This paper describes the application of the USC Center of Future Research's Cross-Impact Model to the simulation of the evolution of an emerging industry. The Data Processing Industry (DP) and its emerging Network and Decentralized Systems sub-industries are used to illustrate the application of the model. This application of the USC cross-impact forecasting model to the simulation of the evolution of an emerging industry and its component parts will constitute its first micro-economic application.

During recent data processing forecasting projects, it became apparent that a model describing the behavior of complex emerging industries was needed to facilitate forecasting the structure of such industries as data processing, calculators, solar energy and copiers. The author, therefore, embarked on applying the methodology of the Interactive Cross-Impact Model to demand analysis of such technology-based emerging industries using the DP industry as a case study.

The objective of this paper is, therefore, to present a methodology for forecasting and analyzing the effect of technology innovations and policy interventions on the growth and structure of an emerging industry, such as Data Processing. The Network Information Services and the Decentralized Systems sub-industries were chosen as illustrative of the class of technology-based emerging industries, since they are currently at a crossroads in their development due to the emergence of new competing and enabling communications and mini-computer technologies and services. More detail on the model and methods can be found in *A Cross-Impact Simulation of an Emerging Industry: A Case Study of Data Processing* (Rosenthal, 1979).

**EMERGING INDUSTRY STRUCTURE**

The model described defines three levels of emerging industries: the independent new industry, the primary industry segment—offering new services, and the subindustry segment—offering new production methodology for supplying an existing service.

**Independent industry**

The independent or total industry level is illustrated by the Data Processing Industry. It would have been considered by Lynn (1966) a "new" industry and its evolution would be measured by its total sales growth or its change in GNP percentage.

**Primary industry segments**

The primary industry segment level is illustrated by Centralized Data Processing Services, Decentralized Data Processing Services, and the Networked Services Industries. This level is characterized by product differentiation. For example, to the user these industries offer different services and products for use in different applications. For the technologist, these industries are simply varying delivery vehicles using varying mixtures of the same or different technologies.

**Subindustry segments**

The subindustries level is illustrated by the Network Information Services (NIS) Industry and its competitive, internally provided Dedicated Networked Systems Industry. This level is characterized by multiple delivery or technological approaches to providing the same product.

It is this level that much of the technological change literature approaches. For example, Gold, Pierce, and Rosegger (1975), in their paper on "Diffusion of Major Technological Innovations," measure the proportion of total output accounted for by fourteen major production process innovations in the Iron and Steel Industry.

Both this level and the primary industry segment level are measured in the model by share of market and by penetration rate (rate of change of market share). This approach cancels out the effects of total industry growth.
THE CROSS-IMPACT FORECASTING PROCESS

The cross-impact forecasting system consists of: a generic simulation system that operates in either scenario or statistical mode, a user-provided application model, and a data base of input parameters and output variables.

Data base contents

The data base consists of an input parameter file defined and completed by the user and FORTRAN COMMON tables defined by the generic cross-impact system. The input parameter file created by the user contains the following information for the emerging industry model:

a. Event Probabilities—cumulative innovation event probabilities. These probabilities are normally derived through interviews with experts.
c. Event/Event Impacts—event-on-event odds multipliers and delay/decay time periods that determine a period of applicability. These values are also derived from experts’ interviews.
d. Event/Trend Impacts—event-on-trend multipliers and delay/decay time periods that determine the length of time an innovation takes to impact product cost/performance. These values are determined from forecasts and technology diffusion equations.
e. Elasticity/Substitutability Coefficients—a series of parameters determining the impact level of cost/performance on total industry growth and industry segment market shares. Their values are derived by tuning the model to historic data or to short-term forecasts.
f. Initial Industry Structure—A series of initial values for the current size of the industry. These values are derived from historic data.
g. Names—names are given for events, trends, and output variables for clarity of output reports.

The output variable tables are used during computations to store intermediate values and are formatted and printed on the output reports. Data is maintained for the three levels of emerging industries: independent total industry, primary industry segments, and subindustry segments. The data maintained for each level includes:

a. Current Dollar Sales—dollar sales in current dollars for each time period of the simulation.
b. Real Dollar Sales—dollar sales for each time period in real dollars using the initial year of simulation as the base year.
c. Growth Rate—percent growth rate for each time period of real dollar sales.
d. Market Share—market share of segment for each time period.
e. Technology Index—the value for each time period of its cost/performance trend associated with the segment. This value is the nominal value adjusted by the impact of innovation (event) occurrences.
f. Penetration Rate—the rate of change of market share for each time period of the segment.

Scenario mode simulation

Figure 1 outlines the structure of scenario mode simulation. Each interaction with the forecasting system defines and performs a single simulation forecasting run covering the time period of the forecast. The process consists of the following five steps.

Define industry model

This step involves the creation of the input data base and an initial interactive intervention defining the simulation period and the initial random number seed.

Initialize industry variables

This step consists of an emerging industry FORTRAN application program that moves initial values from the input file to the variable tables.

![Scenario mode simulation diagram](Figure 1—Scenario mode simulation.)
Cross-impact generic model

This step is performed by a large and complex generic cross-impact FORTRAN program. A Monte Carlo simulation methodology is used to determine event occurrence based on accumulative event occurrence probabilities. When an event (an innovation or policy intervention) occurs, odds multipliers are used to modify other event probabilities and multipliers are used to modify cost/performance trends. The impact of an event occurrence can be spread over several periods through its use of delay and decay coefficients. A later step is also performed by the same program consisting of output printing of scenario results and a test for last time period.

Emerging industry model

This step is performed by an emerging industry FORTRAN application program that computes variable values based on adjusted trends and prior period variable values.

Interactive intervention

This generic cross-impact FORTRAN program allows interactive modification of probabilities, input trends, and non-structural emerging industry parameters such as elasticities and substitutability coefficients. The program will not change the number of events, trends, or variables.

Statistics mode simulation

Figure 2 outlines the structure of statistics mode simulation. Each interaction with the forecasting system causes multiple scenario simulations to be run, each utilizing a different set of random numbers to generate event occurrences. At the completion of each scenario, sums and sums of square are accumulated for use in computing means and standard deviations of trend values. Means are also computed for event occurrence frequency and segment variable values including current dollars, real dollars, growth rate, market share, technology index, and penetration rate.

INDUSTRY STRUCTURE

A straightforward, two-level tree emerging industry structure is utilized as shown in Figure 3.

The total Industry is defined as the sum of its segments each characterized by highly differentiated product or service types. The share of market of each industry segment is assumed to be based primarily on the cost/performance of their products with moderate substitutability over time. The simulation model is used to trace the sudden emergence of a new subindustry product or service within one of the industry segments because of substantially improved product cost/performance. The parent segment then ex-
pands, creating expanded research and development activities in the related industry segments, which may then produce innovations increasing the cost/performance of related products.

As an example of this type of industry evolution, the data processing industry structure in Figure 4 will be used.

Note that the Data Processing Industry is defined not by sales of components or services (such as minicomputers, printers, leased lines, or software packages), but instead by end user deliverable systems which meet specific information processing needs. Because of this structural approach, technological innovations in components and services, such as micro-electronics and communication services, will differentially impact all industry segments with one segment often receiving the primary benefit.

MARKET STRUCTURE

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 Data Processing Industry forecasts.

The equation used for total industry demand was:

\[ S_t = S_o(1 + r)(1 + \eta \sqrt{T_t - T_{t=0}}) \]

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.
- \( \eta \) = 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.

\[ S_{\text{constant dollar sales}} = S_o(\text{Initial Industry Size}) \times f(\text{real GNP}) \times g(\text{cost/performance}) \]

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).

\[ \frac{dS}{dT} = aT^{-b} \text{ where } 0 < b < 1 \]

Solving the differential equation and substituting for initial conditions at \( t = 0 \), gives the basic production/demand function used.

\[ S = nT^d + 1 \text{ where } d = -b + 1 \text{ implying } 0 < d < 1 \]

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:

\[ M_{i,t} = M_{i,t-1} \left(1 + (1 + 0) \frac{nT_{i,t}}{\sum T_{i,t}}\right) \]

where

- \( M_{i,t} \) = market share of \( i \)th industry segment during period \( t \).
- \( \theta = 0 \) there is complete substitutability, no product differentiation. If \( \theta = 1 \) there is no substitutability, isolated markets. (The data processing industry simulation used values of .25 and .15 for the primary and subindustry 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.

Nominal scenario

A nominal scenario was run utilizing existing variable inter-relationships, nominal trends, and excluded the impact.

<table>
<thead>
<tr>
<th>Total Industry</th>
<th>Data Processing Industry</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Parent Segment</td>
<td>Network Systems</td>
</tr>
<tr>
<td>A. Emerging Sub-Industry</td>
<td>Network Information Services</td>
</tr>
<tr>
<td>B. Competitive Sub-Industry</td>
<td>Dedicated Networked Systems</td>
</tr>
<tr>
<td>2. Related Segment</td>
<td>Centralized Systems</td>
</tr>
<tr>
<td>3. Related Segment</td>
<td>Decentralized Systems</td>
</tr>
</tbody>
</table>

Figure 4—Illustrative market structure.
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.

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.

Alternate policy scenarios

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

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).

OPINION SURVEY INPUT DATA

A survey was performed of several 'experts' in the data processing industry who specialize in technology forecasting. The experts were requested to supply the following subjective estimates:

a. Cumulative probability of innovation/regulation event occurrences (assuming non-occurrence of the other events).

b. Event-on-event cross-impact coefficients.

c. Event-on-trend impact coefficients and delay/decay periods.

Figure 5, "Event-on-Event Impact Structure," shows the structure of the event-on-event odds modifiers derived from the opinion survey and used in the Data Processing Industry simulation.

The impact of the occurrence of an innovation or policy event on the value of a cost/performance trend is multiplicative in nature. The results of the opinion survey produced the structure shown in Figure 6.

RESULTS OF THE SIMULATIONS

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

<table>
<thead>
<tr>
<th>Type</th>
<th>Service</th>
<th>Nominal Scenario</th>
<th>Stochastic Scenario</th>
</tr>
</thead>
<tbody>
<tr>
<td>Centralized</td>
<td>24</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>Decentralized</td>
<td>16</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>NIS</td>
<td>5</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>Dedicated Network</td>
<td>10</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>Total DP Industry</td>
<td>$55</td>
<td>$79</td>
<td></td>
</tr>
</tbody>
</table>

The introduction of the innovation occurrences significantly expanded the market while changing its structure. The detailed results of the simulation forecasts are shown in Figures 7, 8, and 9. These charts include stochastic simulation results, nominal simulation results, calibration forecasts, and the IDC fortune forecast. The charts indicate that IDC appears to include some level of new technological innovations impact in their five-year and over forecasts.

The Primary Industry Segment Forecast Chart illustrates an unexpected trend resulting from the cross-impact patterns. By 1990, the growth of the data communication-oriented industry segments can be expected to reduce centralized and decentralized stochastic forecasts to the same level or less than the nominal forecasts.

<table>
<thead>
<tr>
<th>Type</th>
<th>1990 Statistical Results</th>
<th>1990 Alternate Results</th>
<th>Percent Reduction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total DP Industry</td>
<td>$214M</td>
<td>$194M</td>
<td>-9.3%</td>
</tr>
<tr>
<td>Networked Systems</td>
<td>100M</td>
<td>97M</td>
<td>-3.0%</td>
</tr>
<tr>
<td>Centralized</td>
<td>64M</td>
<td>56M</td>
<td>-12.5%</td>
</tr>
<tr>
<td>Decentralized</td>
<td>45M</td>
<td>41M</td>
<td>-8.9%</td>
</tr>
<tr>
<td>N.I.S. Segment</td>
<td>78M</td>
<td>68M</td>
<td>-12.8%</td>
</tr>
<tr>
<td>Dedicated Segment</td>
<td>28M</td>
<td>29M</td>
<td>+3.6%</td>
</tr>
</tbody>
</table>

The introduction of the solution significantly expanded the market while changing its structure. The detailed results of the simulation forecasts are shown in Figures 7, 8, and 9. These charts include stochastic simulation results, nominal simulation results, calibration forecasts, and the IDC fortune forecast. The charts indicate that IDC appears to include some level of new technological innovations impact in their five-year and over forecasts.
Figure 6—Event-on-trend impact structure.

TABLE III.—Total DP Industry Forecasts (in 1975 billions of dollars)

<table>
<thead>
<tr>
<th>Model Results</th>
<th>1980</th>
<th>1983</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$</td>
<td>ERROR</td>
</tr>
<tr>
<td>Nominal Scenario</td>
<td>50.0</td>
<td>1%</td>
</tr>
<tr>
<td>Stochastic Scenario</td>
<td>53.9</td>
<td>6%</td>
</tr>
<tr>
<td>Validation Data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Composite Forecasts</td>
<td>50.5</td>
<td>69.5 (11.2% growth rate)</td>
</tr>
<tr>
<td>IDC Forecasts</td>
<td>&gt;50</td>
<td>78 (15.6% growth rate)</td>
</tr>
</tbody>
</table>

From the collection of the Computer History Museum (www.computerhistory.org)
ALTERNATE FUTURES FORECASTS

Researchers and planners are often interested in a forecasted industry profile based on the occurrence of selected exogenous events. Such simulations use the impact coefficients of technical and policy event occurrences on cost/performance trends. Typical exogenous events include governmental policy interventions or union activism. An illustrative simulation was performed that used "White Collar Unionization Expansion" as an exogenous event.

The alternate simulation modified industry segment growth in Table II.

All segments and industries other than Dedicated Networked Systems were reduced by approximately 10 percent, losing almost one-third of their gain over the Nominal Simulation. Since the Dedicated Segment lost sales late in the simulation when stochastic event occurrence was introduced, the reduction in their impact due to the policy variable would be expected to give the gain shown in the simulation.

ACCURACY OF FORECASTS

The forecasts produced by simulation are compared to the independently obtained composite forecasts in Table III.

IMPLICATIONS OF DP INDUSTRY FORECAST

The results of the DP Industry forecast are summarized in Table IV and show the forecasted growth rates of each industry segment. These growth rates were derived from the constant dollar expenditure forecasts summarized in Table V.

The figures are based on the results of the stochastic simulation. The expenditure data has been adjusted to 1980 dollars from the 1975 dollar figures used in the prior calculations.
CURRENT DOLLARS
(000,000)

I
I
I
I
I

I
I
I
I
I

60 - DeJica teJ Services'

---
NIS

50
40
30
20
10
0

Figure 9—Data communication sub-industry segment forecasts.

Trend summary

The results shown in Tables IV and V indicate that the 1980's will be a period of rapid growth for the data processing industry with real growth running at least twice that of the total economy. The early 1980's will see rapid growth for decentralized minicomputer-based systems as well as for NIS services, while the late 1980's will see a continuation of the NIS growth but a significant slowing in the growth of decentralized systems.

While real dollar sales of the DP Industry will grow at 6 percent to 8 percent during the decade, its largest and oldest segment, centralized large-scale computer-based systems, will remain constant in size during that period. Approximately two-thirds of DP growth will be accounted for by NIS expansion with the remaining one-third accounted for by expanding decentralized systems.

These DP Industry forecasts, although surprise-free in the sense that no unusual trends developed, did generate some unexpected results and contributed to an understanding of the dynamics of DP Industry growth. The forecasts indicate that the growth of network services during the 1970's through dedicated networks will, during the 1980's, switch to shared network utilization via NIS-type services. The next decade, therefore, should see a rapid growth of Value Added Network carriers that offer shared data communication services, permitting electronic mail and data base access.

The forecast also unexpectedly showed a slowdown in the erosion of centralized data processing services. The 1970's were a decade of rapid transfer of centralized batch applications to decentralized and networked services. This erosion will slow, and centralized services will grow slightly during the 1980's as expected improvements in large-scale computer productivity make integrated management information systems practical for most firms.

Also unexpectedly, minicomputer-based decentralized services growth will slow to the industry average by the mid-1980's. This indicates that increasingly the minicomputer vendors must, within the next several years, look to network-oriented products (not computational products) if they wish to maintain their rapid growth.

Impact on management

The primary impact of the continued expansion of decentralized information systems over the next five years will be the ability of small and satellite organizations to automate such day-to-day applications as: order entry, inventory con-

<table>
<thead>
<tr>
<th>TABLE IV.—Forecasted Growth Rates</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Annual Compounded Real Dollar Growth Rates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total DP Industry</td>
<td>6-1/2%</td>
</tr>
<tr>
<td>Centralized Services</td>
<td>2%</td>
</tr>
<tr>
<td>Decentralized Services</td>
<td>18%</td>
</tr>
<tr>
<td>Networked Services</td>
<td>8%</td>
</tr>
<tr>
<td>NIS</td>
<td>27%</td>
</tr>
<tr>
<td>Dedicated</td>
<td>0%</td>
</tr>
</tbody>
</table>

TABLE V.—Forecasted Constant Dollar Expenditures (in billions of 1980 dollars)

<table>
<thead>
<tr>
<th></th>
<th>1980</th>
<th>1985</th>
<th>1990</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total DP Industry</td>
<td>$54</td>
<td>$74</td>
<td>$109</td>
</tr>
<tr>
<td>Centralized Services</td>
<td>32</td>
<td>36</td>
<td>52</td>
</tr>
<tr>
<td>Decentralized Services</td>
<td>7</td>
<td>16</td>
<td>25</td>
</tr>
<tr>
<td>Networked Services</td>
<td>15</td>
<td>22</td>
<td>54</td>
</tr>
<tr>
<td>NIS</td>
<td>3</td>
<td>10</td>
<td>40</td>
</tr>
<tr>
<td>Dedicated</td>
<td>12</td>
<td>12</td>
<td>14</td>
</tr>
</tbody>
</table>
control, billing, and office automation. The productivity of the small firm or office will be improved and the current trend toward managerial decentralization should continue into the mid-1980's.

The primary impact of the rapid and continuing growth of NIS through the 1980's will be the growth of integrated systems linking vendors, suppliers, remote offices, and homes (Diebold, 1977). This trend should start to reverse the managerial decentralization trend of the mid-1980's as corporate management and staff link their decentralized systems through NIS. They will then demand compatibility and simultaneously achieve the capability for day-to-day monitoring and control of remote operations. In summary, during the 1980's, electronic communication and storage of information will be less expensive than paper-based systems, expanding the pace of white-collar automation. This direction is well summarized by Charles P. Lecht (1977, page 178).

While it may yet be possible to argue that if you destroyed all the computers in existence today it wouldn't seriously affect your life, it seems equally obvious that by the early 1980's we would be drowning in paperwork. Progress would necessarily retreat without the powerful computer systems and networks upon which a service-oriented society is—we would say inescapably—dependent.

The expanding computer/communication environment forecasted will, therefore, expand the impact of data processing on the economy. Management must plan for this expanding role of Information Automation, and assure its profitable use.

BIBLIOGRAPHY

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