THE NEW YORK TIMES NEWSPAPER PLANNING MODEL

Leonard Forman Ph.D.

Director Of Economic Research
The New York Times

ABSTRACT

The New York Times model is a large-scale model which forecasts sales and earnings for the N.Y. Times newspaper. Structurally, it is composed of two major blocks; a demand module and a production, cost, and revenue module. The demand module, the heart of the model, is a set of simultaneous non-linear econometric equations which forecast physical volume, approximately 35 categories of advertising lines and 10 categories of circulation. The second block is recursive and contains roughly 300 equations, some of which are stochastic behavioral equations. This block converts the volume forecasts into paging, newsprint consumption, newsprint distribution, and manning requirements. These physical flows are then monetized,using price and wage forecasts, to produce estimates of revenue, fixed and variable costs, and operating profit. This paper summarizes the development of the model, with emphasis on the advertising and circulation model , and provides some applications of the model's use. It should be noted that the structure of the model is constantly evolving. Consequently, emphasis is placed on the conceptual underpinnings of the model not on a detailed presentation of the current structure.

INTRODUCTION

The New York Times model has been developed as a planning tool to assist in evaluating the impact on of changes in in internal and external conditions on newspaper earnings. The modeling strategy followed at the Times differs from the typical corporate modeling effort. In contrast to most systems, which follow a "top-down" modeling strategy, the Times' approach is a "bottoms-up" strategy. There is no front-end corporate model. Instead, each division is modeled separatly. The output from each model is then aggregated to produce a set of consolidated corporate financial reports.

A second and perhaps more interesting departure is the attempt to fully integrate the financial, production, and marketing activities of each division into a truly integrated planning model. The major effort to achieve this has been at the New York Times newspaper. This division accounts for approximately one-half of corporate revenues. The Newspaper model is considerably larger and more complex than other division models.

The modeling approach at the newspaper is based on the premise that in the short run the state of the economy is the primary determinant of the demand for our product. Although other factors influence demand, we believe that the economy is the dominant factor. Accordingly, a significant amount of time has been spent in the construction of a marketing model, referred to as the demand model below. This model, fueled by forecasts of key economic indicators, is the driving force in the newspaper model.

In what follows, no attempt is made to describe the model in all its detail. Rather, emphasis is placed on the structure of the model and the methodology followed in constructing it. Section I gives a brief overview of the model. Section II describes its structure and the theory underlying that structure. The third section highlights the econometric methodology and discusses some of the economtric equations. Most of the effort is devoted to the demand model which contains almost all of the econometric equations in the planning model. The fifth section describes some of the model's applicatins in planning and budgeting

II AN OVERVIEW

The development of the model followed a pragmatic strategy: each sector was built according to need. Initially, the most pressing problem was to produce reliable forecasts of advertising and circulation. The well publicized economic problems of the N.Y. metropolitan region necessitated a a more rigourous approach to supplement the traditional sales forecasts produced by the sales staff. The decision was made to build an econometric model to forecast volume.

While normal econometric methodology was followed, the supreme test was its forecasting precision. If the advertising and circulation model was to be an operational tool for improving planning and short term budgeting, forecast accuracy was essential. The preliminary version of the demand model was completed in late 1974. Since then, the model has been used with a great deal of success for forecasting and simulation. It is an integral part of the formal planning process at the Times.

The second phase of the modeling project was the development of the cost and production model, which when linked to the demand model, provided a mathematical description of the newspaper's operations. The immediate problem was to eliminate many of the time consuming manual calculations performed by the various departments in preparing the plan. The manual procedures ruled out sophisticated planning since even a single change in the underlying assumptions, either changes in exogonous variables or changes in control variables, required a great deal of effort and time to work through. Naturally, two or more efforts were in many cases impossible to handle.

The first step was the restructuring of the production and financial relationships into a systematic and internally consistent set of mathematical equations. In this manner a working model was available, in a reasonably short time, for simulation experiments. Subsequent analysis revealed that many of the parametric assumptions and relationships were unsatisfactory and new relationships were developed. Some of these were stochastic .

Currently, the model contains over 300 equations. Approximately 65 are econometric equations. There are 30 exogonous variables. The model's specification produced a mixture of recursive and simultaneous linear and non-linear difference equations. It is, however, not a finished product. Apart from the normal management of an econometric model, we are constantly

example, many of the production and cost econometric equations will eventually be supplemented by a more sophisticated optimization analysis. The model is an evolving organism, undergoing continuous change and refinement.

Furthermore, although model development responded to particular problems, conceptually, a unified structure was conceived prior to model construction. In Jay Forrester's terminology, this structure represented a systems view of the New York embodying "the dynamic feedbacks and linkages between information, materials, orders, people, and equipment" for all the components of the newspaper. Our modeling effort is based on the rather ambitious goal of describing total system behavior: how decisions are made in various parts of the organization and how they contribute to the total behavior of the system.

II THE ELEMENTS OF A NEWSPAPER MODEL

Before describing the structure of the model a description of some of the salient characteristics of the newspaper industry, with the emphasis on their relevance for analysis rather than on comparisons with other industries, is useful.
A newspaper really produces two different products and sells them in to different markets. It sells copies to readers and space to advertisers. By itself, there is nothing unusual in this. However, for newspapers, this joint product presents unusual marketing problems. The demand for advertising space is greatly affected by the sale of copies. There is also evidence of reverse causality although the relationship is weaker. A newspaper becomes a more attractive vehicle for advertisers the larger its circulation base. Similarly, many people purchase a paper for its advertising content which provides readers with an information alternative to news content. Furthermore, it is even more unusual to find that the demand for sapce is affected by the character of the newspaper's readers. In this case, class does matter.

Conceptually, the model is based on the micro-economic theory of demand and cost modified by the peculiarities of the newspaper industry and internal financial requirements. The production function is vector-valued with output defined as the number of copies produced and the number of pages per copy.

$$Q_t = \begin{bmatrix} C_t \\ P_t \end{bmatrix} = F_1(K_t, L_t, TC_t) \qquad (1)$$

where
- Q_t is output
- C_t is copies produced
- P_t is pages per copy
- K_t is capital stock
- L_t is labor
- TC_t is technological change

The cost equations derived from this production function differ from the traditional microeconomic approach. In order to insure the model's acceptance and use the cost equations were designed to be consistent with the financial and accounting conventions at the Times. In these terms, the model can be viewed as a rather large and sophisticated income statement or operations model. Consequently, capital stock appears only indirectly through the depreciation equations. Thus, while the fixed-variable cost specification is retained capital stock costs do not explicitly appear.

Variable costs are defined as those costs which vary directly with paging and copies produced; newsprint and production labor costs. The number of men working and the number of hours worked vary each day depend on paging size and copies produced. Although capital is held fixed, capital utilization is not. Theoretically, depreciation is a function of capital utilization and should be treated as a variable cost. Because accounting conventions do not handle depreciation in this fashion, in the model, depreciation does not vary with output.

Newsprint costs are computed by calculating the newsprint tonnage required to produce a given number of pages and copies and then multiplying by the appropriate cost per ton. There are equations for different grades of newsprint as well as equations for the many different sections of the paper.

Econometric cost equations are used to determine variable labor costs for each of the variable labor departments. While the equations differ, in general they look as follows:

$$LC_{it} = F_2(C_t, P_t, TC_t, W_{it}) \qquad (2)$$

where LC_{it} is unit labor cost for the i'th group in period t

C_t, P_t, TC_t are as defined above

W_{it} is the appropriate wage rate

Changes in production technology are rapidly making obsolete historical production and cost relationships. Eventually, these econometric cost equations, as mentioned earlier, will be replaced by an optimization analysis. For example, a given volume forecast of pages and copies can be produced in a variety of press configurations, each press configuration completely determines labor requirements. Programming techniques will be used to determine the least-cost press configuration. The labor associated with this press configuration will then be monetized at the appropriate wage rates.

To complete the cost side, non-capital fixed costs are divided into labor and non-labor fixed costs. Because the accounting system does not track these costs historically econometric methodology is not used. Fixed labor costs are determined by equations which compute manning requirements and wage rates in 32 different categories.

In the current version of the model, manning requirements are determined exogonously. They are policy variables controlled by management. Research is currently underway to determine the linkages between fixed labor, output, and profitability; and formalize the decision rules management intuitively follows in determining annual fixed manning levels. The objective is to fully endogenize these linkages.

Non-labor fixed costs (excluding capital) are treated in a very simplistic way. Dollar cost for each category is increased each period by a growth factor. These costs account for a very small percentage of fixed costs.

The demand model is a simultaneous block of non-linear econometric equations for advertising lines and copies sold (circulation). Both outputs are sold in oligopolistic markets with varying product differentiation and intense non-price competition. Because the Sunday and daily papers are distinct products, Sunday and weekday advertising and circulation are treated separately. Conceptually, the demand for each output is treated as a function of its price, the prices of its competitors, indicators of market activity, and quality. Symbolically, the demand for circulation looks like the following:

$$CIRC_{ti} = F_3(p_i, p_{ji}, Prom, M_1 \ldots, M_n, Q) \qquad (3)$$

where $CIRC_{ti}$ is average circulation in period t for the i^{th} market

p_i is the unit price per copy in the i^{th} market.

p_{ji} is the price of the competitive product in the i^{th} market

Prom is promotion expenditures

M_i are indicators of market activity

Q is quality

The variable quality is a measure of the appeal of the paper to its readers. Potential readership depends on attitudes which in turn depend on education, household income, occupation, orientation to NYC, politics.

Similarly, the demand for advertising lines;

$$L_{ti} = F_4(p_i, p_{ji}, Prom, Q, M_1, \ldots, M_n) \qquad (4)$$

where all variables are defined as above and L_{ti} is advertising lines for the i^{th} category in period t. The variable Q represents the quality of a unit of linage to the advertiser as measured by the purchasing power of its readers, prestige of the newspaper, and readership Generally, circulation can be used as a proxy variable for advertising quality. The conceptual framework for the model is summarized in fig. 1. This flow diagram depicts the dollar flows between the various components of the system. The underlying physical flows are

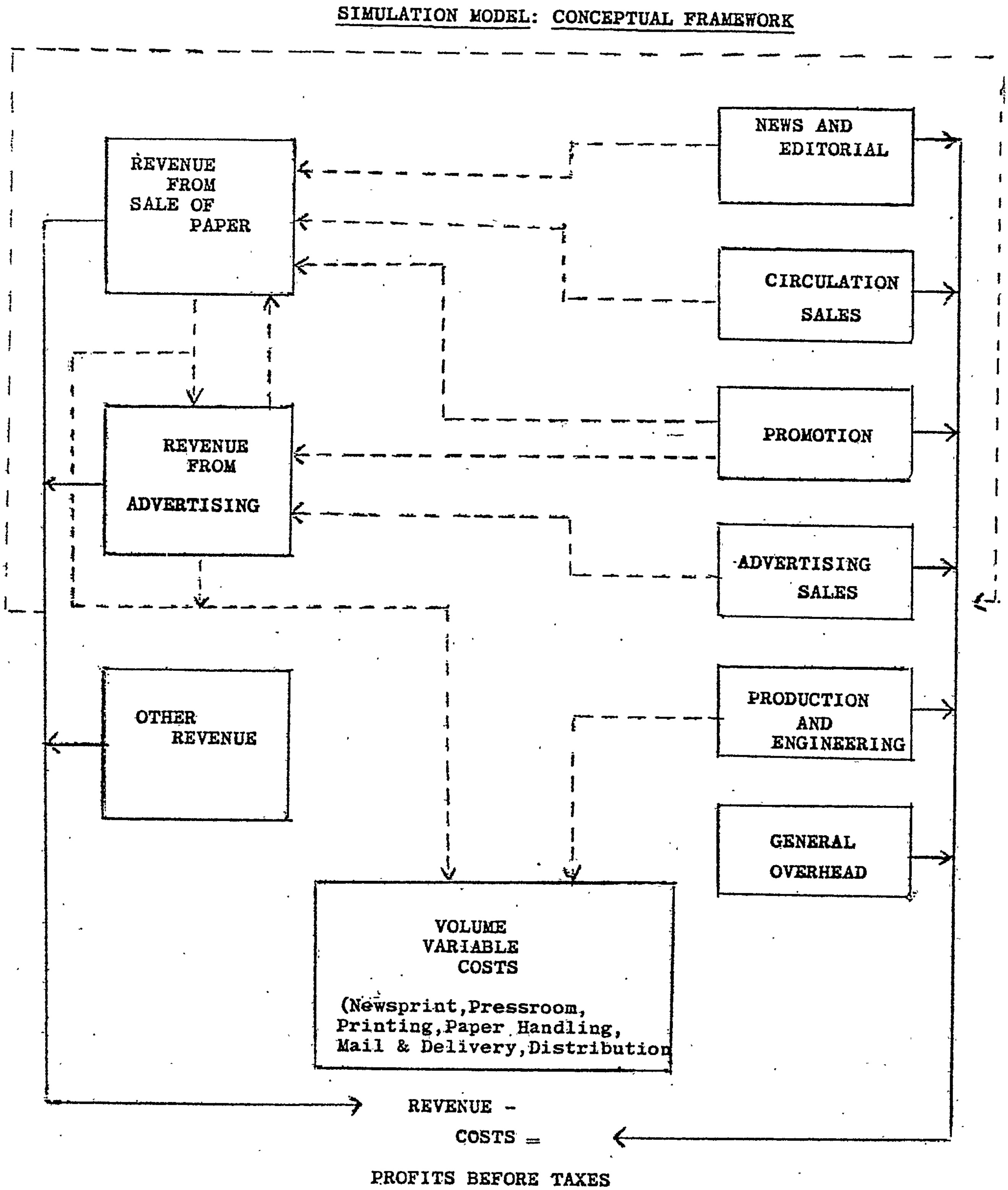

SIMULATION MODEL: CONCEPTUAL FRAMEWORK
REVENUE FROM SALE OF PAPER
REVENUE FROM ADVERTISING
OTHER REVENUE
NEWS AND EDITORIAL
CIRCULATION SALES
PROMOTION
ADVERTISING SALES
PRODUCTION AND ENGINEERING
GENERAL OVERHEAD
VOLUME VARIABLE COSTS
(Newsprint, Pressroom, Printing, Paper Handling, Mail & Delivery, Distribution
REVENUE -
COSTS =
PROFITS BEFORE TAXES

FIG. 1

NEW YORK TIMES MODEL STRUCTURE

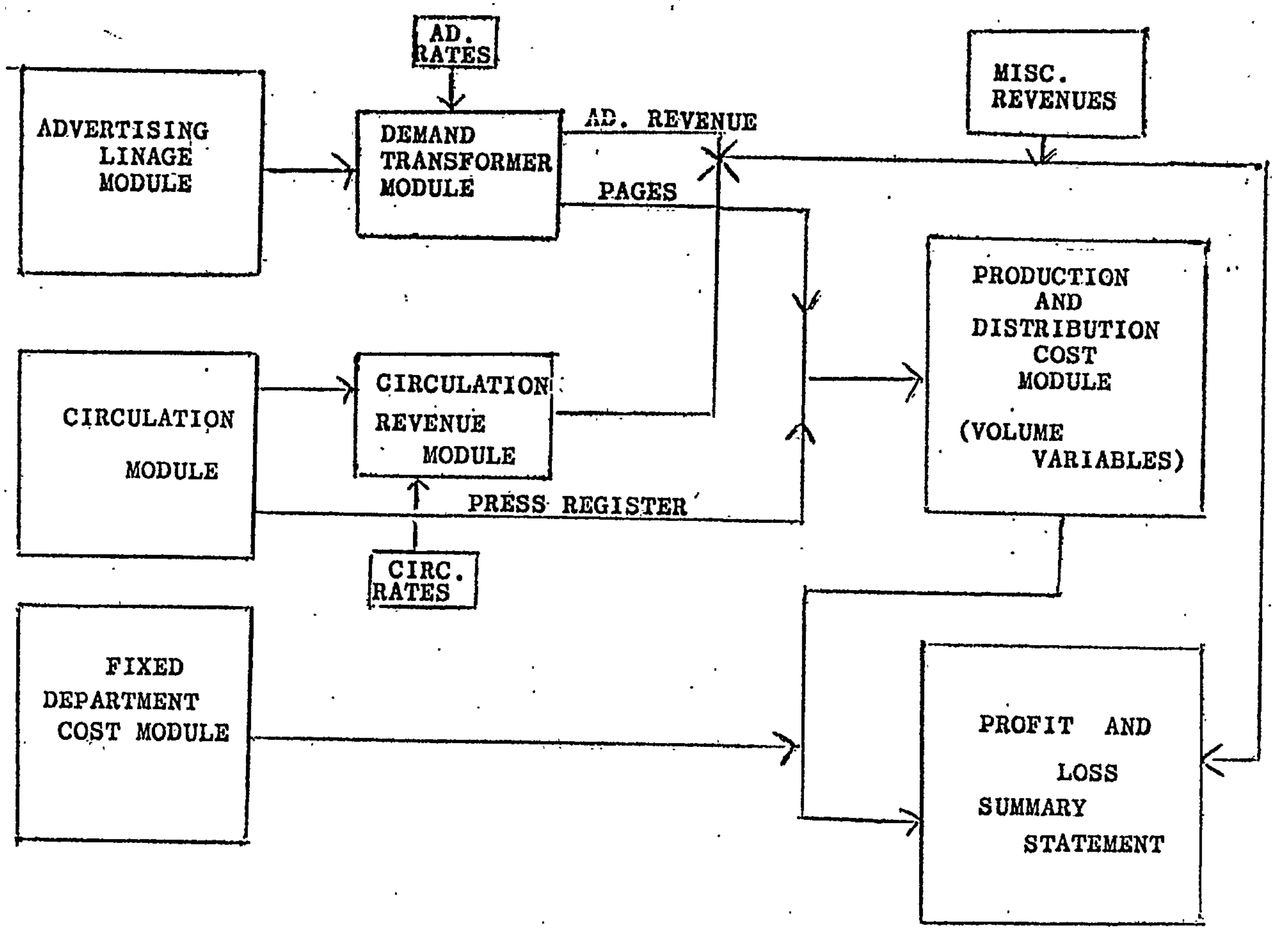
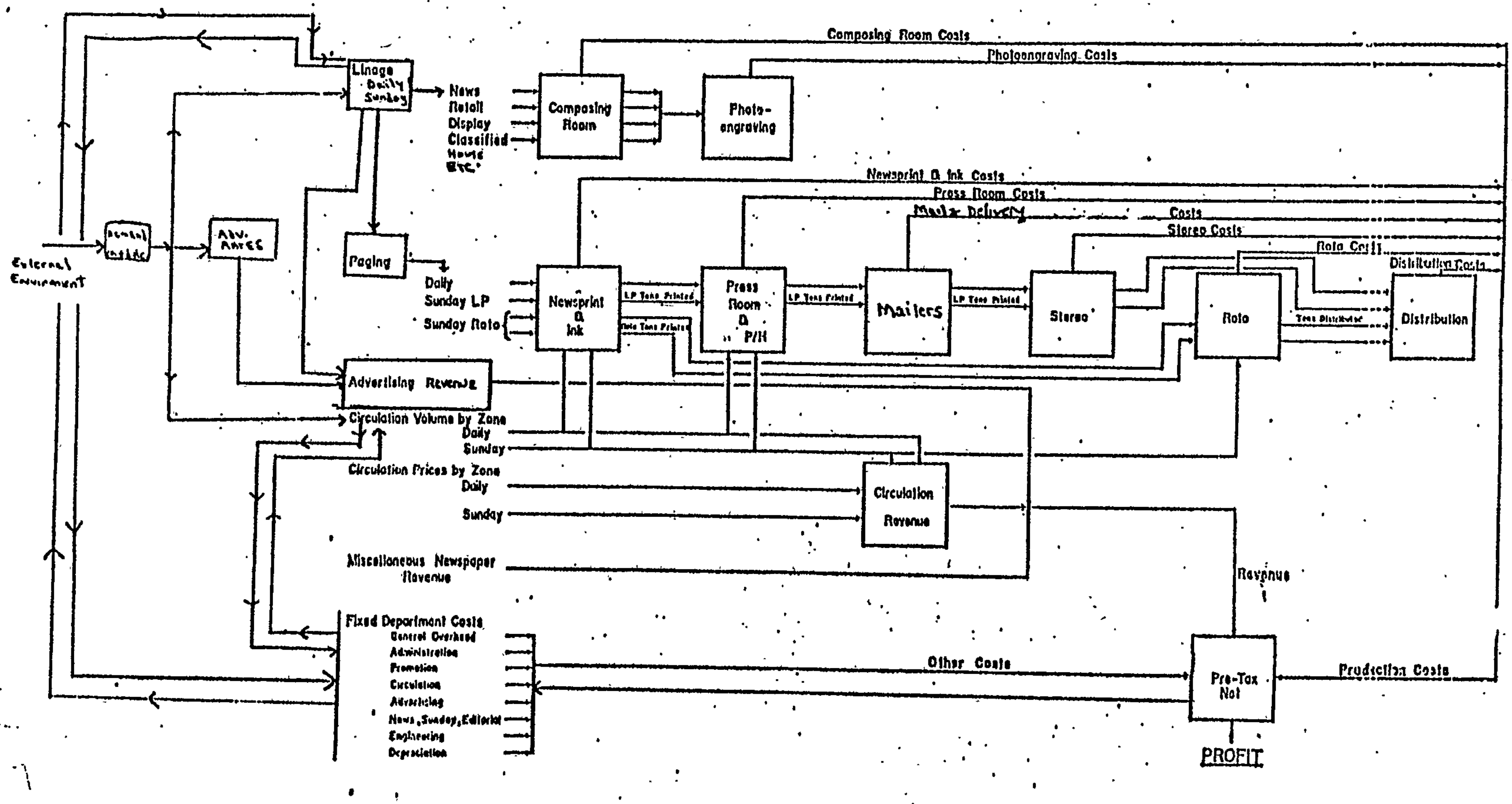


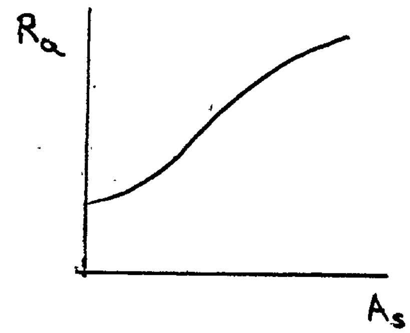
FIG. 2

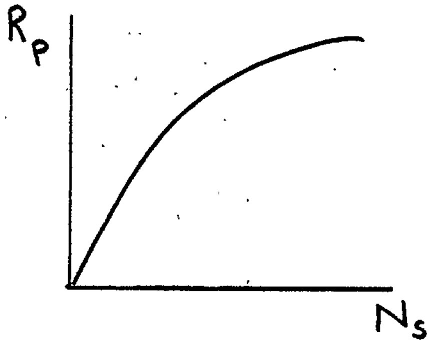
FIG. 3

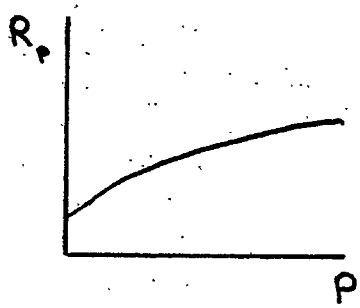
PRODUCTION - DISTRIBUTION COST MODULE

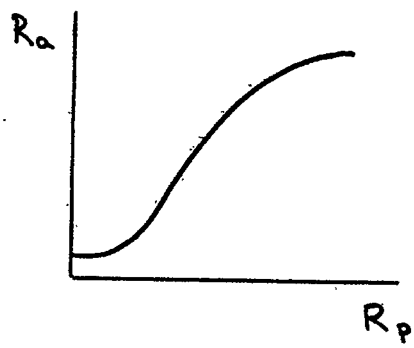


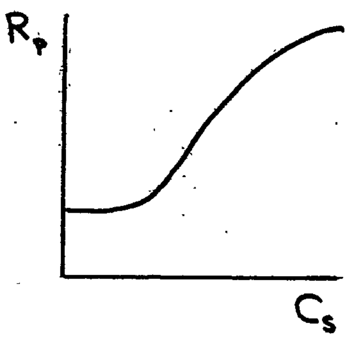
shown in figs. 2 and 3. The solid lines in fig. 1 represent dollar flows, revenues on the left and costs on the right. The dotted lines portray the interaction between these dollar flows with the arrows indicating the presumed direction of the interaction. For example, expenditures for promotion are assumed to influence advertising and circulation revenues. Conversely, advertising and circulation revenues impact the dollars available for promotion expenditures. The likely nature of some of these non-linear interactions are shown in the accompanying graphs.

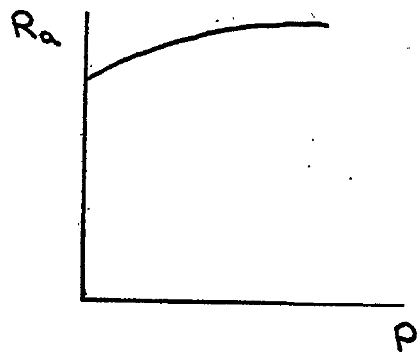












where R_p is circulation revenue
R_a is advertising revenue
P_c is production costs
N_s is news and editorial costs
C_s is circulation sales costs
P^s is promotion costs
A_s is advertising sales costs
E^s is engineering costs

Thus, real increases in news and editorial expenditures are likely to have a positive effect on the sale of newspapers although diminishing returns are probable. Similarly, if circulation sales expenditures were reduced to zero revenues from paper sales would probably drop but not to zero. The revenue loss would occur gradually not instantaneously. The flow diagrams and graphs are conceptual simplifications of complex time-dependent interrelationships.

None of the fixed department relationships are as yet built into the model. I suggested earlier that management has a rational belief concerning the nature of these interactions which they act on in determining manning and expenditure levels. As the graphs above show, these relationships can be defined mathematically and built into the model. Simulation experiments could then be performed to test the effects of various fixed department hypothesises and determine those relationships which cause simulated performance to agree with actual performance.

Shown in fig. 2 is modular structure which is a convenient basis for thinking about the model although this structure does not correspond exactly to the way the system is modeled. For purposes of discussion, each of the boxes can be defined as a module. The modules are:

1. Demand module. This represents forecasts of the demand for advertising by category. Either manual forecasts by the advertising sales force or projections from the model can be used.
2. Circulation module. This represents the manual circulation department forecasts or the projections generated by the model.
3. Overhead and fixed department module. The current exogonous manning forecasts are used to compute fixed labor costs. Non-labor fixed costs are computed in this module also.
4. Demand Transformer. Given advertising lines and rates by category, this module generates news lines, paging and advertising revenue.
5. Circulation revenue module. Forecasts of circulation by region are used to generate revenue and copies produced.
6. Miscellaneous revenues. Exogenous projections of other revenues.
7. Production and distribution module. Given total pages and copies produced, this module builds up the costs of printing and distributing the paper.

A more detailed flow diagram, which corresponds more closely to the mathematical structure of the model is shown in fig. 3. This diagram details the major physical flows of the model and the financial output which results. This financial information is summarized in an income statement for management.

III THE CIRCULATION AND ADVERTISING MODEL: SPECIFICATION AND ESTIMATION

The generalized specification of the advertising and circulation equations shown earlier briefly outline the theoretical structure of the demand model . The major departures from that structure are the absence of relative prices in the advertising equations and demographic variables in the circulation equations.

Preliminary analysis indicated that within the normal range of price change there is a high degree of price inelasticity. This finding is consistent with the economic theory of product differentiated markets. Competition between the Times and other newspapers and electronic media exists but it is based on factors other than price. I do not mean to suggest that price changes do not affect advertising sales. They do. However, price competition is minimal.

Unfortunately, data problems prevented us from including the competitive factor in the model. Ideally, market share type equations would be preferable. However, because of data problems their reliability would be suspect. In any event, changing economic conditions proved to be the major cause of cyclical swings in advertising lines.

In the circulation equations, the lack of meaningful quarterly time series demographic data precluded their inclusion as explanatory variables. There are plans to remedy this defect by linking cross-section demographic data with the available time series data

While the extent of specification bias caused by the exclusion of relevant variables cannot be precisely determined, the nature of the N.Y. market suggests that it is probably small. Over the last 10 years, population shifts have resulted in a gradual change in the spatial distribution of our circulation sales with suburban sales increasing relative to city sales. Since these changes have not been abrupt proxy variables, which mirror these changes, have been constructed and used in the circulation equations.

Another long-term problem, well documented and related indirectly to demographics, is the nation-wide decline in newspaper readership. The N.Y times is not unaffected by this trend. A number of theories purport to explain this phenomenon. It is however, not a cyclical problem. Consequently, time trends were used to capture this process of erosion.

The quarterly fluctuations in circulation, apart from seasonal and random variation, seem to be explaind by price dynamics and economic changes. Circulation in each category is explained by economic indicators, advertising lines,prices and circulation from other categories. There is a good deal of substitution and complementarity between the various categories of circulation. For example, people who buy the daily paper tend to also purchase the Sunday paper. A weaker reverse relationship also holds. There is also some substitution between home delivery sales and newsstand sales. The equations were designed to take these relationships into account.

Advertising in each of the equations is related to various economic indicators, circulation, and the substitute-complement relationships which hold for advertising.

The parameters of the demand model were estimated according to the following methodology:

1. Preliminary analysis indicated those variables which explained quarterly fluctuations in volume. Data problems and multicollinearity eliminated some of the variables which belong in the equations.
2. Next, the appropriate functional form for each of the equations was determined. The usual summary statistics were used in exploring various specifications. In latter versions, multiple time series techniques were used to determine dynamic lag structures.
3. Each equation was then estimated by ordinary least squares. In many cases, serial correlation was present and the standard procedures were used to correct it. Unfortunately,in dynamic simulations this correction magnifies specification biases producing explosive out-of-sample forecasts. The usual approach,variable adjustments to the constant term with these adjustments diminishing over time was applied. I am currently trying another technique which allows for more complex stochastic specification then the typical first-order auto-correlation adjustment. It involves applying time series techniques to the residuals from each equation. The forecast for each variable then consists of a deterministic part and a stochastic part. The use of time series techniques to determine stochastic and dynamic specification is a generalization of the transfer-model approach outlined by Box and Jenkins and studied by Zellner and Palm in the context of simultaneous econometric equation models.
4. Finally, simultaneous estimation tecniques were applied and the estimated model was subjected to a number of validation tests.

IV THE INTEGRATION OF THE MODEL INTO THE PLANNING AND BUDGETING CYCLE: SOME APPLICATIONS

The model outlined in the preceeding pages is fully integrated within the budgeting and planning cycle. Several examples will perhaps illustrate this.

A. THE ANNUAL BUDGET

In the fourth quarter, the newspaper prepares its annual budget for the coming year. The most important components of the budget are the volume forecasts of advertising and circulation. The advertising and circulation sales departments spend several weeks preparing preliminary estimates Independently, the model is used to produce several forecasts corresponding to a most likely scenario, a more optimistic scenario, and a pessimistic scenario.

Management then reviews each set of forecasts assesing their strenghts, weaknesses, and reliablility. An analysis of the differences between the model's forecasts and the sales forecasts begins the process of producing a final set of volume forecasts in which the major differences have been reconciled. These forecasts are then distributed to the managers of the operating departments and the budgeting process continues. Between the begining and end of the budget cycle, a period of about four months, the demand model is periodically run to determine whether or not there are any departures from the original forecasts.

The procedure just outlined was followed in bothe the 1975 and 1976 budget cycle. model's 1975 forecasts,produced in December 1974, were quite accurate. The forecast error for total linage was under one percent and the circulation error slightly over one percent. Individual equation errors were larger but cancelled out when aggregated.

At the end of the cycle, the full model is run and the budget is reproduced. A set of sensitivity experiments is then performed to determine the impact of of volume changes and policy options in response to those changes. A set of income statements summarizes these simulations. As a result of these exercises, contingency planning is greatly facilitated.

B. THE FIVE YEAR PLAN

The planning and budget cycles at the Times overlap: the annual budget is the first year of the five year plan. The planning cycle begins in March and ends in the early part of the fourth quarter when the budget cycle begins.

In March, a complete run of the model is produced. The volume forecasts and income statement summary are compared to the previous year's plan and the current operating goals set by management. The results of this analysis are distributed to the operating managers for comment and analysis. The advertising and circulation departments, using the model's volume forecasts as a guide, then prepare their own forecasts. This process eventually results in a set of consensus volume projections.

With this new set of volume numbers, the production-cost model is re-run and a preliminary five year income statement is produced. If at this time the financial results still fall short of the plan's objectives a set of action plans, strategies designed to improve performance, are developed and simulation experiments are performed, to determine their impact on profits. This process continues until the five year plan is finalized.

Note the pivotal role that the model has in the planning process. Indeed, in the absence of a model the planning process, in the true sense of the word, does not really exist. The manual procedures followed by the various departments prior to the model's development limited the process to essentially one pass. The model has given management the capability to explore alternative plans quickly and react promply to changes in external conditions which result in a departure from the established plan.

C. NEW PRODUCTS

The cost and revenue implications of current and planned product changes have been analyzed using the simulation model. For example, responding to shifting population, special regional editions are being produced. In the 1976 plan the model was used to determine profit and loss statements for each of the new products. Following a base run, each new product change was introduced separately. In this manner the incremental costs and revenues for each product were generated. From this information a break-even analysis could be prepared and an appropriate pricing strategy developed. This analysis provided the information management needed to determine the economic feasibility of each product, information which would not have been available had manual procedures been followed.

V CONCLUSIONS

The preceding examples document the integration of newspaper model into the planning process at the New York Times. The model is used to set goals, design strategies to achieve these objectives, forecast, and monitor changes which might prevent those goals from being reached. Management uses the model and believes its output.

The model, however, is far from finished. Some of the changes, which will eventually improve the model, were suggested in the preceding pages. New information and changing technology require constant modification of the model's structure. It is a living organism, constantly growing and changing. In this way we hope to insure its continued usefulness to management.

BIBLIOGRAPHY

1. Box,G . and Jenkins,G.., Time Series Analysis: Forecasting and Control, Holden-Day, 1970
2. Zellner,A. and Palm, F., "Time Series Analysis and Simultaneous Equation Models", Journal Of Econometrics. Vol.2, May 1974, p 17-54