Cross-Impact Simulation of an Emerging Industry: The Case of Data Processing

Paul H. Rosenthal
Gottfried Consultants Inc.
Los Angeles
and
Paul Gray
Cox School of Business Administration
Southern Methodist University

Abstract

Emerging 'High Technology' and 'New' industries present complex forecasting problems since their growth patterns are correlated to GNP growth, to nominal improvements in cost/performance from routine R & D, and to potential innovations or policy interventions. This paper describes the application of the USC Center for Future Research INTERAX System to forecasting the size and structure of Data Processing Industries. Results indicate that some proprietary planning organizations may be factoring into their forecasts subjective probability-based estimates of the impact of technology innovations on product cost/performance and industry sales. The emerging industry simulation methodology is data-driven, thereby permitting rapid definition of industry structure and derivation of elasticity/substitutability coefficients by tuning the model to historic or short-term forecast data.

INTRODUCTION

This paper presents a methodology for forecasting the long-term growth and structure of emerging industries. Such industries present complex forecasting problems since their rapidly changing technologies and product mix do not lend themselves to the traditional growth rate extrapolation methods, which assume stable product and technology interrelationships. Described is an easy-to-use computer simulation system that forecasts the growth of multiple types of inter-related 'high technology' and 'new' industries.

Our hypothesis was that a straightforward econometric model could be produced that would incorporate as parameters "natural" measures such as cost/performance trends, and would accurately forecast long-term industry growth and structure. This approach and hypothesis is a continuation of the long-range forecasting and modeling efforts of the University of Southern California's Center for Future Research. The generic cross-impact simulator program used was developed and is extensively used by the Center. The Data Processing Industry was chosen as the illustrative case study for the emerging industry application model incorporated in the generic program, since extensive traditional short-term forecasting data was available from a recent multi-client research study in Network Information Services performed at the Center.

METHODOLOGY

The forecasting model uses five types of data: elasticity coefficients, cost performance trends, event probabilities, event on trend impact coefficients, and event on event odds multipliers.

a. Elasticity coefficients were derived for total industry growth characteristics and for sub-industry penetration (market share) characteristics. The derivation of the elasticity coefficients includes the "tuning" of the model to fit short-term forecasts produced by various proprietary market research firms. For emerging industries where such forecasts are not available, historical results could also be used.

CH1437-3/79/0465-0471$00.75
1979 Winter Simulation Conference

b. Cost/Performance trends were derived directly from technology forecasts by various laboratories. Such forecasts are normally available in the literature of high technology industries and follow the "S" type curve of increasing then decreasing rates of cost/performance growth.

c. Event probabilities are cumulative probabilities of each independent event occurring through period t. These probabilities were derived from published sources (Delotta et. al., 1976; Lecht, 1979; Withington, 1975) and from in-depth Delphi interviews with several experts in data processing technology forecasting. This methodology limits the practical number of events severely (Eymard, 1977, page 228). Therefore, the model used five innovation events plus an individually introduced alternate scenario policy event.

d. Event on trend impact coefficients are multiplicitive coefficients that modify cost/performance trends over a specific period. The values of these coefficients were also derived from the Delphi interviews.

e. Event on event odds multiplication coefficients are odds of occurrence multipliers that modify normal event probability distributions (see Alter, 1976, for methodology) and are used to define event cross-impacts. These cross-impacts specify the impact of occurrence of an event on the probability of other events occurring in subsequent periods. The values of these coefficients were also derived from the Delphi interviews.

The emerging industry application model was incorporated into the USC generic cross-impact model program, and the elasticities tuned to fit 1975 through 1983 traditionally derived industry forecasts.

The equation used for total industry demand was:

$$S_t = S_{t=o}(1 + r)^t (1 + n\sqrt{\overline{T}_t - \overline{T}_{t=o}})$$

where S_t = total industry constant dollar sales at period t.

r = basic annual real growth rate of the economy (a value of .035 was used for the data processing industry simulations).

T = weighted mean technology index of cost/performance trends.

n = technology/demand elasticity coefficient (simulations used a value of .015 for the data processing industry.)

The demand equation can be derived as the product of three factors: initial industry size, real GNP correlate, and emerging technology correlate.

S_t (constant dollar sales) =
S_o (Initial Industry Size) x
f(real GNP) x g(cost/performance)

A long-term view of GNP growth is used for the second term of the demand equation.

$$f(\text{real GNP}) = (1 + r)^t$$

The traditionally high correlation between GNP and industry demand is widely used in forecasting (Spencer and Siegelman, 1964, page 150). Its use in the demand equation accounts for growth due to an expanding total economy.

The emerging technology term is derived from the simplest functional form of a declining marginal utility curve (Tintner, 1965, page 54).

$$\frac{dS}{dT} = aT^{-b} \quad \text{where } o < b < 1$$

Solving the differential equation and substituting for initial conditions at t=o, gives the basic production/demand function used.

$$S = nT^d + 1 \text{ where } d = b + 1$$
$$\text{implying } o < d < 1$$

The last step in the derivation is the determination of the exponent (d) of the technology term T. The mid-point of the range for the exponent (.5) was used as an initial value, as did Wahi (1972). Based on the close fit of the derivative of the simulation result curves to the calibration data, no later adjustment was made.

The equation used for determining market share at both the primary industry level (product differentiation level) and at the sub-industry segment level (technology or delivery system differentiation level) was:

$$M_{i,t} = M_{i,t-1} (\emptyset + (1 + \emptyset) \frac{n\, T_{i,t})}{\sum T_{i,t})}$$

where $M_{i,t}$ = market share of i th industry segment during period t.

$\emptyset$ = substitutability coefficient

if $\emptyset = 0$ there is complete substitutability, no product differentiation. If $\emptyset = 1$ there is no substitutability, isolated markets. (The data processing industry simulation used values of .25 and .15 for the primary and sub-industry levels.)

The industry segments used were:

i = 1	All Network Services (sum of i = 4 and 5)
i = 2	Centralized Services
i = 3	Decentralized Services
i = 4	Network Information Services
i = 5	Dedicated Network Services

These econometric demand functions are similar in format to the equations used in numerous management simulations games; see, for example, the Yale University game (Shubik, 1964) and the IBM game (Wahi, 1972).

The simulation model was used to produce a fifteen-year forecasting run of the data processing industry from 1975 through 1990. Three types of simulation runs were performed.

a. Nominal Scenario
A nominal scenario was run utilizing existing variable inter-relationships, nominal trends, and excluded the impact of future events. This run operated the simulator model in the same deterministic format as traditional short- and medium-term simulation models and excluded the long-range forecasting cross-impact elements.

b. Stochastic Scenario
A stochastic scenario was run utilizing cross-impact probabilities and impacts. This run used Monte Carlo methods to approximate the impact of occurrence of abrupt events and changes.

c. Alternate Policy Scenarios
Several alternate stochastic scenarios were run incorporating the specialized policy or technology event, "White Collar Unionization Expansion." These runs demonstrated the methodology for using the simulator model for studying the impact of potential policy interventions.

SIMULATION RESULTS

The INTERAX simulation model in the companion paper by Enzer (INTERAX: An Interactive Simulation Model for Futures Research) at this conference produced results closely matching the published forecasts of various experts for the next decade in data processing.

Clarity of Model

The trade-off between realism and simplicity defines the clarity and generality of the model. Based on limitations due to complexity requirements (Armstrong, 1978, page 204) and expert opinion survey requirements (Eymard, 1977, page 228), the data processing industry simulation was performed with the following limited number of variables.

o. Cost/Performance Trends
Four independent trends were used: maxi-computer, mini-computer, shared network, and dedicated network cost/performance. A dependent trend, data communications cost/performance was also used.

- o Exogenous Events
  Four independent abrupt impact events were used: computer architecture, micrologic, shared network, and dedicated network innovations. A dependent event, data communications innovation was also used. A policy intervention event was used in the alternate scenario runs.

Although, the four event and four trend model would appear very simple, the research, interviews, and computation required to reliably derive the required variables turned out to be very complex. Several hours of interviewing were required to explain the model and the parameters to each participant and to obtain a full set of parameter estimates. Four "experts" and the authors submitted a full set of parameters. Several other persons interviewed either would not take the time to understand and create the complex opinion set or felt that their expertise did not cover the full range of technologies and products.

The authors feel that the sample size is small but adequate, since the data processing forecasting community is small and utilizes commonly available basic forecasts.

Comparisons of Forecasts

The forecasts produced by simulation are compared to independently obtained forecasts in the following table:

The deterministic forecast tracked a composite forecast combining such diverse groups as: Frost and Sullivan (ComputerWorld, January, 1977), INFORUM research project (Almon et. al., 1974) and Quantum Sciences Corporation (MAPTEK, 1975). The stochastic simulation forecasts tracked the International Data Corporation (IDC) forecasts (Fortune, March, 1977). The IDS forecasters appear to incorporate some level of innovation expectations in their forecasts.

APPLICATION OF MODEL

The model produces annual real dollar and current dollar figures for each segment of the emerging industry. The following table presents the 1990 forecasted constant dollar results for the nominal and stochastic scenarios.

Forecasted 1990 Data Processing Industry Sales
(in billions of 1975 dollars)

Type Service	Nominal Scenario	Stochastic Scenario
Centralized	24	23
Decentralized	16	16
NIS	5	28
Dedicated Network	10	12
Total DP Industry	$55	$79

The introduction of the innovation occurrences significantly expanded the market while changing its structure.

TOTAL DP INDUSTRY FORECASTS
(in 1975 billions of dollars)

	1980		1983	
Model Results	$	MAPE	$	MAPE
Nominal Scenario	50.03	1%	69.9	1% of compromise
Stochastic Scenario	53.89	6%	78.8	1% of IDC
Validation Data				
Composite Forecasts	50.5		69.5 (11.2% growth rate)	
IDC Forecasts 750			78 (15.6% growth rate)	

Econometric models of the type used in the emerging industry simulation are well suited to testing sensitivity and accuracy (Armstrong, 1978, 343). By inserting several values of the parameter under investigation, a sensitivity index can be computed giving the planner a quantitative feel of how results will vary as the inputs are changed.

As an example, the following results were produced during a sensitivity analysis of an alternative policy scenario.

Sensitivity Analysis Data
(in millions of 1975 dollars)

Value of Event-On-Trend Coefficient	Total 1990 DP Industry Sales
1.0*	79
.9	78
.8	77
.7	75
.6	73
.5	71
.4	68
.3	66
.2	64
.1	61

*Has no effect giving same results as stochastic scenario.

As the value of the policy multiplier is reduced, the nominal technology trends are reduced and with it Data Processing Industry Sales. A good way of analyzing these results is to compute, as illustrated, a sensitivity index: the ratio of percentage change of the results to the percentage change of the parameter.

Sensitivity Results

Parameter Deviation	Sensitivity Index
-10%	.08 = (-.8%/-10%)
-20%	.09
-30%	.13
-40%	.16
-50%	.18
-60%	.20
-70%	.21
-80%	.23
-90%	.24
Average	.17

The interpretation of the index is as follows: on the average, a percent change in the value of the event-on-trend coefficient will create a .17 as large percentage change in current dollar data processing sales.

CONCLUSIONS

The objectives of the model were to (1) apply cross-impact modeling to micro-economic planning, (2) develop quantitative technology diffusion and demand functions, (3) incorporate constant dollar cost/performance growth rates, and (4) create a usable system, balancing simplicity and generality and producing realistic results. In general, these objectives have been met, but with several important limitations.

Micro-economic Application of Cross-impact Modeling

The model should be applicable to technology-based emerging industries such as calculators and videographics. The data processing industry simulation forecast used as an example of the system appears successful in the sense of producing consistent and realistic results.

Development of Technology Diffusion and Demand Functions

The model is econometric in nature, forecasting cost/performance trends under technology diffusion assumptions and deriving market share and industry sales based on elasticity-based demand equations. The model accurately forecasted data processing industry structure through forecasting underlying trend variables and derived sub-industry and industry sales.

Incorporation of Constant Dollar Growth Rates

The model is capable of using any form of cost/performance growth curve. The data processing industry forecast uses constant dollar growth rates varying from 4.5% to 9% per year. These growth rates reflect expert industry forecasts (Withington, 1975) and should hold true during the emerging phase of the data processing industry.

Usable System

The model is based on a generic cross-impact program (Alter, 1976) which should permit easy modification or expansion of the model. The system is interactive and runs on a PDP-10 system utilizing the FORTRAN language. As discussed in prior paragraphs, the model, like any cross-impact system, requires the specification of many probabilities, trends, and multipliers. Even with four events and trends, the number of parameters is quite large.

Based on estimates of several weeks to learn and produce the structure of a new emerging industry, and a day for each expert's interview, a month's work is probably required to get a model running. This implies the model is straight-forward, but still complex to implement in practice.

BIBLIOGRAPHY

AFIPS. Information Processing in the United States: A Quantitative Summary. Montvale, NJ: AFIPS Press, 1976.

Almon, C., Buckler, M. B., Horwitz, L. M., Reimbold, T. C. 1985: Interindustry Forecasts of the American Economy. Lexington, Mass.: D. C. Heath and Company, 1974.

Alter, S. The Computational Mathematics of Time-Dependent Cross-Impact Modeling. M26, Center for Futures Research, University of Southern California, October, 1976.

Alter, S. "The Evaluation of Generic Cross-Impact Models." Futures. (2), April, 1979, 132-149.

Armstrong, J. S. Long-Range Forecasting: from Crystal Ball to Computer. New York: John Wiley, & Sons 1978.

Chow, Gregory C. "Technological Change and the Demand for Computers." The American Economic Review. (57), 5, December, 1967, 1117-1130.

Diebold Research Program, The. The Implications of Converging Computer and Communications Technology. M40, New York: The Diebold Group, Inc., August, 1977.

Dolotta, Bernstein, Dickson, France, Rosenblatt, Smith, and Steel. Data Processing in 1980-1985. New York: John Wiley & Sons, 1976.

Duval, A., Fontela, E., and Gabus, A. Cross-Impact - A Handbook on Concepts and Applications. Geneva: Battelle, 1974.

Eymard, J. A. "Markovian Cross-Impact Model." Futures. (9), 3, June, 1977, 216-228.

Frost and Sullivan, Inc. "Time-sharing Seen Losing Ground as Service Choice." (XI), 3, ComputerWorld, January 17, 1977, 47.

International Data Corporation. "Business Communications: The New Frontier." Fortune. (98), 7, October 9, 1978, 25-80.

International Data Corporation. "Distributed Processing/ Data Communications." Fortune. March, 1977, 31-82.

Lecht, C. P. The Waves of Change: A Techno-Economic Analysis of the Data Processing Industry. New York: Advanced Computer Techniques Corp., 1977.

Linstone, H. A. and Turoff, M. (Ed.) The Delphi Method: Techniques and Applications. Reading, Mass.: Addison-Wesley, 1975.

MAPTEK Strategy Report. San Jose, CA: Quantum Sciences Corporation, December, 1975.

McCarter, P. M., "Where is the Industry Going." Datamation. (23), 2, February, 1977, 63-65.

McLean, M., "Does Cross-Impact Forecasting Have A Future." Futures. (8), 4, August, 1976, 345-349.

Mitchell, R. B. and Tydeman, J. "Subjective Conditional Probability Modeling." Technology Forecasting and Social Change. (11), 2, 1978, 133-152.

Shubik, M. A Business Game for Teaching and Research. New Haven, CT: Cowles Foundation, Yale University, 1964.

Spencer, M. H. and Siegelman, L. Managerial Economics. Homewood, IL: Richard D. Irwin, Inc., 1964.

Tintner, G. Econometrics. New York: John Wiley & Sons, Inc., 1952.

Turn, R. Computers in the 1980's. New York: Columbia University Press, 1974.

Wahi, P. N. "A General Management Business Simulation in APL." IBM Systems Journal. (11), 2, 1972, 169-180.

Withington, F. G. "IBM's Future Large Computers." Datamation. (24), 7, July, 1978, 115-120.