A CORRELATED DISTRIBUTION SIMULATION MODEL FOR CAPITAL BUDGETING

Charles L. Crabtree
Crown Zellerbach Corporation
One Bush Street
San Francisco, California 94119

Lyman C. Dennis II*
Management Optimization Systems, Inc.
510 Pere Marquette Building
New Orleans, Louisiana 70112

Abstract

A simulation model for the analysis of capital projects is presented. A mechanism included in the model for approximating dependencies between variables is described and the effects of the mechanism's operation are analyzed and documented.

1. INTRODUCTION

This paper describes the key features of a recently completed capital investment simulation model, jointly developed by Crown Zellerbach Corporation and Tulane University. The model (named FRANCE for Probability and Risk Analysis of Capital Expenditures), while operational only five months, already has been used in evaluating the risk associated with capital projects and potential acquisitions totaling some $75 million.

The FRANCE model parallels to a large degree the procedure outlined by David Hertz in the article "Risk Analysis in Capital Investments" (Harvard Business Review, Jan-Feb, 1964). As the reader may recall, this approach involves a Monte Carlo simulation technique operating upon empirical distributions for the key variables affecting the outcome of a capital investment: price, volume, operating cost, etc. FRANCE incorporates a number of refinements to the Hertz model, thus increasing

*The authors would like to thank those who aided in the preparation of this paper: Dr. Joseph L. Balintfy, Dr. James T. Murphy, Messrs. Roger D. Eck and Fred H. Dorner, all of the Graduate School of Business Administration of Tulane University; Dr. Erich A. Helfert and Mr. James A. Oswald of Crown Zellerbach Corporation. Thanks also are due to the Tulane Computer Laboratory for furnishing the many hours of computer time required for model building and testing.

the model's usefulness to management as a decision-making tool.

One frequent flaw in capital investment simulation models is the failure to account for the interdependency between key variables. This paper will describe an expedient approximation method for expressing dependency between the random variates of two or more distributions. The technique produces results considered significantly more realistic than those achievable if variable dependencies are ignored. In addition, model refinements will be shown which allow the input distribution modes to be altered yearly at the user's discretion to reflect anticipated market or contractual effects.

Provisions for asset acquisition, replacement and retirement, depreciation and probabilistic asset life are also described. Lastly, this paper will discuss techniques sucessfully employed in communicating to management the usefulness (and limitations) of the Monte Carlo simulation approach to capital investment decisions.

2. THE PROBLEM

In recent years, capital projects at Crown Zellerbach have been evaluated using conventional discounted cash flow techniques based upon single estimates of key variables. Sensitivity analysis normally was employed to determine the project yield and net present value under varying price, volume, and cost assumptions. It was recognized, however, that at times a more sophisticated method was needed to evaluate the risk associated with major projects; in short, a procedure was required which would enable management to make full use of available information regarding possible values of key variables. This was especially needed when analyzing expansion alternatives in the forest products area (particularly lumber products) where a wide range of price fluctuations frequently occurs.

Top management, aware of the potential of the simulation technique, requested that a procedure for evaluating capital projects be developed, which would meet the following requirements:

- It must provide a measure of the risk involved in the prospective capital project.
- It must allow for all cash flows to be on an after-tax basis, including the tax effect of both capital gains and capital losses on asset retirements.
- It should allow for the inclusion of additional capital investments in later years, something often required on major projects.

To these three requirements, the model developers added a fourth stipulation (which will be discussed in a later section):

- It must allow for an expression of the interdependency between certain input variables (operating cost, price, volume).

It was in response to this management request that development began in August, 1968, of the simulation model to be described. The actual nucleus of the model was a skeletal program based upon the Hertz article and programmed in FORTRAN by a former Tulane MBA student, Lawrence T. Restall. Though Restall's model was excellent for academic purposes, extensive modification and development were needed to make it suitable for corporate decision making. All features described in this paper were conceived and added to the basic framework by the authors.

3. EXISTING MODELS

As the reader is doubtlessly aware, the use of Monte Carlo simulation techniques to analyze the risk associated with capital investments is hardly without precedent. Nevertheless, a survey of literature published on this subject, as well as existing simulation models, resulted in the following observations:

- Provisions for calculating annual tax depreciation for assets having varying tax lives frequently are not incorporated into the models.
- Most models lack a means of expressing interdependency among variables (for example, between variable cost and price).

- Provisions for inclusion of future capital additions are often lacking.
- Options for discounting cash flows on end-of-year or on continuous bases normally are not available.
- A provision for working capital buildup throughout a project's life seldom is included.
- Many existing models lack the ability to analyze several competing alternatives and to display the results in a readily comparable format.

In order to satisfy the requirements posed by management, it was felt that PRANCE would have to overcome most of these limitations.

4. CHARACTERISTICS OF THE MODEL

PRANCE is written in FORTRAN because (1) the Restall model from which PRANCE was developed was in FORTRAN, (2) the computer time requirements for a GPSS model were believed to be substantially higher (SIMSCRIPT was not available with the Tulane computer configuration), (3) numerous model features would necessarily be in FORTRAN, thus the GPSS framework offered little attraction, and (4) all desired output was carefully user-designed making GPSS standard output of no benefit.

PRANCE allows the user to enter distribution data on up to 15 variables. (These are tabulated in Figure 1 and show the frequency with which each variable is selected per simulation iteration. Normally, 500 iterations are sufficient.) Empirical distributions for each variable are entered by specifying the possible values which the variable can assume, along with the corresponding probability of occurrence. For example, the range of values for market price (with probabilities) could be as follows:

Value ($ per 1,000 units)	$0.80	$1.00	$1.20	$1.40	$1.60
Probability	5%	10%	25%	50%	10%

Figure 1
VARIABLES HAVING EMPIRICAL DISTRIBUTIONS

	Variable	One Value Selected for Each *
(1)	Investment in land	Project
(2)	Investment in building	Project
(3)	Investment in original equipment	Project
(4)	Additional equipment investment	Equipment replacement
(5)	Years to next rebuild	Project, addition, replacement, rebuild
(6)	Rebuild expense	Rebuild
(7)	Useful equipment life	Project, equipment addition, rebuild or replacement
(8)	Project life	Project
(9)	Sales volume	Year of project life
(10)	Fixed cost	Year of project life
(11)	Variable cost per unit	Year of project life
(12)	Market price	Year of project life
(13)	Land salvage value	End of project
(14)	Building salvage value	End of project
(15)	Equipment salvage value	Equipment replacement, end of project

* This column shows the frequency with which the variable is selected per simulation iteration. (A simulation iteration represents one complete project analysis.) For example, investment in building is selected once per project (or simulation iteration), while the sales volume distribution is sampled during each year of the project life.

One significant feature of the model is the fact that these empirical distributions are entered, for each variable, only once. Unless instructed otherwise, the model assumes that these input distributions apply to Period 1 and, that the shapes of the input distributions apply to successive years of a project's life, with distribution values revised by user-entered distribution modes for each year of project life. Modes are employed rather than means because it was reasoned that the users of the model when assessing the value for a particular variable, think in terms of its "most likely value" (the mode) rather than the mean. FRANCE then holds the shape of the distribution constant and, for each succeeding year after the first, generates distributions having the same shape as the original, but with the mode supplied by the user. These modes are termed level variables.

Another feature of the model is that all empirical distributions are "smoothed." This technique was incorporated in order to transform each histogram-like input empirical distribution into an essentially continuous distribution with the same mean and the same general shape. Figure 2 below depicts the smoothing of the previously listed market price distribution. This smoothing of empirical distributions is important in conjunction with the technique for handling correlated distributions.

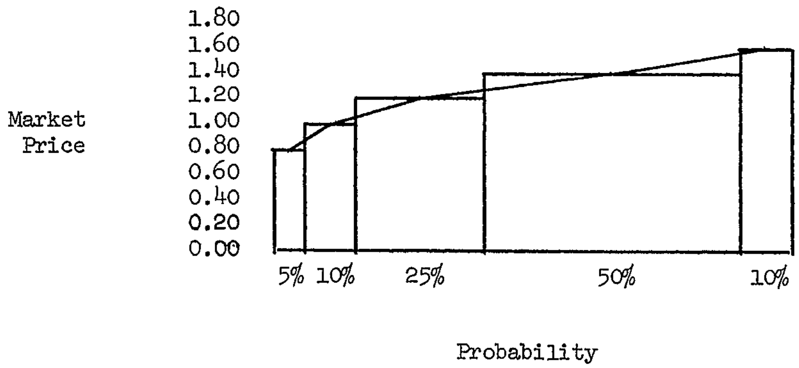
Additional variables utilized by the model but not having empirical distributions are shown in Figure 3. Most of these are self-explanatory with the exception of the disposal recovery factor. The

Figure 3
VARIABLES NOT HAVING EMPIRICAL DISTRIBUTIONS

(1) Tax rate applicable to normal income
(2) Tax rate applicable to capital gains
(3) Depreciation method (double-declining-balance, sum-of-the-years'-digits, or straight line)
(4) Minimum return standard (cost of capital or opportunity rate)
(5) Working capital (expressed yearly by amount or as a percentage of sales)
(6) Disposal recovery factor for land
(7) Disposal recovery factor for buildings
(8) Disposal recovery factor for equipment

disposal recovery factor is a major element of the method used by the model to express asset terminal values and represents the fraction of tax book value (tax basis) estimated to be recovered at project termination. For example, a machine having a tax basis of $5,000 at project termination and a disposal recovery factor of 0.8 would result in a recovery of $4,000 ($5,000 x 0.8) being realized. In this manner, terminal values can realistically change in inverse proportion to project life (with the exception of land which normally appreciates).

Figure 2
EMPIRICAL DISTRIBUTION SMOOTHING



The terminal value arrived at by using a disposal recovery factor differs from salvage value. Salvage value presently is input as an empirical distribution and is used solely to determine the depreciation basis (original cost minus salvage).

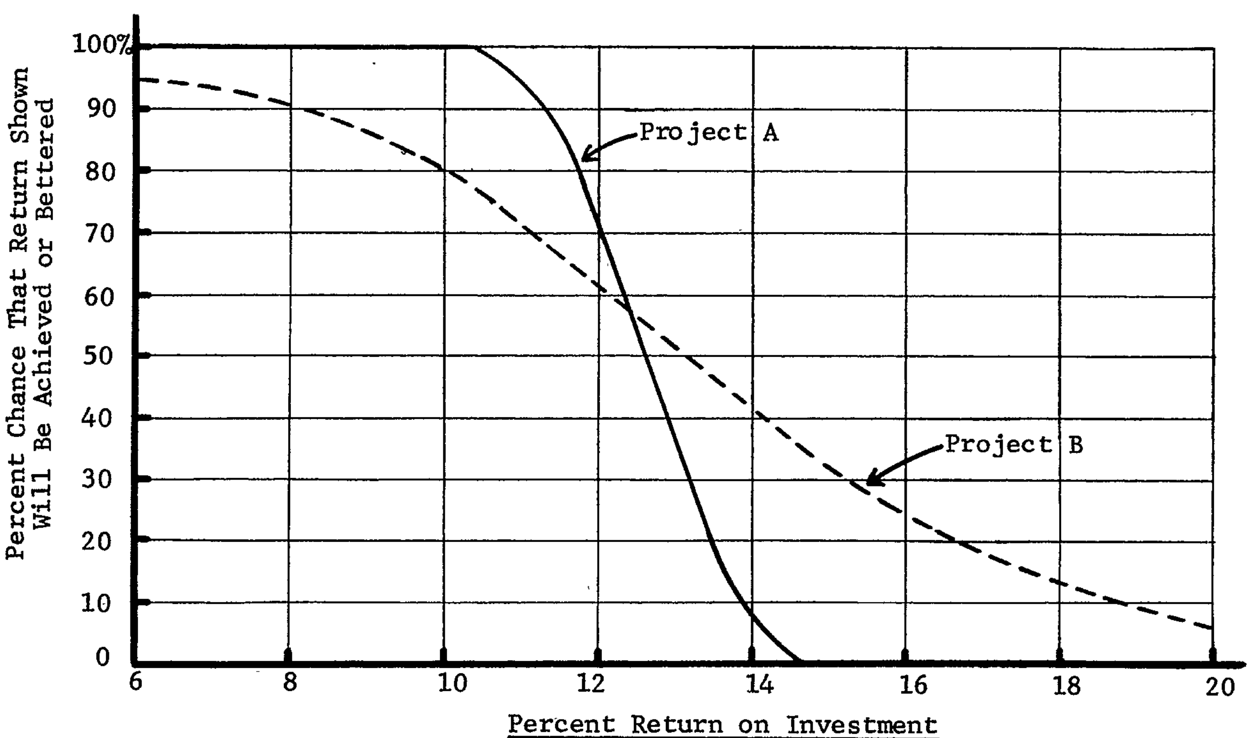
The tax effect of a capital gain or capital loss upon asset retirement at project termination (as discussed above) is taken into account by the model. In addition, extensive programming has been incorporated to allow for equipment additions, rebuilds and replacements throughout the project's life. This procedure works as follows:

The PRANCE input variables affecting equipment additions, rebuilds and replacements are "useful equipment life," "additional equipment investment," "years to next rebuild," and "rebuild expense." As the project is simulated, PRANCE adds new equipment, rebuilds existing equipment (the rebuild expense is capitalized), replaces equipment that has reached the end of its useful life, and, of course, retires the equipment at the end of the project life.

Frequently, a major capital project can generate negative cash flows during the early periods of operation before full capacity is achieved. PRANCE accounts for these negative flows in either of two ways; the resulting tax credits can be carried forward up to five years, or they can be taken as positive cash flows in the year that they occur. This latter condition assumes that the tax credits from losses benefit other corporate projects, which, for large organizations, is generally true.

A conventional analysis would show only a single

Figure 4
SUMMARY OF 500 SIMULATIONS



discounted cash flow rate of return from a major project or acquisition. An analysis via the PRANCE model, however, portrays for management the entire range of returns which the project can generate, and, most important, the probability that various returns will be equaled or bettered. As an example, Figure 4 (preceding page) illustrates a cumulative distribution similar to one generated and graphed by PRANCE on two competing alternatives (Projects A and B).

5. THE PROBLEM OF CORRELATED VARIABLES

An appreciable number of capital investment simulation models now in use or for sale by major U.S. firms contain a significant logical flaw. While the models profess to measure the variability inherent in the resulting cash flows of projects analyzed, failure to consider correlations between the variates of the distributions, in fact, contributes substantially to the variability of the simulation analysis.

For example, consider the following hypothetical case in which price and variable cost are correlated, but are not in any way linked in the simulation model. The manager supplying variable costs and prices has presented the distributions depicted in Figure 5-a. Each value is equiprobable. The manager has reasonably assumed that each value of price has some relation to the values for variable cost. For example, he would probably rate as extremely unlikely the possibility that a price of 32 would arise when the variable cost was 50. But, if such information regarding price-variable cost correlations is not embedded in the simulation model, this combination is as likely as any other.

Figure 5-b shows the margins (price less variable cost) which the manager might have thought likely for five years of the project's life, and to the right are shown the margins which might have arisen in a simulation. (The lines in Figure 5-a show how simulation margins were selected. The manager's margins are represented by the differences between prices and variable costs appearing in the same row.)

Note that although the two sets of figures have the same means, the variability of the simulated set is more than 15 times that of the manager's set of data. Certainly, the manager's expected margins will not always be realized, but they would in most cases be much more likely than the

Figure 5-a
EXAMPLE INPUT DISTRIBUTIONS

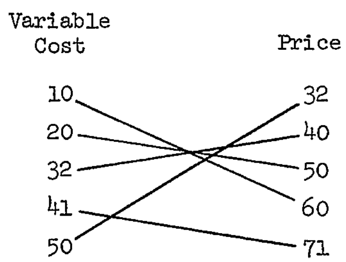


Figure 5-b
MARGINS
(price less variable cost)

Manager Expected	Simulation Picked
22	50
20	30
18	8
19	30
21	-18
Avg. 20	Avg. 20
S.D. 1.41	S.D. 23.19

simulated set of margins.

The variability of the margin in the above example would carry over to the net cash flow for each of the five years. In this instance the variability in the simulation of margins is perfectly masked as the mean over the five years is exactly the expectation of the margin. If, however, the project had been a four year one and the values for the margin had omitted the "50," the resulting average margin would be only 12.5 and the effect would be to introduce spurious variability into the net present value (NPV) criterion for project evaluation. This is a serious problem -- especially for an analysis method designed to measure variability.

But even if excessive variability in margins is dampened out of final NPV to some extent by compensating "errors," variability in annual results often has another detrimental effect. In evaluating projects in which losses are carried forward for their tax shield effects, excessive variability can affect the average simulation value or NPV as well as its dispersion. Figure 6-a illustrates the simulated cash flows from two highly variable 3-year project lives (simulation iterations) labeled (1) and (2) in comparison to the expected cash flows labeled (3). Because the project lives (1) and (2) have the same cash flows, but in a different order, one would expect the average NPV of the two to be the expected NPV. However, because of the carry forward of tax loss from year two, the after-tax cash flows are not "reversly symmetric" as they were before taxes (Figure 6-b), and Figure 6-c shows that this results in the average NPV of (1) and (2) being considerably below the expected NPV in (3).

Again we see that unreasonable variation in margin can deleteriously affect the simulation outcome. Spurious variability can likewise arise if the correlation between price and sales volume is neglected in simulating a price elastic market.

Figure 6-a
CASH FLOW BEFORE TAXES

Year	Iteration (1)	Iteration (2)	Expectation (3)
1	$ 500	$1000	$ 333
2	-500	-500	333
3	1000	500	333
Avg.	$1000	$1000	$1000

Figure 6-b
CASH FLOW AFTER TAXES

Year	Iteration (1)	Iteration (2)	Expectation (3)
1	$ 250	$ 500	$ 167
2	-500	-500	166
3	750	500	167
Avg.	$ 500	$ 500	$ 500

Figure 6-c
PRESENT VALUE
OF AFTER-TAX CASH FLOWS AT 10 PERCENT

Year	Iteration (1)	Iteration (2)	Expectation (3)
1	$ 277	$ 455	$ 152
2	-413	-413	138
3	563	375	125
Avg.	$ 377	$ 417	$ 415

Avg. of iterations (1) and (2) = $ 397 ≠ $ 415

6. DEVELOPMENT OF A STOCHASTIC CORRELATION FUNCTION

If the decision is made to introduce the effects of variable dependencies into a simulation model, several problems soon become apparent. An observed correlation between the variables of two distributions does not distinguish which variable is the cause and which is the effect (or even that a cause-effect relation exists). Even if the cause-effect problem can be solved, some stochastic function must be specified which, for each possible value of the independent variable, specifies a dependent variable in an appropriate range: a particular price can be expected to elicit a volume of sales within a given range, but not a specific value of sales for each specific price. The function relating the variables of the correlated distributions is a function of a random variable and, hence, is not analytic.

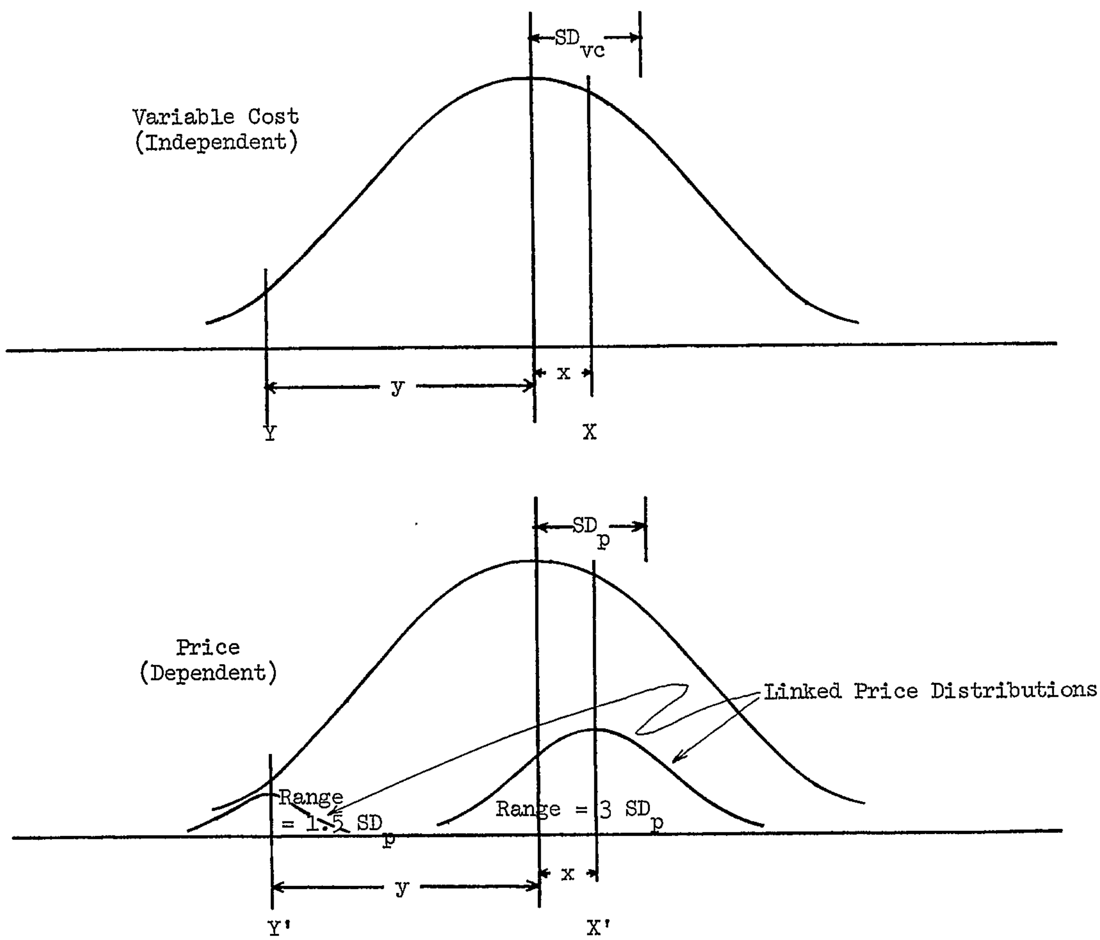
In FRANCE the cause-effect relations between correlated variables were determined on the basis of corporate experience in the relevant markets. In the case in which the variates of two distributions are correlated and a cause-effect relation has been established, we say that the two distributions are linked.

7. THE MECHANISM OF LINKING

Each distribution in FRANCE is defined to be representative of the complete spectrum (or population) of all possible values of the variable it represents. For dependent distributions such as price (linked to variable cost), this information is put to use in a rather interesting fashion. Figure 7 shows for the purpose of exposition two curves, each having the shape of a normal distribution, representing respectively variable cost and price.

If the simulation process selects a variable cost X which is x standard deviations from the median of the variable cost distribution, the linkage

Figure 7
DISTRIBUTION LINKING MECHANISM



mechanism will locate a price which is x standard deviations in the positive direction from the median of the price distribution. About that price, the model will establish a linked price distribution which spans a subset of the original price distribution. In Figure 7, the linked price distribution corresponding to a variable cost of X is centered on the price X' and is indicated by the normal-shaped curve over that point. The span or range of the linked distribution is a user-supplied parameter which controls the variability that is allowed in the linkage mechanism. Inverse relationships such as one might expect between price and sales volume can readily be expressed by translating positive displacements from the median in the independent distribution into negative displacements on the dependent distribution and vice-versa.

It is easy to see that the linking procedure has no effect upon the expected or average outcome of a simulation analysis if the following example is considered: Price is dependent upon (linked to) variable cost, and the range of the linked price distribution is taken as zero (this has no effect upon the generality of the following argument). Suppose that both the price and variable cost distributions are divided into centiles and each centile of the variable cost distribution is selected exactly once in an experiment of 100 simulations with no random error. Then, in the price distribution, each centile value will be correspondingly selected exactly once because of the direct linkage between the distributions. At the conclusion of the experiment, the average values of variable cost and price for the experiment will be exactly the means of the two distributions, as alleged above, i.e.,

$$\sum_{i=1}^{100} x_i f_i$$

where x_i is the value in the i-th centile and f_i is its frequency in the experiment, namely 1. Thus, the most restrictive linking conditions preserve the expected value of the distributions associated by the linkage and, hence, preserve the expected NPV.

8. QUADRATIC RANGE FOR LINKED DISTRIBUTIONS

Referring back to Figure 7, one sees that the range of the linked price distribution centered at X' is approximately $3SD_p$ or 1/2 the range of the main price distribution. Had the linked price distribution been centered at a point a distance y (say, $y = -2.5SD_p$) from the median of the main price distribution, it would be reasonable to expect that the range of the linked price distribution would be considerably less than $3SD_p$. More central values of variable cost cause little limitation upon corresponding values of price, whereas exceptionally high or low values for variable cost seem to predicate correspondingly high or low prices.

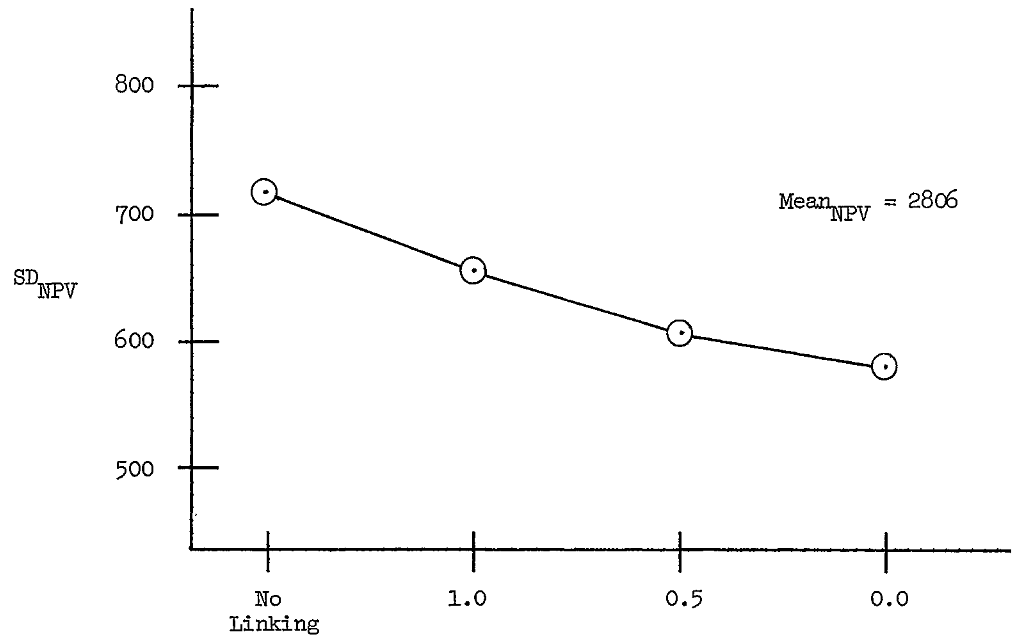
The PRANCE model embodies such a regulation of range in linked distributions. The range of a linked distribution is computed as an inverse quadratic function of the displacement of the linked distribution mean from the median of the main distribution. In Figure 7, if $x = SD_p/2$ and the range of the linked distribution at X' is $3SD_p$, the corresponding range of the linked distribution at Y' with $y = 5SD_p/2$ is approximately $3SD_p/2$.

9. MODEL VALIDATION

A large number of experiments were run with PRANCE to determine that it was functioning properly and to ascertain the sensitivity of the model to variations in selected input parameters. For example, it was verified that the average simulation value for NPV both with and without linking (and without tax loss carry forward) corresponds closely to expected value. Tests were made to determine that the centiles of empirical distributions were, in fact, picked on the average one time in 100 trials. The multiplicative congruential random number generator had been previously tested and found superior to other generators available.

The linkage technique just described was the subject of extensive testing and sensitivity analysis. Figure 8 below provides some interesting information relating the standard deviation and various degrees of variability in the input distributions. It is not possible on the basis of results to date to forecast the decrease in NPV variability with any degree of precision. In terms of Figure 8, a "fraction" of 0.5 is generally used for PRANCE simulation analyses.

Figure 8
EFFECT OF LINKING ON NPV VARIABILITY



linking procedure between variable cost and price. The degrees of linking shown on the abscissa vary from "no linking" to "linking with linked distribution having the same range as the main distribution" (fraction = 1.0) to "linking with a range of zero for the linked price distribution" (fraction = 0.0).

The data plotted in Figure 8 are from an actual project and represent the typical effect that linking has on the outcome of a project; it reduces the total variance. The magnitude of the effect is a function of the variability of the

Because PRANCE is a relatively new model, we have not yet been able to compare simulation results with actual project outcomes. However, the results of decisions made on the basis of PRANCE analyses are being closely studied as actual performance data is obtained.

10. OPERATING CHARACTERISTICS

PRANCE operates off-line in batch mode, reading user-supplied input data for one project, performing an expected value analysis using the mean (calculated by the model) of each input distribution, and then simulating the life of the project for the number of times indicated in the input data.

If the project group contains more than one project, only output pertaining to the first project alone is generated when that project is run. Subsequent projects in the group are run in the same manner as the first, and the processing of the last project in the group causes stored summary information on all projects to be printed.

Running time for a single project is directly proportional to the number of simulation iterations and the number of years in project life. For 500 iterations with a 10-year project, running time on the IBM 7044 is approximately 8 minutes.

11. OPERATING EXPERIENCE

Collecting the necessary input for FRANCE analyses has not proven difficult. Specifically, we have found that management can more readily supply input distributions expressed empirically, as opposed to supplying a distribution mean and standard deviation (assuming a normal distribution). The model has the ability, however, to accept input distributions in either form. It was felt that management could better represent empirical distri-

Figure 9
TYPICAL DATA COLLECTION FORM

FIXED COST ($ 000's)

Fixed costs represent those costs which are relatively independent of volume and consist largely of the costs which result from the possession of plant, equipment and a basic organization. Property taxes, insurance, salaries of key personnel and certain sales expenses are examples. A portion of the labor expense should be regarded as "fixed" since, as a mininum, a one-shift operation is necessary.

Depreciation should be excluded (from both fixed and variable categories) as the economic benefits from the tax shield are automatically handled by the model.

Year 1

Value ($ 000's)	____	____	____	____	____
Probability (%)	____	____	____	____	____

Successive Years

Year	____	____	____	____	____	____	____	____	____	____
Most Likely Outcome	____	____	____	____	____	____	____	____	____	____
Year	____	____	____	____	____	____	____	____	____	____
Most Likely Outcome	____	____	____	____	____	____	____	____	____	____

butions in histogram form by specifying five to eight values, as contrasted to supplying only the most likely, most pessimistic and most optimistic values and fitting a beta distribution to these points. Simplified data collection forms were designed for users of the model, of which Figure 9 (previous page) is an example.

The PRANCE model has been implemented throughout Crown Zellerbach. This widespread utilization has been achieved through a series of presentations to management describing the simulation approach to investment analysis, as well as the specific operation of the model. These presentations were results-oriented as opposed to technically-oriented.

A Users' Manual, in the final stage of preparation, will include sample analyses, simplified data input coding sheets, and other material describing the model.

12. FUTURE MODEL DEVELOPMENTS

At present, no explicit provision is included to relate sales volume to capacity so the user must carefully consider capacity in setting up level variables (annual distribution modes) for sales volume. The only linkage currently affecting sales volume is that with the price distribution, and this linkage frequently has not been employed because many of Crown's products are of a commodity nature and, thus, relatively price inelastic. We envision that modifications may later be incorporated so as to place capacity constraints on sales volume and also relate the machinery addition and rebuild schedules to these capacity constraints.

One modification currently being programmed is the inclusion of empirical distributions for terminal values (expressed as disposal recovery factors) on buildings and equipment.

13. CONCLUSIONS

This paper has described a Monte Carlo capital investment simulation model developed in order to meet specific objectives established by management. We believe that these objectives have been met. The method developed for linking dependent distributions has proven its usefulness by contributing decision-making data free of spurious variability for consideration by top management. The capital addition schedule, level variables and the other features described have resulted in more detailed and complete analyses than are presently available from other known capital investment simulation models.

Mr. Crabtree is a financial analyst in the corporate planning department, Crown Zellerbach Corporation, San Francisco, California. He joined Crown Zellerbach in 1968, the year he received his MBA from the Tulane University Graduate School of Business Administration. At Crown, Mr. Crabtree's duties have included acquisition studies and special project analyses (both frequently employing the capital investment simulation model described in this paper), and also teaching assignments in this connection. Mr. Crabtree formerly was a marketing representative with Rohm and Haas Company of Philadelphia, Pennsylvania. His undergraduate degree, in industrial engineering, was received from the University of Tennessee in 1959.

Mr. Dennis is President of Management Optimization Systems, Inc., a newly formed operations research consulting and software company. His activities outside corporate administration have included development work with Dr. Joseph L. Balintfy and C. Ralph Blackburn II on a soon to be announced integer programming code. Mr. Dennis was formerly Project Director of Computer System Research at Tulane University where he was concerned with applications of mathematical programming. At Tulane also, Mr. Dennis received an MBA in 1966 and is a Ph.D. candidate in operations research, with a thesis topic in mathematical programming in capital budgeting. He holds a BS from Purdue in physics.