Proceedings of the 1988 Winter Simulation Conference
M. Abrams, P. Haigh, and J. Comfort (eds.)

A graphical simulation environment for stochastic investment analysis

Jong S. Huang
Department of Information Systems
Auburn University at Montgomery
Montgomery, AL 36193

1. ABSTRACT

This paper presents a graphical simulator, RISK, for stochastic investment analysis. Issues concerning the user interface, the design concept, and the modeling process are discussed. An example is given to demonstrate the capability of the simulator.

2. INTRODUCTION

A typical investment may involve several factors such as the initial cost, expected life of the investment, the market share, the operating cost and so on. Values for factors are projected at the time the investment project is first proposed and are subject to deviations from their expected values. Such variations in the outcomes of future events, often termed risk, have been of primary concern to most decision makers in evaluating investment alternatives. RISK (Responsive Investment Simulation Kit) was developed to demonstrate how a special purpose simulator could simplify investment analysis under uncertainty.

3. MODELING PROCESS

The modeling process of RISK consists of four different stages. First, factors affecting the Net Present Value (NPV) are identified in the factor definition stage. Second, if correlated factors exist, their correlations are specified in the correlation definition stage. Third, factors are grouped together by arithmetic operators to form modules in the module definition stage. Fourth, modules are used to set up the simulation model in the model building stage. Finally, the simulation stage facilitates the execution of the defined model and the statistics are collected and displayed.

3.1 Factor Definition

The factor definition stage is the stage for identifying factors (variables) that are involved in the calculation of the NPV of a project. The user can define a factor by providing its name, distribution, initial seed value, growth factor, and attribute.

3.1.1 Screen Layout. When we enter the factor definition stage, we see the screen shown in Fig. 1. On the left side of the screen is the icon area, where icons of eight distributions are displayed. The defined factors are displayed in the factor area below. The big blank window in the middle is the information area where a user can enter (display) information associated with the factor he/she is going to define (review). The window at the lower left hand corner is reserved for displaying feedback messages.

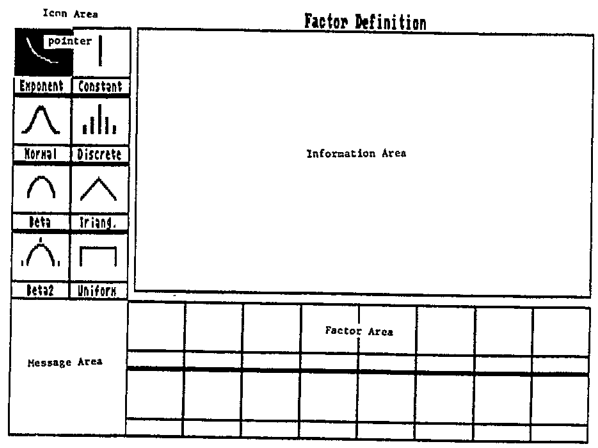

Figure 1: The Screen Layout for the Factor Definition Stage

3.1.2 Input Information. The input required for the factor selected can vary depending on the type of distribution chosen. There are, however, five pieces of information common to all factors defined.

A. name: Each factor is given a unique name to remind the user of its nature.

B. seed value: Each factor, except a constant, has a seed value. This will be the starting value used when generating random numbers.

C. identifier: The identifier is generated by RISK automatically and will be used to define modules later in the module definition stage. Identifiers for factors are designated by lower case letters.

D. attribute: The attribute value determines how often a random number has to be generated for a particular factor. There are two possible values we can specify for a given factor:

   1. L: Entering a letter "L" as the attribute means the factor is generated once throughout the duration of the investment project.

   2. P: Entering a letter "P" as the attribute means the factor is generated once for each time period.

E. growth: Growth factors are used to specify whether a factor will grow (or fluctuate) over time. For instance, rather than generate a new value for market share in every period, we would generate a value at the beginning of a simulation run, then assign a growth factor to that initial value. The growth factor can be either a constant number or another factor (i.e. a distribution.) By entering a "0" means no growth (fluctuation) is perceived.

The rest of the parameters required are necessary to identify a unique probability distribution.

3.2 Relation Definition

The relation definition stage is the stage where correlations between factors are specified. A correlation is measured by the correlation coefficient, for which a value between -1 and 1 is required. The user can either specify the correlation coefficient directly or answer questions asked by RISK, which will help find the correlation coefficient.

3.2.1 Screen Layout. The relation definition screen is demonstrated in Fig. 2. The screen is divided into three areas--the information area, the relation area, and the factor area. Factors defined in the factor definition stage are displayed in the factor area. The correlated pairs of factors are displayed in the relation area. The two factors involved are specified, and the correlation coefficient is estimated in the information area.

3.2.2 Input Information. The information required in this stage is to identify the two correlated factors and provide the program with their correlation coefficient if the user already knows it. If the user has no idea of what value he/she should use for the correlation coefficient (which is the common case), he/she should answer the question(s) provided by the program. The question asked, for example, would be like: " What is the most likely value for x, when the value for y is 100?" The correlation coefficient is then estimated through the answer(s) given by the user.

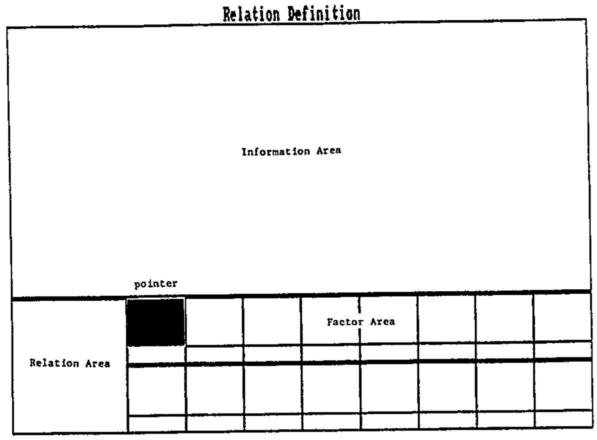

Figure 2: The Screen Layout for the Relation Definition Stage

3.3 Module Definition

In this stage we combine the defined factors into larger modules with arithmetic formula.

3.3.1 Screen Layout. The screen of the module definition stage is divided into four areas--the icon, module, factor, and information areas. (See Fig. 3.) There are sixteen different icons from which a user can choose to represent the module he/she defines. The user could develop icons that are different from those shown in Fig. 3. by using the icon editor in the simulator. Factors and modules defined in previous stages are displayed in the factor area and the module area, respectively. The information area is used by the information editor to obtain necessary information when defining a module.

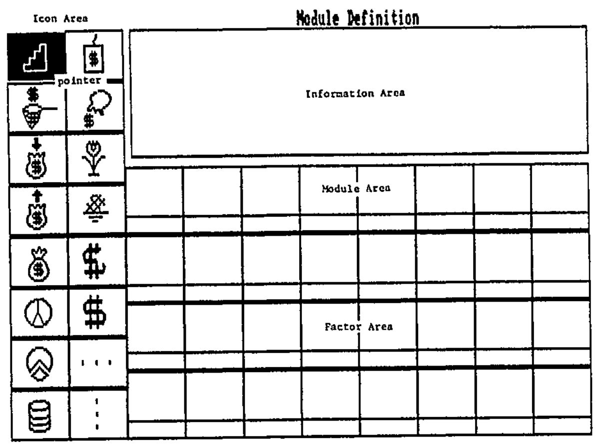

Figure 3: The Screen Layout for the Module Definition Stage

3.3.2 Input Information. There are four pieces of information required in this stage:

A. name: A name of 1 to 7 characters is assigned to each module defined to remind us of its nature.

B. identifier: This is an upper case letter assigned by the program. It has the same function as the identifiers assigned to factors.

C. attribute: There should be one attribute associated with each module defined. The attribute is used to specify the relation of the module to the NPV of the project.

1. L: This attribute tells RISK that the module is used for the project life.
2. I: An "I" is used to identify the module as the minimum attractive rate of return.
3. +: A "+" sign means the value of the module realized should be added to the NPV.
4. -: A "-" sign means the value of the module realized should be subtracted from the NPV.

D. formula: A formula which combines constants, factors, or modules with arithmetic operators.

3.4 Model Building

This is the stage where modules are combined to set up a simulation model. A user can move the defined modules into one of three areas: life-time, modeling, and life-end. Each of these areas can accommodate only specific modules. A model is interpreted by the way modules are assembled in these three areas and the attributes associated with each module.

3.4.1 Screen Layout. The screen for the model building stage is divided into four areas: life-time, modeling, life-end, and module areas. (See fig. 4.) The life-time area is where the user would put modules in effect throughout the entire simulated life of the project, such as project life or MARR. Modules that would be in effect only at the end of the project life, such as residual value, are put in the life-end area. The module area is used to display all previously defined modules. Modules representing the net cash flow in each year are put in the modeling area. The user can only see six modules at a time in the modeling area, but there can be as many as 50 modules defined in this area by scrolling.

A period number is associated with each net cash flow module put inside the modeling area. The difference between successive period numbers determines how many times a module will be used in the simulation. For example, if the first module (say G) is stored under period 5 and the next module is stored under period 10, module G will be used for periods 5 through 9 in the simulation.

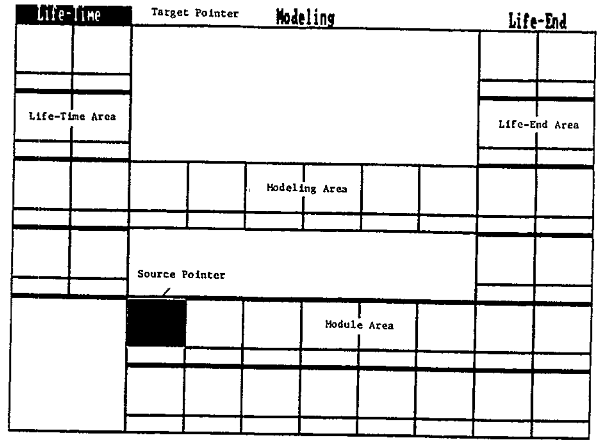

Figure 4: The Screen Layout for the Model Building Stage

3.5 Simulation

The simulation stage executes the model constructed in the model building stage and collects model statistics. The way the model is going to be simulated is controlled by use of the control panel.

3.5.1 Screen Layout. Along with the life-time, life-end, and modeling areas retained from the model building stage, the simulation screen has a graphics display area, a control panel, and a statistics display area (Figure 5).

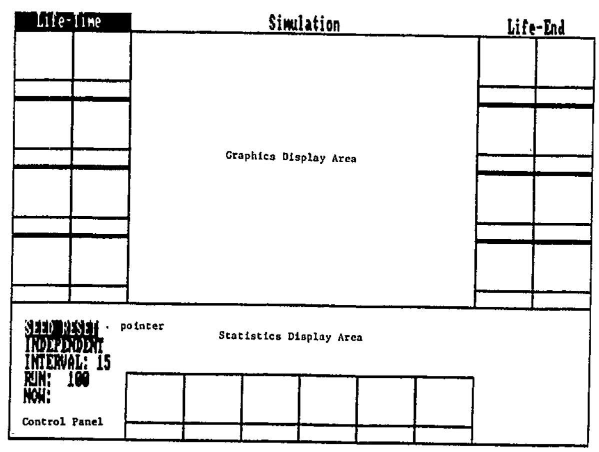

Figure 5: The Screen Layout for the Simulation Stage

The graphics display area displays the histogram of the net present values obtained in the simulation runs. The histogram is updated automatically for every hundred runs. The control panel controls some parameters used in the simulation, such as the run number, and the number of intervals used in the histogram. Statistics collected in the simulation runs are displayed in the statistics area. These statistics are:

A. NPV: The average net present value obtained up to this point from the simulation run.

B. STD: The standard deviation of the NPV obtained.

C. MIN: The minimum NPV obtained up to this point.

D. MAX: The maximum NPV obtained up to this point.

E. POS. NPV: The probability the the project will have a positive NPV.

4. EXAMPLE

After describing the modeling concepts and user interfaces, now is the time to look at a real world problem, which will be modeled in RISK.

4.1 Problem Definition

A medium-sized industrial chemical producer is considering an expansion of its processing plant. The nine input factors that the management has decided to use in evaluating the project, with the ranges and the most likely estimates for these variables, are listed in Table 1. The independent probability distributions selected by the management are shown in Table 2.

The management expects that the market growth rate (G) and share of market (S) will be dependent upon the particular value selected for the initial market size (M), that the useful life (L), residual value (R), and fixed cost (F) will be dependent upon initial investment (I), and that selling price (S) will be dependent upon operating cost (V). The correlation coefficients are given in Table 3. (This is the same example used by Hertz (1964), Eilon (1973), and Hull (1977).) We assume the Minimum Attractive Rate of Return (MARR) is 10%.

Table 1: Best Guess and Range for Each Factor Considered in the Example

Factor	Best Guess	Range
M Initial market size (thousand tons)	250	100-340
G Market growth rate (% p.a.)	3	0-6
P Selling price ($ per ton)	510	385-575
S Share of market (%)	12	3-17
I Initial investment (million $)	9.5	7.0-10.5
L Useful life (years)	10	5-15
R Residual value (million $)	4.5	3.5-5.0
V Operating cost ($ per ton)	435	370-545
F Fixed cost (thousand $)	300	250-375

Table 2: Independent Probability Distributions Assessed by the Management

M	100	0.05†	145	0.12	190	0.23	240	0.40	290	0.20	340
G	0.0	0.15	1.5	0.20	2.5	0.30	3.5	0.20	4.5	0.15	6.0
P	385	0.03	420	0.07	460	0.30	500	0.55	540	0.05	575
S	3	0.05	5	0.15	8	0.25	11	0.35	14	0.20	17
I	7.0	0.08	8.0	0.30	9.0	0.30	9.5	0.20	10.0	0.12	10.5
L	5.0	0.04	7.5	0.26	9.5	0.40	11.5	0.26	13.5	0.04	15.0
R	3.50	0.15	4.00	0.15	4.25	0.33	4.50	0.23	4.75	0.14	5.00
V	370	0.05	405	0.55	440	0.30	475	0.07	510	0.03	545
F	250	0.15	275	0.30	300	0.30	325	0.15	350	0.10	375

† The probability that M will lie between 100 and 145 is estimated to be 0.05.

Table 3: Correlation Coefficients between Correlated Factors

Dep/Indep Var.	G/M	S/M	L/I	R/I	F/I	P/V
Corr. Coeff. ρ	-.49	-.58	.62	.74	-.59	.50

4.2 Factor Definition

To define factors, we move the pointer to the distribution we want and press <ENTER> key. When prompted by the program, simply type in factor definitions using the information editor and press the <END> key when we're through with each definition. There are ten factors defined in this stage (see Fig. 6). A constant value is assumed for the MARR, and the rest of the factors are all estimated by distributions. The required input for each factor is listed below.

A. name: Msize
seed: 219
attribute: L
growth factor: Mgrowth
distribution: enter the intervals and probabilities for the discrete distribution listed in Table 2.

B. name: Invest.
seed: 23127
attribute: L
growth factor: 0
distribution: enter the intervals and probabilities for the discrete distribution listed in Table 2.

C. name: Salvage
seed: 219
attribute: L
growth factor: 0
distribution: enter the intervals and probabilities for the discrete distribution listed in Table 2.

D. name: Life
seed: 4131
attribute: L
growth factor: 0
distribution: enter the intervals and probabilities for the discrete distribution listed in Table 2.

E. name: Mshare
seed: 5555
attribute: L
growth factor: 0
distribution: enter the intervals and probabilities for the discrete distribution listed in Table 2.

F. name: Fcost
seed: 60331
attribute: L
growth factor: 0
distribution: enter the intervals and probabilities for the discrete distribution listed in Table 2.

G. name: Vcost
seed: 71349
attribute: L
growth factor: 0
distribution: enter the intervals and probabilities for the discrete distribution listed in Table 2.

H. name: Price
seed: 8713
attribute: L
growth factor: 0
distribution: enter the intervals and probabilities for the discrete distribution listed in Table 2.

I. name: Mgrowth
seed: 27967
attribute: L
growth factor: 0
distribution: enter the intervals and probabilities for the discrete distribution listed in Table 2.

J. name: Ror
value: 0.10
growth factor: 0

4.3 Relation Definition

There are five relations specified in this stage (see Fig. 7), as listed in Table 3. Since all five correlation coefficients are known, we simply enter the two correlated factors and their correlation coefficient.

A. b>f: -0.59 (f is dependent on b)

B. b>d: 0.62 (d is dependent on b)

C. b>c: 0.74 (c is dependent on b)

D. a>i: -0.49 (i is dependent on a)

E. a>e: -0.58 (e is dependent on a)

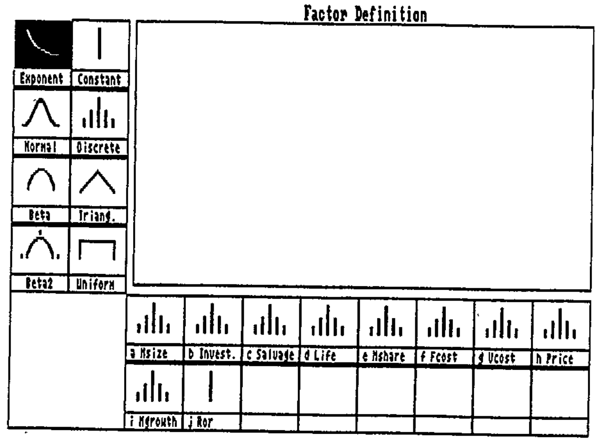

Figure 6: The Factor Definition Stage for Example

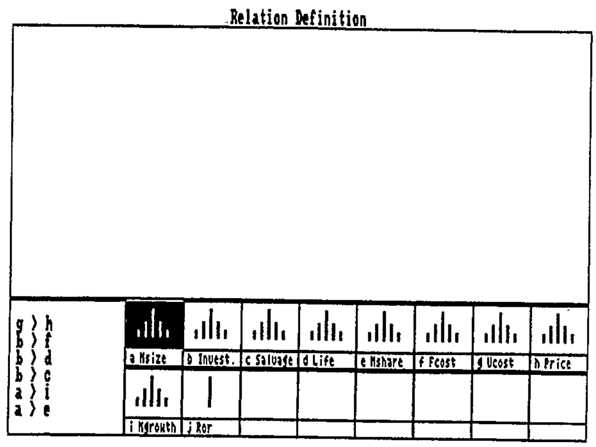

Figure 7: The Relation Definition Stage for Example

4.4 Module Definition

After identifying the factors involved in calculating the NPV, we need to combine mathematically related factors into modules, which will then be used as building blocks in model construction. To define the modules, the user must first select the icon he/she would like to use, then enter associated information when prompted. There are eight modules defined in this stage, as shown in Fig. 8.

A. name: Volume
attribute: +
formula: a * e / 100.0

B. name: Inflow
attribute: +
formula: A * h

C. name: Outflow
attribute: +
formula: g * A + f

D. name: Rate
   attribute: i
   formula: j

E. name: Life
   attribute: L
   formula: d

F. name: Icost
   attribute: -
   formula: b

G. name: Net
   attribute: +
   formula: B - C

H. name: Salvage
   attribute: +
   formula: c

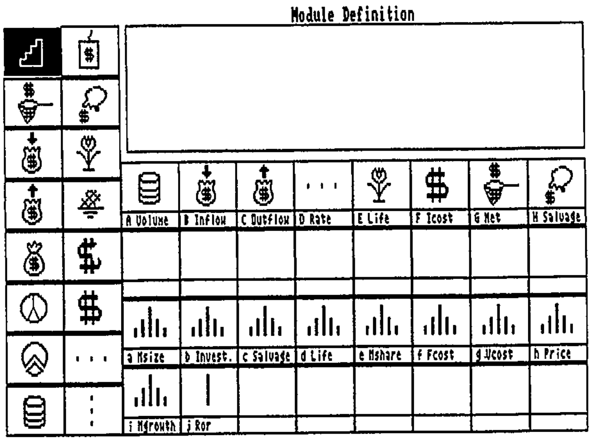

Figure 8: The Module Definition Stage for Example

4.5 Model Building

When the module definition stage is completed, we have eight different modules. To set up a model, we need only move the modules into one of the three areas on the screen--the life-time, the modeling, and the life-end area (see Fig. 9). Modules that are in effect throughout the life of the project should be put inside the life-time area, such as MARR and Life; modules that are used only at the end of the project life are moved into the life-end area, such as Salvage; modules that are used in certain periods over the life span of the project are placed in the modeling area. The numbers in the modeling area specify when a module is evaluated. For example, in Fig. 9, the module Icost is evaluated in period 0 whereas the module Net is evaluated from period 1 through the life of the project as determined by the module Life.

4.6 Simulation Results

The results obtained from the simulation are shown in Fig. 10 for the independent case and in Fig. 11 for the dependent case.

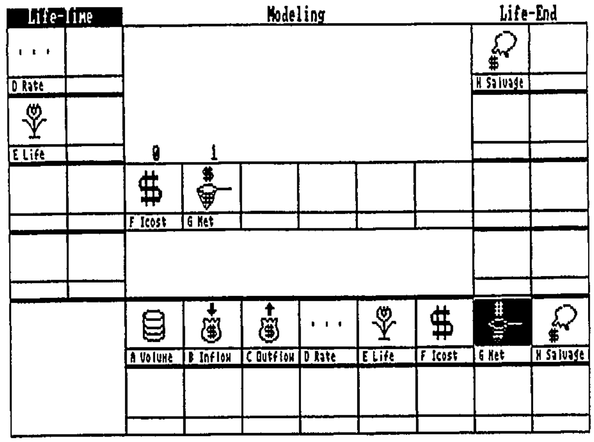

Figure 9: The Model Building Stage for Example

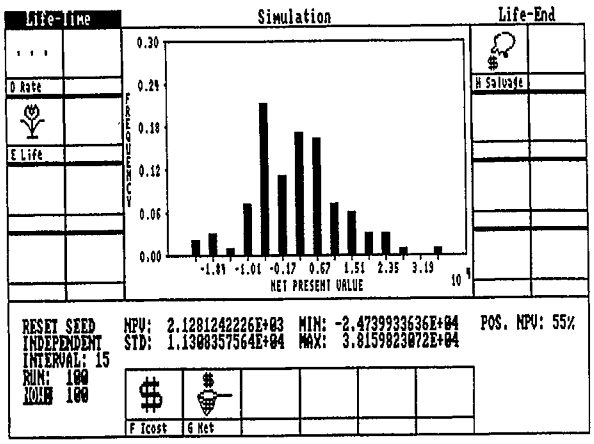

Figure 10: The Simulation Results for the Independent Case

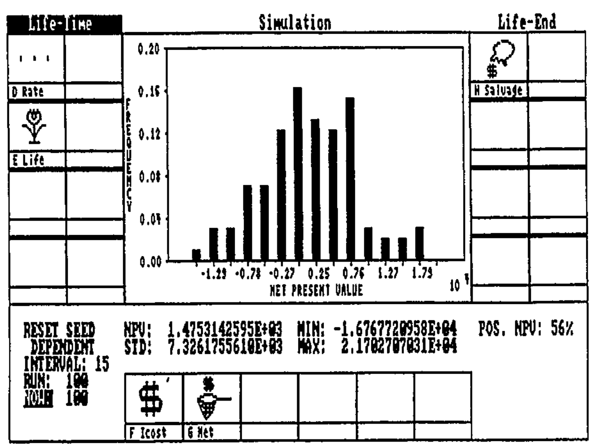

Figure 11: The Simulation Results for the Dependent Case

5. SUMMARY

The simulation process of traditional simulation language often involves modeling and coding as well as the computer simulation itself. More often than not, coding becomes the most time-consuming part of the whole simulation process. For those who are not familiar with the simulation languages, it is almost impossible to obtain simulation results in a short period of time. The development of RISK serves as a demonstration of how interactive graphic modeling can speed up the simulation process. It does not impose the coding task on the user; therefore, it allows the user to concentrate on modeling and analysis of results rather than programming.

6. REFERENCES

Banks, J. and Carson J.S.(1984). *Discrete-Event System Simulation.* Prentice Hall, 358-367.

Bonini, C.P.(1975). Risk Evaluation of Investment Projects. *Omega,* Vol. 3, No. 6, 735-751.

Bonini, C.P.(1975). Comment on 'Formulating Correlated Cash Flow Streams. *The Engineering Economist,* Vol. 20, No. 3, 209-221.

Bussey, L.E. and Stevens G.T.(1973). Formulating Correlated Cash Flow Streams. *The Engineering Economist,* Vol. 18, No. 1, 1-30.

Economos, A.M.(1968). A Financial Simulation for Risk Analysis of a Proposed Subsidiary. *Management Science,* Vol. 15, No. 12, 675-682.

Eilon, S. and Fowkes T.R.(1973). Sampling Procedure for Risk Simulation. *Operations Research Quarterly,* Vol. 24, 241-252.

Hertz, D.B.(1964). Risk Analysis in Capital Investment. *Harvard Business Review,* Vol. 42, 95-106.

Hertz, D.B.(1968). Investment Policies that Pay off. *Harvard Business Review,* 153-165.

Hull, J.C.(1977). Dealing with Dependence in Risk Simulation. *Operations Research Quarterly,* Vol. 28, 201-213.

Newendorp, P.D. and Root P.J.(1976). A Method for Treating Dependencies Between Variables in Simulation Risk-Analysis Models. *Journal of Petroleum Technology,* 1145-1150.

BIOGRAPHICAL SKETCH

JONG S. HUANG is an assistant professor of Information Systems at the Auburn University at Montgomery. He holds a Mater's degree in Industrial Engineering from Stanford University and a Ph.D. degree in Industrial Engineering from Auburn University. His research interests include simulation, object-oriented programming and database, computer graphics, and artificial intelligence.

Dept. of Information Systems
Auburn University at Montgomery
Montgomery, AL 36193