THE DRESSER PLANNING MODEL

Charles H. Hatfield, Jr.
Bryant K. Kershaw

ABSTRACT

The Dresser Planning Model is a large scale econometric model to forecast sales and earnings for a domestic Division of Dresser Industries. The model uses demand and cost functions. Solution is a mixture of recursive and simultaneous equations.

SECTION I.0

INTRODUCTION

The Dresser Planning Model is a large scale econometric annual forecasting model of a major operating division of Dresser Industries. The structure follows the neo-classical economic theory of an oligopolistic firm with product differentiation. The model contains 342 behavioral equations, 638 identities and 151 exogenous variables. The specification of the model resulted in a mixture of recursive and simultaneous linear and nonlinear difference equations. The model is structured as a series of demand and cost functions. Constant dollar sales are forecast by the demand equations. An income statement is generated from the cost functions. The cost functions are derived from production functions which use output (sales in 1967$) from the demand equations. The solution is done in 1967$ and inflated to current dollars. This approach allows Dresser to track factors of production and their cost even though the model does not have a supply side.

The approach was pragmatic and focused on specific short- (one year) and long-range (five year) marketing and financial problems of Dresser Industries. The model is an <u>operational</u> tool for improving the accuracy and quality of short- and long-term profit planning. The model definition was begun May 24, 1973, the data collected, equations estimated, computer program built and installed at the Division on September 26, 1974; one day ahead of schedule. The model was subsequently updated and re-estimated and a more efficient computer solution program was designed in 1976. Currently (late 1978) the model is being re-estimated and re-orientated to match the changing strategic thrust of the Division's business.

The development was strictly an applied econometric exercise. Our foremost criteria was the logical validity of the structure of the model as viewed from the marketing viewpoint. The second criteria was the statistical and theoretical validity as assessed in the normal manner. Third was potential forecast accuracy. Section 2.0 contains the approach used to develop the model structure. Section 3.0 investigates the demand functions. Section 4.0 investigates the cost functions. Section 5.0 covers the econometric analysis. The conclusions are in Section 6.0.

SECTION 2.0

APPROACH

The major objective of the Dresser Planning Model was to forecast sales and earnings for a domestic division of Dresser Industries. Demand and cost functions derived from micro-economic theory were used as the basis of the model. The resulting model had to meet three objectives. The model was to improve the accuracy and information content of the annual and long-term profit plans. Second, it was to be a tool for management to evaluate alternative market strategies through simulation of the division. Third, it was to be a tool for evaluating the Division profit plans at the corporate level.

Production was divided into fifteen product categories. A product category is defined as a homogenous group of products from a marketing viewpoint. (This happens to coincide with a production definition). For purposes of this paper, these products are identified as A, B, C,.....O. This represented the production from three plants. All the products are basic industrial goods that go into construction and maintenance of plant, equipment and pipelines.

The product categories were further divided into three markets; called X, Y, and Z. That is, Division output was divided into forty-five categories based on a market definition. Thus, the errors of aggregation for a demand relationship were minimized. The solution was done in 1967$.

The financial data consists of an income statement for each product category. The items in the income statement were considered to be costs of two types of actions. The first set of items are things management does to increase sales or reduce cost. These are market effort, engineering development, engineering support of sales and average net investment level. In the model, this data is deflated so that it measures a level of effort. These items are exogenously determined and are used in the demand and production functions. The first set of indicators are called management decision variables. Management interacts through the management decision variables to influence demand and cost.

The second set of income statement items are considered to be factors of production. This includes direct labor and variable overhead, direct materials, manufacturing capacity cost, and administrative and general cost. The first two are obvious. Manufacturing capacity costs include depreciation, rental, maintenance, various taxes, etc., applicable to the plant, equipment and manufacturing function. It is used to measure the flow of capital services. The majority of administrative and general cost support the production process. A & G is somewhat arbitrarily called production inputs. Admittedly, some of this cost supports the market effort. As a consequence, market variables (marketing or engineering expense) often appear in the A & G equations. These production variables are plant oriented. Therefore, most are in simultaneous sets consisting of each plant.

The data collection task was tremendous. Data on all the variables mentioned were collected annually and, for the first model, quarterly from 1955 to 1972. Due to acquisitions, inconsistent product line data, and lack of data the actual start dates for the various series ran from 1955 to 1963. We literally had people opening boxes in the warehouse to put the data together. Quarterly forecasts were never more than spreads by seasonal adjustment factors. All quarterly forecasts were dropped with the 1966 update.

The model was linked to the New Wharton Annual and Industry Forecasting Model developed at Wharton EFA. The objective was to choose variables forecasted in this model to measure Dresser market forces. These variables were used three ways. First as actual indicators. That is, the level of new housing investment is a direct indicator of how much product A goes through market X to be used in new housing construction. Second, economic variables are used as proxies for things for which we have no data and/or were not forecasted. Third, they can serve as indicators of economic conditions. For example, durable goods output is a good overall indicator of the level of economic activity in all investment goods industries. Keep in mind that all variables discussed are in 1967$.

The next step was to review each product market with experienced marketing personnel. They gave Operations Research a review of each product-market category; how the products are used, when they are used, and who uses them. The potential economic indicators and management decision variables were chosen. The statistical analysis was then begun. The equations were again reviewed after estimation.

SECTION 3.0

DEMAND ANALYSTS

The products contained in the Dresser Planning Model are marketed in an industry best described as oligopolistic with product differentiation. The degree of product differentiation varies from practially none to very strong.

The price elasticities are very low in most cases. The sales are primarily to utilities. The price elasticity is lowest in those products-markets where product differentiation is highest. Prices do not even appear in many of the demand functions. (This is also due to the multicollinearity problem and the available proxy for industry price).

The demand function would be stated as:

$$Q_i = D(P_i, M_1, \ldots, M_k, E_1, \ldots, E_n, Q_j, \ldots, Q_{j-1}, Q_{j+1}, \ldots Q_j) \tag{3-1}$$

Where: Q_i = quantity demanded of the ith product-market items.

P_i = relative price of ith product-market to fabricated metal prices.

M = the management decision variables.

E = economic indicator of market activity.

Q_j = quantity demanded of complementary products sold by Dresser in the same market where $j \neq i$.

The demand functions are developed according to this hypothesis given in equation 3-1. The general choice of economic and management decision variables are based on Division market personnel's description of the markets.

3.1 Demand Equations

The larger products in markets with stable histories give the best results as expected. Table 3-1 shows a representative equation for a major product-market category.

TABLE 3-1

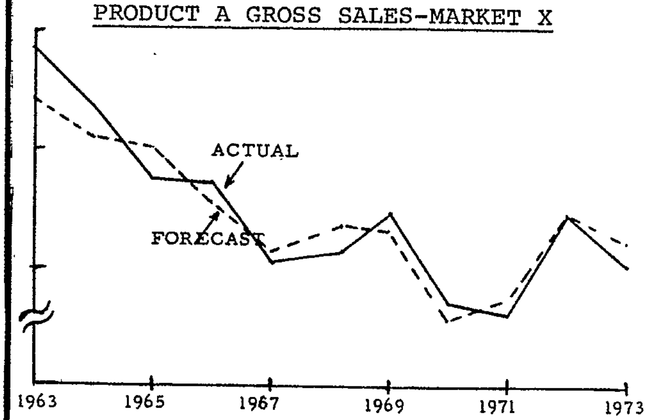
PRODUCT A MARKET X GROSS SALES (1967$)

R^2 = 93.809% σ = 3.582%

$$\frac{AXGS}{AXPI} = 2.63152e^{-.1436155*TIME} NF^{.5336525} GPO29^{.356440} HOS^{1.975547}$$

AXPI =	Product A Market X Price Index	1967 = 1.0
TIME =	Time Trend	
NF =	Investment, Structures, Nonfarm Residential	1958$
GPO29 =	Output, Fabricated Metal Products	1958$
HOS =	Personal Consumption, Household Operating Services	

PRODUCT A GROSS SALES-MARKET X



SECTION 4.0

COST FUNCTIONS

Cost functions are used to determine the factors of production. The term cost function is used to denote cost expressed as a function of output. This is opposed to a cost equation which denotes cost in terms of input levels and input prices. The cost function is used to determine input levels in constant dollars. A cost equation is then used to express cost in current dollars. Each product production function is expressed as:

(4-1) $Q_i = P(VLO_i, DM_i, MF_i, AG_i, INV_i, TECH_i)$

Where: Q_i = output quantity of the ith product.

VLO_i = direct labor required for Q_i for ith product.

DM_i = direct material required for Q_i for ith product.

MF_i = manufacturing cost required for Q_i for ith product.

AG_i = administrative and general cost required for Q_i for ith product.

INV_i = total investment (capital stock) available for use in production of Q_i for ith product.

$TECH_i$ = technological change in production process of the ith product.

All values in 1967$.

Each product was produced in only one plant. This avoided having more than one production function for any of the products.

It is assumed that management will always expand and contract according to an expansion path. The expansion path is determined by the plant's technological constraints. The expansion path is defined by:

(4-2) $Q_i = E(VLO_i, DM_i, MF_i, AG_i, INV_i, TECH_i)$

Where: The variables are defined as for the production function.

Since all outputs and inputs are defined in constant 1967$, these two equations can be reduced to a single cost equation in 1967$.

(4-3) $C_i = G(Q_i, INV_i, TECH_i)$

Where: C_i = 1967$ vector of production cost items for the ith product.

The cost equation is stated as:

$$(4\text{-}4)\quad C_i^* = VLO_i{*}PI_{vlo} + DM_i{*}PI_{vlo} + MF_1{*}PI_{mf} + AG_i{*}PI_{ag}$$

Where: C_i^* = current \$ cost of the ith product.

PI_{vlo} = price index of direct labor 1967 = 1.0.

PI_{dm} = price index of direct material 1967 = 1.0.

PI_{mf} = price index of manufacturing cost 1967 = 1.0.

PI_{ag} = price index of administrative and general cost 1967 = 1.0.

The specification of equation 4-3 is satisfactory for variable cost categories of labor (VLO) and material (DM). The VLO and DM cost functions were first solved, printed out in constant dollars, multiplied by price and printed out again. This gives management the ability to track real labor and material requirements. The specifications of 4-3 and 4-4 are not satisfactory for manufacturing cost (MF) and administrative and general cost (AG).

The accounting system just doesn't track AG and MF costs as well as it does labor and material. Manufacturing costs and administrative and general costs are allocated costs. This makes the categories sensitive to levels of production of other products in a given plant. MF contains fixed costs that are functions of plant size and are period costs. Other areas vary with levels of production. AG contains period costs as does MF. In addition, it has costs that are related to the market and engineering functions. All these elements must be taken into consideration in equation development.

Individual input cost equations for each product can be derived from the product cost equation (with the modification for MF and AG).

$$(4\text{-}5)\quad VLO_i = C_{vlo}(Q_i,\ INV_i,\ TECH_i)$$

$$(4\text{-}6)\quad DM_i = C_{dm}(Q_i,\ INV_i,\ TECH_i)$$

$$(4\text{-}7)\quad MF_i = PI_{mf}\left[C_{mf}(Q_i,\ INV_i 1, (MF_1, \ldots MF_{i-1},\ MF_{i+1}, \ldots,\ MF_k))\right]$$

$$(4\text{-}8)\quad AG_i = PI_{ag}\left[C_{ag}(Q_i,\ TME_i,\ ES_i,\ ED_i,\ (AG_1, \ldots, AG_{i-1},\ AG_{i+1}, \ldots,\ AG_k))\right]$$

Where: VLO_i = direct labor required for Q_i for ith product.

C_i = cost functions where i stands for VLO, DM, MF, and AG.

Q_i = output quantity = quantity demanded of product i

DM_i = direct material required for Q_i for ith product.

MF_i = manufacturing cost required for Q_i for ith product.

AG_i = administrative and general cost required for Q_i for ith product.

INV_i = total investment (capital stock) available for use in production of Q_i for ith product.

TECH = technological change in the production process of the ith product.

k = number of products produced in a given plant.

PI_{mf} = price index of manufacturing cost 1967 = 1.0.

PI_{ag} = price index of administrative and general cost 1967 = 1.0.

TME_i = total market expense of the ith product.

ES_i = engineering support of the ith product.

ED_i = engineering development of the ith product.

In the case where two or more products share facilities, their production cost becomes interdependent and this must be recognized. This was not a consideration in the demand equations. The output or sales would exhibit interdependence only near capacity levels.

The cost functions are long-run functions constructed by specifying short-run functions, then shifting them through changes in the INV or capital stock variable. Representative equations are shown in Tables 4-1 through 3.

TABLE 4-1

PRODUCT A DIRECT LABOR AND VARIABLE OVERHEAD (Current $)

$R^2 = 92.15006\%$ $\quad\sigma = 2.543237\%$

$$AVLO = (2914.81+0.1704761*(ANS/API)+103.4059*(W13*3.673)-315.4397*\ln(AINV))*W13$$

ANS = Product A Net Sales

API = Product A Price Index 1967 = 1.0

W13 = Wage Rate, Fabricated Metal Products 1967 = 1.0

AINV = Product A Average Investment

PRODUCT A DIRECT LABOR AND VARIABLE OVERHEAD

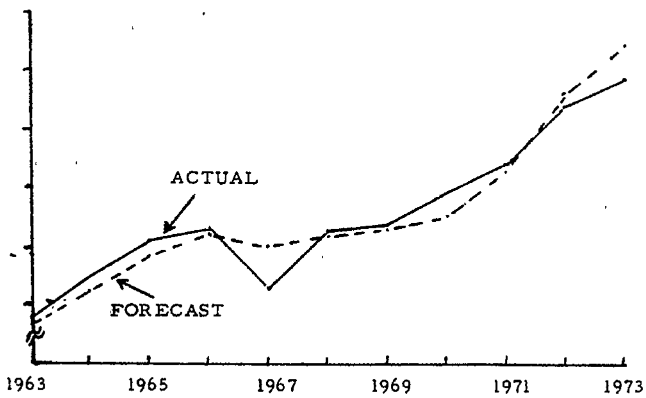


TABLE 4-2

PRODUCT C ADMINISTRATIVE AND GENERAL (Current $)

$R^2 = 97.97633\%$ $\quad\sigma = 6.068488\%$

$$CAG = -28.06905+0.01728005*CNS+0.3923547*CES+.08603203*(AAG+BAG+DAG+EAG)$$

CNS = Product C Net Sales

CES = Product C Engineering Support

AAG = Product A Administrative and General

BAG = Product B Administrative and General

DAG = Product D Administrative and General

EAG = Product E Administrative and General

PRODUCT C ADMINISTRATIVE AND GENERAL

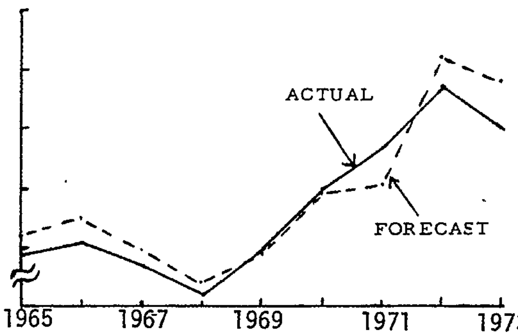


TABLE 4-3

PRODUCT J MANUFACTURING CAPACITY COST (Current $)

$R^2 = 97.31861\%$ $\quad\sigma = 4.008326\%$

$$JMF = 482.0115+JNS*.0249047*JNS - 20.62074*TIME*0.5663776*(KMF+LMF+MMF+NMF)$$

JNS = Product J Net Sales

TIME = Time Trend

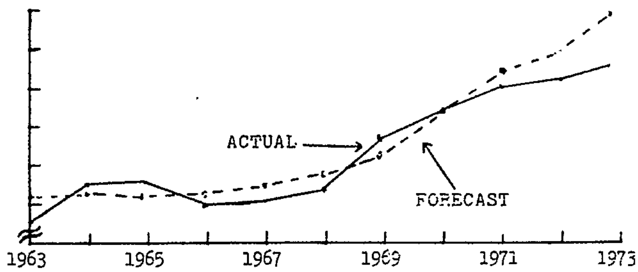
KMF = Product K Manufacturing

LMF = Product L Manufacturing

MMF = Product M Manufacturing

NMF = Product N Manufacturing

PRODUCT J MANUFACTURING CAPACITY COST



SECTION 5.0

ECONOMETRIC ANALYSIS

The estimation objective was purely a pragmatic one. No apprehension was felt against adding a lot of "art" to the equation estimation and specification. In deciding between alternative specifications; three factors were always considered: (1) the validity of the logic of cause and effect implied; (2) the values of the elasticities; and (3) the equation statistics, such as the t statistics of coefficients and Durbin-Watson.

The number of observations available was a definite limiting factor on specifications. The data began between 1955 and 1963, depending upon the series, and went through 1973 for the first model. This was more of a constraint in the demand functions than in the cost function.

The functional form of the vast majority of the equations is nonlinear in variables. Most of these are double Ln functions. None are nonlinear in coefficients. The model is primarily estimated from stochastic regressors. These arise from the methods of definitions and collections of accounting data, methods of definition and collection of national income data and the model structure.

Multicollinearity dominated the econometric analysis. Sample size was the second most important factor. The model structure was defined to minimize both effects through minimizing the number of variables necessary in each equation. Each product category and market was distinct and homogenous. The products would have only one or two general applications in their markets. They would either be repair products, consummable and/or investment items for new plant, equipment or construction.

The market could then be described with only three or four economic indicators and management decision variables. The management decision variables (market expense and engineering support of sales) were highly correlated for those products where engineering support was important. This was resolved by treating them as one variable.

Autocorrelation was not a significant factor. The Durbin-Watson statistics were watched closely. Autocorrelation did not cause serious problems, except in a very few cases. These were solved by using alternative specifications of the equations. The lack of a serious autocorrelation problem was attributed to three things: (1) the nature of the business, (2) the fact that it was an annual model, and (3) lack of lagged variables. Every product-market runs through a complete manufacturing and market cycle in one year. Orders are taken in late winter and early spring, manufactured and shipped through the spring, summer and fall. A complete new cycle is started each year.

"Errors in variables" (in an econometric sense) were both worked around and lived with. In the case where this was a problem, an instrumental variable was used. That is, another specification of the equation was used. In the case of manufacturing costs and administrative and general costs, the equations simply had to live with it. Through the estimation procedure, the value of elasticities were constantly evaluated. Alternative specifications would be tried when the coefficients acquired what we considered to be an unacceptable bias.

The stochastic effects due to the regressors being jointly determined along with regressand was evident in some cases. The structures of the model had minimized the problem in the demand function but not the cost function. Sections of the model were estimated with two-stage least squares and also the general linear model. The two-stage least squares advantage of consistency does not seem to help that much with our small sample sizes. It must be kept in mind that the coefficients used are manually scanned for gross bias. However, this is a point which will be watched very closely during the model update when more sample points become available.

In summary, the estimation process was determined by: (1) the logic of the cause and effect, (2) the value of the elasticities, and (3) the equation statistics such as the t statistics of the coefficients and Durbin-Watson. Using the structure developed in Section 2.0 and the statistical approach developed in this section, the equations needed to contain only three to five variables to achieve good specifications with respect to variables.

SECTION 6.0

CONCLUSIONS

A large simulation model cannot be built without the enthusiastic support of top management. The resources required to collect the data properly are too great. Second, such a model must answer questions about alternatives in which top management is interested. That is, questions about short-term and strategic planning. The model must also have the support of the department which uses it. They must devote the time and effort to understand the structure of equations and the solution program.

The model was built according to traditional microeconomic theory. Detailed explanations of market behavior obtained from market managers and the marketing services department were translated into these terms. Statistically this explanation fits the data very well. Historical simulations gave a mean square sales error of about five percent on all except relative new products that have not stabilized in a market.

Earnings before taxes (EBT) estimates are small numbers when compared to sales and major cost items. Any biases built into the simulation routine will create large EBT errors. Errors in EBT forecasts by product were about ten to twelve percent. Aggregation to the Division level reduced all errors considerably. Division sales errors of about two percent and Division EBT errors of seven percent are expected. This error analysis results from a fourteen-year simulation from 1960 to 1973.