-based samples, does *not* lead to an unbiased estimate of the population acquisition probability. The resulting bias can be calculated as follows:

$$p' - p = \frac{(n_1/N_1 - n_2/N_2)\ p(1-p)}{(n_1/N_1)\ p + (n_2/N_2)(1-p)}.$$

Since usually N_1 is much smaller than N_2 and n_1 is equal to n_2,

$$(p' - p) > 0,$$

except for the uninteresting cases of p being equal to 0 or 1. In other words, the estimated acquisition probability always overstates the true value.

The magnitude of the bias varies across samples as a function of the sample design. The bias is directly proportional to the difference in the sampling ratios of the targets and non-targets. For a given sample design, the bias varies across firms as a function of the true acquisition probability.[6]

The bias in the estimated probabilities does not alter the relative ranking of firms in terms of their acquisition probabilities. It is simple to show that if the true acquisition probability of firm A is greater than that of firm B, the estimated probability of A would also be greater than that of B. Hence, if the purpose of the estimated model is only to *rank* probabilities, the above bias is unimportant. However, if the estimated parameters are to be used to test hypotheses, the bias and inconsistency become important. If the objective is to use the model to predict targets, which is the case with the earlier acquisition prediction studies, the bias leads to erroneous inferences.

When the biased estimates of the acquisition probabilities are used to predict targets and non-targets, the observed prediction accuracies do not reflect the true predictive ability of the model. It is simple to show that the observed error rates understate the model's true error rate in predicting targets and overstate the true error rate in predicting non-targets. To see this, consider the classification scheme where a firm is classified as a target if the acquisition

[6] To see the seriousness of the bias, let us consider the estimation sample used by Stevens (1973) which consists of 40 targets and 40 non-targets. Let us assume for the purpose of illustration that the total population consists of 1000 firms. Since Stevens samples all the targets, the sampling ratio for the non-targets is 40/960, or approximately 0.042. Now, let us consider a firm whose true acquisition probability is 0.1. Given the above sampling scheme, the acquisition probability of the firm estimated by the model will be approximately 0.73.

probability is greater than a pre-specified value $\bar{p}$. If a firm has a true acquisition probability p (or an unbiased estimate of the true probability), this classification would lead to one of the following four outcomes: (1) correct prediction of a target if p is greater than $\bar{p}$ and i is actually a target, (2) incorrect prediction of a non-target as a target (a Type II error) if p is greater than $\bar{p}$ and i is actually a non-target, (3) correct prediction of a non-target if p is less than $\bar{p}$ and i is actually a non-target, and (4) incorrect prediction of a target as a non-target (a Type I error) if p is less than $\bar{p}$ and i is actually a target.

Now, suppose p', a biased estimate from a state-based sample, is used instead of p. How would the outcome of the prediction change? Note that since p' is always at least as large as p, whenever p is greater than $\bar{p}$, so would p' be. Thus, the prediction does not change for cases (1) and (2) above. However, this is not true for cases (3) and (4) above. When p is smaller than $\bar{p}$, p' can either be smaller or larger than $\bar{p}$. If p' is also smaller than $\bar{p}$, the prediction would not change; if, on the other hand, p' is larger than $\bar{p}$, the prediction would shift from (3) to (2) or from (4) to (1). A shift from (3) to (2) would introduce an additional Type II error; a shift from (4) to (1) would eliminate a Type I error.

Using p' as the estimated acquisition probability, therefore, understates the error rate in predicting targets and overstates the error rate in predicting non-targets. The net effect on the overall error rate is uncertain and is determined by the cutoff probability employed and the 'true' acquisition probabilities of the firms in the sample.

The effect of state-based sampling on inferred prediction error rates has been examined above analytically. In a recent study, Zmijewski (1984) examines the same issue empirically. He estimates a single bankruptcy prediction model using several alternative samples with varying proportions of bankrupt and non-bankrupt firms. The results reported by him are consistent with the conclusions of the above analysis.

The biases pointed out above, which result from the use of state-based samples in model estimation, can be avoided. To do this, the estimators used have to be appropriately modified to recognize the nature of the sample. Manski and McFadden (1981) discuss two possibilities: the conditional maximum likelihood estimator (CMLE) and the weighted maximum likelihood estimator (WMLE).[7] Both these estimators are obtained from relatively simple modifications to the ordinary maximum likelihood estimators commonly em-

[7]Ohlson (1980) avoids the biases discussed in this section by using the entire population rather than a state-based sample to estimate his bankruptcy model. While this is a strategy which requires no modification to the ordinary likelihood estimator, there are two drawbacks associated with it: (1) significant increase in computational cost and (2) loss of a 'hold-out' sample to test the predictive usefulness of the model.

ployed.[8] In the empirical study in this paper, the CMLE approach is employed. The specific details of this procedure are discussed in section 3.

2.1.2. State-based sample as a prediction test sample

The previous subsection discusses the bias arising out of using a state-based sample in the estimation of a dichotomous state prediction model. State-based samples have also been employed to test the predictive ability of the model. This practice results in yet another source of error in assessing the true predictive ability of the model.[9]

In judging the forecasting usefulness of a model, the statistic of interest usually is the expected error rate when the model is used to forecast the firms in the population as targets and non-targets. Since a state-based sample is non-random by definition, the error rate inferences based on it are not directly generalizable to the population. The very unequal distribution of targets and non-targets in the population, which justifies the use of a state-based sample in model estimation, argues strongly against its use in prediction testing. Since only a small fraction of the firms are targets, predicting them is like searching for a needle in a haystack. The use of a contrived sample with a large proportion of targets tends to obscure this difficulty.

To understand the specific nature of the distortion, again consider a population of N_1 targets and N_2 non-targets. If an acquisition prediction model is used to classify a sample of n firms consisting of n_1 targets and n_2 non-targets, and if m_1 and m_2 are numbers of misclassified targets and non-targets, the sample forecast error rate is

$$e' = \frac{m_1 + m_2}{n_1 + n_2} = \frac{m_1 + m_2}{n}.$$

The expected prediction error rate in the population is

$$\begin{aligned} e &= \frac{N_1(m_1/n_1) + N_2(m_2/n_2)}{N_1 + N_2} \\ &= \frac{m_1(N_1/n_1) + m_2(N_2/n_2)}{n} \frac{n}{N_1 + N_2}. \end{aligned}$$

[8] It is necessary to modify the estimator only if one is interested in obtaining unbiased and consistent estimates of the model parameters. If the objective is merely to use the estimated probabilities for prediction, an alternative to modifying the estimator is to adjust the cutoff probability appropriately to take into account the bias introduced by the state-based estimation sample.

[9] Eisenbeis (1977) also points out this problem.

Since for a state-based sample, by design,

$$\frac{n_1}{N_1} \neq \frac{n_2}{N_2} \neq \frac{n}{N_1 + N_2},$$

e is generally not equal to e'. Yet, the practice in the acquisition prediction literature is to use e' as the estimate of e. The bias from this can be calculated as

$$e' - e = \frac{(n_1 N_2 - n_2 N_1)}{n(N_1 + N_2)}(m_1/n_1 - m_2/n_2).$$

The sign of the right-hand part of the equation is determined by $(m_1/n_1 - m_2/n_2)$ since $(n_1 N_2 - n_2 N_1)$ is always positive. (Note that n_1 and n_2 are usually equal whereas N_1 is much smaller than N_2.) Since m_1/n_1 is the sample error rate for targets and m_2/n_2 is the sample error rates for non-targets, the bias is positive or negative depending on their relative magnitudes. The size of the bias is proportional to the difference in the two types of sample error rates as well as the difference in the ratios of population and sample shares of targets and non-targets.[10]

As pointed out already, unlike the case of model estimation, there is no econometric justification in employing state-based samples in prediction tests. Since actual use of a model involves the entire population, it is desirable to make the prediction test sample resemble the population as closely as possible. Once the parameters are estimated, the computation cost of state probabilities for prediction tests is relatively low. Hence, it is not unrealistic to suggest that a large sample, or even the entire population of firms at a given time, be employed in prediction tests to avoid the bias pointed out above. This, in fact, is the method employed in the empirical study described later in this paper.

2.2. *Optimal cutoff probability*

If a researcher is only concerned with testing whether a set of variables bears a significant statistical relationship to the acquisition probability of a firm, testing the predictive accuracy of a model is not necessary. Instead, the focus

[10] To see how serious the bias is, let us once again consider the study by Stevens (1973) which uses a sample of 40 targets and 40 non-targets and reports a prediction error rate of 15% for targets, 45% for non-targets, and an overall prediction error rate of 30%. He considers the population of firms listed on the COMPUSTAT tape. Let us assume that this consists of 1000 firms. Since Stevens samples all the available targets, the numbers of targets and non-targets in the population would be 40 and 960, respectively. The expected population error rate, based on the reported sample error rates would be 44%. Hence, the expected prediction accuracy in the population would be 56% and not 70% as reported by Stevens.

of interest in that case would be the significance of the overall explanatory power of the estimated model and, presumably, the significance of the estimated coefficients of the variables. Prediction testing becomes important only if the objective is to develop a statistical model to predict potential targets.

The prediction tests typically involve classifying a group of firms into targets and non-targets based on the estimated acquisition probability. To classify a firm, the estimated acquisition probability of the firm is compared with a predefined cutoff probability. If the estimated probability is less than the cutoff probability, the firm is classified as non-target.

The appropriate cutoff probability to be employed in the prediction tests is determined by the decision context in which the model's predictions are to be used. To derive the 'optimal cutoff probability', it is necessary to specify the decision context of interest, an appropriate payoff function, and the prior state probabilities. The standard decision theory methodology can then be applied to derive the optimal classification scheme.

The earlier acquisition prediction studies develop statistical models with a view to predicting takeover targets, and hence all of them conduct tests to examine the predictive ability of the estimated models. In performing the prediction tests, however, they do not derive an optimal cutoff probability as outlined above. Instead, they employ an arbitrary cutoff probability, usually 0.5. Since the decision context in which the estimated model's predictive ability is judged is not explicitly stated, the results of prediction tests in these studies are difficult to interpret.[11,12]

In order to rectify this problem, the empirical study in this paper uses the optimal cutoff probability derived below in a well-defined decision context. It is assumed that the purpose of the estimated acquisition model is to provide predictions which are to be used in a stock market investment strategy. As discussed earlier, the stock market does not seem to do well in predicting targets far in advance of the actual announcement of takeovers. The question examined in the prediction tests is whether the estimated acquisition model provides a superior mechanism not available to the market for predicting targets. Stated differently, we examine the hypothesis that it is possible to earn

[11] This is because, for a given set of estimated probabilities, the results of a prediction test are determined by the cutoff probability used. If the cutoff is derived within a decision context, the observed prediction accuracies indicate the extent to which the model's predictions are useful in that decision context. Otherwise, it is not clear what the observed prediction accuracies indicate. Further, if the results are statistically significant with one cutoff probability but not with another, it is not possible to choose between the two sets of results.

[12] The use of arbitrary cutoff probabilities in prediction tests has been pointed out as a problem by others. Eisenbeis (1977) points out that the use of arbitrary cutoff probabilities is one of the serious 'pitfalls' in using discriminant analysis. In the bankruptcy prediction literature, Ohlson (1980) discusses the problem and attempts to deal with it by presenting prediction error rates for a number of cutoffs rather than for a single cutoff probability. Altman, Haldeman and Narayanan (1977) address the issue by selecting bank loan classification as the decision context and by estimating from historical data the prior probability and misclassification costs.

abnormal returns by investing in firms that are predicted by the model to be potential targets.

To determine the classification scheme that maximizes the expected payoff, let us consider a firm i in the test sample. Let

q = the market's assessment of the probability that the firm becomes a target,
S_1 = the stock price if the firm becomes a target, and
S_2 = the stock price if the firm does not become a target.

The variables q, S_1 and S_2 are assumed to be common knowledge. Assuming market efficiency with respect to this information, the current stock price S would be such that

$$S = qS_1 + (1-q)S_2. \tag{1}$$

Denoting C_1 $(=S_1 - S)$ as the payoff if the firm becomes a target and C_2 $(=S_2 - S)$ as the payoff if it does not, the price S in eq. (1) would ensure that the *expected* payoff, based on market probability q, is zero. That is,

$$qC_1 + (1-q)C_2 = 0. \tag{2}$$

Now, suppose we develop a statistical model which predicts a probability of acquisition p for firm i. We hypothesize that the model's prediction is new information unavailable to the market and seek to exploit this 'private information' to earn abnormal returns. We agree with the market's assessment that S_1 would be the stock price if firm i becomes a target, and S_2 if not. Since the model prediction is hypothesized to be unknown to the market, the current stock price S, and hence the state payoffs C_1 and C_2, remain unchanged. The expected payoff from investing in firm i, however, changes (for us) in light of the new information from the model, p.

Given the market prior q and the model prediction p, the posterior probability q' can be computed using Bayes' formula:

$$q' = \frac{qf_1(p \mid i = \text{target})}{qf_1(p \mid i = \text{target}) + (1-q)f_2(p \mid i = \text{non-target})}, \tag{3}$$

where $f_1(p|i = \text{target})$ is the conditional probability density of observing p if i is in fact a target, and $f_2(p|i = \text{non-target})$ is the conditional probability density of observing p if i is a non-target.

The expected payoff from investing in firm i, given the posterior probability q' and the state payoffs C_1 and C_2, is $[q'C_1 + (1-q')C_2]$. Hence, firm i is expected to have a positive payoff if

$$q'C_1 + (1-q')C_2 \geq 0. \tag{4}$$

Using eq. (3), (4) can be rewritten as

$$\frac{f_1(p|i=\text{target})}{f_2(p|i=\text{non-target})} \geq \frac{-(1-q)C_2}{qC_1}. \tag{5}$$

Any firm with a predicted acquisition probability p which satisfies condition (5) has an expected positive payoff.

Assuming that there is no budget constraint, the scheme that maximizes expected payoff is to classify all firms that satisfy condition (5) as potential targets and invest in them. Firms that fail to satisfy condition (5) are to be classified as non-targets.[13] The relation between q, C_1 and C_2 implied by eq. (2) allows us to rewrite condition (5) as

$$\frac{f_1(p|i=\text{target})}{f_2(p|i=\text{non-target})} \geq 1. \tag{6}$$

The above condition implies that the optimal classification scheme is to classify a firm as a target if the predicted acquisition probability is such that the marginal probability of observing p *if* the firm is actually a target is greater than the corresponding marginal probability *if* the firm is a non-target. The optimal cutoff probability is the value where the two conditional marginal densities are equal.[14]

To use condition (6) to determine the optimal cutoff probability in an empirical context, we need to know the conditional probability density functions $f_1(\cdot)$ and $f_2(\cdot)$. Recall that $f_1(\cdot)$ is the distribution of the acquisition probability among targets; $f_2(\cdot)$ is the corresponding distribution for non-targets. By plotting the distribution of the estimated probabilities for the targets and non-targets in the same sample that is used to estimate the model parameters, we can obtain empirical approximations of $f_1(\cdot)$ and $f_2(\cdot)$. The cutoff probability is the value where the two plots intersect. This method is illustrated in section 3.

The empirical tests described in the next section use the above classification scheme and examine whether it is possible to earn abnormal stock returns by investing in the firms predicted by the model to be predicted targets. The tests thus assess whether the acquisition model provides predictions which are superior to the stock market's prediction. It is important to interpret the prediction test results in this context.

[13]The optimality of the classification scheme based on condition (5) is a well-known result from decision theory literature [see, for example, Press (1972, p. 371)].

[14]Note that eq. (2) implies that the *expected* costs of Type I and Type II errors are equal. The same assumption underlies an alternative procedure to determine the cutoff probability under which the probability which minimizes the overall sample error rate is chosen as the optimal cutoff. Thus, our procedure is equivalent to this alternative procedure.

2.3. Summary

Three methodological flaws have led to erroneous estimates of prediction error rates of extant acquisition prediction models. These are related to (1) sampling for model estimation, (2) sampling for prediction tests, and (3) specification of the cutoff probability. The nature of the biases introduced by each and ways to avoid them have been discussed above.

3. The empirical study

This section describes the data and methodology used in this paper to estimate and test an acquisition prediction model.[15] The discussion includes the specification of an econometric model, the selection of a set of potentially interesting variables, the sample selection, and the estimation methods employed.

3.1. Acquisition likelihood model

The following probability model is employed in this study to specify the exact functional relationship between the firm characteristics and its acquisition likelihood in a given period. Let $p(i, t)$ be the probability that the firm i will be acquired in period t, $x(i, t)$ a vector of measured attributes of the firm, and β a vector of unknown parameters to be estimated. Then,

$$p(i, t) = 1/[1 + e^{-\beta x(i,t)}].$$

In other words, $p(i, t)$ is a logit probability function of the measured attributes of the firm.

The intuition behind the above model is as follows. Whether or not a firm is acquired in a particular time period depends on the number and type of acquisition bids it receives in that period. This, in turn, depends on the firm's own characteristics as well as the motives and attributes of the bidders. In the above model, the relevant characteristics of the target which can be quantitatively measured are denoted by $x(i, t)$ and enter the model explicitly. The qualitative characteristics of the target which influence its attractiveness and the characteristics of the target–bidder combination are modeled as stochastic random variables. It is the probability distributions of these random variables, which are endogenous to the acquisition process, that determine the specific functional form of $p(i, t)$.

Under certain economic assumptions which include, among others, that there are a large number of active bidders in the market and that the shareholders of

[15] The acquisition model estimated in the study attempts to predict firms which are targets of successful takeover bids. This is consistent with the models developed by the earlier acquisition prediction studies.

a target accept the most profitable bid among those which offer a premium over the current market value of their stock, it is possible to show that the random variables in the above model follow the Type I extreme value distribution.[16] This, in turn, implies that $p(i, t)$ is a logit probability function of $x(i, t)$.[17]

3.2. Variables

The variables to be included in the acquisition likelihood model are specified on the basis of six hypotheses, frequently suggested in the academic and/or popular financial literature, on the types of firms that are likely to become acquisition targets.[18] The hypotheses and the variables implied by them are discussed below.

(1) *Inefficient management hypothesis*: Firms with inefficient managements are likely targets.

This hypothesis is based on the finance theory premise that acquisitions are a mechanism by which managers of a firm who fail to maximize its market value are replaced. The excess return on a firm's stock, averaged over an extended period of time is used as a proxy for management efficiency in this study. The excess stock return on a firm is calculated using a market model and daily stock return data and is averaged over a period of four years. As an alternative to the excess return measure, accounting profitability is also used as a proxy for management performance. The profitability is computed as the return on stockholders' equity averaged over a period of four years.[19] The

[16]The cumulative density function of the standard Type I extreme value distribution has the form: $1 - \exp(-\exp(x))$. The probability density function of a standard Type I extreme value distribution is very close to that of a log-normal distribution. The difference of two independent random variables, each with the same Type I distribution, has a logit distribution. For further discussion on this, see Johnson and Kotz (1970).

[17]The usual practice followed by the other acquisition prediction studies is to specify an *a priori* statistical model exogenously. Here, an attempt is made to take the economics of the acquisition process into consideration to arrive at an appropriate statistical model. A more formal development of this model is presented in Palepu (1982).

[18]Most of the earlier studies have not chosen the variables to be included in their model on the basis of a set of pre-specified hypotheses. Instead, the practice has been to start with a large number of popular financial ratios and then let the empirical analysis determine a subset of variables to be retained on the basis of their statistical significance in a step-wise procedure. For example, Simkowitz and Monroe (1979) start with a set of 24 ratios and finally retain 7 variables. This method of variable selection is arbitrary and leads to the statistical 'overfitting' of the model to the sample at hand. We attempt to avoid this problem by choosing variables based on a set of hypotheses from the literature.

[19]Accounting profitability measures only current performance. The excess return measure reflects, in addition to the current performance, the market's expectation of future performance. Hence, the excess return measure is probably a better proxy.

computation method for these variables, as well as for all the others, is described in the appendix.

(2) *Growth–resource mismatch hypothesis*: Firms with a mismatch between their growth and the financial resources at their disposal are likely targets.

This hypothesis implies that two types of firms are likely targets: low-growth, resource-rich firms and high-growth, resource-poor firms. The notion that low-growth, resource-rich firms are natural acquisition targets is commonly put forward in the popular financial press as well as in corporate finance textbooks. The hypothesis that high-growth, resource-poor firms may be attractive targets is suggested by the recent finance literature that analyzes the investing and financing decisions of firms under asymmetric information. [For example, see Myers and Majluf (1984).]

The growth–resource imbalance hypothesis indicates that growth and resource availability are important variables in determining a firm's acquisition likelihood. In this study, growth is measured as the average sales growth of a firm. Liquidity, measured as the ratio of net liquid assets to total assets, and leverage, measured as the debt to equity ratio are used to proxy the financial resource availability. A dummy variable, denoted as the growth–resource dummy (*GRDUMMY*) is employed to indicate the growth–resource mismatch. The *GRDUMMY* is assigned a value one for the combinations low growth–high liquidity–low leverage or high growth–low liquidity–high leverage. The dummy is set to zero for all other combinations. Each of the variables is defined as 'high' if its value for a firm is greater than the population average, and it is defined as 'low' otherwise.

The growth–resource dummy is included in the acquisition model to test the imbalance hypothesis. In another version of the model, in addition to the dummy, the three variables growth, liquidity, and leverage are also included. This is done to see if one of the two imbalances discussed earlier is predominant in our sample. No specific sign is hypothesized for the three variables since *a priori* it is not known which imbalance is predominant.

(3) *Industry disturbance hypothesis*: Firms that are in an industry subjected to 'economic disturbances' are likely acquisition targets.

The above hypothesis is suggested by the 'economic disturbance theory' proposed by Gort (1969) to explain observed variations in merger rates both across industries and over time. Gort argues that mergers are caused by valuation differentials among market participants which are triggered by economic shocks like changes in technology, industry structure, and regulatory environment.

The economic disturbance theory suggests that acquisitions cluster by industry. A factor that signals the acquisition likelihood of a firm is, therefore, the recent history of acquisitions in its industry. In this study, the variable industry dummy (*IDUMMY*) is used for this purpose. The industry dummy is assigned a value one if at least one acquisition occurred in a firm's four-digit SIC industry during the year prior to the year of observation.

(4) *Size hypothesis*: The likelihood of acquisition decreases with the size of the firm.

The above hypothesis is based on the premise that there are several size-related 'transaction costs' associated with acquiring a firm. These include the cost associated with the absorption of the target into the acquirer's organizational framework as well as the costs associated with fighting a prolonged battle that a target may wage to defend itself. These costs are likely to increase with the target size and hence the number of potential bidders for a firm is likely to decrease with size. To test this hypothesis, the size of a firm, measured by its net book assets, is included as a variable in the model.

(5) *Market-to-book hypothesis*: Firms whose market values are low compared to their book values are likely acquisition targets.

The proponents of the hypothesis assume that firms with low market-to-book value ratios are 'cheap' buys. Since the book value of a firm need not reflect the replacement value of its assets, the economic validity of this assumption is suspect. However, since this explanation of takeovers appears frequently in the popular press, it may be interesting to test it empirically. Hence the market-to-book ratio is included as a variable in this study. The market-to-book ratio is defined as the market value of the common equity divided by its book value.

(6) *Price–earnings hypothesis*: Firms with low P/E ratios are likely acquisition targets.

This is another popular explanation of acquisitions whose economic logic is questionable. According to the proponents of this hypothesis, bidders with high P/E ratios seek to acquire low P/E firms to realize an 'instantaneous capital gain' because of the belief that the stock market values the earnings of the combination at the higher P/E ratio of the acquirer. Once again, since P/E ratio is considered to be an important determinant of a firm's acquisition attractiveness, it is included as a variable in this study.

The above discussion leads to the identification of nine potential determinants of a firm's acquisition probability. The six hypotheses and the variables they imply are summarized in table 1. The hypothesized sign of each variable

Table 1

Acquisition likelihood hypotheses and independent variables.

Hypothesis	Variable(s)	Expected sign[c]
1. Inefficient management hypothesis[a]	Average excess return (*AER*) Accounting return on equity (*ROE*)	– –
2. Growth–resources imbalance hypothesis[b]	Growth–resources dummy (*GRDUMMY*)	+
3. Industry disturbance hypothesis	Industry dummy (*IDUMMY*)	+
4. Firm size hypothesis	*SIZE*	–
5. Asset undervaluation hypothesis	Market-to-book value (*MTB*)	[illegible]
6. Price–earnings magic hypothesis	Price–earnings ratio (*P/E*)	–

[a]Accounting return on equity is used as an alternative to the average excess return, a stock price based measure, in some versions of the model.

[b]The variables growth, liquidity and leverage are also included in some versions of model. These three variables are used in constructing the *GRDUMMY*. No definite signs are hypothesized for these three variables.

[c]A positive sign hypothesizes that the variable increases the likelihood of acquisition and a negative sign implies the opposite.

shows whether the acquisition likelihood is expected to go up (+) or down (–) with that variable.

3.3. Sample

A sample of 163 firms that were acquired during the period 1971–1979 and a random sample of 256 firms that were not acquired as of 1979 are used for the estimation of the acquisition model. Both the targets and non-targets (1) belong to the manufacturing or mining sectors, (2) are listed on either the New York or the American Stock Exchange, and (3) have data in the COMPUSTAT and CRSP files.

A list of targets from the period 1971–1979 was prepared from two sources: (1) The Statistical Report on Mergers and Acquisitions of the Federal Trade Commission, 1979 (published in July 1981) and (2) the delistments from the stock exchanges due to mergers and acquisitions, obtained from the CRSP files and confirmed by the Wall Street Journal Index. A total of 277 targets were

Table 2

Composition of the estimation sample.[a]

Year acquired	Number of firms
Targets	
1971	9
1972	5
1973	8
1974	12
1975	14
1976	20
1977	29
1978	35
1979	31
Total targets	163
Non-targets	
Firms not acquired during 1971-1979	256
Total sample	419

[a]All the firms in the sample (1) belong to the manufacturing or mining industries, (2) are listed on either the New York or the American stock exchange, and (3) are on the COMPUSTAT and CRSP data files.

initially identified. Of these, 163 were included in the estimation sample after screening for data requirements.

The population of 2054 firms, which were not taken over as of 1979 and satisfied the criteria for inclusion in the sample as non-targets, was first arranged in alphabetical order. Every sixth firm was selected from this list to generate a random group of 343 non-targets. Of these, 256 firms met the data requirements and were included in the sample. The composition of the estimation sample is summarized in table 2.

To test the predictive ability of the estimated model, a separate group of firms is used. This includes all the targets from the year 1980 and all the non-targets, other than those used in the estimation sample, listed on the COMPUSTAT tape in 1980. After screening for the criteria for inclusion in the study listed earlier, and the data requirements, this group consists of 30 targets and 1087 non-targets. Notice that the targets form only about 2.6% of this group. This is a more realistic group to test the true predictive ability of the model than the type of hold-out samples used by the earlier studies.

3.4. Model estimation

The parameters of the model are estimated by a maximum likelihood procedure using the statistical package QUAIL [Berkman et al. (1979)]. As pointed out earlier, since the estimation sample is a state-based sample, the population acquisition probability, $p(i, t)$, cannot be used to compute the sample likelihood function. Instead, we have to use the conditional probability that a firm is a target given that it is included in the sample. This probability, denoted as $p'(i, t)$, can be computed in the following manner.

In choosing the estimation sample, all the available targets in the population are selected. However, out of the 1384 non-targets which met the selection and data requirements, only 256 (or 18.5%) are included in the sample. Hence, the probability that a firm in the population is in the sample is one if it is a target and only 0.185 if it is a non-target. Under this sampling, suppressing the arguments i and t for convenience, we have

$$p' = \frac{(1)(p)}{(1)(p) + (0.185)(1-p)}.$$

Since

$$p = 1/(1 + \mathrm{e}^{-\beta x}),$$

we have

$$p' = 1/(1 + 0.185\mathrm{e}^{-\beta x}) = 1/(1 + \mathrm{e}^{\ln(0.185) - \beta x}).$$

Notice that the functional form of p' is also logistic. This is a convenient feature of the logistic probability model. The likelihood function to be maximized in the estimation uses the above expression for p'. Subsequent to the estimation, the parameters that determine the population probability p can easily be recovered since all the parameters other than the constant term are unaffected and the constant terms in the two models differs by a know value, $\ln(0.185)$ or -1.68.

In estimating the model, the dependent variable is assigned a value one for the targets and zero for the non-targets. Two versions of the logit model are estimated. In one version, the raw values of the independent variables are used. In the second version, each of the independent variables of an observation drawn from a given year is rescaled by its population average in that year, the population being defined as all the COMPUSTAT firms that met the criteria for inclusion in this study. Since the sample contains observations drawn from several different years, such a rescaling is likely to make these observations more homogeneous by eliminating the mean shift in the population characteristics that may have occurred from year to year during the period 1971–1979. The results of the two versions of the model are similar. The results presented in the next section correspond to the rescaled data.

4. Results

4.1. Logit model estimates

Four different versions of the logit model are presented in table 3. Models 1 and 3 include six independent variables each in addition to a constant term. These six variables correspond to the six acquisition likelihood hypotheses discussed in section 3. Model 1 uses average excess return, a stock-price-based performance measure, and model 3 uses accounting return on equity in its place. Models 2 and 4 are re-estimates of models 1 and 4, respectively, with the inclusion of three additional variables: growth, liquidity and leverage. Recall that the three variables are used in defining the variable growth–resource dummy. The reason for re-estimating the model with these variables in addition to the growth–resource dummy is to examine which specific type of growth resource mismatch is predominant in the present sample.

The parameter estimates of the logit acquisition models and the associated *t*-statistics are presented in table 3. Also presented in the table are the likelihood ratio index for each version of the model which provides an indication of the overall explanatory power of the model, and the likelihood ratio statistic that tests its statistical significance.

In model 1, the variables average excess return, growth–resource dummy and size are statistically significant and have expected signs indicating that inefficiency, growth–resources imbalance and smaller size are likely to increase a firm's probability of becoming a target. The variables market-to-book ratio and price–earnings ratio have statistically insignificant coefficients indicating that they are not important determinants of acquisition likelihood. The coefficient of industry dummy is significant but has a negative sign which is contrary to the industry disturbance hypothesis. The negative sign of the industry dummy implies that if a firm is randomly chosen from an industry in which at least one acquisition occurred during the previous year, that firm is more likely to be a non-target than a target during the current year. This is inconsistent with the existence of industry 'acquisition waves' which last for more than one year.[20]

The above conclusions remain unaltered when the three variables growth, liquidity and leverage are also included in model 2. The coefficients of growth and leverage are negative and statistically significant, and the coefficient of liquidity is insignificant. This indicates that the targets in the sample are

[20] One possible interpretation of this result is that the acquisition waves triggered by the industry disturbances have a life of less than one year. Under this scenario, an industry effect may cause a group of firms in an industry to become desirable targets. Given an active acquisition market, all these potential targets are acquired by bidders in a short period of time. The following year, in the presence of the new equilibrium, there will be few likely targets in that industry. If the evidence is interpreted this way, it is consistent with the industry disturbance hypothesis with the modification that the industry effects are usually short-lived.

Table 3

Estimates of logit acquisition likelihood models.[a]

Variables[b]	Expected sign	Estimates[c,d]			
		Model 1	Model 2	Model 3	Model 4
Average excess return	−	−1.332 (−2.53)[g]	−1.338 (−2.50)[g]		
Return on equity	−			0.003 (0.086)	0.005 (0.11)
Growth–resource dummy	+	0.5467 (2.47)[g]	0.4432 (1.86)[h]	0.4616 (2.32)[g]	0.4024 (1.88)[h]
Growth			−0.0245 (−2.65)[g]		−0.0261 (−3.18)[g]
Liquidity			−0.005 (−0.49)		−0.008 (−0.85)
Leverage			−0.0035 (−2.07)[g]		−0.0034 (−2.17)[g]
Industry dummy	+	−0.7067 (−2.97)[g]	−0.6900 (−2.86)[g]	−0.5802 (−2.75)[g]	−0.5608 (−2.61)[g]
Size	−	−0.0005 (−2.61)[g]	−0.0005 (−2.62)[g]	−0.0004 (−2.52)[g]	−0.0004 (−2.63)[g]
Market-to-book ratio	−	−0.0044 (−0.17)	0.0117 (0.33)	−0.0051 (−0.2)	0.0126 (0.36)
Price–earnings ratio	−	0.0065 (0.78)	0.0099 (1.08)	0.0031 (0.51)	0.0041 (0.636)
Constant		−2.1048 (−2.49)[g]	−2.1096 (−2.45)[g]	−2.1533 (−3.35)[g]	−2.1898 (−3.47)[g]
Likelihood ratio index[e]		0.1010	0.1245	0.0695	0.0979
Likelihood ratio statistic[f]		58.65	72.32	47.78	67.29

[a] From a sample of 163 target firms that were acquired during the period 1971–1979 and 256 non-targets that were not taken over as of 1979. All the firms (1) belong to mining manufacturing industries, (2) were listed on the New York or the American stock exchange, and (3) have data on COMPUSTAT and CRSP tapes. For more details on the sample, see section 3.3.

[b] The independent variables are measured as of the end of the fiscal year prior to the year of takeover for targets and as of the end of the fiscal year prior to 1979 for non-targets. For a complete description of how these are computed, see the appendix.

[c] Four different versions of the model are estimated. Model 1 consists of six variables corresponding to the six hypotheses in table 1. Model 2 is a re-estimation of model 1 with the three additional variables growth, liquidity and leverage. Models 3 and 4 are re-estimations of models 1 and 2, respectively, with accounting return on equity replacing average excess return, a market performance measure. The constant term in all the four models is corrected for the sampling bias.

[d] The t-statistic, computed to test the null hypothesis that the estimated coefficient is equal to zero, is shown in parentheses for each coefficient estimate.

[e] The log likelihood ratio index is defined as (1 − log likelihood at convergence/log likelihood at zero). It is similar to the R^2 statistic in the case of a multiple regression model and provides an indication of the logit model's explanatory power.

[f] The likelihood ratio statistic is computed to test the hypothesis that all the parameters in the model are simultaneously equal to zero. Under this null hypothesis, the statistic has an asymptotic distribution which is a chi-square with the degrees of freedom equalling the number of parameters in the model. The statistic is significant at the 0.01 level for all the models.

[g] Significant at the 0.05 level, two-tailed test.

[h] Significant at the 0.10 level, two-tailed test.

characterized by low growth and low leverage; there is no significant difference between the targets and non-targets in terms of liquidity. In models 3 and 4, the coefficient of accounting return on equity, which is used as an accounting proxy for management efficiency, is insignificant; the coefficients of all the other variables are consistent with those in models 1 and 2, respectively.

The likelihood ratio index for the four models ranges between 6.95% and 12.45%. The associated likelihood ratio statistic, which is asymptotically chi-square distributed, is statistically significant for all four models. This implies that the models provide a statistically significant explanation of a firm's acquisition probability. However, the magnitude of this explanation is quite small since a maximum of only 12.45% of the variation in a firm's acquisition probability is explained by the models.

Of the four estimated versions of the acquisition model, model 2 has the largest explanatory power as measured by its likelihood ratio index of 12.45%. Further analysis of the predictive ability of the model, therefore, employs model 2. The coefficient estimates of model 2 from table 3 are used to compute estimated acquisition probabilities for the 163 targets and the 256 non-targets in the sample. The sample median probability for the targets is 0.144 and that for non-targets is 0.087. The values of the estimated probabilities are generally small, even for firms that subsequently become targets. Thus, while the model is statistically significant, it provides only a weak signal as to whether or not a firm will become a target in the future.[21]

4.2. Prediction tests

4.2.1. Estimation of cutoff probability

To test the predictive usefulness of the estimated model, the optimal cutoff probability to be used has to be estimated. As pointed out in section 2.2, the optimal cutoff probability is determined by the distributions of acquisition probability for targets and nontargets. Empirical approximations of these distributions are obtained below using the computed estimated acquisition probabilities for the 163 targets and 256 non-targets in the estimation sample.

The maximum estimated probability value in the estimation sample is 0.46, and all the probabilities except one are found to fall within the range 0 to 0.40. To obtain the sample distributions of the acquisition probability, the range 0 to 0.4 is divided into ten equal intervals. The number of (and the percentage of the total) targets that fall within each of these intervals is tabulated and shown

[21] If the bias arising from the state-based sample is ignored, the median takeover probability would have been 0.478 for the targets and 0.325 for the non-targets. Notice that the biased probabilities would lead us to conclude that the model provides a fairly strong signal regarding the takeover probability.

Table 4

Distribution of estimated acquisition probability for targets and non-targets in estimation sample.[a]

Estimated acquisition probability		Target firms		Non-target firms		
Range	Mid-value (p)	Number	Percent $f_1(p)$	Number	Percent $f_2(p)$	$f_1(p)/f_2(p)$
0.000–0.039	0.02	6	3.7%	37	14.4%	0.26
0.040–0.079	0.06	23	14.1	74	28.9	0.49
0.080–0.119	0.10	35	21.5	60	23.5	0.92
0.120–0.159	0.14	36	22.1	48	18.8	1.18
0.160–0.199	0.18	23	14.1	19	7.4	1.19
0.200–0.239	0.22	22	13.5	12	4.6	2.94
0.240–0.279	0.26	8	4.9	4	1.6	3.06
0.280–0.319	0.30	5	3.1	0	0	
0.320–0.359	0.34	2	1.2	1	0.4	3.0
0.360–0.399	0.38	2	1.2	1	0.4	3.0
> 0.4	—	1	0.6	0	0	
Total		163	100	256	100	

[a]The acquisition probabilities are computed for the 163 targets and 256 non-targets in the estimation sample using the coefficient estimates of model 2 in table 3. The maximum estimated probability value is 0.46 and all but one probability values are in the range 0 to 0.40. The range 0 to 0.4 is divided into ten equal intervals. The number of firms that fall within each of these intervals are tabulated separately for the targets and non-targets. The figures in the column under $f_1(p)$ are calculated by dividing the number of targets in each probability interval by 163 and expressing the result as a percentage. Similarly, the figures under $f_2(p)$ are calculated by dividing the number of non-targets in each interval by 256 and expressing the result as a percentage. In fig. 1, $f_1(p)$ and $f_2(p)$ are plotted against p.

in table 4. To obtain discrete approximation of the distribution of the acquisition probability for the targets, the percentage of targets in each probability interval, shown in table 4, is plotted against the mid-value of that interval. Similarly, the percentage of non-targets in each interval is plotted against its mid-value to get a discrete approximation of the density function for the acquisition probabilities of non-targets. These plots are shown in fig. 1.

The graphs show that while there is a considerable overlap between the two probability distributions, there is also a systematic difference. The two distributions intersect at a probability value 0.112. At probabilities below this, the distribution function for the non-targets is greater than that for the targets; at probabilities above this, the distribution function for the targets dominates. At 0.112, the values of the two distribution functions are equal.

It is shown in section 2.2 that the optimal cutoff probability, when the estimated probabilities are used for predicting targets with a view to investing in their stocks, is that value at which the distribution functions for the targets and non-targets are equal. Since this value is 0.112 in the present case, it is used as the cutoff probability in the prediction tests discussed next.

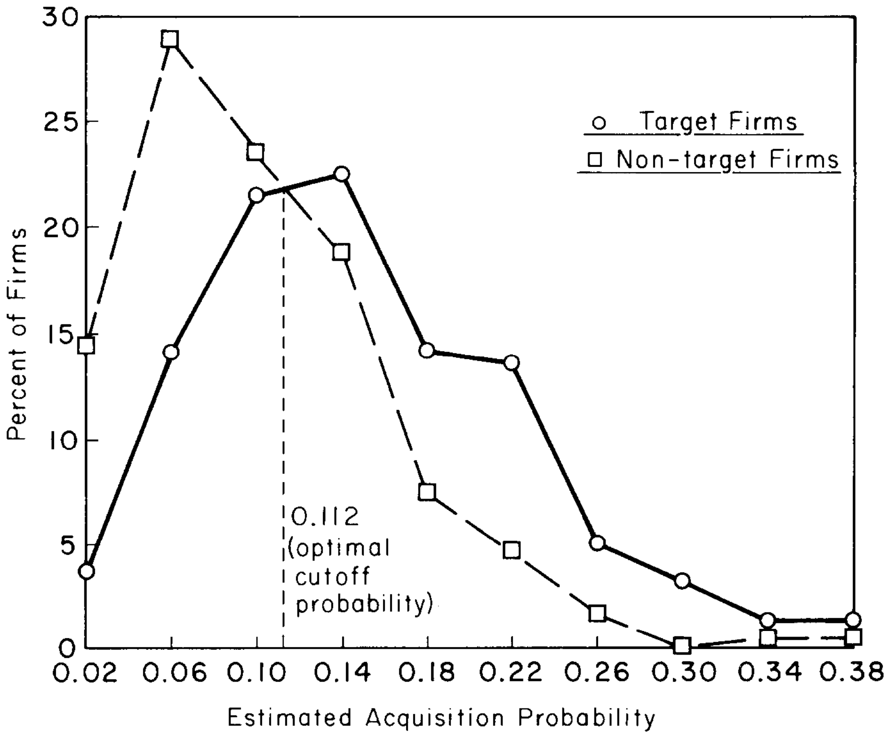

Fig. 1. Empirical probability density function of acquisition probability.

4.2.2. Predictions in a hold-out sample

To examine the ability of the model to predict targets in advance, the model has to be tested on a group of firms. Since the model parameters as well as the cutoff probability are obtained from the estimation sample, any test based on this sample is likely to be biased. Hence, the test described below uses a separate group of firms. This includes 30 firms which were actually taken over during 1980 and 1087 non-targets from the same year. These firms represent all those that are listed on the 1980 COMPUSTAT tape, meet the criteria for inclusion in this study, and have the required data. None of these firms were used in the estimation of the model parameters.

The values of the independent variables in the acquisition model are computed for each of the above 1117 firms following the methodology in the appendix.[22] The estimated parameters of model 2 from table 3 are then used to compute for each firm the probability that it will be a target in 1980. Using the

[22] The values of the independent variables are rescaled by their respective population averages to conform with the procedure used in the estimation sample.

estimated optimal cutoff acquisition probability, each firm is predicted to be a target if its acquisition probability is equal to or more than 0.112. Firms with an estimated acquisition probability less than 0.112 are predicted to be non-targets.

The above exercise results in classifying 625 firms to be targets and 492 to be non-targets. Of the 625 firms predicted to be targets, 24 are in fact targets in 1980. Of the 492 firms predicted to be non-targets, 486 are in fact non-targets. Stated differently, out of the 30 firms in the group that are actually targets, 24 (or 80%) are predicted by the model. However, in achieving this, the model misclassifies a large number of non-targets: of the 1087 non-targets, only 486 (45%) are correctly predicted.[23]

4.2.3. Excess returns from model predictions

The model predictions described above are used to examine the possibility of earning abnormal returns from investing in the stocks of predicted targets. The tests described in this section use daily excess returns drawn from the Center for Research in Security Prices (CRSP) excess return file. The CRSP excess returns are computed as

$$XR_{it} = R_{it} - \mathrm{E}(\tilde{R}_{it}),$$

where

XR_{it} = excess return on asset i for day t,
R_{it} = return on asset i for day t,
$\mathrm{E}(\tilde{R}_{it})$ = expected rate of return on asset i for day t.

$\mathrm{E}(\tilde{R}_{it})$ is estimated by grouping annually all securities listed on the New York and the American stock exchanges into ten equal control portfolios ranked according to their betas. The observed return to the control portfolio which has approximately the same beta as security i is then used as the estimate of $\mathrm{E}(\tilde{R}_{it})$.

The acquisition model identifies 625 firms as potential targets during 1980. The excess returns for these firms are examined over a holding period of 250 trading days beginning on January 2, 1980, which is the first trading day of the year. The excess returns of these firms on a given day are unlikely to be

[23] The overall accuracy rate, assuming for a moment that the two types of errors are additive, is 510/1117 or 45.6%. Notice that this is considerably worse than the accuracy rates reported by the earlier acquisition prediction studies. This is largely attributable to the fact that the prediction test here uses a very large number of non-targets. To see this, let us assume that the above test is conducted on 30 targets and 30 randomly drawn non-targets. With the prediction accuracy rates of 80% and 45% for the targets and non-targets, the overall accuracy rate would be 62.5%, which is closer to the rates reported by others.

independent. To take this contemporaneous cross-sectional correlation into account, the stocks of all the firms are formed into an equally weighted portfolio, and the tests are performed on the portfolio excess returns.

The average excess return for the portfolio on each relative day t is calculated as

$$AXR_t = \frac{1}{N} \sum_{i=1}^{N} XR_{it},$$

where N is the number of securities with excess returns during day t.[24] Average cumulative excess returns, CER, is the sum of the average excess returns over time,

$$CER = \sum_{t=1}^{k} AXR_t,$$

where CER is for days 1 to k.

To test the statistical significance of the excess returns, the daily portfolio excess returns are standardized by their standard deviation. The portfolio standard deviation is computed using the portfolio daily excess returns for 250 trading days preceding January 2, 1980 (relative days -250 to -1). Since the composition of the portfolio may vary over time, the portfolio standard deviation is recomputed for each day. The standardized portfolio daily excess return is computed as

$$SXR = \sum_{i=1}^{N} XR_{it} \Big/ SD_t,$$

where SD_t is the standard deviation of portfolio on day t.

A t-statistic which tests whether the portfolio cumulative excess return is significantly different from zero is calculated as

$$t = \sum_{i=1}^{k} SXR_t \Big/ \sqrt{k},$$

where k is the number of days over which the portfolio excess return is cumulated.

[24] Since some of the firms in the portfolio are taken over during the course of the holding period, the number of firms in the portfolio varies over time.

Table 5

Average daily cumulative excess returns over a holding period of 250 trading days beginning January 2, 1980 for 625 predicted target firms, 24 of which became targets during 1980, and for 492 predicted non-targets and 6 predicted non-targets which became targets during 1980.

	Predicted targets		Predicted non-targets	
Day	All firms (625) CER(%)	Actual targets (24) CER(%)	All firms (492) CER(%)	Actual targets (6) CER(%)
1	0.10	−0.61	−0.16	−0.59
10	0.63	−1.38	0.90	5.55
20	0.79	−3.01	0.82	4.57
30	0.79	−3.58	1.49	4.31
40	0.96	−1.68	1.50	4.61
50	0.50	−1.44	1.23	3.47
60	0.47	0.32	0.32	1.18
70	−0.79	−2.12	0.02	7.41
80	−1.62	−1.40	−0.56	11.03
90	−2.17	1.71	−1.39	16.80
100	−1.81	6.17	−1.22	14.35
110	−2.27	6.82	−1.08	15.15
120	−2.94	6.56	−1.74	13.79
130	−3.27	6.51	−1.53	14.63
140	−2.88	7.43	−1.40	14.98
150	−1.74	11.55	−1.44	16.23
160	−0.54	16.99	−1.18	14.27
170	0.16	17.03	−1.65	17.37
180	0.81	17.85	−1.64	22.30
190	0.51	18.77	−1.37	33.31
200	0.17	19.26	−1.40	31.98
210	0.05	19.52	−1.35	31.59
220	0.05	20.20	−1.12	31.64
230	−0.85	20.79	−0.26	32.96
240	−0.67	20.93	−1.53	35.16
250	−1.62[a]	20.98[b]	−1.51[a]	36.24[b]

[a] Not significant at the 5 percent level.
[b] Significant at the 5 percent level.

Table 5 reports the average daily cumulative excess return, CER, starting with the first trading day of 1980 in intervals of 10 trading days. In addition to the group of 625 firms which are predicted by the model to be potential targets, CERs are reported for three other portfolios: the 24 predicted targets which in fact became targets, the 492 firms which are predicted to be non-targets, and the 6 predicted non-targets which in fact became targets.

The results are not consistent with the hypothesis that it is possible to earn significant positive excess returns by investing in firms identified as potential targets by the model. The average cumulative excess return for the 625 predicted targets over the 250 days is -1.62% ($t=-0.77$). The average CER for the predicted targets is small and hovers around zero throughout the holding period. In contrast, the average *CER* for the subgroup of 24 predicted targets which became targets during the holding period is large and impressive. The 250-day *CER* for this group is 20.98% ($t = 2.28$). Thus, the reason for the insignificant returns for the predicted targets as a whole lies in the fact that this group consists of 601 firms which are incorrectly predicted by the model as potential targets. The large number of firms erroneously classified as targets significantly reduces the overall economic usefulness of the model's predictions.

For comparison, the *CER*s for the predicted non-targets are also presented in table 5. The *CER*s for the predicted non-targets mirror those of predicted targets. The 250-day *CER* for the 492 predicted non-targets is -1.51% ($t = -0.79$), and that for the 6 predicted non-targets which became targets during this period is 36.24% ($t = 2.46$). This indicates that, on average, excess returns for predicted targets are not very different from those of the predicted non-targets.

4.2.4. Discussion

The above prediction test results show that investing in the potential targets identified by the model does not yield significant excess returns. This implies that the model's ability to predict takeover targets is not superior to that of the stock market.

As in the previous acquisition prediction studies, the data used in the estimation and prediction in this study is on average six months old relative to the takeover announcements. As argued earlier, the pre-takeover stock price movement of target firms indicates that the stock market does not identify takeover targets very accurately six months prior to their takeover. Hence, the above results imply that the model, just like the stock market, does not predict targets with a high degree of accuracy long before the takeover announcements. This conclusion differs from that reported by the earlier acquisition prediction studies.

There are some limitations to the generalizability of our conclusion. First, the set of independent variables included in our model is not an exhaustive set of all possible variables. The conclusions, which are based on the limited set of variables considered, cannot therefore be interpreted to imply that targets are unpredictable from all public data. Two factors tend to mitigate this limitation. The variables in this study are selected based on a set of frequently stated acquisition hypotheses. Thus, the results at the minimum indicate that these six hypotheses do not enable the prediction of targets with a high degree of accuracy. Also, while no attempt has been made to try all the variables considered by the other studies, the set of variables in this study includes most of the variables found to be important by others. Hence, the difference in the conclusions of this study and the others can not be totally attributed to the variables considered.

A second limitation is that the data used in the model estimation and prediction are on average six months old relative to the takeover announcements. This reduces the ability of the model to identify potential targets accurately. In a realistic application, one would want to use the most recent data, which are likely to improve the predictive ability of the model. However, this is unlikely to alter the predictive ability of the model relative to the stock market.

5. Summary and conclusions

A number of studies develop statistical models to predict takeover targets. They claim the ability to identify targets with high accuracy rates ranging from 60% to 90%, using information which is publicly available six to twelve months prior to the takeover announcements. The pre-takeover stock price movement of target firms reported by Dodd and Ruback (1977) and others, however, indicates that the stock market does not identify potential targets with a high degree of accuracy even three months prior to the takeover announcements. Thus, the results reported by the earlier acquisition prediction studies imply that their models have a superior ability than the stock market in identifying takeover targets. To probe this issue further, this paper undertakes a methodological and empirical analysis.

An examination of the methodology used by the earlier acquisition prediction studies shows that there are three principal methodological flaws which make the reported prediction accuracies unreliable. First, the use of non-random samples in the model estimation stage, without appropriate modifications to the estimators, leads to inconsistent and biased estimates of the acquisition probabilities. This results in overstating the model's ability to predict targets. Second, the use of non-random samples in prediction tests leads to error rate estimates that fail to represent the model's performance in the population. Third, the use of arbitrary cutoff probabilities in prediction tests makes the computed error rates difficult to interpret.

This paper adopts methodological modifications to avoid the above problems. The empirical study described in the paper estimates a binomial logit model with the independent variables selected on the basis of a set of six frequently stated hypotheses on the determinants of a firm's acquisition probability. To obtain unbiased and consistent estimates from a state-based sample of 163 targets and 256 non-targets, the conditional maximum likelihood estimator proposed by Manski and McFadden (1981) is employed. The prediction ability of the model is tested on a large group of firms which resembles the population in a realistic use of the model. The cutoff probability is derived to test the possibility of earning excess returns by investing in potential targets identified by the model.

While the estimated model is found to be statistically significant, its explanatory power is quite small. The magnitudes of the acquisition probabilities are in general very small. When the model is tested on a group of 1117 firms, 24 of the 30 (80%) actual targets and 486 of the 1087 (45%) actual non-targets are correctly classified. The strategy of investing in the 625 firms identified by the model to be potential targets is found to result in statistically insignificant excess returns. Hence, the estimated model's ability to predict targets is not superior to that of the stock market. Since the market does not seem to identify targets very accurately long before the takeover announcements, it is concluded that the model also does not predict targets accurately.

The methodological issues addressed in this paper are relevant to other areas of research which involve the use of dichotomous state models with the population proportions of the two states skewed. Methodological critiques by others, including Eisenbeis (1977) and Zmijewski (1984), point out some of the same problems in the context of other applications. While the problems are well-known in the methodological literature, the applied research seems to lag behind in addressing them. This paper demonstrates that ignoring these problems can lead to serious biases in inferences. It is also shown that the problems can be avoided by using relatively simple modifications to the current methodology.

Appendix A

Definitions and computations of variables

(1) *Average excess return* (*AER*): The excess return on a firm's stock is defined as the difference between the firm's actual return and the expected return from a two-parameter market model. The data are drawn from the CRSP daily stock return file. The parameters of the market model are computed for each firm using one year's data, those of the fifth year prior to the observation year. The excess returns are computed over a period of four years prior to the observation year. (For example, consider a firm which was a target in the year 1975. Data from 1970 are used to estimate the market model, and the excess returns are computed over the period 1971–1974.) *AER* is computed as the average excess return per day over this four-year period. The unit of measurement of *AER* is percent per day.

(2) *Return on equity* (*ROE*): Return on equity is defined as the ratio of net income before extraordinary items and discontinued operations to the common and preferred equity of a firm. COMPUSTAT data items 10, 11 and 18 are used for net income, common equity and preferred equity, respectively.

The ratio is computed and averaged over a period of four years prior to the year from which an observation is drawn.

(3) *Growth* (*GROWTH*): Growth of a firm is defined as the annual rate of change in the firm's net sales. COMPUSTAT data item 12 is used in the computations. The annual sales growth is computed and averaged over the three fiscal years prior to the observation year. (For example, consider a target firm from the year 1975 with a December 31 fiscal year. The sales data from the period January 1, 1972 to December 31, 1974 are used to compute the sales growth during the three fiscal years 1972, 1973 and 1974, and the average growth rate for these three years is used as the growth variable.) The unit of measurement of the growth variable is percent per year.

(4) *Liquidity* (*LIQUIDITY*): Liquidity is defined as the ratio of the net liquid assets of a firm to its total assets. The net liquid assets are defined as the cash plus the marketable securities less the current liabilities. COMPUSTAT data items 1 and 2 are used to compute the net liquid assets and data item 6 is used for the total assets. The liquidity ratio is computed for the three fiscal years prior to the observation year, and the average is used as the liquidity variable. The unit of measurement for the liquidity variable is percentage per year.

(5) *Leverage* (*LEVERAGE*): Leverage is defined as the ratio of the long-term debt of a firm to its equity. The equity is defined as the sum of the preferred and common equity. COMPUSTAT item 9 is used for the long-term debt, and the data items 10 and 11 are used to calculate the equity. The debt/equity ratio is computed for the three fiscal years prior to the observation year, and the average is used as the leverage variable. The unit of the leverage variable is percent per year.

(6) *Growth–resource dummy* (*GRDUMMY*): The growth–resource dummy is a 0/1 variable defined on the basis of the three variables growth, liquidity and leverage defined above. The dummy variable is assigned a value one if the firm has a combination of either low growth–high liquidity–low leverage or high growth–low liquidity–high leverage. The dummy is set to zero for all the other combinations. Each of the three variables growth, liquidity and leverage is defined as 'high' if its value for a firm is larger than the average for all the COMPUSTAT firms, otherwise, it is defined as 'low'.

(7) *Industry dummy* (*IDUMMY*): The industry dummy is a 0/1 variable. It is assigned a value one if at least one acquisition occurred in a firm's four-digit SIC industry during the year prior to the observation year; otherwise, it is given a value zero.

(8) *Firm size* (*SIZE*): The variable *SIZE* is defined as the total net book value of a firm's assets. COMPUSTAT data item 6 is used to measure the total book assets. The variable is measured as of the fiscal year end immediately prior to the observation year. The units are millions of dollars.

(9) *Market-to-book ratio* (*MTB*): *MTB* is defined as the ratio of the market value of the common equity of a firm to its book equity. COMPUSTAT data items 24, 25 and 60 are used in computing the ratio. Both the market value and the book value are measured at the end of the fiscal year preceding the observation year. The variable is expressed as a ratio.

(10) *Price–earnings ratio* (P/E): The price–earnings ratio is defined as the ratio of a firm's stock price per share to its earnings per share. COMPUSTAT data items 24 and 58 are employed in the computations. The P/E ratio is computed as of the fiscal year end preceding the observation year.

Note: The observation year is defined for a target firm as the year in which it was acquired; for the non-targets, it is defined as 1979, the year as of which they are observed to be not acquired.

References

Altman, E.I., R.G. Haldeman and P. Narayanan, 1977, ZETA analysis: A new model to identify bankruptcy risk of corporations, Journal of Banking and Finance 1, 29–54.

Asquith, P., 1983, Merger bids, uncertainty, and stockholder returns, Journal of Financial Economics 11, 51–83.

Belkoui, A., 1978, Financial ratios as predictors of Canadian takeovers, Journal of Business Finance and Accounting, Spring, 93–107.

Berkman, J., D. Brownstone and Associates 1979, Qualitative, intermittent, and limited dependent variable statistical program (Department of Economics, MIT, Cambridge, MA).

Castagna, A.D. and Z.P. Matolcsy, 1976, Financial ratios as predictors of company acquisitions, Journal of the Securities Institute of Australia, Dec., 6–10.

Cosslett, S.R., 1981, Efficient estimation of discrete-choice models, in: C.F. Manski and D. McFadden, eds., Structural analysis of discrete data with econometric applications (MIT Press, Cambridge, MA).

Dietrich, J. Kimball and E. Sorensen, 1984, An application of logit analysis to prediction of merger targets, Journal of Business Research 12, 393–402.

Dodd, P. and R. Ruback, 1977, Tender offers and stockholder returns: An empirical analysis, Journal of Financial Economics 5, 351–373.

Eisenbeis, R.A., 1977, Pitfalls in the application of discriminant analysis in business, finance, and economics, Journal of Finance 32, 875–900.

Gort, M., 1969, An economic disturbance theory of mergers, Quarterly Journal of Economics 83, 624–642.

Holthausen, R.W. and R.W. Leftwich, 1983, The economic consequences of accounting choice: Implications of costly contracting and monitoring, Journal of Accounting and Economics 5, 77–117.

Jensen, M.J. and R.S. Ruback, 1983, The market for corporate control: The scientific evidence, Journal of Financial Economics 11, 5–50.

Johnson, N.I. and S. Kotz, 1970, Distributions in statistics: Continuous univariate distributions – I (Houghton Mifflin, Boston, MA) 272–295.

Manski, C.F. and S.R. Lerman, 1977, The estimation of choice probabilities from choice based samples, Econometrica 45–8, 1977–1988.

Manski, C.F. and D. McFadden, 1981, Alternative estimators and sample designs for discrete choice analysis, in: C.F. Manski, and D. McFadden, eds., Structural analysis of discrete data with econometric applications (MIT Press, Cambridge, MA).

Ohlson, J., 1980, Financial ratios and the probabilistic prediction of bankruptcy, Journal of Accounting Research 18, 109–131.

Palepu, K.G., 1982, A probabilistic model of corporate acquisitions, Unpublished Ph.D. thesis (MIT, Cambridge, MA).

Press, S.J., 1972, Applied multivariate analysis (Holt, Rinehart and Winston, New York).

Simkowitz, M. and R.J. Monroe, 1971, A discriminant analysis function for conglomerate targets, Southern Journal of Business, Nov., 1–16.

Stevens, D.L., 1973, Financial characteristics of merged firms: A multivariate analysis, Journal of Financial and Quantitative Analysis 8, 149–165.

Zmijewski, M.E., 1984, Methodological issues related to the estimation of financial distress prediction models, Journal of Accounting Research 22, suppl., 59–82.