University of Pennsylvania
ScholarlyCommons

Marketing Papers

1-1-1968

Long-Range Forecasting For International Markets: The Use of Causal Models

J. Scott Armstrong
University of Pennsylvania, armstrong@wharton.upenn.edu

Suggested Citation:
Armstrong, J.S. (1968). Long-Range Forecasting For International Markets: The Use of Causal Models. In *Marketing and the New Science of Planning.* by Robert L. King. American Marketing Association.

© 1968 American Marketing Association

This paper is posted at ScholarlyCommons. http://repository.upenn.edu/marketing_papers/172
For more information, please contact repository@pobox.upenn.edu.

Long-Range Forecasting For International Markets: The Use of Causal Models

Abstract
Many researchers appear to operate under the impression that causal models lead to more accurate forecasts than those provided by naive models (or "projections"). This study was based on the premise that causal models lead to better forecasts than do naive models in certain situations. The key element of these situations is that there are "large changes." One situation where large changes might be expected is that of long-range forecasting—and, in particular, long-range forecasting for international markets. Recent improvements in the quality and availability of international data have substantially reduced the cost of developing causal models in this situation. A study of camera markets in seventeen countries indicated that the margin of superiority of causal models over naive models is of great practical importance.

Disciplines
Business | Marketing

Comments
Suggested Citation:
Armstrong, J.S. (1968). Long-Range Forecasting For International Markets: The Use of Causal Models. In *Marketing and the New Science of Planning.* by Robert L. King. American Marketing Association.

© 1968 American Marketing Association

This book chapter is available at ScholarlyCommons: http://repository.upenn.edu/marketing_papers/172

Long-Range Forecasting For International Markets: The Use of Causal Models

J. SCOTT ARMSTRONG *

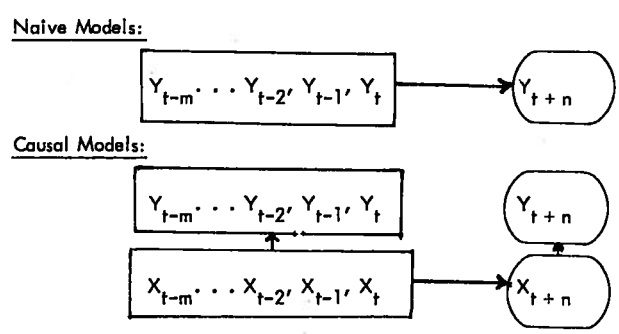
Many researchers appear to operate under the impression that causal models lead to more accurate forecasts than those provided by naive models (or "projections"). This study was based on the premise that causal models lead to better forecasts than do naive models in certain situations. The key element of these situations is that there are "large changes." One situation where large changes might be expected is that of long-range forecasting—and, in particular, long-range forecasting for international markets. Recent improvements in the quality and availability of international data have substantially reduced the cost of developing causal models in this situation. A study of camera markets in seventeen countries indicated that the margin of superiority of causal models over naive models is of great practical importance.

Models for sales forecasting may be divided into two categories—naive and causal. Native models attempt to forecast by using only historical sales data. Causal models [1] attempt to go beyond sales data to utilize other variables (*i.e.*, the causal variables) in making sales forecasts. Figure 1 provides illustrations of each type of model where "Y" represents sales of a given product, "X" represents the set of causal variables, and "t" represents the current year. The objective of each approach is to forecast sales in year t+n. The naive model does this directly. The causal model relates sales to other variables which, in essence, makes the problem one of forecasting "other variables" and using these forecasts to estimate sales.

These two approaches are actually extreme points of a continuum. Development of a forecasting model may borrow from both the naive and the causal approaches. Examples of the naive model approach would be use of simple trend projections or use of moving averages (such as exponential smoothing).

FIGURE 1

BASIC APPROACHES TO FORECASTING



Examples of the causal model approach would be those based on "econometric" models with no lagged sales variables as predictors. [2] In practice, these models are not so pure: the naive approach is often tempered by a subjective appraisal of underlying causal factors, while the causal model generally uses knowledge about current sales rates and attempts only to predict *changes* in sales.

The causal and naive approaches are contrasted in Figures 2 and 3. It should be noted that the develop-

[1] Causal is to be interpreted in its common-sense meaning. A causal variable, X, is one which is necessary or sufficient to the occurrence of an effect, Y, and X precedes Y in time. An interesting interpretation of the literature on causal models may be found in Hubert M. Blalock, Jr., *Causal Inferences in Nonexperimental Research* (Chapel Hill: University of North Carolina Press, 1964).

* Assistant Professor of Marketing, Wharton School, University of Pennsylvania

[2] Econometricians prefer the word "structural" to "causal."

Published in Robert L. King, Marketing and the New Science of Planning. American Marketing Association, 1968.

ment of the causal model is much more complex than the development of the naive model. This added complexity would seem to be a disadvantage in that poor measurement at any stage of development would adversely affect the performance of the model. The causal approach rests on the assumptions that the important causal variables are known, that the causal relationships can be measured adequately, that the changes in the causal variables may themselves be forecast with accuracy, and that errors in these predictions of the causal variables will not lead to serious errors in the prediction of sales.

FIGURE 2

DEVELOPMENT OF A NAIVE MODEL

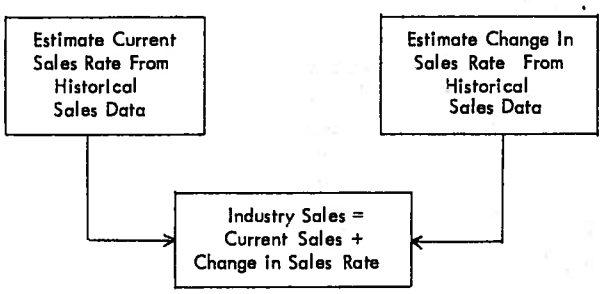
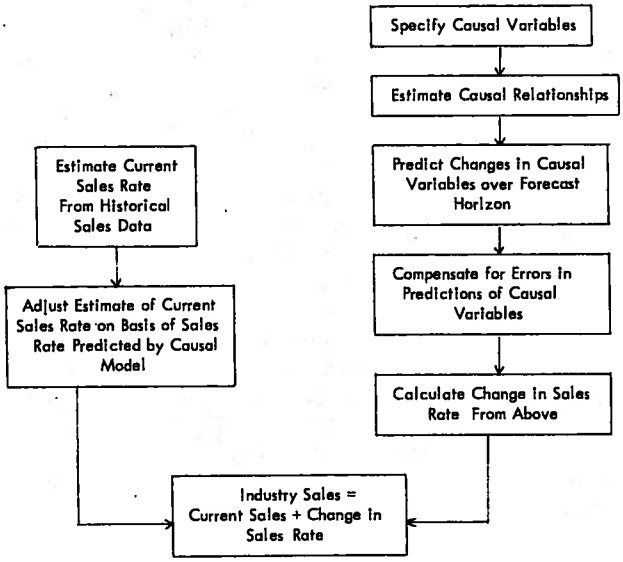


FIGURE 3

DEVELOPMENT OF A CAUSAL MODEL



A COMPARATIVE STUDY

While the primary concern of this study was with improving the firm's forecasts, it was assumed that an improvement in industry sales forecasts would lead to a corresponding improvement in the firm's sales forecast. This assumption, which appears to be made by most companies, simplifies, the problem in that it allows one temporarily to ignore competition among firms in the industry. It appears to be a reasonable assumption for industries which are "somewhat competitive" in the sense that the actions of any one firm have little effect upon industry sales.

It was hypothesized that the causal approach will lead to substantially better forecasts than the naive approach in situations where "large changes" are expected. By "substantially better forecasts," it is meant that large improvements can be made in forecast accuracy with relatively small expenditures.

One rather common situation in which large changes are expected to occur is that of long-range forecasting—and, in particular, long-range forecasting for international markets. For example, events such as the Kennedy Round tariff cuts or the formation of the European Common Market are expected to lead to large changes in sales rates for many types of goods. Improved accuracy in long-range international market forecasts would seem to be of substantial value to firms, since the level of international trade has been gaining rapidly and since there is much room for improvement in forecast accuracy.

If it is true that the causal approach does offer substantial benefits, then it should be possible to demonstrate this in a particular case. The still-camera industry was selected as a test case due to the author's previous experience in this industry. Unconditional [3] forecasts were desired for unit camera sales by country. The objective was to examine what improvements in long-range (5 to 10 years) forecasting result from the use of the causal approach over what could be achieved by the naive approach.

A Naive Model

A naive model which purported to be representative of current practice was developed an *a priori* grounds, and was based on an average of a "no-change" model (*i.e.*, sales in the future in country A will be the same as today's sales) and of a "constant percentage trend model" (*i.e.*, sales in country A will continue to grow at the same percentage rate as they have in the recent past). There are, of course, many other formulations which could have been used; the concern here was to settle on the formulation which seemed most reasonable in the light of current practice. This selection of a naive model was based on an impressionistic analysis of the literature on "how business firms forecast" and on my experience in forecasting in a number of industries. Whether this is a reasonable representation will be discussed further below.

Data on still camera sales were obtained from trade and production figures for the period 1960-65. These data were available from the U.S. Department of Commerce or from the United Nations. Much effort and substantial amounts of subjective interpretation were required; however, to obtain data which were comparable across countries. The current sales rate was based on the average yearly sales from 1960-1965, while the trend for each country was based on the percentage change from 1960-1965 in each country.

A Causal Model

As indicated earlier, the development of the cau-

[3] By unconditional, we mean that the values of the *causal* variables are not known — they must be forecast.

sal model is substantially more complex. The model which was developed utilized much of the causal approach, but it also depended upon naive components. In other words, this model was intermediate between the extremes of causal and naive models.

Figure 4 summarizes the causal model which was developed to forecast camera sales by country. The details of this model are not important since this is only one of many ways in which the model could have been structured, and since the model is a rather typical econometric formulation. The important parts of the model are conventional—measures of current sales rates and of the effects of percentage changes in what are essentially population, income, prices and quality. The basic form of the model, the multiplicative (or "log-log") is the most commonly used econometric formulation. It yields exponents which are easily interpreted [4] and allows one to utilize previous research results most conveniently.

The development of the causal model followed the outline of Figure 3 above: current sales were estimated on the basis of 1960-1965 data, and this estimate was adjusted by use of a causal model which attempted to explain the level of sales for each country; the causal relationships were specified completely on the basis of an *a priori* analysis, and these estimates were updated by a series of econometric measurement models; changes in causal variables were based on naive projections except in a few cases where outside information was available (*e.g.*, camera price changes and scheduled tariff changes due to the Kennedy Round); the effects due to errors in predictions of the causal variables were based primarily upon subjective estimates. Details of the model development are developed elsewhere by the writer. [5] Clearly it was much more difficult to develop the causal model than the naive model.

The data for the causal variables were obtained primarily from secondary sources. It was necessary, however, to make a survey in order to obtain data on camera prices in each country. Recent improvements in the quality and availability of the data proved to be extremely helpful. Russett [6] was one of the most useful data sources.

EVALUATING PREDICTIVE VALIDITY

While there are many criteria by which to judge the causal model, the most important is the test of predictive validity. Unfortunately, however, a substantial time lag would be involved if a direct test were made of the ability of the causal model to provide long-range unconditional forecasts. In order to eliminate this time lag one could use old data, say from 1950 to 1955, to estimate the parameters of the model, and then forecast 1965. The major drawback to this approach is that it is very difficult to obtain these old data. Data for more recent periods are generally more readily available and, if they are not available, one may run a survey to obtain the needed data. This situation is especially pronounced in the case of international marketing where the quality and availability of data have improved rapidly over the past decade.

An approach which avoids the time lag and which utilizes the most recent data for model development is to forecast backwards in time. For want of a better term, we call this backcasting. In this study, data from 1967 to 1960 were used to estimate parameters of the forecasting model. It was then assumed that nothing was known prior to 1960, and unconditional backcasts of 1954 were obtained for seventeen countries. These backcasts were then compared with actual 1954 camera sales. The results are presented in Table 1.

It may be noted from Column C in Table 1 that large changes did occur in camera sales over the 8.5 year time span. On the average, sales in 1954 were about half those in 1962. It may also be noted that the updated causal model (column A) was superior in predictive ability to the naive model purporting to represent current practice (column D). The mean absolute percentage error was reduced by twenty percentage points (from 43% to 23%). Such a margin of superiority would appear to be of great practical significance if one could generalize to the forecasting situation.

DISCUSSION

The results were consistent with the hypothesis that the causal approach leads to substantial improvements in accuracy. There are, however, alternative explanations for these results which should be considered.

Chance

It is possible that the predictions from the causal model could have been generated from a process which was no better than that used for the naive model. To examine whether the superiority of the causal model was due to chance, the role of the classical statistician was adopted. A significance level of 5% was specified and the student t-test was used to test the null hypothesis that there was no difference between the mean absolute percentage errors from each model. The null hypothesis was rejected (calculated level of significance = .03; one-tail test). Since "unusual events don't happen to me," I conclude that the performance of the causal model was superior to that of the naive model.

Fortunate Selection of a Causal Model

The question here is whether the results are replicable. Could other researchers develop useful causal models? To answer this question, various tests of sensitivity were performed on the model parameters. The overwhelming conclusion was that the results were not very sensitive to variations in the parameters. In other words, given the structure of the model, there was a great deal of latitude available in the choice of coefficients.

One indication of the lack of sensitivity to the value of the coefficients is found in an examination of the forecasts obtained from the *a priori* version of causal

[4] The exponent (which is the coefficient in the log-log formulation) gives the percentage change in sales due to a one percent change in the given *causal* variable.

[5] J. Scott Armstrong, *Long-Range Forecasting for a Consumer Durable in an International Market*, (Unpublished Ph.D. Thesis, M.I.T.,) June, 1968.

[6] B. M. Russett, *et. al.*, eds:, *World Handbook of Political and Social Indicators* (New Haven: Yale University Press, 1966).

FIGURE 4

A CAUSAL MODEL TO PROVIDE UNCONDITIONAL FORECASTS

(t = current year; n = forecast horizon)

STAGE I

Prediction of Camera Prices

$$P_{t+n} = (P_t)(0.984)^n \left(\frac{T_{t+n}}{T_t}\right)^{0.8} \left(\frac{C_{t+n}}{C_t}\right)^{1.5} \left(\frac{Q_{t+n}}{Q_t}\right)^{0.5}$$

where P_t is retail camera price index as measured by survey the 0.984 is a constant representing the effects of technological change and

T is a taxation index: (tariffs x imports + sales taxes)

C is resale price maintenance index

Q is non-tariff trade controls index

STAGE II

Market Penetration

$$R_{t+n} = (R_t)(1.02)^n \left(\frac{E_{t+n}}{E_t}\right)^{1.0} \left(\frac{P_{t+n}}{P_t}\right)^{-1.0}$$

when R_t is current sales rate per potential buyer (a weighted average of an estimate based on trade and production figures and a prediction from a causal model explaining sales levels in each country);

the 1.02 is a constant representing the effects of quality change and E is personal consumption expenditure per capita. P is retail price index for camera goods.

STAGE III

Market Size

$$M_{t+n} = (T_{t+n})(L_{t+n})(A_{t+n})(N_{t+n})^{0.3}$$

where M is the number of potential buyers
T is the total population
L is the proportion literate for people 15-64
A is the proportion of the population 15-64
N is the proportion non-agricultural employment.

STAGE IV

Unit Camera Sales By Country

$$L_{t+n} = (R_{t+n})(M_{t+n})$$

where L is the "long-run" rate of unit camera sales

TABLE 1

BACKCASTING CAMERA SALES FOR 1954 (ALL SALES RATES GIVEN IN THOUSANDS—NUMBERS ROUNDED FOR PRESENTATION)

Country	Actual Sales (1954)	A Predicted Sales (Percentage Errors in Parentheses[a]) Causal Model: Updated Version		B Causal Model: A Priori Version		C Naive Model: No Change		D Naive Model: Current Practice	
Austria	53	49	(10)	42	(24)	84	(-45)	78	(-38)
Bel.-Lux.	86	94	(-8)	82	(5)	177	(-69)	104	(-19)
Denmark	103	66	(44)	56	(60)	128	(-21)	134	(-26)
Finland	20	19	(6)	14	(36)	37	(-59)	24	(-18)
W. Germany	1320	790	(53)	720	(59)	1690	(-25)	1128	(16)
Netherlands	106	116	(-8)	118	(-11)	301	(-96)	198	(-61)
Norway	48	28	(56)	24	(68)	51	(-5)	46	(4)
Portugal	20	11	(57)	10	(63)	25	(-24)	17	(16)
Sweden	146	145	(1)	126	(15)	278	(-62)	174	(-18)
Switzerland	94	116	(-20)	118	(-23)	264	(-95)	178	(-62)
Mexico	48	26	(60)	24	(68)	69	(-36)	62	(-25)
Peru	6	7	(-9)	7	(-23)	18	(-100)	10	(-50)
Venezuela	12	15	(-24)	15	(-25)	43	(-114)	32	(-91)
New Zealand	32	34	(-5)	36	(-12)	66	(-68)	53	(-49)
Thailand	4	5	(-8)	5	(-3)	13	(-99)	8	(-67)
Israel	3	3	(12)	3	(9)	13	(-121)	7	(-80)
S. Africa	33	38	(-14)	36	(-9)	89	(-92)	92	(-94)
Totals	2134	1562		1436		3346		2345	
Mean Absolute Percentage Deviation			(23)		(30)		(67)		(43)
Average Percentage Error			(+12)		(+18)		(-67)		(-39)

[a]Percentage Error = $\frac{\text{Actual} - \text{Predicted}}{1/2(\text{Actual} + \text{Predicted})}$

model (column B). This model utilized parameter estimates which differed somewhat from those in the updated model, yet its performance in backcasting was superior to all of the naive models.

The sensitivity of the model to changes in the structure of the model was not examined. The plan here was to try to follow conventional practice and to keep the structure simple.

Unfortunate Selection of a Naive Model

It is possible that the poor performance of the naive model was due to an unfortunate choice of a model to represent current practice. To test this possibility, additional formulations of naive models were considered. These included models based on no change, on a constant trend for the total market, and on a constant trend by country, as well as various combinations of these three basic models. In addition, some consideration was given to the use of more sophisticated formulations of the naive models. Thus, instead of using simple moving average and trend calculations, consideration was also given to exponential smoothing models which weight more recent data most heavily.

The best version of a naive model which resulted from this *a posteriori* search was based on a simple average of the forecasts generated by each of the three basic naive models. Still, the performance of this naive model did not match that of the updated model; nor, in fact, did it measure up to the performance of the *a priori* version of the causal model. Table 2 summarizes the results for some of the naive models which were considered. It is important to note that none of the naive models did as well as either of the the two versions of the causal model.

Other Explanations

It is possible that still-cameras are different in such a manner as to favor the use of causal models. Or it might be argued that the time period used for the test of predictive validity is favorable to the use of causal models. Answers to these questions were not provided by this study.

TABLE 2

PERFORMANCE OF VARIOUS MODELS IN BACKCASTING SALES

FOR 1954 (n = 17 COUNTRIES)

Naive Models	Mean Absolute Percentage Error	Average Percentage Error
No Change	67	-67
Constant Trend by Country	51	+23
Constant Trend Over All Countries	37	+28
"Current Practice"	43	-39
"Best" Naive Model	33	-23
Causal Models		
A Priori	30	+18
Updated	23	+12

SUMMARY

The use of causal models was shown to be of substantial value in a situation which is currently of great importance—long-range forecasting for international markets. The causal model, based on conventional economic relationships, proved to be superior to a series of naive projections. Evidence of the predictive value of the causal model was provided by obtaining unconditional backcasts of 1954 based on data from 1967 to 1960.