5-1994

A comparision of alternative forecasting techniques : an application of scan data

Kent L. Wolfe

To the Graduate Council:

I am submitting herewith a dissertation written by Kent L. Wolfe entitled "A comparision of alternative forecasting techniques : an application of scan data." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Agricultural Economics.

David B. Eastwood, Major Professor

We have read this dissertation and recommend its acceptance:

John R. Brooker, S. Darrell Mundy, Pratibha A. Dabholkar

Accepted for the Council:

Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

To the Graduate Council:

I am submitting herewith a dissertation written by Kent L. Wolfe entitled " A COMPARISON OF ALTERNATIVE FORECASTING TECHNIQUES: AN APPLICATION OF SCAN DATA." I have examined the final copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Agricultural Economics.

David B. Eastwood

Dr. David B. Eastwood, Major Professor

We have read this dissertation
and recommend its acceptance:

John R. Brooker

Dr. John R. Brooker

S. Darrell Mundy

Dr. S. Darrell Mundy

Pratibha A Dabholkar

Dr. Pratibha A. Dabholkar

Accepted for the Council:

C W Minkel

Associate Vice Chancellor
and Dean of the Graduate School

A COMPARISON OF ALTERNATIVE FORECASTING TECHNIQUES: AN APPLICATION OF SCAN DATA

A Dissertation
Presented for the
Doctor of Philosophy Degree

The University of Tennessee, Knoxville

Kent L. Wolfe

May 1994

Acknowledgments

I am indebted to all those who have helped me in the writing of this dissertation. I would like to extend a special word of thanks to Dr. David B. Eastwood for his patience and guidance during the completion of this research. Dr. Eastwood genuinely cared about me as an individual as well as my education and research which made this monumental task a lot easier. I would also like to express a word of thanks to Dr. Morgan Gray for assisting me in the necessary computer programming.

Thanks also to Dr. John Brooker, Dr. Pratibha Dabholkar, and Dr. S. Darrell Mundy for their insight, suggestions, and encouragement. I would also like to thank Dr. Handy Williamson for providing me the opportunity to further my education through a research assistantship.

To my wife, Heidi, and my two sons, Jacob and Jonathan, thank you for standing by me through the completion of this research. My completion of this dissertation required great sacrifice, encouragement, and love all of which you greatly supplied.

Abstract

Scan data have provided a suitable data source for estimating consumer demand relationships at the retail level. These data capture actual quantity and price information which can be combined with additional explanatory variables, such as advertising, promotion and seasonality, to estimate the retail demand function for specific products. Despite the availability of scan data, relatively little research has focused on forecasting at the retail level.

The competitive nature of the supermarket industry, both the encroachment of warehouse food retailers and generic private label products, have lead to an increased interest in consumer demand analysis at the retail level. The increased competition from nontraditional retail outlets has eroded the traditional supermarket's market share. The nontraditional grocery outlets are perceived to be less expensive than their traditional grocery outlets. Thus, traditional grocery outlet managers have become increasingly interested in reducing operating costs. One method of reducing operating costs is to reduce inventory levels via implementation of an efficient consumer response (ECR) strategy, a version of just-in-time delivery. The ECR strategy has the potential to reduce inventory levels which can directly lower inventory costs. The rise of ECR has created a need for accurate product demand forecasts at the supermarket level to maintain adequate inventory levels. The ability to forecast weekly demand in response to changes in seasonality, holidays, and promotional and advertising campaigns is very important to retail

managers for implementation of an ECR strategy.

The objectives of this study are to 1) develop alternative forecasting methods that are suitable for scan data, 2) estimate and compare the alternatives with respect to food groups and individual products in terms of their forecast accuracies using a scan data base, and 3) estimate and compare the alternatives with respect to food groups and individual products in terms of two week trial forecasts.

The theoretical forecasting model was developed utilizing economic theory and previous consumer demand research. The model described weekly product item movement as being a function of own- and cross-prices, own- and cross-advertising (television, radio, and newspaper), holidays, and seasonality. The theoretical model for the brand product also included point of purchase and the start of the Knox County, Tennessee, school year.

The second forecasting model specification was developed using the Box-Jenkins methodology. This technique does not incorporate structural explanatory variables, but rather, identifies and replicates underlying patterns in the data series utilizing past item movement and disturbances in the series.

The third forecasting method combines the structural variables contained in the theoretical model with the pattern identification and replication ability of the Box-Jenkins model to produce a composite model known as a transfer function.

This study utilized weekly scan and advertising data (television, newspaper, radio, and point of purchase) which was supplied by a multi-regional supermarket chain. The data consisted of weekly UPC-level prices, item movement, and chain-initiated television, radio, and newspaper advertising. The data were pooled across five stores that catered to average to above average income food shoppers.

The data were divided into two subgroups. The first subgroup of data was used to estimate the alternative forecasting models and generate product backcasts for technique evaluation and comparison. The second subgroup of data, the last 26 weeks for each product, was used to generate a two week trial forecast. Again, the models and their forecasting abilities were evaluated and compared across alternative methods.

The three alternative forecasting models were estimated using the historic subgroup data. The alternative forecasting models were evaluated individually by the evaluation criteria to choose the "best model" to represent each technique. These model estimates were then used to generate backcasts of the data series for each of the three food products, brand b, group g, and steak. The alternative techniques were then evaluated and compared.

The results of the backcast forecast evaluation and comparison suggested that the transfer function forecast was superior to the Box-Jenkins and theoretical forecast in predicting weekly item movement for brand b and steak. Group g's weekly item movement was best forecast utilizing the Box-Jenkins methodology.

The results of the two week trial forecast evaluation suggested that the transfer function technique was superior to the theoretical and Box-Jenkins techniques in accurately forecasting weekly item movement for each

of the three products, a highly process brand, its associated group, and steak a variable weight perishable product. This study has found that the transfer function is the best of the three techniques for use in forecasting weekly retail item movement for both brand and category peanut butter and the steak category. However, the results also indicate that each of the forecasting models was relatively accurate in forecasting actual item movement but performed poorly in predicting directional change.

TABLE OF CONTENTS

TABLES

TABLES

FIGURES

Chapter I

Data, Problem Statement, and Objectives

A. Introduction

The competitive nature of the retail grocery industry has lead to an increased interest in consumer demand analysis at the retail level. Store managers need to understand the fundamentals of consumer demand in order to respond appropriately to changes in prices, seasonality, holiday periods, and promotional campaigns. In addition to understanding the consumer demand relationships, the supermarket manager needs access to reliable information on future demand levels to use in making inventory level decisions.

B. Scan Data

Government agencies are responsible for generating and providing the vast majority of public data about the agricultural sector and food supply and demand in particular. The public time-series data are based on aggregate annual, quarterly, or monthly consumer purchase data. Because the data represent aggregate consumer purchases, information on consumer purchases of individual products and their prices are not available. Consumer panel and survey data contain specific product and socioeconomic information at various points in time. These data sources do not include prices. Instead they have to be imputed from reported quantity and expenditure information (Capps 1989). The cross-section data do not permit dynamic analyses of food demand for specific foods. Private, proprietary, time-series and cross-sectional data are available for selected foods, but these data are very expensive and usually lack the rigorous sampling and/or measurement designs implemented by federal government agencies. In general, the private and public data are not designed for retail food demand analysis for individual foods.

Scan data are a relatively new alternative source of data that began to emerge after the introduction of scanners in the 1970's. Scanner systems read a product's UPC bar code. A UPC code consists of a machine readable bar code (a series of bars and spaces) and a corresponding human readable UPC number (the numbers directly beneath the bar code). A product's UPC bar code is assigned by the Uniform Code Council. The scanner system uses lasers to read a product's UPC bar code which is then matched with a master file to identify the product and it's price. The price, quantity, and a brief product description are recorded for each product scanned, interpreted by computers, and used to generate customer bills. Because the information is automatically entered, it can be captured and incorporated into a data bank. Scan data provide supermarket management with an efficient method of monitoring actual product sales as opposed to using warehouse product movement or secondary data.

Scan data comprise an alternative data source to obtain new estimates of food demand relationships (Capps, 1987 Micro-Data Base). A scan database may contain records that track individual products across time, and if pooled with data from different stores and/or customers, may have characteristics of pooled time-series and cross-sectional data. Product specific information allows one to estimate relationships between substitute and complimentary products as well as consumer demand responses to changes in price, seasonality, holiday periods, and/or promotional campaigns. These relationships provide useful information regarding the trade offs consumers make when purchasing grocery products.

Scan data are primary data that reflect actual item movement over time and provide supermarket managers with a tool for tracking consumer purchases. Monitoring scan data allow supermarket managers to track actual product movement over time and observe what is going out the "front door" as opposed to monitoring what is coming in the "backdoor." Comparing actual product sales to warehouse product item movement provides supermarket managers with a tool for estimating product shrinkage.

Part of the supermarket manager's profit maximizing goal is to increase total grocery sales. Product tracking provides a tool for estimating a product's actual sales performance. The sales performance can provide managers with information to be used in a variety of applications such as discontinuing a product, allocation of shelf space, and reordering. A product that is performing poorly may be discontinued or allotted less shelf space and/or a less desirable shelf location. This frees shelf space for the addition of new products or to expand the shelf space for existing higher performance products.

Supermarket management has become more concerned with reducing costs as retail competition intensifies. Nontraditional retail outlets (e.g., SAMS, KMART) have started to offer grocery products and have captured market share from supermarkets (Supermarket Business 1993). Wholesale grocery outlets have also entered the retail industry and are perceived as being less expensive relative to supermarkets. The loss of market share has increased supermarkets' concern with cost reduction and improved management.

Direct product profit (DPP) accounting methods were introduced in the 1980's in response to supermarket managers' interest in cost reduction. DPP is a method of assigning direct costs to a product. It involves the allocation of the total cost of a product from it's transportation cost through the display cost. Calculating a product's DPP allows managers to compare the profitability of brands and food groups. Aside from the theoretical and empirical problems associated with the cost allocation, the unit price obviously has an impact on profit. The effect of price changes depends on the price elasticities, which can be estimated with scan data.

Efficient consumer response (ECR) is a new strategy being implemented by supermarket managers to reduce costs. ECR is an inventory management technique that emulates just-in-time delivery. The approach requires the supplier to restock the distributer continuously and the supplier and distributer to restock the store continuously . This type of inventory management can reduce inventory levels which can lower inventory costs.

But full implementation requires accurate forecasts of sales in order to avoid stockouts or excess inventories. The desire to reduce operating costs through efficient inventory management necessitates the need for accurate forecasts of consumer demand in response to changes in price, seasonality, holidays, and promotional campaigns. If the inventory level is too low in relation to consumer demand, the supermarket will experience a stockout. If inventory levels are too high, additional costs are incurred in maintaining the excess inventory. Thus, there is a real need for accurate sales forecasts because of the small margin for error, and an incorrect forecast could result in substantial costs to the supermarket.

Another concern of supermarket management is increasing category sales as a method of increasing total sales revenue. Raju suggests that increased brand sales do not necessarily result in increased category sales. For example, a sales increase for one particular product in a category might lower the sales of competing products as consumers switch brands. The substitution between the promoted product and competing products may result in no significant change in total category sales. Product category sales are also important to supermarket managers because of the lack of brand specific products in certain departments (e.g. produce, fresh meats). These products are generally perishable which increase the need for accurate sales predictions to avoid losses from spoiled products.

Consumers, both domestically and globally, have changed their attitudes regarding brand and private lable products. The consumer is more conscious of price rather than brand. (Schiller). For example, in Britain 32 percent of consumer expenditure are on private label products. This trend also is expected to continue in the United States (Oster, Savery, and Templeman). In the United States this has led to increased sales of store brands which are cheaper than their competing national brands. This situation provides another motivation for using scan data to evaluate food demand at the product specific level.

C. Problem Statement and Objectives

The current literature focusing on forecasting applications includes a wide diversity of forecasting techniques and evaluation criteria. There has been considerable interest in forecasting commodity and livestock prices, bond and interest rates, economic indicators, and demand for natural resources. Published studies involving forecasts of retail grocery sales have been very limited. However, the private sector has vendors who have developed forecasting algorithms. Concerns with these methodologies include the lack of information about the techniques employed and the accuracy of the forecasts. The recent emergence of scan data, the ability to track the sales of individual foods, and supermarket manager's need for accurate sales predictions necessitates the need for additional applied research in retail sales forecasting.

The main objective of this research is to explore the application of alternative forecasting methods in the context of supermarket scan data. Specific objectives are to 1) develop alternative forecasting methods that are suitable for scan data, 2) estimate and compare the alternatives with respect to food groups and individual products in terms of their forecast accuracies using a scan data base, and 3) estimate and compare the alternatives with respect to food groups and individual products in terms of two week trial forecasts. Interest encompasses food groups as well as individual brands within a commodity group. Demand is considered to be represented by quantity sold per thousand customers per week.

Chapter II

Types of Forecasts and Evaluation Criteria

A. Forecasting Techniques

A forecast is an estimate of the future value of a variable. Quantitative models using historical and current information are used to predict, or forecast, future events. There is an extensive literature associated with forecasting research. For example, extrapolation, econometric, time-series, and composite forecasting techniques are frequent topics. The different forecasting techniques can be grouped into three broad categories: theoretical, statistical, and composite. A theoretical forecast combines economic theory and logic to identify a set of determinants that describes the variation present in a dependent variable. The estimated relationship is then used to generate a forecast of the data series. Statistical forecasting techniques are not concerned with the causal relationships that produce a data series, as are the econometric techniques, but instead, they are used to identify patterns without considering any causal relationship. Composite forecasting techniques combine the theoretical and statistical techniques. The estimated model is then used to generate a forecast.

Various types of forecasts can be generated regardless of the technique employed (conditional, unconditional, ex post, ex ante, point, or interval). Thus, a researcher must decide on the most appropriate forecasting technique as well as the type of forecast he is going to generate. Once a forecasting technique and type of forecast have been chosen, the researcher must decide on what evaluation criteria he is going to employ. The following section reviews the many different types of forecasts, forecasting techniques, and evaluation criteria available to researchers.

Forecasts can be classified into two broad categories: unconditional and conditional. An unconditional forecast is generated using independent variables for which the values are known with certainty or which can be estimated accurately. The independent variables included in the forecast may be current values of independent variables or lagged values of independent and dependent variables. The conditional forecast is generated using independent variables for which the values are not known with certainty. The independent variable values must be estimated and then included in the conditional forecasting model.

An ex post forecast is considered to be unconditional because both the exogenous and endogenous variables are known with certainty. The observed data set is divided into two periods. The first subperiod is used to estimate the relationships. The estimated model is then used to forecast the remaining subperiod values which can be compared to the actual data to obtain a measure of forecast accuracy.

An ex ante forecast may either be conditional or unconditional. A forecast using data which are known with certainty represents an unconditional ex ante forecast. The estimated relationship is used to generate a forecast of the dependent variable beyond the estimation period. The predicted values are then compared to the observed data as they become available.

An interval or single value forecast can be generated depending on the needs of the forecaster. A point forecast is used to predict a single value. An interval forecast generates a confidence interval at a specified level of significance. The significance represents the probability that the actual value of a variable lies within the interval.

Because a forecast is an estimate, it may not be identical to the actual value. Different situations may lead to the introduction of error into the forecasting process. First, errors may be generated because the independent variables may have been estimated and could, therefore, be different from their actual values as in a conditional forecast. On average, forecasted values of the independent variables are assumed to equal the actual values, but if the estimated and actual independent variable values are not identical, the model generates a forecast using the incorrect data resulting in forecast error. Second, error may be introduced by having to estimate the causal relationship. Even if the estimated parameters are unbiased, differences between the parameters and estimated values can cause forecasts to differ from actual values. Specification error can attribute to forecast error. Omitting important variables from a function may lead to biased estimators. Implementing the incorrect functional form may also introduce error into the process. The error results from the functional form not accurately capturing the true process, thereby generating errors in the system.

1. Theoretical Forecasting Techniques

The simplest forecasting model is the naive or no change model. This forecast approach assumes that there is no change in the future value or direction of the dependent variable. It is commonly used as a measure of forecast accuracy by which more complex forecasting models are compared.

The forecast accuracy of a complex model should be equal to or greater than the naive forecast.

Econometric and statistical models are represented by a variety of notations which can lead to reader and researcher confusion. This section contains a large number of modeling techniques and evaluation criteria which necessitates the inclusion of a common variable list and definitions. The following is a list of commonly used variables and their definitions.

Common variable definitions:

ε_t = the random population disturbance term in time period t.

Y_t = dependent variable in time period t.

x_t = vector of k independent variables in period t.

b = vector of k+1 population coefficients, including the intercept.

t = time period t.

Regression models can be used to generate forecasts. Before a forecast can be made, an explicit functional form must be specified. This consists of identifying the determinants of the dependent variable and the way they relate to the dependent variable. Economic theory and knowledge of the industry are used to derive the determinants and specify the form of the model. For example, economic theory implies that the demand for a good is a function of income and own- and cross-prices. The model might include a trend variable if the dependent variable exhibits an upward or downward tendency over time. After the model has been specified, it is estimated. A forecast of the dependent variable is obtained by plugging values of the independent variables into the estimated regression equation. Evaluation involves using the goodness of fit criterion associated with regression analysis (described in a later section).

A method of extrapolating a series using a simple regression is by using a dummy or binary variable (e.g. equation 1). This procedure accounts for the seasonal fluctuations present in the series by representing each of the different seasons. If a particular season is present its dummy variable will equal one while the remaining seasonal variables are set equal to zero. The parameter estimates for each of the dummy variables provide an estimate of the relationship between a given season(s) and the dependent variable. The inclusion of the dummy variable allows each of the particular seasons to impact the intercept of the forecast differently.

Dummy variable extrapolation

(1) $$Y_t = b_0 + b_1 x_1 + \ldots + b_{k-1} x_{k-1} + b_k S_i + \varepsilon_t.$$

S_i = seasonal dummy variable, where S_i is assigned either a one or zero depending on the season in time t.

A simultaneous equation system is a more complex econometric forecasting problem. The procedure uses a set of multiple equations to describe the economic relationships and linkages among the variables. The simultaneous equation system generates forecast values of more than one endogenous variable. For example, it may represent a simple supply and demand relationship (e.g. equations 2 and 3).

Simultaneous equation system

(2) $\quad Supply\text{:}\quad Q_t = b_1+b_2P_t+\varepsilon_{St}.$

(3) $\quad Demand\text{:}\quad Q_t = b_1+b_2P_t+b_3Y_t+\varepsilon_{Dt}.$

P_t=price in period t.

Y_t=income in period t.

Q_t=quantity in period t.

ε_{ST} = supply equation random disturbance in period t.

ε_{DT} = demand equation random disturbance in period t.

Single equation estimation techniques cannot be used because of the correlation between the error terms and the endogenous variables present on the right hand side (in the above case P_t). The equation system can be manipulated so that each endogenous variable is expressed as a function of predetermined variables and is referred to as a reduced form equation (e.g. equations 4 and 5). The reduced form equations for the above supply and demand model are:

Reduced form equations

(4) $\quad q_t = \pi_{12}y_t+v_{1t}.$

(5) $\quad p_t = \pi_{22}y_t+v_{2t}.$

$$\pi_{12} = \frac{\alpha_2 b_3}{\alpha_2-b_2}Y_t.$$

$$v_{1t} = \frac{\alpha_2\mu_t-b_2\varepsilon_{St}}{\alpha_2-b_2}.$$

$$\pi_{22} = \frac{b_3}{\alpha_2-b_2}.$$

$$v_{2t} = \frac{\mu_t-\varepsilon_{Dt}}{\alpha_2-b_2}.$$

Each reduce form equation can be estimated separately yielding consistent and unbiased estimates assuming no mutual correlation between the equations. Predicted values of the independent variables are plugged into the equation system to generate a forecast using the simultaneous equation system.

A recursive model (e.g. equations 6 and 7) resembles the simultaneous equation system in that it utilizes a set of equations to explain the relationships and linkages between variables.

Recursive model

(6) *Supply*: $Q_t = b_1 + b_3 P_{t-1} + \varepsilon_{St}$.

(7) *Demand*: $P_t = b_1 + b_2 Q_t + b_3 Y_t + \varepsilon_{Dt}$.

There is unidirectional dependency among the endogenous variables. For example in equation (6) the endogenous variable is a function of exogenous variables and can be estimated independently. In equation (7) the endogenous variable is a function of the preceding equation's endogenous variable plus an exogenous variable. Or, in general each of the endogenous dependent variables can be estimated sequentially, given the preceding equations' estimated endogenous variables.

The vector autoregressive model (VAR) assumes that each variable in the equation system is endogenous and can be written as a linear function of its own lagged values and the lagged values of all the other variables contained within the system (e.g. equation 33). In principle, the VAR is a set of reduced form equations derived from the system of structural equations. Each of the equations is estimated separately. The VAR modeling technique can be viewed in terms of the restrictions imposed on the structural equations. The VAR technique is able to display the relationships that the structural model is trying to estimate. The VAR has a problem of overparmeterizing a model. A restriction can be placed on the model which is designed to restrict the number of parameters and choose the lag length by minimizing a prespecified goodness of fit criteria, for each series (Park). The restricted VAR is represented as RVAR.

The Bayesian vector autoregressive (BVAR) forecasting technique is a variant of the VAR model. The VAR model can overparameterize a model which leads to a very good in-sample fit and a poor out-of-sample fit. The BVAR allows the researcher to employ his prior beliefs about coefficients and whether or not to include them in a model. The researcher has two choices regarding coefficient beliefs. The first is that they are not zero and should be included in the model. The second is the estimated coefficient should be zero and therefore excluded from the model. The Bayesian approach instead of assuming that a large number of the coefficients is equal to zero, assumes that a large number of coefficients is close to zero. The more recent values have greater effects on the dependent variable and are weighted more heavily than earlier values. Thus, the BVAR can incorporate the original information by including large lags and at the same time specify that the more distant the lag the more likely its coefficient is to be equal to zero (Litterman).

A modeling technique known as the state space model is a generalization of the linear regression model where the state of a system is estimated using noisy measurements. For example, estimating the state of a satellite given knowledge of the various parameters which change over time. The dependent variable is expressed as a linear function of observed independent variable values, a time varying parameter, and an error term. The unique feature of the time varying parameter is a vector or the state variable. The procedure estimates the state of the dependent variable given the state of the time varying parameter and the independent variable values.

Once the model has been estimated, a procedure is available to enhance the parameter estimation. The procedure is referred to as the Kalman Filter. The Kalman filter is a recursive estimation process used to calculate optimal parameter estimate given the information at time t for a state space model. A state space model is a method of modeling a process in which all of the pertinent information at time t about the dependent variable is contained in an alpha matrix. The state space function's alpha matrix contains all the relevant information on the system at time t while using the least number of elements as possible. For example, it contains the previously estimated parameter coefficients and the mean and variance of the dependent variable. The Kalman filter minimizes the mean squared errors and provides a mean squared error forecast estimate. The process updates the state of the model as new observations become available. The procedure assumes a normal distribution of the original model's error term. Given the assumption of normality, the likelihood function can be represented by the one-step-ahead prediction errors which are provided by the Kalman filter. Using the likelihood function, it is possible to estimate unknown parameters in the model as well as supplying a foundation for statistical testing and specification of the model.

The Kalman filter can smooth a series and generate predictions about a variable at a specific point in time. The procedure generates the optimal estimates for each of the elements in the alpha matrix one time period at a time, starting in period n. The filter then uses the t-n estimates and the information contained in t-n+1 to generate new optimal estimates of the elements contained in the alpha matrix. This iterative process is continued until all of the observations have been processed.

Discriminant analysis can be used as a forecasting technique when a dependent variable is classified into mutually exclusive increasing or decreasing categories. The model can be used to predict the direction of change. The dependent variable is expressed as a linear function or index (e.g. equation 8).

<u>Discriminant analysis model</u>

(8) $$Z_t = b_0+b_1x_{1t}+b_2x_{2t}+\ldots+b_kx_{kt}+\varepsilon_t$$

Z_t = discriminant score.

The estimated coefficients represent the discriminant weighting coefficients for each of the independent variables. The estimation procedure provides an estimate of Z_t, which is referred to as the discriminant score and can be compared to a critical discriminant score, Z_{cv}, to determine the predicted change in the dependent variable's movement. If $Z_t>Z_{cv}$, the model predicts the dependent variable will move upward, and if the $Z_t<Z_{cv}$ the model predicts the dependent variable will move downward. (Menkhaus and Adams)

2. Statistical Models

Statistical models are deterministic, meaning there is no attempt to consider the underlying determinants responsible for creating the series. These techniques try to identify patterns in a data series. Observed values of a variable over time are assumed to be the result of a random process. A time-series is assumed to contain all of the relevant information needed to make predictions about the future values. (Beilock and Dunn) The time-series used to produce a forecast is assumed to have been created by a stochastic process. This means that the observed values of the dependent variable have been generated by a probability distribution function that assigns a probability with each possible dependent variable value. (Pindyck and Rubinfeld)

The random walk is one of the simplest forecasting techniques and uses a probability distribution with a zero mean (e.g. equation 9).

Random Walk

(9) $y_t = y_{t-1} + \varepsilon_t.$

This forecasting technique describes the dependent variable as a function of last period's actual value and the current period's error term. The forecast of the dependent variable is a random selection from the joint probability distribution. A trend variable can be added to the random walk forecast.

Extrapolation is a simple statistical modeling approach that does not try to explain the variation in the dependent variable but tries to reproduce the series over time. There is a variety of simple extrapolation techniques available. They differ with respect to the functional form. If the process is linear, a linear trend model can be used which describes the dependent variable as being a function of time (e.g. equation 10).

Simple extrapolation model

(10) $Y_t = b_0 + b_{1t}(t) + \varepsilon_t$.

The trend (b_{1t}) may be defined as the directional movement of the dependent variable over time. The simple regression model, containing a trend variable, generates a forecast using historical data and regression analysis to fit a linear line to the data. The forecast is obtained by projecting the trend of the dependent variable into the future.

If the process exhibits exponential growth, an exponential extrapolation procedure may be employed (e.g. equation 11). This procedure assumes that the dependent variable changes by a constant percent (b) over time instead of changing by a constant absolute amount.

Exponential growth function

(11) $Y_t = be^{(bt)}\varepsilon_t$.

A third extrapolation procedure is the autoregressive trend model which describes the dependent variable as being a function of its past values (e.g. equation 12).

Autoregressive trend model

$$(12) \quad Y_t = b_1 + b_2 Y_{t-1} + \varepsilon_t.$$

By setting b_1 equal to zero, the b_2 parameter represents the rate of change in the dependent variable. If b_2 equals one and the intercept is not equal to zero, the model extrapolates the series by the same absolute amount each time period. The logarithmic autoregressive trend model expresses the autoregressive model in log form.

The quadratic trend model is a simple extension of the linear trend model in that it incorporates a squared trend variable (e.g. equation 13). This model describes an increasing or decreasing process depending on the values of the b_1 and b_2.

Quadratic trend model

$$(13) \quad Y_t = b_0 + b_1 t + b_2 t^2 + \varepsilon_t.$$

If the process is characterized by an S-shaped curve, a series that exhibits a lower and upper bound, the nonlinear logistic model can be used (e.g. equation 14).

Logistic extrapolation model

$$(14) \quad Y_t = \frac{1}{(b_0 + b_1 x^2)} + \varepsilon_t.$$

These extrapolation models are not able to predict turning points in a data series accurately due to the limited information they contain. A common feature is they generally describe a dependent variable as a function of its past values plus some type of trend variable. Thus, the extrapolation model tends to predict the variable as moving in the same direction as it was in period t-1.

A slightly more complex extrapolation model is the moving average model which predicts future values of the dependent by simply using the average of its past values (e.g. equation 15). The estimated dependent variable, $\hat{Y}$, can be estimated sequentially as data become available.

Moving average extrapolation model

(15) $$\hat{y}_{t+1}=(\frac{1}{n})(y_t+y_{t-1}+\ldots+y_{t-n+1})+\varepsilon_t.$$

In equation (15) all previous values, begining with period t-T, are weighted equally. However, the most recent data may be more important in determining the future value of the dependent variable and could receive a greater weight. This would place an emphasis on the more recent data in predicting future values of the dependent variable. The exponentially weighted average extrapolation model (EWMA) is a procedure that allows recent values to be weighted more heavily than more distant values (e.g. equation 16).

Exponentially Weighted moving average model

(16) $$\hat{y}_{t+1}=by_t+b(1-b)y_{t-1}+b(1-b)^2y_{t-2}+\ldots+b(1-b)^{\tau}y_{t-\tau}+\varepsilon_t.$$

$b=$ the weighting mechanism.

The power to which the weighing mechanism is raised is crucial in determining the slope of the extrapolated function. The larger the power of the weighing mechanism, the smaller the impact of distant observations on the predicted dependent variable. To use the EWMA model, a data series must be detrended. If the trend is not removed and the series contains an upward or downward movement, this method will generally over or underestimate the future dependent variable value. A series that has been detrended can be used to generate an EWMA forecast, and the trend can then be added back to achieve the final forecast.(Pindyck and Rubinfeld).

If a particular variable exhibits seasonal variation, a seasonal variation trend variable can be added to the extrapolation function. The seasonal variation trend variable is obtained by dividing the observed dependent variable value by its estimated value for each season. These ratios are computed for each time period in each season and then averaged to become a seasonal factor. The seasonal factor is then multiplied by the appropriate seasonal trend variable.

Smoothing a time-series may be required if short-run fluctuations are present. A forecast using smoothed data is somewhat arbitrary in that the parameter values derived from the smoothing process are not identical to the actual values. The simple extrapolation model, equation (10), could be used to smooth the data by weighting all of the historical observations of the dependent variable equally. Another method of smoothing data is the exponential smoothing technique (e.g. equation 16). This procedure involves using the EWMA which assigns the most recent values of the dependent variable a greater weight than more distant values in smoothing the data series.

The double exponential smoothing technique (e.g. equation 17), places little emphasis on past data points and is capable of more heavily smoothing a series than the exponential smoothing technique.

Double exponential smoothing technique

(17) $\tilde{y}_t = b\dot{y}_t + (1-b)\tilde{y}_{t-1}$.

$\dot{y}$ = singly smoothed series.

$\tilde{y}$ = doubly smoothed series.

This method assigns greater weights, depending on the size of the weighting mechanism, to the current values of the dependent variable. It involves exponentially smoothing a series and then performing the exponential smoothing procedure to the series again. A centered moving averages process uses predicted and past values of the dependent variable to smooth the series, whereas the other techniques use only historical and present dependent variable values. (Pindyck and Rubinfeld)

A seasonal adjustment may be needed before the model can be estimated. This procedure involves generating a seasonal index to approximate the seasonal variation in a series. Each value in the series is divided by its corresponding seasonal index value to obtain a deseasonalized series. This procedure diminishes and, hopefully, removes all the seasonal and irregular components from the series. It has the capability to remove more than the seasonality, so it should be implemented only when needed.

The SAS X11 procedure is a method of adjusting seasonal additive, monthly, or quarterly time-series data. The X11 procedure requires data that are chronologically ordered and at least three years in length. The procedure removes seasonal fluctuations while leaving the trend, cyclical, trading day, and irregular fluctuations present in the series. The remaining variation is attributed to the long-term trend, business cycles, and cyclical factors. The trading day variable accounts for the variation involved with the passage of time over the year. The irregular component represents the variation present in the residuals. The X11 procedure requires monthly or quarterly observations, so it cannot be used with weekly data.

More complex modeling techniques are the autoregressive and moving average models. These methodologies are based on stochastic and stationary time series. A stochastic time-series is one that has been randomly generated by present and past errors. That is, they are generated by a white noise process which means that the error terms are randomly distributed, independent though time, and contain no recognizable pattern (Nazem). The second important assumption about the data series is stationarity. A data series is defined as being stationary if it has a constant mean, variance, and covariance over time. Plotting the series against time can provide insight into whether or not a series is stationary. Inspection of the estimated autocorrelation function (ACF), (e.g.equation 18), provides a measure of the level of correlation present between neighboring data points in a time-series.

Autocorrelation function

$$(18) \qquad ACF = \hat{\varrho}_k = \frac{\sum_{t=1}^{T-K} (y_t - \overline{y})(y_{t+k} - \overline{y})}{\sum_{t=1}^{T} (y_t - \overline{y})^2} .$$

$\hat{\varrho}_k$=the autocorrelation coefficient with k leads.

$\bar{Y}$= the sample mean.

k= the number of lead periods.

An ACF that does not die out quickly indicates that the data series is nonstationary. Nonstationarity occurs when the mean and variance are not constant, there is a seasonal pattern, increasing or decreasing variability is present, or structural change may have occurred.

Determining cyclical components or explaining the variance in a time series can be accomplished by using spectral analysis. Spectral analysis estimates how much of the variance in a series can be attributed to a cycle. The cyclical variance and corresponding frequency is referred to as the spectral density function. Observation of the density spectrum function provides insight into whether a process is considered a white noise process or not. If the density function is flat, the process is assumed to be a white noise process, whereas an increasing or decreasing function indicates seasonal variation.

A homogenous series is a nonstationary series which can be made stationary through differencing. The order of the homogeneity is determined by the number of times a nonstationary series must be differenced in order to make it stationary (Cleary). Stationarity is important because a statistical model can be built to describe a stationary series and used to generate a forecast.

The most commonly used time-series forecasting procedure was developed by Box-Jenkins. The Box-Jenkins methodology is a procedure that systematically allows the identification, estimation, and diagnostic checking of time series models like the autoregressive, AR(p), and moving average, MA(q), models (eg. equations 19 and 20).

Moving Average model

(19) $MA(q)\ Y_t=\mu_t+\epsilon_t+\theta_1\epsilon_{t-1}+\theta_2\epsilon_{t-2}+\ldots+\theta_q\epsilon_{t-q}.$

q = the order of the process.

θ = the parameter estimate.

μ_t= the mean of the process.

Autoregressive model

(20) $AR(p)=Y_t=\phi_1 Y_{t-1}+\phi_2 Y_{t-2}+\ldots+\phi_p Y_{t-p}+\delta+\varepsilon_t.$

ϕ = parameter estimate.

δ = a constant term representing the series trend.

p = the number of lagged periods.

The first step in specifying a time-series model is the identification of the series. Identification of the series involves determining the order, the number of lagged disturbances, and/or the number of autoregressive terms to be included in the model. The ACF and PACF are used to determine at which point the correlation between the dependent variable and past disturbance and/or autoregressive terms truncates. The partial autocorrelation function measures the indirect relationship between two variables regardless of the direct relationship found between them. For example, if there are three variables: A, B, and C, assume that there is a direct and significant relationship between A and C and a direct and significant relationship between A and B. No direct relationship between B and C exists, but there is a relationship between B and C due to the relationships found between A and C and A and B. The PACF measures this indirect relationship (e.g. equation 21).

Partial autocorrelation function

(21) $$PACF=\phi_{kk}=\frac{\varrho^*_k}{\varrho_k}.$$

ϱ_k = matrix of autocorrelation coefficients, obtained from the matrix of Yule-Walker equations with diagonal elements as unity.

ϱ^*_k= matrix of autocorrelation coefficients, obtained from the Yule-Walker equations, except that the last row contains a column vector of all the autocorrelation coefficients.

The PACF is derived from the ACF function and is used to determine the order of the autoregressive process. The variable p represents the order of the autoregressive process, while the variable q represents the order of the moving average process. The researcher must first decide which modeling technique, AR(p) or MA(q), most accurately describes the series under consideration. Identification of a series, or determining the (p,q) values, is accomplished by inspection of the ACF and PACF functions.

The Box-Jenkins methodology assumes the data series to be used in generating a forecast is stationary or homogenous nonstationary. The series is then visually inspected to determine if trends and/or seasonality are present. If trends or seasonality are present, they must be removed before the Box-Jenkins procedure can be used. Once the model has been identified, it is estimated using a nonlinear least-squares procedure. Because of autocorrelated error terms in the moving average process, the assumption of no autocorrelation in OLS is violated and cannot be used to estimate the model. Estimates of the variance and covariance are also needed so significance tests can be used to evaluate the model.

Diagnostic checking is the last step in the Box-Jenkins procedure and provides insight into the accuracy and specification of the forecasting model. Examination of the residuals and the statistical properties of the residuals are used in diagnostic checking to test the statistical significance of the estimated parameters. There are many types of diagnostic testing and evaluation criteria commonly used in evaluating the forecast and model. (These various evaluation criteria are discussed in further detail in the forecast evaluation criteria section of this dissertation). One method of determining if the model is correctly specified is by over fitting the model. This requires adding an additional variable(s) to the identified model and testing the significance of the additional variable(s). The estimated variance should decrease if the additional variables improve the model's fit and provides an indication of whether the model is correctly specified. A chi-squared statistic is calculated and used to test the significance of the estimated ACF to determine if any systematic error is present in the residuals. Diagnostic checking of the data series is also important to determine if the process was generated by a white noise process (Cleary).

Equation (19), is a deterministic model that uses the weighted averages of past error terms plus the mean of the moving average process to predict future events. The correlation coefficient, or autocorrelation function, is a measure of the interrelationship between a variable and its own past values. In regression analysis, the correlation coefficient is a measure of the causal effect of an independent variable on the dependent variable. The statistical model assumes the observed time series is influenced by its past values.

Equation (18) is calculated for different lags and used to determine the order of a pure moving average process. The lag number for which the ACF truncates or approximates zero determines the order of the MA(q) process. For example, if the ACF truncates after two lags, the moving average process is identified as being second order, or MA(2). Inspection of the standard normal value calculated for the autocorrelation coefficient, is used to test the null hypothesis that the ACF coefficients are equal to zero. The statistical significance of the estimated mean and coefficients can be evaluated using the t-statistic at a specified level of significance.

The chi-square test can be used to test the specification of the model. Specifically, the chi-square test is used to test the null hypothesis that the estimated autocorrelation coefficients are equal to their actual values. A significant chi-square value indicates that one or more of the ACFs is significant. The estimated ACF value and its corresponding standard errors, one for each lag, are need to determine the order of the moving average process. The last significant ACF value determines the order of the moving average process. (Nazem)

The MA(q) model is estimated using a recursive process. The error term is estimated by subtracting the actual from the estimated dependent variable values. The forecast is then generated by multiplying the parameter estimate by the estimated error and subtracting this from the mean of the process. The MA(q) model is capable of forecasting ahead q periods. This process continues to generate each q period ahead forecast as soon as the previously forecast value is known.

Equation (20) is a nondeterministic statistically based time-series model. The dependent variable in the AR model is a function of its weighted past values, a disturbance term, and the error term. It very closely resembles a multiple linear regression model because the dependent variable is regressed on its past values.(Cleary)

p represents the order of the process which is determined by first inspecting the ACF to determine if the process is a pure AR process. Inspection of the PACF indicates whether the process is autoregressive of order p. PACF values are calculated for different lags. When the PACF truncates or becomes zero, the lag beyond the one that caused the truncation of the PACF is assumed to exhibit no true relationship between the variable value and its past values. p is the lag number that caused the PACF to truncate and indicates the order of the AR process. For example, if the PACF truncates after three lags, the autoregressive process is AR(3).

If the dependent variable is a function of all the past and current values of an independent variable (e.g. equation 22), further adjustment is required.

(22) $y_t=b_0x_t+b_1x_{t-1}+...+b_px_{t-p}+\varepsilon_t.$

Difficulties arises because of collinearity between the lagged values of the independent variable or because of the large number of regressors in relation to the number of degrees of freedom. By assuming these distributed lag coefficients emulate a specific pattern, these problems may be avoided. The Koyck distributed lag specification, equation (23), is a commonly used procedure that assumes the coefficients decline geometrically.

Koyck distributed lag specification

(23) $y_t=\beta x_t+\beta\lambda x_{t-1}+\beta\lambda^2 x_{t-2}+\ldots+\beta\lambda^k x_{t-k}+\varepsilon_t.$

Mathematical manipulation of equation (23) allows the dependent variable, lagged one period, to be described as equation (24).

(24) $\lambda y_{t-1}=\beta\lambda x_{t-1}+\beta\lambda^2 x_{t-2}+\ldots+\lambda\varepsilon_{t-1}.$

This produces a manageable autoregressive equation (e.g. equation 25).

(25) $y_t=\lambda y_{t-1}+\beta x_t+(\varepsilon_t-\lambda\varepsilon_{t-1}).$

There are a variety of statistically-based models that combine moving averages and autoregressive processes in order to forecast a data series. One composite model is the autoregressive moving average, ARMA, process of order (p,q). Equation (26) contains both moving average and autoregressive components.

Autoregressive Moving Average model

$$(26) \quad ARMA(p,q)\; Y_t=\phi_1 Y_{t-1}+\phi_2 Y_{t-2}+\ldots+\phi_p Y_{t-p}+\delta+\epsilon_t+\theta_1\epsilon_{t-1}+\theta_2\epsilon_{t-2}+\ldots+\theta_q\epsilon_{t-q}.$$

An ARMA model generally contains different orders for each of the different components. Inspection of the ACF and PACF, as well as their respective standard errors, are used to determine p and q. A crucial assumption in generating an ARMA forecast is that the data series used in estimating the model has to be stationary.

The ARMA(p,q) can rewritten using the back shift operator B for simplicity (e.g. equation 28). The operator B imposes a one period time lag every time it is applied to a variable (e.g. equation 27).

$$(27) \quad B_t^n = \varepsilon_{t-n}.$$

ARMA(p,q) with Back shift operator

$$(28) \quad \phi_p(B)\,y_t = \delta + \theta_q(B)\,\varepsilon_t.$$

$$\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p.$$

$$\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q.$$

If a data stationary series y_t is a homogeneous nonstationary series, it can transformed into a stationary series by differencing the series one or more times. Equation (29) is a general representation.

Homogeneous nonstationary series of order d

$$(29) \quad z_t = \Delta^d y_t.$$

$$\Delta y_t = y_t - y_{t-1} = z_t.$$

d represents the number of times a homogenous series must be differenced to make it stationary. For example, if a series has to be differenced twice to make it stationary it could be represented by equation (30).

$$(30) \qquad z_t = \Delta^2 y_t.$$

After transforming y_t into a stationary series z_t, where z_t is an ARMA(p,q) process, it is referred to as an integrated autoregressive moving average or ARIMA model of order (p,d,q) (e.g. equation 31).

Integrated Autoregressive Moving Average model

$$(31) \qquad ARIMA(p,d,q)\ Y_t=\phi_1 Z_t+\phi_2 Z_{t-1}+\ldots+\phi_p Z_{t-p}+\delta+\epsilon_t+\theta_1\epsilon_{t-1}+\theta_2\epsilon_{t-2}+\ldots+\theta_q\epsilon_{t-q}.$$

$$Z_t = Y_t - Y_{t-d}.$$

The ARIMA model differs from the ARMA model in that it incorporates a differencing component (d) to obtain homogenous stationary data. p and q determine the order of the autoregressive and moving average components, and d represents the number of times a homogenous series must be differenced to make it stationary.

The ARIMA(p,d,q) can be rewritten using the backshift operator (e.g. equation (32).

ARIMA using backshift operator

$$(32) \qquad \phi_p(B)\,Z_t = \delta + \theta_q(B)\,\varepsilon.$$

The constant, δ, can be eliminated by adjustment $\tilde{y} = y_t - \delta/(1-\Phi_1 \cdots - \Phi_p)$.

If there is a system of MA(q) equations to be estimated a VMA(q) model can be used. The VMA(q) is obtained by simply representing the MA(q) model in a vector generalization (e.g. equation 33).

VMA model

(33) $VMA = \theta_q(B)\varepsilon_t.$

A variant of the ARIMA model is the vector autoregressive moving average (VARMA) model (e.g. equation 34). The VARMA (p,q) model is simply a vector generalization of the univariate ARIMA model.

VARMA model

(34) $VARMA = \phi(B)y_t = D + \theta(B)\varepsilon_t.$

$\phi = 1 - \phi_1 B - \ldots\ldots - \phi_p B^p.$

$\theta = 1 - \theta_1 B - \ldots\ldots - \theta_q B^q.$

D = a vector of constants which relates to the mean of the process.

ϕ_p, θ_q = matrices of the estimated AR(p) and MA(q) coefficients.

The VARMA is identified using the sample cross-correlation matrices and the partial autocorrelation function.

The seasonal integrated autoregressive moving average model (SARIMA) is a variation of the ARIMA forecasting technique and is of order (P,D,Q) (e.g. equation 35).

SARIMA model

(35) $$SARIMA(P,D,Q)\ Y_t = \phi_s Y_{t-s}+\ldots+\phi_{ps}Y_{t-ps}+\delta+\epsilon_t+\theta_s\epsilon_{t-s}+\ldots+\theta_{qs}\epsilon_{t-qs}.$$

Seasonality may be broken down into two components. For example the supply of corn in the United States contains two forms of seasonality. The first is recurring seasonality and may be measured with some certainty. The supply of corn is large after harvesting and is assumed to be so every year. The supply of corn is also influenced by random factors such as the weather. The first type of seasonality can be removed from the series by differencing, but the second form of seasonality, the stochastic seasonal component, needs to be captured by statistical parameters. SARIMA uses unadjusted data, meaning the stochastic seasonal component is not removed from the data series before being used in generating a forecast. If seasonality is important, it may be useful to generate a forecast using the SARIMA forecasting technique. Seasonal adjustment enables the underlying trends and cycles to be forecasted more precisely. In some instances, unadjusted data are needed so that forecasts of actual values can be recovered. (Cleary)

Another time-series forecasting technique is the Holt-Winters procedure (e.q. equation 36), which is a generalization of exponential smoothing to incorporate a trend and seasonal variation.

Holt-Winters model

(36) $$Y_{t+j}=(\mu_t+\beta_t j)\,S_{t+j}+\varepsilon_{t+j}.$$

This procedure requires that the mean (μ), slope, and seasonal factors (S) of a process be calculated. A set of smoothing parameters is used to generate estimates of the mean, slope, and seasonal factors. The model minimizes the sum of the squared one-step-ahead forecast errors. (Rashivar)

C. Composite Forecasting Techniques

There is variety of forecasting methods that combine different forecasting techniques in an attempt to increase forecast accuracy. Combining theoretically and statistically based models allows the introduction of more information into a forecast than the use of a single method. The theoretical components provide an explanation as to why a data series occurs, while the statistical components try to replicate the series using past values of the dependent variable.

In combining different forecasting techniques, a weighting system is employed to produce the composite forecasts. There are several ways of assigning weights to the different components of a combination forecast. One method is simple weighting which combines different forecasting techniques by equally weighting them in the composite model. This method is commonly used when there is no prior forecast accuracy information available. If there was a forecast history, logic would dictate that the most accurate forecasting technique receive a greater weight in producing the composite model. Another method, adaptive weighting, assumes that forecasting techniques may not perform consistently over time and that the weights assigned to each forecasting technique may need to change.

Adaptive weighting uses the forecast error of each forecasting model as a basis for determining the weight assigned to the individual forecasting techniques. The procedure is adaptive in that forecasts can change over time with the accuracy of the individual forecasting techniques. A third method is the minimum variance weighting procedure. It assumes a constant forecast performance from each model. The weights for each model are determined by assigning weights to the forecasting model in a manner that minimizes the forecast error of the composite model over the historical period.

The Bayesian composite forecasting technique is similar to the simple average and adaptive forecasting techniques in that it assigns different weights to the individual components of the composite model. The weights are assigned with regard to the forecaster's prior belief about the accuracy of the individual forecasts and the relative performance of the individual components over time. This forecasting technique allows the forecaster to have a greater influence on how the individual forecasts are weighted in the composite forecasting procedure compared to other methods. The researcher assigns prior weights to each of the individual forecasting techniques contained in the composite model. If there are multiple forecasts, the priors may be arranged in a m by m matrix of pairwise beta distributions containing elements that measure the likelihood of one forecast out performing all others contained in the composite model. The beta distribution is depicted by two parameters that provide the mean and variance of the distributions. The distribution is depicted by m, the number of forecasts being examined. This priors matrix is then merged with a performance matrix which contains rankings of the performance, based on trial forecasts, on how well each of the individual forecasts performed compared to the others. The combination of the two original matrices yields a matrix beta from which weights for each of the individual forecasts are assigned (McIntosh and Bessler).

The time-varying coefficient-autoregressive (TVC-AR) model is a forecasting technique in which the movement of the coefficients, as well as the initial conditions, are previously specified to depend on a two dimensional vector of parameters (e.g. equation 37).

Time-varying coefficient- autoregressive model

(37) $$y_t=Y_t b_t+c_t+\varepsilon_t.$$

Y_t = vector of lagged dependent variables.

C_t = time varying intercept.

ϵ_t = the uncorrelated conditional Gaussian disturbance term.

The TVC-AR forecast has four advantages over the Bayesian VAR forecast (Carnova). The first advantage is that the TVC-AR technique does not impose beta priors for the parameter values and their underlying distributions. Secondly, the TVC-AR technique does not impose a constraint on the frequency domain. The third advantage of the TVC-AR technique is it can endogenously account for several forms of seasonality and their evolution over time. The fourth advantage is the TVC-AR technique's ability to account for the uncertainty present when specifying the functional form of the seasonality.

The Swamy-Tinsley (S-T) or stochastic coefficient model is a composite forecasting technique which combines theoretical and statistical models (e.g. equation 38). A stationary process is described by deterministic stochastic and random indeterministic components. The statistical model tries to reproduce the indeterministic components of the process. The theoretical model describes the endogenous variable as a function of a conditional mean function and a nonstationary error term which follows an ARMA process.

(38) $y_t = x_t' b_t.$

y_t = univariate, nonstochastic process in period t.

x'_t = a k component vector of fixed independent variables in period t.

b_t = unobservable coefficient vector which contains an additive disturbance term.

b_t consists of an additive disturbance term and a time dependent intercept and can be represented as equation (39).

(39) $b_t = \bar{b} z_t + \varepsilon_t.$

$\bar{b}$ = matrix of fixed parameters.

z_t = vector of fixed variables, the intercept, deterministic trend and seasonal variables.

$$\varepsilon_t = \sum_{i=1}^{p} \phi_i \varepsilon_{t-i} + \sum_{j=1}^{q} \theta_j \alpha_{t-j} + \alpha_t.$$

If b_t is stationary, the function can be written as equation (40).

(40) $b_t = b_t^d + \varepsilon_t.$

b^d_t = the deterministic component of the process.

ϵ_t = the stochastic component.

α_t = a random k-variate innovation process.

Substituting equation (39) into equation (38) yields the S-T model (e.g. equation 41).

Swamy-Tinsley

(41) $y_t = x_t' \bar{b} z_t + x_t' \varepsilon_t.$

The S-T model allows each of the coefficients in the stochastic model to vary around its mean by an error term that is correlated with it's past error term as well as the past error terms of the other coefficients. If $\bar{b}_t$ is equal to zero, the technique reduces to a univariate ARMA model. Each of the error terms is assumed to have a white noise component that is currently correlated with the white noise components of the other error terms (Conway).

The autoregressive components (AC) model is a variant of the AR(p) model. It is an autoregressive model that incorporates some economic variables that are considered determinants of the dependent variable so that it is a composite model. The model contains variables that are considered to be determinants of the dependent variable along with the time series information provided by the autoregressive component (Fair and Shiller).

A transfer function, or multivariate integrated autoregressive moving average, MARMA, model is a composite forecasting model (e.g. equation 42).

Transfer function model

(42) $$y_t=b_0+b_1x_{1t}+b_kx_{kt}+\theta_q^{-1}(B)\ \phi_p(B)\,z_t.$$

ϕ_q= autoregressive backshift operator.

θ_p= moving average backshift operator.

z_t = the transformed data series.

It contains an econometric forecasting model with a time-series ARIMA model. The model describes the dependent variable as a function of explanatory variables, lagged values of the dependent variable, and a time-series generated error term. The transfer function model uses the econometric structural model to describe the variation found in the dependent variable. The error term of the structural model represents the unexplained variation present in the dependent variable. Forecasting with the structural model is accomplished by plugging predicted values of the dependent variable into the estimated equation. The transfer function model contains economic information that explains part of the variation in the dependent variable plus a time-series component to explain the unexplained variance present in the dependent variable. The logic associated with using a transfer function model is that this modeling technique incorporates more information with the objective of generating a superior forecast when compared to each of the individual models. (Pindyck and Rubinfeld). The problem is that there is no way to include the true error term in the model. The transfer function model uses an ARIMA model to forecast the error term. The ARIMA model is then inserted into the structural model in place of the error term.

Helmer and Johansson describe how the Box-Jenkins procedure can be used in specifying a transfer function model. There are nine substeps that fall under the three general steps, identification, estimation and diagonsitc checking.

The identification procedure consists of seven sub steps. The first procedural step of the Box-Jenkins identification procedure requires some prior knowledge so the appropriate functional form can be used to avoid the problem of misspecification. The second through seventh procedural steps contained within the identification phase are used to identify the transfer function. The second procedural step involves prewhitening the series to remove all the systematic or predictable components present in the data series. The third procedural step differences the data to remove the variation caused by seasonality, trends, and other causes of nonstationarity. The fourth step uses the prewhitening procedure to ensure the correlation between the independent and dependent variables remains in the series.

Step five analyzes the estimated dependent and independent relationships from the second and fourth steps, respectively. The correlation between the dependent and independent variables is taken to represent the response of the independent variables on the dependent variable. The correlation is then used as a measure of the impulse response coefficient. The sixth step uses a parsimonious transfer function to substitute the estimates of the direct effects of the independent variable on the dependent variable. This step generates at least one model used to describe the relationship(s). The seventh step regards the residual series generated from step five as being a distinct time series. A Box-Jenkins univariate time series model is used to identify the transformation which is responsible for removing the situational and remaining unspecified factors referred to as noise. This step generates at least one model to describe the residual series.

The eighth step, or the second general step, involves the estimation of the models. A maximum likelihood procedure is used to derive estimates of the independent variable coefficients as well as estimates of the autocorrelation and moving average coefficients. The ninth and final step, referred to as the third general step, is the diagnostic checking of the estimated transfer function models. This step selects the transfer function model that best describes the time series.

B. Forecast Evaluation Criteria

A wide range of evaluation criteria is available to measure the accuracy of a forecast. An ex post forecast uses the first subset of data from a historical data set, which has been divided into two periods, to generate a model. The model is then used to generate a forecast using the second subperiod. The comparison of the predicted to the remaining actual values is used as a measure of forecast accuracy. These data can then be used in the measures described in this section.

Forecasts are commonly evaluated by their abilities to predict turning points in a time series. Kaylen argues that this particular evaluation criterion should be taken one step further. The ability of a forecast to predict turning points should be redefined to include whether a forecast is capable of accurately predicting peaks, troughs, and no-turns. Based on the general evaluation, if a peak is forecast by the model instead of the trough that actually occurred, the model accurately predicted a turning point. The ability of a forecast to predict turning points under Kaylen's definition would provide more useful information because it indicates the direction the series is moving instead of just the fact that the series in changing direction.

Kaylen re-evaluated the forecasts reported in the studies of Brandt and Bessler (1981), Brandt and Bessler (1984), and Harris and Leuthold (1985) using his new definition of accurately predicting turning points. He found that using actual observations in place of previously forecasted observations with his new definition of turning points actually improved the forecast accuracy of the different models.

A simple method of choosing a forecast among alternatives is the parsimony rule. The parsimony rule implies that when a tie exists between two comparable models, the model with the least number of estimated coefficients should be chosen. This reduces the amount of information necessary to generate a forecast which should provide a better forecast. The rule is consistent with the scientific method. (Cleary)

The bias criterion is a measure of the mean difference between the actual and forecasted values of the dependent variable. It is a criterion used to determine if there is a tendency to over or underestimate a series. A related criterion is the comparison of the means of the forecasted values to the means of the actual series.

The following set of evaluation criteria is based on the error component of the forecast. These criteria use the error terms generated by a forecasting model to evaluate forecast accuracy. The standard error test is used to determine the significance of an error estimate (e.g. equation 43). T equals the number of observations in the forecast.

Standard Error

$$(43) \qquad \hat{\sigma}_Y = \sqrt{\frac{\sum_{t=1}^{T}(Y_t - \hat{Y}_t)^2}{T - 2}}.$$

The standard error is a measure of how the estimates are dispersed around their means and is used in calculating the t statistic. A critical value is used to test null hypothesis that each of the estimated coefficients is not significantly different from zero. If the computed t-value is large enough, the estimate is statistically significant at the specified level of significance.

The sum of squared errors, SSE, provides a measure of accuracy between the actual and predicted values of the dependent variable (e.g. equation 44).

Sum of Squared errors

$$(44) \qquad SSE = \sum_t \varepsilon_t^2.$$

Mean error, ME, is provides a measure of the average size of the model error or the difference between the actual and predicted dependent variable values (e.g. equation 45).

Mean Error

$$(45) \qquad ME = \frac{1}{T}\sum_t \varepsilon_t.$$

The mean squared error (MSE) criterion provides a measure of the individual forecast errors. A small MSE,(e.g. equation 46), indicates a good fit or accurate forecast.

Mean Squared Error

(46) $$MSE=\frac{\sum(\hat{Y}_t-Y_t)^2}{T}.$$

The mean absolute error (MAE) is another measure of the size of the forecast error (e.g. equation 47), and represents the absolute difference between the actual and forecasted errors. The smaller the MAE is, the better the forecast.

Mean absolute error

(47) $$MAE = \frac{1}{T}\sum_{t=1}^{T}|\hat{Y}_t-Y_t|.$$

The mean absolute percentage error (MAPE) (e.g. equation 48), criterion expresses the forecast error as a percent of the actual value of the variable.

Mean Absolute Percentage Error

(48) $$MAPE=\frac{1}{T}\sum_{t=1}^{n}\frac{|\hat{Y}_t-Y_t|}{Y_t}.$$

Another common measure of the deviation is the mean percentage error, (MPE) (e.g. equation 49).

Mean percentage error

(49) $$MPE = \frac{1}{T}\sum_{t=1}^{T}\frac{\hat{Y}_t-Y_t}{Y_t}.$$

The root mean square error (RMSE) criterion provides insight into the difference between the actual and forecasted values of the dependent variable (e.g. equation 50). The square root of the mean squared error term is compared to the observed dependent variable values to provide an indication of the forecast accuracy.

Root Mean Square Error

(50) $$RMSE=\sqrt{MSE}.$$

The RMSE is the square root of the MSE. It shows how much on average the predicted values deviate from the actual values. The percentage root mean square error is similar to the RMSE except that it expresses the forecast error in a percentage term (e.g. equation 51).

Root mean percentage error

(51) $$RMPE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\frac{(\hat{Y}_t-Y_t)^2}{(Y_t)^2}}.$$

The squared correlation criterion is a measure of the correlation between the actual and predicted value for the dependent variable. The closer this value is to one, the better the forecast fit.

The coefficient of variation, V, is a measure of the relative dispersion between series and is calculated by dividing the standard deviation of the series by the mean of the series (e.g. equation 52).

Coefficient of Variation

(52) $$V=\frac{[\Sigma(Y_i-\bar{Y})^2/T^2]^{1/2}}{\bar{Y}}.$$

The resulting value provides a criterion for comparing the relative variation between two or more series.

The Chi-squared statistic, Q, is used to determine whether a residual series is nonwhite. The null hypothesis states that the series is nonwhite, meaning that the series has an implied structure, and the estimated model is inadequate (e.g. equation 53).

Chi-Squared Statistic

(53) $$Q=T\sum_{k=1}^{K}\hat{\varrho}^2{}_k.$$

$\hat{\varrho}_k$ = the estimated autocorrelation function equation (18).

A large Q or Chi-squared statistic means that residual autocorrelation is significant in the series. The Chi-squared test is used to test the null hypothesis that the estimated autocorrelation coefficients are equal to the actual values. The autocorrelation coefficients are nonzero until q. After point q, the estimated autocorrelation coefficients are be equal to zero.

A common criterion known as Theil's inequality coefficient, U, provides an index of relative forecast accuracy (e.g. equation 54). This test is based on the ratio of the forecast RMSE to the no change forecast RMSE. Theil's U can take on any nonnegative value. When the U=0, the forecast RMSE is zero meaning the forecast is perfect. If $U \geq 1$ the model is a poor predictor because at U=1, the model predicts as well as the nochange model.

Theil's inequality coefficient

$$(54)\qquad U^2 = \frac{\frac{1}{T}\sum_{t=1}^{T}(\hat{Y}_t - Y_t)^2}{\frac{1}{T}\sum_{t=1}^{T}(\hat{Y}_t)^2 + \frac{1}{T}\sum_{t=1}^{T}(Y_t)^2}.$$

$\hat{Y}$ = estimated dependent variable value.

Y = actual dependent variable value.

The Theil's U can be decomposed into three different proportions. The first proportion is the bias, U^M, which provides an indication of how much the actual and simulated averages deviate from each other (e.g. equation 55).

The Bias Component of Theil's Coefficient

$$(55)\qquad U^M = \frac{(\bar{\hat{Y}} - \bar{Y})^2}{(1/T)\sum(\hat{Y}_t - Y_t)^2}.$$

$\bar{\hat{Y}}$ = mean estimated dependent variable value.

$\bar{Y}$ = mean actual dependent variable value.

T = number of time periods.

A U^M larger than .2 indicates the presence of bias.

The second proportion represents the variance, U^S, and it is an indication of how well the model is able to simulate the variability found in the dependent variable (e.g. equation 56). A large U^S value suggests that the model is incapable of simulating the variation found in the dependent variable.

The Variance Component of Theil's Coefficient

(56) $$U^{S}=\frac{(\hat{\sigma}-\sigma)^{2}}{(1/T)\sum(\hat{Y}_{t}-Y_{t})^{2}}.$$

$\hat{\sigma}$ = estimated standard deviation.

σ = actual standard deviation.

The third proportion is the covariance, U^{C}. Equation (57) provides a measure of the remaining error present in the forecast after the deviations from the average values and variations have been considered. A large U^{C} suggests a good correlation between the forecasted value and the actual value of the dependent variable. The ideal values for the U^{M} and U^{S} would be zero with the U^{C} equalling one.

The Variability Simulation component of Theil's coefficient

(57) $$U^{C}=\frac{2(1-\varrho)\hat{\sigma}\sigma}{(1/T)\sum(\hat{Y}_{t}-Y_{t})}.$$

ϱ= the correlation coefficient.

The Gauss-Seidel procedure is an iterative process which is used to derive estimates of the goodness of fit measures for a simultaneous equation system. There are two variants of the Gauss-Seidel procedure. The first derives the estimated endogenous variable estimates using historical values of the independent variables and predetermined endogenous variable values. This information is then used to calculate different goodness of fit measure like the MAPE, squared correlation between the estimated and actual dependent variable values, and the Theil's statistic. The second uses historical endogenous variable values which are substituted in place of the estimated values in the first option.

The Akaike Information Criterion (AIC) can be used to compare statistical models containing different numbers of parameters (e.g. equation 58). According to the AIC, the order of the statistical process should be chosen so that the AIC estimate is minimized.

Akaike Information Criterion

(58) $AIC = \tilde{\sigma}^2 \exp[2(n+d)/T]$.

d = the number of nonstationary elements in the state vector.

It has been shown that the AIC can overestimate the order of the autoregressive process, and a Bayesian Information Criterion (BIC) is a better method of determining the order of the autoregressive process. The BIC (equation 59) assesses a greater penalty per parameter.

Bayesian Information Criterion

(59) $BIC = \tilde{\sigma}^2 \exp[\log T(n+d)/T]$.

The BIC is diverges from the AIC in that it multiplies the AIC by one-half the log of T, which means that it leans more toward models with fewer parameters (Schwarz). As with the AIC criterion, the model with the smallest BIC estimate should be chosen.

Ashely developed a theorem to defend his logic of eliminating forecasted variables from econometric models. If a model is misspecified, the variance of an estimated explanatory variable ($\hat{x}$) may be larger than its true value. Therefore, the MSE of a dependent variable forecast in which the forecasted explanatory variable is excluded may not be larger than the MSE of a dependent variable forecast in which the forecasted explanatory variable is included. Ashely developed a test where the MSE $(\hat{x})$ is divided by the Var (x_t). If this ratio is larger than 0.7, he concluded that including the forecasted explanatory is not guaranteed to increase forecast accuracy.

Chapter III

Literature Review

A. Introduction

This review of pertinent forecasting literature describes studies that provide insight into the application and performance of the various forecasting techniques and their uses in food demand research. These studies evaluated and compared the various forecasting procedures and determined which techniques provided the most accurate forecasts for the given situation. The information obtained in reviewing the following articles is instrumental in determining the most suitable type of forecasting technique(s) and evaluation criteria to forecast weekly supermarket sales of food groups. The literature contained with this section employed a wide range of forecasting techniques, from the naive nochange to the intricate composite model.

B. Forecasting Studies

Leitch and Tanner discuss the logic behind firms spending millions of dollars on economic forecasting in spite of their large forecast errors. The authors compared simple forecasts with those provided by a forecasting service to determine which generated a better forecast. Summary statistics used to evaluate and compare forecasts in this study are the: MAE, RMSE, and Theil's U.

The study forecasted the interest rate for a three-month treasury bill using six different forecasting techniques: (1) a professional service; (2) an ARIMA model; (3) a forward rate model; (4) a nochange model; (5) a constant rate of growth model; and (6) a survey forecast. Forecasts were generated using monthly data which started in January of 1982 and continued through December of 1987. The six different forecasting techniques were used to predict directional changes in the interest rate. These four interest rate forecasts were (1) whether the interest rate would simply increase or decrease; (2) whether the forecasted interest rate was larger or smaller than the futures market interest rate; (3) whether the interest rates were expected to change; and (4) whether the forecasted interest rate moved positively or negatively compared to the futures market interest rate. They were used to determine the position a firm should take in the treasury bill market. Profits generated from each position were calculated by subtracting the transaction costs from the revenues generated by each method.

The most accurate forecasting methods were the nochange, forward rate, and ARIMAs given the summary statistics. Forecasts were also evaluated on their abilities to generate profits. The most profitable forecast was generated by the professional service forecast, given the revised evaluation criterion. Various correlations were calculated which revealed that the only forecast that was correlated with profitability was the directional accuracy measure. The size of the MAE and RMSE were inversely correlated to profits in three out of the four times. Theil's U was positively correlated with profits but was not statistically significant three out of the four times.

The authors concluded that the forecast error criterion did not provide an accurate indication of a forecast's ability to generate profits. The research found no systematic relationship between profits and the size of the forecast error, and the authors suggested that the only significant evaluation criterion is directional accuracy which is attributed to the strong relationship found between directional accuracy and profits. Thus, firms are acting rationally when they purchase economic forecasts. The authors also suggested that least squares may not be an appropriate estimation technique in conducting economic research because the results indicate that there is no systematic relationship found between the error criteria and profitability.

Seven ex post forecasting techniques were used by Vere and Griffith to forecast three important variables in the South Wales prime lamb market: slaughterings, real saleyard prices, and per capita consumption of lamb. The forecasting techniques were the (1) single-equation regression, (2) ARIMA, (3) structural, (4) restricted reduced form, (5) composite, (6) NSWMPFC (New South Whales Meat Production Forecasting Committee), and (7) nochange models. Both dynamic and static forecasts were generated and compared based on applicability and forecast accuracy. The accuracy of each forecast was evaluated according to the MSE, MAE, Theil's U, and directional accuracy. Ex post forecasts were used to compare the accuracies of each different forecasting technique.

Quarterly data starting with the first quarter of 1969 and ending in the last quarter of 1984 were used to generate the various forecasts. Twelve periods starting in the first quarter of 1985 through the fourth quarter of 1987 were included to generate the ex post forecasts. The exception was the NSWMPFC forecast which used data through September of 1987.

The most accurate forecasts of slaughterings and consumption were provided by the single-equation regression and the composite models, for both the static and dynamic forecasts. The ARIMA model provided the most accurate forecast of the real saleyard prices. Evaluation of each individual forecast revealed that no one model generated superior forecasts for each of the variables under study. The authors suggested that combining different forecasting methods may lead to an increase in forecast accuracy.

Harris and Leuthold conducted a study to evaluate and compare the ex post forecast accuracies of five short-run livestock, live cattle and hog, price forecasting methods using quarterly data from the first quarter of 1969 continuing through the last quarter of 1980. The authors compared the individual four quarter ahead forecasts with composite forecasts to determine if a gain in forecast accuracy was realized by combining individual forecasting techniques. The five different forecasting techniques were (1) single-equation econometric, (2) univariate Box-Jenkins analysis, (3) composite forecast in which the econometric and univariate models are equally weighted, (4) integrated model referred to as the serial correlation regression model which combined an econometric and an ARMA model for the residual series, and (5) integrated multivariate time series model.

The different forecasting techniques were compared and evaluated based on the RMSE and directional accuracy criteria. The single-equation price dependent cattle and hog models were built by the University of Illinois Price Forecasting and Sales Management team (PFSM) and were estimated using OLS. The composite forecast consisted of equally weighted

ARIMA and econometric forecasting models. The serial correlation model was an integrated model that combined the structural model provided by the PFSM and an ARIMA model used in forecasting the residual series. The structural portion of the model was estimated using GLS.

The ARMA and ARIMA forecasts of live cattle prices were the most accurate based on the root mean square error evaluation criterion, whereas none of the models generated a superior hog price forecast. Changing the evaluation criterion to the models' abilities to forecast turning points correctly, the results indicated that the models with the lowest RMSE did not outperform the other models having larger RMSEs. Contrary to theoretical thought that combining forecasting techniques increases forecast accuracy, the results indicated that the individual forecasting methods and the serial correlation regression model performed as well as the integrated models for forecasting cattle and hog prices. The ARIMA model provided a superior forecast for predicting live cattle and hog prices, using the RMSE and turning point criteria, when compared to the integrated models. The results revealed that the integrated models provided an inferior forecast, except for one interval, when compared to the ARIMA model forecasts. Two possible explanations for this particular result were that the single-equation demand model was misspecified or that the demand for beef shifted in the 1980s.

Fair and Shiller examined structural and VAR models with the intention of comparing the two forecasts. The structural econometric models contained a large number of predetermined variables. These models were usually restricted in that in some instances the predetermined variables exceeded the number of observations used in the generation of the forecast. The VAR and AC models, on the other hand, have fewer variables thus requiring less information in generating a forecast than the structural econometric model. Six VAR and eight AC forecasting models were each evaluated and compared to the Fair structural model.

Instead of using the RMSE criterion for evaluating the two forecasts of real GNP, the actual change in each model was regressed on the forecasted changes. This procedure allowed the researchers to draw an inference about whether the larger structural models contained the same information found in the smaller atheorectical VAR and AC models. The evaluation was performed on quasi ex ante forecasts. The exogenous variables found in the structural model were replaced by autoregressive equations to eliminate all of the exogenous variables in the model. The model and exogenous variables were estimated through t-1 for each new forecast period t, which is referred to as a rolling estimation procedure. The rolling estimation procedure was used to estimate the structural, VAR, and AC forecasts.

The quarterly data used to generate the one-quarter and four-quarter-ahead forecasts ranged from 3(1976) through 2(1986) and from 2(1972) through 2(1986), respectively. The results suggested that there was a significant difference in the information contained in the structural econometric, VAR, and AC models. The results also suggested that the Fair structural model provided a better forecast than either the VAR or AC models.

Paul et al constructed a multi-equation end use model of Australia's demand for sawntimber. A three equation demand model was estimated simultaneously by the maximum likelihood method. The three demand equations were dwelling starts, sawntimber used per housing start, and sawntimber used in nondwelling purposes. These three variables were felt to be the most significant in determining the consumption of Australian sawntimber.

The data used in this study were obtained from various Australian Bureau of Statistics publications and covered the period 1951-1986. Value and price data were deflated using the consumer price index. The evaluation of both static and dynamic forecasts was based on the ability to predict the actual series accurately. The static simulation, where the endogenous variables were not updated annually as they became available, was estimated using data from 1951-1986. The dynamic simulation, which incorporated previously estimated values of the endogenous variables, was estimated using data from 1972-1986. Theil's U associated with the demand function for total sawntimber suggested that there might have been an error related to using the sawntimber used per dwelling equation. The model was respecified and estimated using actual wood per dwelling instead of the simulated variable for the same time period.

The RMSE, RMPE, and Theil's U suggested that the dynamic forecast in which actual wood per dwelling was substituted for sawntimber provided a suitable forecast of total sawntimber consumption. Sensitivity analysis was preformed using a one percent change in each independent variable. The analysis suggested that a change in any variable, except the overdraft variable, would significantly influence the consumption of Australian sawntimber.

Eastwood, Gray, and Brooker used the Box-Jenkins methodology to forecast demand for two variable-weight food items, ground beef and beef roasts. The models were generated using weekly scan data obtained from five supermarkets in a southern metropolitan area. The data covered the time period May 14, 1988 through November 11, 1989. Two-week trial forecasts were estimated using the data from October, 7 through November, 11.

The forecasts were evaluated using the forecast means, coefficients of variation, absolute errors, turning points, direction of change, and Theil's U. The results revealed that the mean forecasts were very close to the actual means for both the ground beef and beef roast forecasts. The forecast for ground beef estimated 37 turning points while the actual series only had 34. The beef roast forecast predicted 39 turning points while the actual series had 48. However, the forecasts were only able to predict actual turning points 3 and 9 times for ground beef and beef roast respectively. In predicting the direction of change, the forecasts were accurate 27 times for ground beef and 29 for beef roast. Examination of Theil's U for both forecasts revealed that the ground beef forecast was preferred to the nochange forecast while the beef roast forecast was slightly better than the forecast generated by the nochange forecast. The trial forecast for beef roast did not perform well in estimating roast demand. This result was not unanticipated due to the marginal results obtained from the Box-Jenkins forecast. The trial forecast for ground beef did not match the actual item movement but accurately forecasted the directional changes in the series. The inability to forecast item movement accurately was attributed to the large variability in demand.

Lin *et al* forecasted the price of farmed Atlantic Salmon in the United States and the Euroean Community. A dynamic simultaneous equation model was estimated to determine the factors that affected the supply and demand of Atlantic Salmon in the United States and European Community. The system of equations contained three structural equations and four identities. The structural equations represent the demand for Norwegian Atlantic salmon in the United States and the European community and the supply of Norwegian salmon in the United States. The significant exogenous variables contained in the system were forecast using the Box-Jenkins procedure. The predicted exogenous variables were then substituted into the dynamic equation system to derive a forecast of future supply and demand levels. The demand function included own-price, substitute-prices, income, monthly indicator, and lagged dependent variables. This combination of modeling techniques allowed the researchers to forecast Norwegian Atlantic Salmon supplies and prices.

Monthly data from January 1983 through December 1987, were used to estimate via two-stage least squares the simultaneous equation system. The exogenous variables were production of Norwegian salmon, exchange rates, incomes, and price indexes. Norwegian Atlantic salmon and Pacific chinook salmon were considered substitute commodities. The Norwegian Central Bureau of Statistics supplied the data on monthly Norwegian salmon exports. Prices were calculated by dividing the import value by the quantity imported for both the United States and European Community. Six-time series models were estimated. Salmon production was assumed to increase at a constant rate of ten thousand metric tons a year. The study was based on a fixed exchange rate because of the volatility found in the monetary exchange rates. The time-series models were evaluated using MPE, RMPE, Theil's U, UM, and US.

The results suggested that the signs associated with the estimated simultaneous equation coefficients were consistent with economic theory. The monthly dummy variable, substitute prices in both the United States and European Community, and the United States and European Community income variables were the only variables not significant at the one percent level. The forecasted exogenous price for U.S. Chinook fluctuated according to the usual seasonal patterns but exhibited an upward trend over time.

Wu et al compared and evaluated the forecast accuracy of the Box-Jenkins, H-W, and track methods for shipments of a carefully managed IBM product called boxes. The track method was a term given to the monthly planning figures which were based on the planner's expectations of manufacturing capacity, product introductions, and price changes. Quarterly data provided by IBM were used to generate two forecasts for 1988. The data were rescaled to maintain confidentiality. The sample period contained 72 monthly observations that began in January of 1982 and continued through December of 1987. A plot of the series revealed a strong seasonal pattern, peaks at the end of each quarter, and increased variability after December of 1984. Seasonality was removed by the seasonal differencing method. Shipments of boxes in July and August of 1987 had to be postponed until September of that year. This problem was accommodated by taking the total third quarter shipments and dividing them by three.

The Box-Jenkins and H-W forecasts were compared with respect to their abilities to forecast special events (eg. the introduction or termination of products or marketing promotions), as well as the features contained in the sales series, and to the accuracies of the one, two, and four quarter ahead forecasts. The best Box-Jenkins forecast was superior to the best H-W forecast in estimating shipments in all the quarters according to the QFC. Each Box-Jenkins forecast was also compared with the track method prepared by the IBM planners. The Box-Jenkins models had a lower QFC than the track for 1988 shipments.

Conway, Hallahan, Stillman, and Prentice suggested that structural changes in the demand for meat led to the over prediction of meat prices due to the fixed coefficients used in the industry's forecasting models. They argued the supply function of meats had also changed due to changing domestic and international economic conditions. This paper evaluated and compared the forecast performance of the stochastic S-T, fixed coefficient, and OLS with ARIMA-fitted residuals models. The models were estimated using price data provided by the USDA's Economic Research Service. The stochastic coeficient model was estimated using the OLS, maximum likelihood, and the Cochrane-Orcutt procedures.

The OLS estimations using the Cochrane-Orcutt and maximum likelihood models were developed under the assumption of first order autocorrelation. Evaluation of the estimated forecasts, using out of sample data, were based on the RMSE, MAPE, and turning point error criteria. The data covered the first quarter of 1980 through the third quarter of 1983. Based on the RMSE and MAPE criteria, S-T was the most accurate forecasting technique for beef and chicken prices. The Cochrane-Orcutt and maximum likelihood models outperformed the other forecasts of pork prices using the same criteria. These findings were attributed to the degree of variation found in the prices of the different meats. Pork prices had little variation while beef and chicken prices varied a lot. Using the turning point criterion, the stochastic coefficient model predicted as well as or better than either the Cochrane-Orcutt or maximum likelihood models. These findings suggested that the stochastic coefficient model provides a superior price forecast for more variable time-series data.

Geurts and Ibrahim evaluated and compared two forecasting methods, Box-Jenkins and exponential smoothing. The two models were evaluated on their respective abilities to forecast tourism in Hawaii accurately. Logarithmically transformed monthly tourist data from 1952 through 1971 were used to generate the forecasts. The reason for transforming the data was to adjust for seasonal fluctuations which exhibited an increasing amplitude. Using the Box-Jenkins methodology and the maximum likelihood procedure, four adequate forecasting models were found based on the fact that the RMSEs for all the models were very close. After examining the one period forecasts and Theil's U, the best Box-Jenkins model was selected to be compared with the exponential smoothing model. One period ex post forecasts were generated using two years of monthly data. Theil's U suggested that neither model was superior in generating accurate forecasts. The authors concluded that the exponential smoothing model provides an equivalent forecast to the Box-Jenkins method, but it is easier to use and cheaper to generate. Consequently the former may be an appropriate statistical model to combine with a theoretical model to obtain a composite model.

Menkhaus and Adams developed a discriminant analysis model to predict the direction of price movements for feeder cattle. The authors compared the forecasts generated by discriminant, regression, and composite methods containing the discriminant price variable in the regression model. The discriminant model forecasted price movement as either increasing or decreasing. Monthly USDA data for October through March from 1925 through 1968 were used. The ex post forecast was generated using data from 1970 through 1980. For comparative purposes, the authors estimated a feeder cattle price regression model. The feeder cattle price was obtained by adding the slaughter price and the cost associated with holding a calf until it reached its slaughter weight.

The results of the study suggested that, in predicting the direction of price movements, the discriminant analysis model preformed only slightly better than the regression model. The discriminant analysis was able to predict 76 percent of the price movements correctly. The regression model predicted 73 percent of the price movements correctly over the same time period. The naive nochange model correctly predicted 70 percent of the price movements. Theil's U for the regression model decreased from 0.48 to 0.40 when the directional change variable was included. Thus, the results suggested that the forecasting accuracy of a price model may be improved by including a directional change variable.

The usefulness of a transfer function model in forecasting the impact changes in the aggregate economic activity on employment levels in a small regional economy was evaluated by Weller. A transfer function model combined econometric and time-series forecasting techniques to generate a forecast.

Weller evaluated employment level forecasts of 25 models, one univariate time-series model and four transfer function models for each of the following categories: manufacturing, durable manufacturing, nondurable manufacturing, nonmanufacturing sectors, and total regional employment. Weller used the Box-Jenkins methodology in building the univariate and transfer function models. Forecasts were updated using an iterative process of including each new data point as the actual value became available. The models were compared on the basis of the standard error, AIC and BIC.

The data used in this research were obtained from the monthly establishment survey administered by the Bureau of Labor Statistics. They covered the time interval of 1958 through the first month of 1983. Ex post forecasts were generated using data from the second month of 1983 through the last month of 1985. Weller concluded that the transfer function models consistently out performed the univariate models in forecasting the employment levels in all five employment categories.

Because there is no rice futures market, businesses involved with the rice industry have to rely on price forecasts provided by the Rice Outlook and Situation (RO&S) to make decisions. The RO&S provides a price forecast for the year starting the first of August and ending the last day of July. Elam and Holder evaluated and compared the accuracy of the price forecasts provided by the RO&S against those generated by a univariate Box-Jenkins model. The authors then evaluated whether RO&S price forecasts reduced price variability and/or increased producers prices.

The comparisons of the RO&S and forecasts were made using the bias, MAE, and RMSE criteria. The forecast bias suggested that both the ARIMA and RO&S underestimated the September price by 48 and 88 cents per hundred weight respectively. The March price forecast bias for the RO&S and ARIMA forecasts were .16 and -.10 respectively. t-tests indicated the bias was insignificant, using the 95 percent level of significance. The MAE for the September ARIMA forecast was lower than the MAE for the RO&S forecast. The RO&S forecast had a lower MAE for the March forecast. The RMSE of the RO&S was lower than the ARIMA's RMSE for the September and March forecasts. These findings suggested that the RO&S forecast was only marginally better than the forecast generated by the univariate ARIMA model.

The RO&S and univariate forecasts were combined to determine if there was any increase in forecasting efficiency. The results suggested that there was no gain. This conclusion was based on the MAE and RMSE evaluation criteria. Using Ashely's test, the results suggest that no significant difference exists between the RO&S and combination price forecasts.

Oliveria, O'Connor, and Smith used an ARIMA forecasting model to generate weekly cattle price forecasts. The data used in this study were gathered from six-cash market price series and the Chicago Mercantile. Two hundred and sixty weekly price observations, starting in January of 1972 and ending in December of 1976, were used to generate the different forecasting models. The cattle prices were forecast from January 1977 through August 1977.

The accuracy of the various ARIMA models' forecasts were evaluated using three different criteria: the standard error, coefficient significance using t-tests at the 95 percent level of significance, and the parsimony rule. Using these criteria, the most accurate of the ARIMA models was selected to generate <u>ex post</u> forecasts. The standard error revealed that the larger models do not clearly generate superior forecasts in comparison to the smaller forecasting models.

Forecast horizons of 1,4,6,8,12,16,20 and 24 weeks were generated from post-sample period data. A comparison between the nochange and the ARIMA models suggested that the nochange model was superior in generating one to eight week forecasts, whereas the ARIMA was superior in providing longer-run forecasts of 16 and 20 weeks. The results indicated that the ARIMA models provided more accurate forecasts of short-run cash prices and long-run commodity future prices. Little if any gain in accuracy was realized by using larger ARIMA models, or even using an ARIMA model, over a nochange model.

Beilock and Dunn used Box-Jenkins to construct a univariate time-series forecasting model. The model was estimated using end-of-month frozen French fry stock data starting with January 1956 and continuing through March 1978. The data were obtained from the Crop Reporting Board of the Statistical Reporting Service. Two SARIMA models were used to generate forecasts which were then evaluated and compared using the SSE and MSE evaluation criteria. The most accurate one year forward forecast was compared with actual frozen French fry end-of-month stock data for April 1978 through March 1979.

The results of the study suggested that the SARIMA forecast accurately predicted 61 percent of the squared deviations from the mean. The forecasts overestimated the stock of frozen French fries for December through March. A positive increase in potato stocks occurred in August due to the harvesting of the crop. The forecast predicted the change to occur in September instead of August which was not surprising after inspection of the data. The data revealed that in eight of the last nine years, the increase in frozen french fry stock levels had occurred in September rather than August.

Spriggs evaluated and compared a Box-Jenkins univariate model with OLS regression analysis on their abilities to predict the price of Indiana corn. The regression model described monthly prices of Indiana corn as a function of the future price for corn. He then combined the two methods to see if forecast accuracy was improved. Spriggs generated 160 forecasts, 80 Box-Jenkins and 80 regression. Forecasts were generated, for September through December for 1959-1978. The price data were obtained from the Agricultural Prices and the Chicago Board of Trade publications.

The MSE revealed that the Box-Jenkins forecast provided a superior forecast for September corn prices, while the regression forecast was superior for the remaining three months. The analysis suggested that the Box-Jenkins method provided a better estimate of historical prices while the regression model predicted future prices better.

Four composite forecasts were compared, using the MSE with the individual Box-Jenkins and the regression forecasts. The four composite forecasts had different weights assigned to the individual forecasting components. Weights were derived by equally weighing each individual forecast. A constrained OLS regression was used to estimate the weight of each individual forecast, and a mixed estimation procedure to determine the different weights. The simple forecast for September and December corn prices was equivalent to the OLS composite forecasts. The simple composite forecast provided a superior forecast for the October and November corn prices.

The regression forecast was superior to the one generated from the Box-Jenkins methodology based solely on the MSE. No gains in forecasting efficiency were achieved by using a composite model. Thus, Spriggs concluded that there was very little independent information contained in the two forecasts.

Craine and Havenner compared the ex post forecasts of interest rates and corporate bond rates using four different forecasting methods. The random walk, univariate time-series, multivariate time-series, and two MIT-Penn-SSRC (MSP) forecasting methods were evaluated and their forecast accuracies compared. The authors believed that the random walk forecasting method was overly simplistic, and they compared it with other models, each of which incorporated additional information not contained in the one preceding it. These four models were used to generate both short and long-run forecasts of corporate bonds and commercial paper interest rates.

Monthly data starting with the third month of 1953 and continuing through the seventh month of 1977 for both short and long-run interest rates were used to estimate the four different forecasting models. Ex post forecasts were then compared to the actual data for the eighth through the tenth month of 1977. The short and long-run forecasts were one and three month ahead forecasts respectively.

One month ahead forecasts for both the interest rate and corporate bond rate were generated using the random walk, univariate, and multivariate models. The forecasts were evaluated using the MSE, MAE, and Theil's U. Comparisons of the MSE revealed that the univariate forecast was superior to the multivariate forecast which was superior to the random walk forecast. The MAE revealed a superior forecast was provided by the multivariate model followed by the univariate model forecast which was followed by the random walk model. Theil's U was smallest for the univariate forecast, followed by the multivariate forecast, and finally the random walk forecasts, 0.865, 0.935, and 1.0, respectively. The one month ahead forecasts of corporate bond rates was evaluated using the same criteria. The results indicated that the multivariate forecast was superior to both the univariate and random walk forecasts. Three month ahead forecasts were generated for both the interest rate and corporate bond rate using the four different models. Results indicated that the multivariate forecast was superior to the other three forecasting methods.

Harvey and Todd compared the forecasting ability of an ARIMA model to a structural forecasting model. The models generated forecasts of consumers' expenditures on durable goods, consumers' expenditures on all other goods and services, investment, inventory investment, imports of goods and services, and GDP or Gross Domestic Product. The univariate structural model was different from the Box-Jenkins univariate model in that the dependent variable was described in the former as a function of a trend, seasonal, and irregular component variables. The ARIMA model was forecast using a differenced and nondifferenced time-series. The forecasts were evaluated using the predicted error variance and goodness of fit measures. The ARIMA model used in this study was developed by Prothero and Wallis. Data from the third quarter of 1957 through 1966 were used.

In forecasting consumer expenditures on durable goods, the structural model performed only slightly better than the ARIMA in both the in-and out-of-sample predictions. In forecasting consumer expenditure on all other goods and services, the structural model performed slightly better than the ARIMA model inside the sample period, whereas the ARIMA performed better outside of the sample period. The ARIMA model provided a slightly better forecast in both the in and out-of-sample periods. The structural model provided a poor forecast of inventory investment. The nondifferenced time-series provided a superior forecast to the differenced forecast, which was poor. The structural forecast was superior to the ARIMA model in forecasting imports of goods and services. The structural and ARIMA models performed equally well in forecasting GDP. These results indicated that the forecasting abilities of the structural and ARIMA models were very comparable for this problem setting.

Liu compared the forecast accuracy of a simultaneous transfer function (STF), regression, and the ARIMA models. The STF model is a simultaneous equation system with an ARMA model. The forecasts were compared on their abilities to generate efficient forecasts of computer parts sales. Liu generated forecasts using three different types of data. The first data set was obtained from a computer component manufacturer and contained sales and order information. The second data set was obtained from the 1970 Box and Jenkins study and contained leading indicator information about sales. The third data set was obtained from the Bureau of Census and contained information pertaining to durable good shipments, new orders, and inventory levels. The time-series data sets began in April 1974 and continued through October 1982.

Data from November, 1982 through October, 1984 were used to compare the one step and two step ahead forecasts from the three different techniques. The three and four period ahead forecasts generated by the different forecasting techniques were then compared and evaluated using the MSE and MAPE criteria. In order to generate a three and four step ahead forecast, the regression models needed the inclusion of two variables. These two variables represented the order status three and four months prior to the actual shipment.

Results using the sales-order data indicated that the regression and STF techniques yielded comparable forecasts, both of which were superior to the forecasts provided by the ARIMA model. In each of the one through four step ahead sales-order forecasts, the STF forecast was superior to the ARIMA forecasts. The increase in efficiency was most noticeable in the shorter forecasts. The author concluded that when major changes in a time series occur, the multivariate model provided a better estimate than the univariate ARIMA model which generally exhibited large deviations.

The one-step to five-step ahead sales forecasts using leading indicator data revealed that the STF model provided a better forecast than the ARIMA model, assuming that the noise was appropriately identified as white noise. If the noise was inappropriately assumed to be white when in fact it was not, the STF forecast was worse than the forecast provided by the ARIMA model.

In forecasting sales using durable good shipments, new orders, and inventories data, the STF model performed slightly better than the ARIMA model, but only if the noise was assumed to be white noise. If the noise was incorrectly assumed to be white, the STF forecast was inferior to the ARIMA forecast. These results indicated that when there was a strong presence of seasonality in a series, the STF model did not increase efficiency over the ARIMA forecasts. The author suggested when building a multivariate model using a time series that exhibits seasonality more attention should be focused on seasonality than on the inclusion of marginally significant variables.

Watson and Pastuszek developed seven forecasting models which were used to forecast electricity sales for two Northeastern utility companies. This study is relevant because it used the SARIMA and compared the forecast accuracies of the SARIMA and ARIMA forecasting techniques. They generated monthly forecasts for two periods. The first period was October 1983 through September 1984. The second period was October 1984 through May 1985. The seven models were ARIMA, a seasonal autoregressive, two state-space, a exponential smoothing, and two econometric models. The forecast accuracies of the seven models were evaluated using the size of their RMSEs. Monthly commercial sales data were supplied by the Massachusetts Electric Company and Narragansett Electric Company. Both companies supplied data from January of 1975 through September of 1983.

The most accurate first period commercial sales forecasts, for both companies, were generated by the two econometric models. The exponential smoothing model provided good forecasts of Mass Electric sales, while the state space models predicted Narragansett Electric sales better. The results indicated that the Box-Jenkins and the seasonal autoregressive forecasts were inferior to the ones provided by the other methods.

The second period forecasts of Mass Electric sales indicated that the exponential smoothing and the two econometric models provided the most accurate forecasts. Each model experienced a decline in its RMSE by roughly forty percent in the second forecast. Both of the seasonal autoregressive models indicated small decreases in the sizes of their

RMSEs. The ARIMA model performed poorly. The analysis also indicated a decline in the performance of the state-space model from the first forecast to the second forecast, the second forecast was comparable to one provided by the Box-Jenkins methodology. The results of the Narragansett forecast revealed that the best forecast was provided by the exponential smoothing model and that all other models generated comparable forecasts. The exception was the poor forecast generated by the ARIMA model.

Hayward et al used a recursive supply and demand model to estimate supply and demand linkages in the Maine and United states apple markets. Forecasts of the endogenous variables are substituted into the estimated model to obtain annual apple prices and production in Maine and the United States for 1981 to 2000. The recursive model contained six components which were regional supply response equations for Maine and the United States, a national demand equation, a market clearing equilibrium condition, and a price linkage equation.

The data used in this study were obtained from various government publications. The population projections were obtained from the Bureau of the Census. Extrapolation of historic U.S. disposable income was used to estimate future U.S. disposable income. These estimates were divided by the Census Bureaus population estimates to derive per capita disposable income.

Two options of the Gauss-Seidel solution procedure were used to forecast the endogenous variables and the goodness of fit measures. The MAPE, squared correlation between the actual and estimated values of the dependent variable, and Theil's U were the criteria used to evaluate forecast accuracy. The first option used a combination of historical values of the exogenous variables and previous estimates of the lagged endogenous variables to estimate the endogenous variables. The second option was similar to the first except the historical observations of the endogenous variable were substituted for the values estimated by the model.

The MAPE for both options were low, and the squared correlation values indicated the model provided a good fit. Both options produced Theil's U that were less than one. Goodness of fit statistics suggested that the estimated model provided a sufficiently accurate forecast. It predicted a 119 percent increase in apple production nationwide from 1981 to 2000, and U.S. apple prices were forecast to increase from $4.45 to $7.59 per bushel over the same period. Maine's apple production and apple prices were predicted to increase 28.7 percent and $3.34, respectively, over the same 19 year time period.

Hudson and Capps evaluated alternative techniques of forecasting the price of ex-vessel blue crabs in the Chesapeake Bay region. The authors compared the forecasting abilities of seven individual and composite forecasting techniques. Three individual techniques were the quadratic trend extrapolation, econometric with seasonality, and SARIMA (1,1,o) x (1,0,1). Four composite forecasts were generated using two different techniques. The first composite forecast was a simple average of each of the three individual forecasts and was referred to as the simple average model. Three other composite forecasting models were developed using the minimum variance weighting procedure. These three models were the SARIMA-quadratic, SARIMA-econometric, and quadratic-econometric.

The seven (three individual, a simple average, and three composite) models were estimated for the period January 1973 through July 1979 and then used to forecast monthly blue crab prices for the period August 1979 through 1980. The forecasts were evaluated according to their MSEs and abilities to predict turning points. The SARIMA model, using the MSE, provided the most accurate technique in forecasting blue crab prices. Examination of the MSEs suggested that the second best forecast was provided by the econometric model. According to the size of the MSE, the SARIMA forecast was determined to be only marginally better than the SARIMA-econometric forecast. The MSE of the simple average composite forecast and the ARIMA-quadratic forecast were approximately equal. The quadratic-econometric forecast had a smaller MSE than each individual forecast, but it was larger than the MSE of the other composite forecasts. The turning point evaluation revealed the same general ranking of forecasts with the exception of the SARIMA-quadratic, which outperformed all the other techniques. The SARIMA and quadratic individual techniques tied for second. The econometric, simple average, SARIMA-econometric, and quadratic-econometric models tied for third place.

Allen alludes to the idea that an econometric model using poorly forecasted values will provide an inferior forecast when compared to a simple SARIMA model forecast. He examined the study conducted by Hudson and Capps. Allen proposed that if an econometric forecast can be improved by the inclusion of a forecasted explanatory variable, it should be included. He used a theorem developed by Ashely to defend his logic of eliminating a forecasted variable from econometric models. For example, Ocean Spray needed yield predictions in order to make storage decisions. Allen generated one step ahead forecasts of cranberry production in

Massachusetts using an univariate ARIMA model and nine variations of an econometric model. The forecasts were generated using cranberry production and temperature data from 1932 through 1979 taken from a study by Morzuch, Kneip, and Smith. He used forecasted values of temperature in his model.

All econometric forecasts had lower MSEs than the ARIMA forecast. The econometric forecast including actual temperatures performed better than the econometric model using time and a dummy variable in place of temperature variables. However, when the econometric forecast was generated using forecasted weather data, the MSE increased. These results supported the idea that inclusion of poorly forecasted explanatory variables will cause a decrease in forecast accuracy.

Brandt and Bessler (1981) evaluated the ex ante forecasting performance of composite and individual forecasts of cattle, hogs, and broiler prices. Three different individual forecasting techniques were used to generate forecasts which were then combined to obtain the two and three component composite models. The three individual forecasting techniques were the econometric, ARIMA, and expert opinion.

The composite forecasts were made up of combinations of two and three forecasting methods. The two component composite forecast consisted of the econometric and ARIMA individual forecasting techniques. The three component composite model included the econometric, ARIMA, and expert techniques. The study used the optimal weighing, adaptive weighing, and simple weighing schemes in generating the combination forecast models. These three procedures were combined into composite models. The expert opinion forecasts were only used in the simple composite forecasts due to their availability.

Evaluation of the individual forecasts using the MSE criterion revealed that no one modeling technique supplied a superior forecast. The ARIMA model did provide a slightly better forecast of hog and cattle prices, whereas it provided the poorest forecasts of broiler prices. The best broiler price forecast was generated by the econometric model. The results were similar using the turning point evaluation criterion.

Evaluation of the two models' composite forecasts revealed that the composite adaptive model provided the best forecast for cattle and broilers while the minimum variance forecast provided the best hog price forecast. The simple average model provided a superior forecast to all the individual models for cattle and broilers. In forecasting hog prices, only the individual ARIMA model provided a better forecast. The three component composite forecasts generated superior forecasts when compared to the two component composite forecasts. However, the inclusion of the expert opinion model did not reduce the forecast MSE.

The authors concluded that the poorest composite forecast was better than the best individual model in terms of the MSE. This allowed them to draw the conclusion that decisions made using composite forecasts were superior to the same decisions made using any of the individual forecasts. The results also indicated that combining the expert opinion models with econometric and time-series forecasts improved forecast accuracy.

Brandt and Bessler (1984) developed a simple hog price forecasting model that could be easily updated. They compared the individual forecasts of the econometric, ARIMA, and expert opinion models with four composite models. The composite models were the two-period adaptive, minimum variance, simple average 1, and the simple average 2 models. The two-period adaptive, minimum variance, and simple average 1 models are combinations of econometric and ARIMA models. The simple average 2 model is a combination of the econometric, ARIMA, and the expert-opinion models. Each individual forecasting technique received an equal weight in developing the composite forecasting models.

The various models were estimated using quarterly data covering the time period 1961 through 1975. Forecasts were generated for fourteen periods starting with the first quarter of 1976. The minimum variance model weighted the econometric and ARIMA models to minimize the variance of the forecast error. The two-period adaptive model used a changing weighting system that reweighted the individual econometric and ARIMA models after each period.

The individual and composite models were compared and evaluated on their MSE's and predicted turning points. The most accurate of the individual models, based on the MSE's, was the ARIMA which was significantly superior to either the econometric or expert opinion models. The ARIMA and expert opinion forecasts accurately predicted price movements in over half the periods, whereas the econometric model accurately predicted price movements in only four periods. The MSEs of the simple average 2 and minimum average models were lower than the MSE of the ARIMA. Actually, the MSE of each composite model was found to be lower than the MSE of the econometric and expert opinion models. The composite and individual models were approximately equal in their abilities to predict price movements accurately.

McNees contends that forecasts can more accurately predict a known data series than future data values as indicated by the MAE, RMSE, and ME. He attributed the increased accuracy of ex post over ex ante forecasts to the researcher's availability of historical information. Using the available information, the data series can be more accurately described. McNees generated three ex post and ex ante GNP forecasts for the first, second, and third months of the quarter. Inspection of the evaluation criteria suggested the superiority of ex post forecast.

McNees also implied that forecast accuracy was dependent on the time period to be forecasted and the variability of the of the dependent variable. He demonstrated that the shorter the forecast period, the more accurate the forecast using four period ahead forecasts of GNP provided by other researchers. These forecasts provided large over and under estimates of future GNP levels compared to the forecasts of shorter time periods. The effects of variability in forecast accuracy was demonstrated using forecasts of the increase in the Consumer Price Index, which was larger during the inflationary 1970's than it was during the following eight years.

McNees compared forecasts submitted to a Wall Street Journal survey to a nochange model and evaluated their accuracies. The submitted forecasts consisted of 12 forecasts of short-and long-term interest rates and 29 forecasts of the CPI growth rate, unemployment rate, GNP lag, and GNP lead. The forecasts generated next-half year and next-year forecasts of short-and long-term interest rates and unemployment as well as next-half-year, following half-year, and next-year forecasts of CPI and GNP. The results indicated that the nochange forecast outperformed all but one forecast in predicting short-and long-term interest rates. The 29 CPI forecasts were all superior to the nochange CPI forecast. Two thirds of the unemployment growth rate forecasts outperformed the nochange model. The GNP forecast were compared to the GNP lag and GNP lead. The GNP lag is the nochange model which predicts GNP growth rate to equal the preceding rate. The GNP lead is the preliminary estimate of GNP growth that was not available until two weeks after the forecasts were received. Two-thirds of the forecasted GNP lag models outperformed the nochange model in the next-half year. In forecasting the next-year GNP lag, only one forecast was inferior to the nochange model. A majority of the forecasts outperformed the GNP lead estimate in predicting the following next-half and full-year and next year. Only a few of the forecasts were superior to the GNP lead in predicting the first-half-year GNP growth rate.

The author concluded that forecasting particular variables requires distinctive knowledge while forecasting other more general variables are not as difficult. He also concludes that comparison of forecast errors is a valuable tool in evaluating forecast accuracies, but the information regarding the forecaster's knowledge is just important to a forecast user. McNees suggested that when comparing forecasts, it is not appropriate to compare <u>ex post</u> and <u>ex ante</u> forecasts because of the difference in information available to the forecaster.

McIntosh and Bessler conducted research that compared different forecasting techniques to determine the most accurate method. Price data on barrows and gilts were used to generate the forecasts. The data were taken from the USDA seven-market average of hog prices and covered the period starting with the third quarter of 1973 through the second quarter of 1986. The individual forecasting models were the expert forecast, the futures market price, and a one quarter lead ARIMA model. The composite forecasts were the Bayesian specification, simple average of the three forecasts, restricted ordinary least squares combination, and an adaptive weighting composite model. The restriction on the regression model was the coherence restriction, meaning that the weights of the individual components add up to one.

The results indicate that each of the composite models provided a superior forecast when compared to the best individual forecast. The Bayesian forecast performed slightly better than the simple average and restricted forecasts based on the level of bias and the MSE.

Grubb built and evaluated a multivariate model which was used to forecast the price of flour. The concept behind using a multivariate model was that it could better model the dynamic relationships found among variables. The six forecasting techniques used in the study were the univariate, transfer function, VARMA, VAR, and VMA models. The univariate model was specified and estimated using the maximum likelihood procedure. The summary statistics indicated that the univariate model provided a good fit. The transfer function technique was a transformation of the univariate model where the data for each of the three cities were included along with the differences between them. This procedure produced a smaller variance of the error series compared to the individual univariate models by accounting for the between series correlation. The VAR (1) process was identified after inspection of the ACF and PACF when trying to identify the VARMA process. The elimination of redundant parameters in building a VARMA forecasting procedure was also used, as specified by Taio and Tsay. The ACF and PACF indicated that the process may be of order VAR

(1) or (3). After differencing the data, a VMA process was identified.

The monthly price of flour, in logarithms, for three U.S. cities was used to generate the forecasts. The data covered the period August of 1972 through November of 1980. Visual inspection of the data revealed that the series was a random walk process without any seasonal variation. The AIC was used to compare the VMARA, MVA, and the AR processes and to determine which method was most accurate at reproducing the series. This evaluation criterion indicated that the AR model provided the best fit. Comparisons of the forecast errors revealed that the transfer function and the vector moving average models provided a superior forecast.

Carnova compared the forecasting accuracy of models not discussed above: a restricted univariate time varying component-autoregressive (TVC-AR) model with the unrestricted autoregressive, multiplicative SARIMA, deterministic and seasonal (TRS), unobservable component (UC), and an additive random walk (RWS) forecasting models. Each of the models was evaluated using Theil's U and MAE. Unadjusted data from the third quarter of 1960 through the second quarter of 1982 were used to generate the forecasts of the various models. The forecasts were of ten different macroeconomic variables: GNP, government deficit (GDEF), demand deposits (DD), labor force (LFOR), housing starts (HS), business inventories (BINV), final sales of products (FSAL), consumption of durables (CDUR), and total fixed investment (IFIX).

Based on Theil's U, the TVC-AR model provided a superior forecast when compared to the five alternative forecasting techniques. When the TVC-AR forecasts were compared with the SARIMA model, the MAE criterion indicated that the SARMA model provided a superior forecast for three of the ten series. Comparing the TVC-AR against the UC and RWS forecasts using the MAE criteria suggested that the TVC-AR forecast was clearly superior to the other two forecasting methods. Half of the superior long run forecasts were produced by either unrestricted or TRS models.

Bessler compared four different forecasting models under the condition of little prior information regarding the data series. The three multivariate forecasting models were the unrestricted VAR, four-leg Bayesian VAR (BVAR), and the restricted VAR (RVAR). A univariate model was also used to generate a forecast. There were seven data series used in this study. Little information was supplied regarding the data series except that they were either a price or quantity series. The data were made stationary before being used to estimate the various models.

The four forecasting techniques were used to generate forecasts of two different data series. The data were broken into two sections, and forecasts were generated for each section. The RMSEs were derived from the out of sample forecasts. The results of the first data series, both parts, indicated that the VAR provided an inferior forecast when compared to the univariate forecast using the RMSE criterion. The Bayesian forecast had a smaller RMSE for the first set of data in series one, while the univariate model had the lower RMSE for the second set of data in series one. The RVAR and unrestricted VAR had higher RMSEs for both sets of data in series one. The results from the second series forecast indicated a superior forecast for the restricted VAR model in the second data set, while the RMSE for the first data set was equal to the RMSE for the BVAR forecast. The univariate forecast had the smaller RMSE for the forecast of the second set of data.

Stobaugh and Townsend present different techniques for forecasting the price of 82 two different petrochemicals. They used a logarithmic price model to forecast petrochemical prices using four independent variables: competition, experience, static sales economies (production per manufacturer), and degree of product standardization. Earlier empirical price forecasting models described petrochemical prices as being a function of only one variable, such as accumulated production, annual production, or the stage in the product life cycle. The petrochemical industry is considered to be oligopolistic, so there was not an obvious method of predicting prices or profits in this industry.

The production volume, sales price, and the number of competitors data for each chemical were obtained from the U.S. Tariff Commission. Only those chemicals whose 1969 sales surpassed ten million dollars and were produced by three or more firms were used in the study.

The forecasting model used as independent variables competition, standardization, experience, and static sales economies. Each of these variables was associated with profit decline, over time, in an oligopolistic industry because as competition increases, firms are forced to reduce prices leading to a reduction in profit margins. Product standardization increases the substitutability of products again leading to increased competition and lower prices. As the experience level of the firm increases, the costs associated with production decrease leading to a lower cost of production. The increased production facilities, resulting from increased sales, will reduce production costs as economies of scale are realized.

The model was estimated using a linear regression procedure. The results indicated the significance of each of the four independent variables. The R-squared was low revealing that a small percent of the variation in price was described by the model. The negative constant term suggested an important variable was excluded from the model. The coefficients were transformed into different units to measure their effects better. The transformation revealed that accumulated experience and other time related factors have the largest impact on the prices of petrochemicals. Competition, standardized quality, and the number of producers have a positive, but lesser, influence on prices. Regression analysis, using the experience variable as the only explanatory variable, was compared to the original model. The results indicate that the multivariate regression model provided a higher R-squared. Petrochemicals were separated into more homogenous categories to determine if improvements in the forecasts could be realized. The R-squared rose in 10 out 12 of the redefined categories versus the aggregate. The authors suggested the increase in predictability was due to using the four explanatory variables but emphasized that additional research is needed to determine the missing explanatory variable(s) that could increase the model's goodness of fit.

Helmer and Johansson used a transfer function model to forecast sales levels of Lydia Pinkham vegetable compound. The objective of this study was to develop a more accurate forecast of consumer demand by using a transfer function model which included advertising effects. The authors generated Box-Jenkins forecasting models, and the two best were compared to previously estimated advertising effect models to determine forecast accuracy. The two best Box-Jenkins models were determined using goodness of fit measures and forecast accuracy.

The Box-Jenkins procedure was used to forecast sales advertising data of the vegetable compound using annual data from 1907 to 1946. The last 15 observations were used to evaluate trial forecasts of the two models. The data series was prewhitened and the nonstationarity, which was attributed to the nonstationary advertising series, was removed. After the two best transfer functions were chosen, they were estimated using nonlinear least squares.

The two Box-jenkins transfer function models were compared to eight previously estimated models. The eight models consisted of four cumulative effects models, predictive testing and cross-spectral analysis models, a carry over effect model, and a Box-Jenkins univariate model. Comparison of the models clearly indicated the superiority of the Box-Jenkins models which outperformed all of the other techniques. The Box-Jenkins models had the smallest MSE and Theil's U values. Thus, the Box-Jenkins transfer function forecasts were the most accurate.

C. Advertising Studies

The advertising section contains a review of literature that included advertising and or promotional campaigns as explanatory variables in empricial demand relationships. The objective is to describe insights into the many methods of including advertising and promotional campaigns in demand analysis as well as the diversity of these campaigns. The review and inclusion of these articles is important because it aided in determining the modeling technique used in estimating the effects of advertising on consumer demand for specific supermarket products.

Kinnucan and Venkateswaran used an eight equation semi recursive equation system to model the effects of a generic advertising campaign. Their model attempted to depict the linkages between the advertising campaign and consumers' attitudes and perceptions of farm-raised catfish. Dependent variables were odor, nutritional value, flavor, attitude or ranking of catfish in relation to other types of fish, observation of the advertisement, home consumption, consumption at a restaurant, and three variables used to incorporate demographics and exogenous variables that impact consumption of catfish at home or at restaurants.

Survey data were collected by randomly sampling four hundred households in each of the nine U.S. demographic regions as broken down by the Census Bureau. To determine if a bias was present due to the deletion, Heckman's two-stage probit procedure was used which led to the inference that the estimates were not biased. Full page color advertisements were placed in ten national magazines. The advertisements emphasized the nutritional aspects of catfish consumption and tried to differentiate farm-raised catfish from wild catfish. This advertising campaign was implemented in April of 1987. Information regarding sociodemographics and awareness of the advertisements were collected along with the respondent's beliefs, attitudes, awareness, and consumption of catfish. The survey averaged about twelve minutes per respondent. Of the 3,600 respondents, 2,172 acknowledged consuming catfish. The 1,428 surveys in which the respondent did not consume catfish were discarded.

The significance of demographics and other exogenous variables that impact consumption of catfish at home or at a restaurant were determined by simple t-tests. The significance of other independent variables was determined using the Bonferroni t-statistic. Results suggested that only thirty-eight percent of the respondents were conscious of the catfish advertising campaign. Nonreporting of income, sex, and Western residence were the only significant variables in determining ad awareness. The failure of the campaign may be attributed to the mixed gender subscribers of the magazines. The fact that the advertisements were targeted at audiences so radically different from the typical catfish consumer suggested that the length of the advertising campaign was too brief to penetrate the target audience sufficiently. The campaign did have a significant impact on increasing consumers' awareness of farm-raised catfish. The ad recognition variable was significant at the one percent level. The results indicated that awareness of farm-raised catfish was twelve percent greater for consumers conscious of the campaign. The high income, education, heartland location, and rural residence variables were found to be significantly related to the awareness of the ad campaign.

Broadbent and Colman modeled the sales of specific confectionery products and estimated whether high sales were correlated with big brands or small (market share), high ad expenditures, new campaigns, styles and content of advertisement, and advertising awareness. Time series data covering the period 1982 and continuing through the middle of 1983 were used to estimate sales for the eighteen different models. The sales data, shares of confectionery products, were obtained from AGB's Personal Purchases Index. A corresponding survey was also conducted to measure the effectiveness of the advertising campaign. The models described the sale of a confectionery product as being a function of a constant, relative price, and advertising variables. The estimates of the sales effects for all brands were generated simultaneously. This provided consistent cross-sectional data regarding the advertising campaign which was available for comparison with the advertising of other brands.

Broadbent and Coleman argued that a commercial that performs poorly in relaying the product's image or message received a high awareness level due to a stimulating visual or verbal commercial. Sales effectiveness was represented by two components. The first was a measure of the advertisement's content and how effective it was at relaying the advertiser's message as well as the advertisement's ability to persuade consumers. The second component measured how well the consumer was able to relate the advertisement to the product brand contained in the advertisement. Advertising effects were measured by the average awareness divided by the average adstock or half life. Half life is the time required for half of the total advertising effect to occur and was estimated from the observations. Advertisement awareness was not a significant predictor of sales. This result was attributed to the information contained in the advertisement. The study revealed no significant relationship between sales and ad awareness across brands.

The authors concluded that testing and monitoring of advertising awareness was the only method that provided insight into adverting effectiveness. The authors felt that guidelines should be established to determine what is considered to be successful and unsuccessful advertisements. Tracking of advertisements is a simple procedure that can provide insight into the sales effectiveness of a particular advertising campaign.

Capps and Lambregts conducted research to estimate the retail demand for several finfish and shellfish products using supermaret scan data. A double log demand function in which the dependent variable, quantity of finfish and shell fish, was described as a function of own-price, cross-prices, seasonality, own-advertising, and cross-advertising variables. The model was estimated using joint generalized least squares. The inclusion of the seasonal variables was to account for the monthly seasonal variation in the demand for the given products. The newspaper advertising was measured in terms of print space, squared centimeters per product per week, and cross-advertising was newspaper advertising for competing or substitute products.

Estimates of own-price coefficients were negative and, with the exception of oysters, were statistically significant. The own-price estimates for finfish were more elastic than the estimates for shellfish. This was attributed to the larger number of substitutes. The results indicated the importance of own-price on demand for the various products. Also, the estimates of own-advertising were consistent with previous research. All of the own-advertising estimates were positive and significant with the exception of shrimp, which was positive but not significant. Estimated own-advertising elasticities were all inelastic, and the highest elasticity was .27. Seasonality was found to be significant with a few exceptions. The binary holiday variable was not significant. The cross-price elasticities for the shellfish were all positive and statistically significant, while only four of the 15 cross-product elasticities were statistically significant. Seventeen out of the 28 cross-price elasticities for finfish were statistically significant, while only seven out of the 42 cross-product elasticities were significant.

Cross-advertising effects were negative, with one exception. Four of the ten cross-cut advertising elasticities were significant for shellfish. Five out of the 28 cross-cut advertising elasticities for finfish were significant. The cross-product elasticities revealed similar results. Only four out of the 15 shellfish cross-product advertisements were significant compared to 14 of 42 cross-product elasticities for the finfish.

Raju conducted a study to determine the effects of promotions on category sales, whereas previous promotional research had tended to focus on particular brands. The objectives were to compare the effects of intermittent deep and regular (moderate) discounts, manufacturer competition, package size, and price level category variability. Two multiplicative models, the restricted and unrestricted, were used to determine the impact of promotional activity and other category characteristics on category sales. The author argued that short-term increases in category sales were necessary but not sufficient to increase long-term category sales. The author argued that the emphasis on category sales is important to a retailer because a long-term increase in brand sales does not necessarily imply a increase in category sales.

Six months of weekly UPC price and item movement data were obtained from a national grocery store chain. Advertising data were not available, but because of the high correlation between display activity and advertising, display activity was used as a proxy for advertising. Sales variability was defined as a function of product expensiveness, bulkiness, competition in the product category, and magnitude and frequency of discounts. It was assumed that the weekly category sales of the most frequently purchased goods were constant, given no promotional or discount activity in week t.

A discount was determined by examining a brand for 25 weeks, and the highest price over that period was considered to be the product's regular price. The discount price and magnitude and frequency of discounts were obtained by subtracting the weekly price from the brand's regular price. Expensiveness of the brand was calculated using the market share weighted average of the regular prices of the different brands found in a particular category. Bulkiness was the market share weighted average of the brand volumes. Competitiveness was measured by market share and the number of brands found in a category.

OLS was also used to estimate the restricted model. The restricted model defined sales variation as being a function of the magnitude of discounts and discount frequency. The results of the linear regression model produced an R^2 of .41, and the goodness of fit measure revealed a nonnormal error term series. The Jackknife procedure is a method of estimating an estimator's variance and was used to estimate the standard errors.

The unrestricted model provided a more statistically significant estimate compared to the restricted model. The linear model using two thirds of the available data, which were randomly selected, were used to predict sales variability 100 times. The average prediction revealed a positive relationship between discount magnitudes and sales variation. Discount frequency, bulk, and competitive intensity variables were negatively related to sales variability. Product category expensiveness was not significant in determining variability. The elasticities of each of the variables indicated an inelastic relationship with the exception of frequency which was elastic.

Lattin and Bucklin estimated the effects of promotions on brand purchases. They compared promotional and nonpromotional price elasticities and separated these effects from the price and promotional advertisement reference effects. A logit response model expressed consumer utility as a function of idiosyncratic preference (brand loyalty), brand price, promotional status, difference between actual and reference prices, and the difference between actual and reference promotional status. This was a nested model, and each of the explanatory variables was described as a function of other variables. Idiosyncratic preference was a function of consumers' brand loyalty, consumers' recent choice behavior, and whether consumers last purchased the brand on promotion. The reference price was a function of the promoted and purchase price in t-1 and was obtained using an exponentially weighted average procedure. Promotion exposure was a function of a binary variable representing the presence of a product promotion plus a smoothing parameter. The probability of a consumer purchasing a product on a given purchase occasion was estimated by a logit model, which was then used to represent the brand choice variable.

Three different models were estimated and compared. The first, a restricted model, used only the direct price and promotion independent variables. The second and third models contained the two reference variables. They differed on the restrictions regarding the consumers' perceptions of whether a product is considered to be a promoted product.

The second model considered consumer perception of the product to be unprompted, while the third model allowed for varying degrees of perception regarding product promotion. A product was considered promoted when it had been frequently, predominately, or currently promoted.

One hundred and thirty five weeks of scan data from six Pittsfield Massachusetts retail stores containing price information on four brands of 16 ounce packages of caffeinated coffee comprised the data. Inspection of the data revealed that roughly 80 percent of the caffeinated coffee sales occurred during 75 of the 135 weeks for ten different brands. Consumer panel data were collected on 577 consumers who purchased three of the coffee brands from any of the six stores over the period.

The three models generated ex post forecasts of household purchases using 277 of the households and were compared to the purchases of the remaining 300 households. Models two and three, which contained the reference points, price, and promotion varaiables, outperformed the restricted model. The forecast accuracy of the third model was not significantly different from the second model but inspection of the Chi-squared statistic indicated that model three provided a better fit of the data than the second model.

Powers conducted research to determine the effects of grocery store advertising on the sales of navel oranges. He estimated consumers' weekly responses to grocery store advertisements for fresh California-Arizona navel oranges. The research also reported the impacts of grocery store advertising on quantity and industry revenue. Powers employed the two-stage least squares regression technique to estimate the demand function. The grocer's demand for fresh navel oranges was expressed as a function of the advertising index, advertising index with a one week lag, deflated free-on-board index price, deflated price index for apples, deflated per capita disposable income, three binary variables representing the Christmas holiday season, January to February (JF), and March to May (MAM) periods.

The advertising data were obtained from a publication by Majers Corporation which provides information regarding weekly grocery advertisements, supplements, and flyers that appear in the papers of major U.S. cities. The advertising information for major grocery stores, as defined by market share, is grouped together by city. The advertising information for three years, 1982-85, was collected from November to May for New York City, Chicago, and Los Angeles. The seven month time period represented the navel orange marketing season. The advertising of navel oranges for each city was found to be positively correlated, so a proxy for national advertising was created. The price information was obtained from the Navel Orange Administrative Committee. The income and population information were obtained from the Bureau of Labor Statistics.

Only the lagged advertising and apple price variables were found to be insignificant. The R^2 was low, 0.60, but comparable to similar studies. The results suggested roughly 95 percent of the increase in demand, happened during the advertising period, while the remaining five percent occurred in the following week. Elasticities were calculated for different levels of sales and advertising. The estimates indicated that the advertising elasticities were larger when sales volume was low. The results also suggest that the advertising elasticity was larger when the amount of advertising was low. The peak response to advertising was sooner than reported in studies using specific product and generic advertising. The advertising elasticities generated in this study were larger than those reported in other literature containing generic advertising. One explanation was the information contained in the two different types of advertising. Grocery advertising provides price information, while generic advertising tries to influence consumers' attitudes toward a particular product. The own-price elasticity was inelastic, the income elasticity was highly elastic, and the cross-price elasticity was inelastic. The binary parameter estimates for holiday, JF, and MAM were negative and inelastic.

Chang and Kinnucan conducted a study on the effectiveness of the Dairy Board of Canada's advertising campaign on increasing the demand for butter and related fats and oils. The quarterly data covered the time period 1973-86. The advertising campaign was not implemented until 1979, but information was needed before the advertising campaign to determine the impact of the campaign. The data included information on per capita consumption, retail prices, and advertising information on the products. The price and quantity data were obtained from Statistics Canada. The advertising data for butter were obtained from the Dairy Bureau of Canada, and the remaining advertising data were taken from Elliot Research and Media Measurement Institutes of Canada.

The authors tested for structural change using two semilog demand models, one for the period prior to the advertising campaign and one for the period containing the advertising campaign to estimate the per capita consumption of butter, margarine, shortening, and salad oils. The models were estimated using seemingly unrelated regression.

These models described each of the dependent variables as being a function of advertising expenditures for each specific good, a seasonal dummy variable, own-price, real total consumer expenditures on each good, and a dummy variable representing the presence or absence of the Dairy Bureau's advertising campaign. An interaction term was included to test the hypothesis that advertising causes structural change.

The goodness of fit measures revealed that the advertising campaign caused structural change. Another test for determining structural change was a switching regression model. A Chow test was then used to determine if the parameter estimates were significantly different. Results suggested that structural change did occur after the advertising campaign started.

The estimated expenditure coefficients for the four demand functions (butter, margarine, shortening, and salad oil) were inelastic and positive. The own-price elasticities were significant and negative. The cross-price elasticities between margarine and butter indicate that margarine was a close substitute for butter but that butter was a weak substitute for margarine. The estimated advertising elasticity for butter was the only statistically significant estimate and was inelastic and positive. The interaction term was only significant for butter. This implies that the advertising campaign made the demand for butter more elastic by stressing its unique characteristics. The Dairy Bureau variable was significant at the one percent level for margarine and oil. This suggests that margarine demand decreased as a result of the campaign while oil demand increased. The results indicate that the advertising campaign was successful in increasing demand for the advertised products.

Jones and Ward estimated the impact of generic and brand potato advertising on producer returns. They used a thirteen equation system to model the supply and demand relationships for potatoes. Demand equations for fresh, frozen, and chip potatoes were estimated. The effects of an advertising campaign depended heavily on the type of advertisement used. Brand advertising was used to gain market share and increase demand, while generic advertising focused on influencing consumer perceptions for the industry as a whole. Annual quantity, price (farm, wholesale, and retail), advertising, and consumption data were collected from 1970-85. The advertising data were obtained from the Leading National Advertisers' publication, and the other data were taken from Department of Commerce, USDA, and the Bureau of Labor Statistics.

The results indicate the ineffectiveness of generic advertising in stimulating demand for fresh potato consumption in spite of the fact that the majority of generic advertising was spent on fresh potatoes. Income growth, increased labor force participation of women, and away from home consumption were statistically significant and had negative impacts on fresh potato consumption. The inability of generic advertising to increase the demand of fresh potatoes was considered to be the result of the low levels of generic advertising which may have been too low to combat the negative perceptions of potatoes as being filling and fattening. Both generic and brand advertising had positive impacts on increasing demand for frozen potatoes. Increased income, more women in the workforce, and increases in away from home food consumption were statistically significant in determining the demand for frozen potatoes.

Unconditional ex ante forecasts of potato demand were generated by the reduced form equations. They were made by increasing the expenditure level of brand advertising by three and six percent while excluding generic advertising. Forecasts of the four different potato forms were generated for the 1985-94 period. The model predicted a 3.6 percent increase in chip demand. A three percent increase in frozen potato advertising was forecast to increases sales by 1.9 percent. The results suggest there was no significant cross-advertising effect between frozen potato and chip advertising and the other product forms. The authors suggested that with additional generic advertising expenditures the campaign could increase the demand for fresh potatoes.

Somers et al used a transfer function to estimate the relationships between manufacturer and retail advertising levels for a furniture manufacturing firm. The authors used manufacturer advertising levels as a predictor of retail advertising levels. They assumed that an increase in manufacturer advertising expenditures would result in increased retail advertising expenditures. The results of the transfer function model were compared with the regression forecast results.

The data used in the study were provided by a nationally recognized furniture manufacturer. Monthly corporate and retail advertising expenditures from 1980 through 1985 were used in the study. Eighty-five percent of the retail advertising was in the form of newspaper advertising with the remaining fifteen percent going to television and radio advertising. Inspection of the data series revealed it was stationary, but seasonality was present and had to be removed through a seasonal differencing of twelve months. Both the retail and manufacturing data were pre-whitened to remove autocorrelation. The direction of causality was checked using cross correlations and checking to see if current manufacturer advertising was correlated with past retail advertising. It was determined that manufacturer advertising influenced retail advertising.

The five month forecasts for both the transfer function and the regression forecasts were compared and evaluated. The transfer function forecast generated smaller forecast errors than the regression forecast. Theil's U for the transfer forecast was smaller than that for the regression forecast, which indicated that the transfer function provided the more accurate forecast. It also implies the transfer function technique was better at describing the relationship between manufacturer and retail advertising.

Hooley et el. conducted research which used the Box-Jenkins methodology to identify and include variables other than advertising that influence sales. A Koyck specification model (adstock), an ARIMA, and composite model of consisting of an ARIMA and regression model (adtrack) were estimated, and their forecast accuracies compared. The Koyck model described the dependent variable as a function of current and lagged advertising effects in which the lagged effects were assumed to decay over time. The adtrack forecasting model is a composite in which the error term was described as being a function of advertising.

Sales and national advertising information were available for 157 weeks. Only 151 weeks were used to generate the different forecasts, while the remaining data were used for trial forecasts. A comparison of each forecasting technique was also performed. The adstock model was estimated using the OLS procedure. The adstock model explained 56 percent of the variation found in the dependent variable. The Durbin-Watson statistic indicated there was no autocorrelation. Comparison of the predicted sales values to the six actual values revealed that the forecast error averaged 16.8 percent. The ARIMA model indicated a first order autoregressive process, and first differencing made it stationary. This ARIMA model explained about 60 percent of the variation found in sales and had an average forecast error of only 2.73 percent. The results suggested that the adtrack model had an average forecast error of less than two percent and explained 60 percent of the variation in the dependent variable. The adtrack model provided the best forecast given the evaluation criteria.

Venkateswaran and Kinnucan conducted research to evaluate Canada's extensive promotion to increase consumer awareness of and demand for milk. The Dairy Bureau of Canada implemented generic advertising and promotional campaigns for fluid milk, butter, and cheese to promote and increase fluid milk consumption among consumers, promote milk as a ingredient to be use in food preparation, and to promote away from home milk consumption. The objectives of this research were to determine if generic milk advertising increased consumption, if the increased revenues from the increased consumption justified advertising expenditures, and if optimal advertising expenditures were being used.

Venkateswaran and Kinnucan used double-log, semi-log, log-inverse, and inverse functions to estimate the per capita per quarter milk consumption, which included low-fat, skim, chocolate, and regular. The per capita milk consumption was considered to be a function own-price, orange juice prices, advertising expenditures for fluid milk in Ontario, income, consumers' age, quantity of milk consumed in periods t-n, and seasonal dummy variables. The models were evaluated using the goodness of fit measures and nonparametric tests. Newspaper, television, magazine, outdoor, and point of purchase advertising media were included in the model. Income was adjusted by dividing income by the consumers' price index. The data covered the period from the first quarter of 1973 through the last quarter of 1984.

The lowest R^2 was 0.75, nearly all of the estimated coefficients were significant at the 10 percent level, and each coefficient had the expected sign. A with-in sample forecast was generated for each model to determine predictive accuracy, and nonparametric tests of the standard residual sum of squares were used to determine the function with the best fit. Using the correlation coefficient, the results indicate that the difference in predictive accuracy of each of the models was negligible. Durbin's d-statistic rejected the null hypothesis of no difference in the goodness of fit measures between the competing models. The goodness of fit criterion suggested that the inverse form model provided the best fit as it generated the smallest sum of squared residuals. The estimated elasticities of demand were inelastic, the cross-price elasticities were negative and inelastic, the income elasticities were positive and inelastic, and the advertising elasticities were positive and inelastic. The elasticity estimates varied, to some degree, depending on the functional form of the demand model.

The cost effectiveness of the adverting campaign was estimated by comparing the cost of the campaign to the increase in farm revenues. The effectiveness of the advertising campaign was calculated for each demand model. The results indicate that the milk demand increased by 16.6 to 39.7 million liters depending on the individual model. The returns, the increased revenue expressed in dollars, to a dollar spent on advertising ranged from 9.28 to 23.6 depending on the model. The research also concluded that the marginal revenue of increased advertising was more than the marginal cost. The authors suggested that increased advertising would lead to further increases in demand. The average expenditure over the study period was $0.07 per person per quart of milk. The study reported that the optimal adverting level should be increased from $0.07 ($/person/quarter) to $0.12-$0.16 ($/person/quarter) depending on which model used.

Wilkinson, Paksoy, and Mason analyzed price changes, newspaper advertising, shelf space changes, and in-store promotions. The goods that were under observation were Camay soap (bath size), White House apple juice (32 ounce size), Manhattan rice (the one pound bag), and Piggly Wiggly frozen pie shells. Quantities of the goods sold were recorded as well as competitors' brands prices and quantities and alternative sizes of the good, advertising, and display space over 24 weeks. The researchers calculated 75 percent of the product's price, and this along with the retail price and the cost to the store were the pricing levels used. The method of display consisted of doubling the usual shelf space and using special displays. The advertising was carried out in the supermarket's weekly advertisements, and each advertisement had the same lettering, height, and mention of price and product name.

Price elasticities, substitution patterns, and price-sales relationships were estimated. Pie shells had an estimated inelastic own-price elasticity while juice, rice, and soap were found to have elastic own-price elasticities. The own-price elasticities, as might be expected, were most elastic for the soap and juice, because these products tend to have many perceived substitutes. With respect to the cross-price elasticities, substitution for the store brands of juice and rice decreased as the prices of the test products were lowered. Newspaper advertising was not very effective because demand did not increase significantly. The use of displays and increased shelf space seemed to be very effective ways to increase the sales of the test products. The analysis of the residuals and predicted variability of the model provided evidence in favor of the estimated promotional effects.

Funk, Meilke, and Huff conducted a study on the effects of retail pricing and advertising on the movement of beef (18 specific cuts). The data were collected over the period beginning January 1974 and ending May 1975. Advertising data were collected from an audit of five stores of a major food chain located in Toronto Canada. Price data were supplied by a pricing service. Data were also collected on the weekly sales levels of beef, prices, and weekly newspaper advertisements. Regression analysis led to an inference of elastic own-price responses, so decreasing the price of beef led to increased revenues. Cross-price elasticities were found to be unimportant in this particular study. The own-advertising elasticities were significant and positive for individual products as well for aggregate beef products, but less elastic than the own-price elasticities. The effects of competing products' advertising were found to be insignificant.

Marion and Walker conducted research pertaining to the response of specific meats to weekly prices at the retail level. This study was concerned with the relationships between prices and quantities on a weekly basis for meat products at the retail level. It tried to isolate the demand relationships in the very short run that affect managerial decisions like pricing, advertising, and inventory control. Five meat categories were followed over a 52 week period in two major supermarkets in Ohio. The own-price coefficients were found to be negative, and a majority of the cross-price coefficients were positive. Ten linear regression equations were used, one for each product. The results indicate that newspaper advertising was not significant in any of the models. The variable representing payday was significant. The study found that there was a difference in the quantities sold depending on the week of the month.

Carman and Figueroa conducted a study to analyze the factors that are associated with weekly food sales variation. Data were collected over a 105 week period that started in July 1978 and ended July 1980. Information was collected on sales by department, number of advertised specials by department, store coupons, advertising media used, and gross margin by department. The data were collected from ten stores in Ohio. The stores had variations in sales from 50,000 to 150,000 dollars a week. The study employed ordinary least squares regression analysis.

The study demonstrated that retail food sales tended to decrease as the time period since the last pay-day increased. There was a significant relationship between the percentage decrease in sales and the income level of the consumers who frequented a particular store. The variability in sales, expressed as a percentage, differed by department with meat experiencing the greatest degree of variability, followed by groceries, while produce had the least amount of variability.

Weeks two through four of a month, had negative signs and were significant at the 95 percent level. The coefficients of the variables increased in size as the week variable increased. The holiday variables representing Easter, the Fourth of July, Labor Day, and Christmas were all positive and significant, while New Years and Memorial Day were insignificant. Seasonal variables had negative coefficients and were significant. Trend variables were found to be insignificant.

The advertising variables all had positive signs, but television advertising for produce was the only variable found to be significant at the 95 percent level. Only produce experienced a sales increase from the use of coupons. The grocery special variable was significant and showed a positive relationship between store specials and store sales. The study revealed price was inelastic for all meats and that produce was a substitute for meat.

Curhan's study (1972) was unsuccessful in rejecting the null hypothesis that changing shelf space affected unit sales in supermarkets. Five hundred grocery products were studied, and shelf space was either increased or decreased for specific test items. Four regional stores, which were part of a chain, were used as test stores, and 24 other area stores were used as controls. Unit sales were monitored for five to twelve weeks prior to and after a change in a product's shelf space, (no date). The changes in shelf space were made on the recommendations of store managers and a computer space management system called COSMOS. COSMOS based it's recommendations on the profitability of a product per unit of shelf space it occupied.

The variables were retail price, brand type, market share, rate of sales, shelf capacity, merchandise variety, availability of substitutes, repurchase frequency, and extent of unplanned purchasing. Considerable preparation, minimum preparation, and ready to use categories were also used to help account for impulse purchasing.

Stepwise multiple regression analysis was used to analyze the data. The adjusted R^2 was 0.12. The independent variables also had large standard errors as did the dependent variable. The impact of a change in the shelf space of a product on unit sales had very little impact in relation to other factors that affect unit sales. The research did lead to insights into the shelf space elasticities for subsets of products like rate of sales, extent of display area, test store, and product category.

In another study by Curhan (1974), the effects of merchandising and promotional activities on the unit sales of fresh fruits and vegetables were estimated. The fresh fruit and vegetables were broken down into four groups: hard fruit, soft fruit, cooking vegetables, and salad vegetables. These four groups accounted for nearly all the fresh fruit and vegetables sold in the two supermarkets. Sales data were obtained through inventory counting and delivery records of two stores while display space, retail price, newspaper advertising, and display location quality data were determined by the researcher. The data were collected for two periods. The data obtained in the first period were collected over 12 weeks in the summer of 1972. The second part of the data were collected over a 17 week period in the following fall and winter.

A 7^2 factorial experiment design using a quarter factorial was used to analyze the data. This analysis provided information on certain variables and combinations of variables. The results suggested that an increase in space increased average unit sales of that category. For example, doubling the shelf space of hard fruit increased the category's average unit sales by 44 percent. The effect on unit sales from increasing shelf space of high priced soft fruit was greater than the effect on low priced soft fruit. Price promotion, a decrease in price, was not statistically significant except for soft fruit. This was unexpected because it is commonly assumed that price reductions increase unit sales. Advertising was significant only for hard fruits and cooking vegetables. The effects of advertising were extremely large for seasonal products. The display location quality was significant for hard fruits and cooking vegetables.

Manufacturers set aside large amounts of capital in an attempt to estimate the profit maximizing prices for their products (McLaughlin and Lesser). They have been unable to get accurate estimates nevertheless. Retailers, in general, do not set aside a marketing budget and tend to price the goods they carry either by judgement calls or rules-of-thumb. These techniques are generally good, but in the long run they may not be accurate.

McLaughlin and Lesser used scan data to study the effects of price variations on potato demand. Round, white, ten pound bags of potatoes were used in the study which lasted four weeks. The last week of the study and the following week were used to collect an exit survey of the customers. They concluded that demand differed by store, price changes caused potato sales to change, and consumers did not reduce weekly purchases after a surge in purchases brought on by reduced prices. Thus, decreasing the price of potatoes caused an increase in consumption, not just a shift in weekly sales. This showed that potatoe sales are responsive to price changes even though they are considered to have an inelastic demand.

The effects of promotional programs were analyzed by Whittnik <u>et al</u> The promotional variables to be investigated were temporary price reductions, displays (end-of-aisle), and feature advertisements (the brand name of the product was in the ad). Ten different markets were used in the study, and data were collected over a 52 week period. Competitors' products were also monitored.

The objectives of the study and model were to estimate the short-term effects of specific marketing programs on branded products using time-series data that showed variation among stores. The product used was tuna fish. Data were collected on three major national brands: Starkist, Chicken of the Sea, and Bumble Bee. The regional, private, and smaller brands were excluded. A 6.5 ounce can of chunk light tuna, which accounts for nearly 80 percent of the tuna sales, was selected for use.

The results of the study show that the own-price elasticities were negative and that the cross-price elasticities were positive. The use of displays as promotional activities increased sales and did not differ much between markets. The use of feature advertisements and displays together increased unit sales by roughly 75 percent in one particular market. The increase from the combined effects was 75 times greater than adding the effect of each variable if it was used separately. Analysis showed that brand switching only accounted for eight percent of the increase in the unit sales of Starkist.

The study also looked at toothpaste. When a display is used to promote the 8.2 ounce size of the product, a portion of the sales increase, 14 percent, came from cannibalism of the product's other sizes. The use of feature ads caused a cannibalism rate of 16 percent. Displays generated the greatest increases in sales and caused the least damage regarding sales loss from cannibalism of other size products. Different effects were observed for different promotional activities. Combining displays and feature ads increased sales more than implementing these strategies separately. It was also observed that displays increased brand switching more than other types of promotional activities.

Retail demands for the following goods were analyzed using scan data: beef (steak, ground beef, roast beef), chicken, and pork (pork chops, ham, and pork loin) by Capps (1989). The data were collected from a food retail firm in Houston and covered the period January 1986 to June 1987. The own-price elasticities were generally significantly different from zero and negative as expected. Ground beef's own-price elasticity was negative but not statistically significant. Ham had a positive own-price elasticity that was significant. The cross-price elasticities were generally significant and positive. The variable payday was insignificant. Seasonality was significant. Advertisement fliers increased sales significantly and had positive own-advertisement elasticities except for pork. Only five cross-advertising effects were significant out of the possible eighteen.

In another study, conducted by Tellis, advertising expenditures and gross rating points were used to measure market structure. Scan data were collected for toilet tissue for a one year period (no date). The number of rolls, dollar volume, coupon use, feature ad use, and display use data were also collected. Product movement was recorded in conjunction with monitoring television advertisements. Consumer response to an ad was stronger for brands for which they were loyal. The behavioral response to the effects of advertising was nonlinear. Advertising tended to be more effective at increasing unit sales through increasing the consumption of that product rather than attracting consumers from other products, or brand-switching. Price changes had the same effects on sales. He also noted that displays, coupons, and feature ads increased consumption of loyal consumers. Brand loyalty was a much stronger determinant of a consumer's purchasing decision than was advertising. The other promotional variables also had greater effects on sales than advertising.

By breaking down scan data, Culputa was able to observe the effectiveness of sales promotions and the origin of the sales increase. Scan data from IRI were used in the study. Data were collected on 2,000 households for a two year period between 1980-1982. The prices of products, promotional programs, household identification, and when the products were purchased were all recorded. Ground coffee was the product studied. Brand-switching accounted for 84 percent of the increase in sales as a result of a promotional program. The increase in sales by consumers purchasing an item early accounted for roughly 14 percent. Stockpiling on the other hand, resulted in two percent of the increased sales. Ninety-eight percent of the increase in sales that is seen following a price reduction is the result of brand-switching.

In a study conducted by Lattin and Bucklin, it was observed that if retailers and manufactures implement price changes too often, decreasing the price will no longer increase sales. This is because consumers no longer see the price reduction as a bargain but expect it. It is believed that consumers establish a base reference price for a good and when a promotional price is enacted, they see the reduction in price as a deal. Promotions if used too often will also loose their effectiveness. Consumers are less likely to purchase an item being promoted if the last purchase of the item was during a promotion. Consumers respond to promotional activities, but there is a better response if the promotion is not used on a regular basis. The data were provided by IRI and included price, value, and promotional programs. They were collected over a 75 week period (no date). Maximum likelihood regression techniques were used.

Lattin and Bucklin found that different promotional activities increased sales by different amounts. For example, a price cut of 10 percent for paper towels increased sales 22 percent, and when an ad and a price reduction were combined, sales increased by 177 percent. Promotional impacts varied by product categories, regions, and even neighborhoods within a region. A combination of price reductions and displays increased sales more than by separate increases of each of these activities. They noted it is important to follow sales for several weeks after a promotional activity is discontinued to see if stock piling did occur which would cause sales to drop.

The effects of advertising are rarely all seen in the present advertising period according to Kluyer and Brodie. There is a carry-over effect that can be seen in other periods. It is very difficult to account for the carry-over effect of advertising and promotional programs in other periods. Chocolate biscuits, liquid detergents, and toothpaste were the products used in this study. Fifteen data sets included 28 bimonthly observations over 1975-1980. The data included market share, relative price distribution intensity, and advertising share for chocolate biscuits, liquid detergents, and toothpaste. Market share and relative price were obtained from the Nielsen audits. Nonlinear regressions were estimated. The study found that other promotional variables (displays, price reductions, etc.) did not seem to carry over into other periods and that the results may be different for lesser developed or new products.

Walters and Rinne show that supermarkets use a variety of promotional programs to attract new customers and increase the supermarket's sales. Loss leader promotions (i.e., a retailer puts an item on sale at a price below retailers' cost) are thought to increase the store's profit by increasing traffic and attracting customers. Another method is to use double coupons. The belief behind these promotional programs is that increased traffic will result in increased sales of the nonpromoted higher margin products. The data used in this study were supplied by a grocery chain. Multiple linear regression was used to analyze the data. The study found that the bulk of sales increases came from the promoted low margin goods but not from the nonpromoted high margin goods. Therefore, the use of these promotional programs did not increase the profit of the stores as might be expected. The authors caution that the results may not apply in general because different regions respond in different ways to various marketing programs.

The effects of point of purchase, P-O-P, signs were used in a study by Archabal, McIntire, Bell, and Tucker to see if they had any effect on consumers' purchases. The signs, some of which related to nutritional values of foods, were under investigation to see if they increased a product's movement. Unbranded produce was used, so there was no brand switching induced by the P-O-P signs. Only six products out of the department's 50 items were issued signs for the study. Scan data from 373 stores over a 12 week period were collected (no date). A 3-way analysis of covariance was used. Consumers seemed to be unaffected by the use of P-O-P signs. This indicates that shoppers avoided foods they did not want instead of shopping to increase their nutritional levels. Hidden cameras were used to see if people looked up at the P-O-P signs. Only 4.5 percent of the shoppers glanced at the signs, and only 30 percent of the people that glanced at the signs looked at them for more than one second.

A study conducted by Moriarty (1985a) found that the use of displays increased sales of supermarkets, chain pharmacies, and independent pharmacies. The study used multiple regression, and the data were gathered from scanners (no date). Weekly unit sales, retail price, and newspaper feature advertisements were recorded. The increase was reported to be approximately 38 percent in the supermarkets and 107 percent in the pharmacies. More shelf movement of products was observed when displays were absent, meaning that products were taken from the shelf instead of the display, but no significant differences were found.

Moriarty (1985b) conducted another study to examine the effects of newspaper feature advertisements and price interactions. Data were supplied from five stores for 92-94 weeks (no date). Only one product, unspecified, was used in the study. Regression analysis was employed. The product in stores 1 and 2 had a large share of that product's market. The price of the product rarely changed so the own-price elasticity was hard to estimate. The product in stores 3-5 was promoted more heavily. The data for stores 1 and 2 were pooled, and the data for stores 3-5 were pooled.

In stores 1 and 2 the price reduction and feature advertisement interaction were not significant. This was felt to be due to the large market share, over 50 percent of the unit sales in stores 1 and 2, while stores 3-5 had market shares of under 30 percent. Stores 3-5 experienced a significant increase in sales by using feature advertisements and price reductions together. The interaction effect between price and feature advertizing was negative and significant. This gives rise to the conclusion that consumers responded more to price reductions in the presence of a feature ad than if no feature ad was present.

In a study conducted by Kumar and Leone, the significance of in-store promotions and brand and store substitution were tested. Sixty weeks of data were collected from ten stores using scanners (no date). The product used in the study was disposable diapers. Three major brands accounted for 95 percent of the market. The data were gathered in a southwestern city and contained information on volume, promotional activities, feature advertisements, and in-store displays. Price promotions, feature advertising, and display activities were all found to increase the sales of the particular brand of diapers they were promoting. The study concluded that the increase in sales came from brand switching, consumers switching stores, and general increased store traffic.

In a study by Jensen and Schroeter, the effects of television advertising were evaluated. Data were collected over a 92 week period, late 1985-mid 1987. Scanners supplied price and quantity data on 2,500 panel households. The households were separated into three groups. The first group was subjected to heavy levels of television advertising for a particular product, beef. The second group, was subjected to "base" levels of television advertising. The third group of households was the control group and was not subject to any product adverting. In the last 28 weeks of the study, both the heavy and base groups were exposed to intermediate advertising levels. Linear regression was used to analyze the data. The regression analysis indicated a strong positive correlation between feature ad prices and expenditures on beef. The coefficients on heavy and base advertising levels were found to be insignificant. A Chow test was used to draw an inference about whether the entire vector of parameters was equal for the three levels of advertising. At the 25 percent level, the hypothesis that there is no difference was not rejected. Thus, television advertising was found to be ineffective in stimulating the demand for beef. The study revealed hispanics consumed above-average amounts of beef while college-educated households that planned meals consumed below-average amounts of beef.

In a study by Gagnon and Osterhaus, scanners were used to collect data on pharmaceutical products and demographics. The data were collected in grocery stores and chain and independent pharmacies (no date). Generalized least squares was used to analyze the data. The author estimated the effect of floor displays on shelf unit sales, all other promotional activities were held constant. Displays were found to be significant in increasing product sales in all three of the retail outlets. In grocery stores the effect of floor displays increased sales 38 percent. In the pharmacies, both independent and chain, floor displays accounted for a increase in product sales of 107 percent. The effects of displays did not seem to have a negative impact on shelf unit sales.

D. Conclusions

Accurate forecasts can provide supermarket managers with useful information to be used in making business decisions regarding manufacturing schedules, financial objectives, and baseline business activity. Increased forecast accuracy could prevent the loss of overstocked perishable items such as fresh meats, fish, produce and deli items. These perishable items account for almost half of supermarket sales (Eastwood) and are prone to large fluctuations in demand. The ability to predict the demand for perishable, as well as other grocery products, accurately could provide a useful tool in developing market strategies. For example, accurate forecasts of consumer's responsiveness to price changes, advertising campaigns, seasonal changes, and holidays could provide useful information to store managers and aid them responding quickly and efficiently. Accurate predictions of consumer demand would help store managers reduce costs associated with large inventories, loss of perishable items, and avoidance of stockouts (Beilock and Dunn) and enable managers to implement just-in-time deliveries. An extensive literature review of pertinent forecasting, advertising, and promotional studies provided insight into the application and performance of different forecasting techniques, advertising, and promotional campaigns.

Reviewing the literature revealed the breadth of completed forecasting research. The research could be classified into a few broad categories. The first category involves research related to forecasting livestock (beef, pork, chicken, and lamb) producer prices. Research involved with nontraditional animal products has been limited. Numerous studies have focused on predicting macroeconomic indicator variables like GNP growth, interest, and bond rates.

Studies evaluating and comparing forecasting techniques on their abilities to predict commodity prices and supply accurately have been reported. The types of products found in this literature are varied.

The literature review provided insight into the lack of research involved with forecasting price or quantity movement of products at the retail level. Specifically, out of the entire literature review only a few studies were involved with product forecasting at the retail level. The survey was also valuable in providing insight into the problem of choosing a forecasting technique to predict the given variable accurately. No specific forecasting technique consistently out-performed the alternatives. No one forecasting technique is consistently superior in it's abilities to provide the most accurate forecast given varying circumstances. Thus, the present research will use the theoretical, statistical, and transfer function forecasting techniques to provide weekly forecasts of consumer demand.

Review of the pertinent advertising literature revealed a variety of advertising and promotional campaigns used in consumer demand research.

Notable among them are television, newspaper, magazine, outdoor, and point of purchase advertising as explanatory variables. The literature review revealed that none of the current research incorporated television, newspaper, and radio advertising in the same demand function.

The data used in both the forecasting and advertising literature were as varied as the estimation techniques. Single-purpose survey data, USDA public data, controlled experiential data, and other data sources were used in previous research. In an article by Tomek, he referred to the inadequacies of existing secondary data for use in retail demand analysis and that scan data could be an important source of information for analyzing retail demand. The literature review revealed a few studies which employed scan data in the estimation of empirical models.

Chapter IV

Methodolgy

A. Introduction

The goal of the present study is to explore the applications of alternative forecasting methods within the context of supermarket scan data. More specific objectives are to 1) develop alternative forecasting methods that are suitable for scan data 2) estimate and compare the alternatives with respect to food groups and individual products in terms of their forecast accuracies using a scan data base, and 3) estimate and compare the alternatives with respect to food groups and individual products in terms of two week trial forecasts.

Because the current literature suggests that no one forecasting technique can provide the most accurate forecast, this study employs three different forecasting techniques to provide sales forecasts. The first is the theoretical forecasting technique. Economic theory and previous research are used to specify the theoretical demand function. The estimated model provides insight into the nature and the significance of the relationships found between the dependent and independent variables. The second is the statistical technique which attempts to identify and reproduce patterns contained in a data series. The third technique combines the theoretical and statistical techniques to obtain a transfer function model. The underlying logic behind using a composite forecasting technique is that combining theoretical and statistical models introduces more information into a forecast than is contained in a theoretical or statistical forecast.

B. Theoretical Model

Analysis of consumer demand for a subset of goods requires the assumption that the consumer's utility function is weakly separable. Weak separability is a condition which implies the marginal rate of substitution, MRS, between goods contained in a subset be independent of the quantity demanded for any other good in a different subset. For example, if there are two goods, X and Z, which are considered to be substitutes and in the same subgroup, weak separability implies the MRS between X and Z is independent of the demand for any other good in another subgroup.

Holdren provides a framework for formulating multiproduct retail demand functions. He describes the retail demand for a particular product as being a function of price and nonprice (advertisements, promotional campaigns, operational hours, customer service, etc.) attributes of the retailer. Following the framework of Holdren, the theoretical demand function includes own- and cross-prices. This is consistent with economic theory and current research. The estimated own-price coefficient should have a negative sign as suggested by economic theory because as the price of good i increases the quantity demanded of that good i decreases as consumers substitute relatively cheaper goods in its place. The sign of the estimated coefficient associated with the price of good j depends on whether it is a substitute or complement. Goods i and j are considered substitute goods if the demand for good i increases as the price for good j increases because consumers substitute the relatively cheaper good i in its place. Thus, substitute goods are expected to have positive coefficient estimates. If goods i and j are complementary, the estimated price coefficient should be negative. Goods i and j are considered to be complementary if the demand for i (j) decreases as the price of good j (i) increases.

Holdren suggests that advertising and promotional campaigns are implemented to increase consumer traffic. Television and radio advertising are common marketing techniques used in relaying nonprice attributes for a given supermarket to consumers. These advertisements attempt to entice those who do not frequent a given supermarket into shopping there as well as to relay product and price information to area consumers. As consumers respond to the advertisement, the store experiences increased consumer traffic.

The increase in consumer traffic can have two effects on store sales. The first is a direct effect as consumers respond to the advertisement and store traffic is increased leading to increased overall sales. The second is an indirect effect referred to as impulse buying.

Impulse buying occurs when consumers purchase a given good at the point of sale when they had no original intentions of doing so. Impulse buying can account for 30 percent or more of a supermarket's total sales (Clover). Previous research suggests that advertising (television, radio, and newspaper) can affect consumers' preferences and, therefore, can affect consumer demand (Capps 1989; Capps and Lambregts; Carman and Figueroa; Curhan 1972; Funk et al; Hooley and Wilson; Jensen and Schroeter; Jones and Ward; Kinnuncan and Venkateswaran; Marion and Walker; Moriarty 1985b; Powers; Somers et al; Tellis; Venkateswaran and Kinnuncan;

Wilkinson et al; and Whittnik et al). The effects of television and radio advertising may last more than one period because of the information they contain. These advertising media may have a significant lag effect on supermarket sales (Chang and Kinnuncan; Eastwood, Gray, and Brooker, 1993).

Newspaper advertising provides consumers with specific product price information for a limited period of time (Powers). Powers' research concluded that lagged newspaper advertising variables were insignificant. He reported that 95 percent of the increase in demand occurred during the advertising period. Similar results were found by Eastwood, Gray, and Brooker (1993). Thus, it is assumed that the majority of newspaper advertising's influence on consumer demand occurs during the advertising period because the advertisement is primarily providing current price information. Therefore, the theoretical demand function will not include lagged newspaper advertising variables.

Economic theory can be extended to include advertising by way of influencing consumer preferences and providing price information. However, the empirical results are mixed. Capps and Lambregts; Powers; Chang and Kinnuncan; Venkateswaran and Kinnuncan; Funk et al; Carman and Figueroa; and Jensen and Schroeter have concluded that advertising significantly increases consumer demand. Eastwood, Gray, and Brooker, 1991b; Kinnuncan and Venkateswaran; Jones and Ward; and Wilkinson et al concluded that advertising is insignificant in increasing consumer demand.

After examination of the data, a seasonal variable may be included in good i's demand function to represent the seasonal fluctuations. Different products are prone to different periods of seasonality. Capps, 1989; Capps and Lambregts; and Carman and Figueroa concluded that seasonality significantly impacted the demand for beef, chicken and pork products.

Eastwood, Gray, and Brooker, 1991a; Powers; Capps (1989); Carman and Figueroa; and Capps and Lambregts all included holiday variables in their estimated demand functions. These studies have concluded that the presence of a given holiday, in period t, is significant in influencing the level of demand in period t. The estimated holiday coefficient represents the change in quantity sold associated with a given holiday, given all other variables held constant.

The estimated coefficients for seasonality and holiday variables are dependent on the season's and/or holiday's effect on the demand for a given good. A product might experience a negative relationship with a particular season and/or holiday while exhibiting a positive relationship with a different season or holiday. For example, roast sales generally decrease during the summer and increase during the winter (Eastwood, Gray, and Brooker, 1990).

Weekly data are used in the study, and several arguments can be made to justify the aggregation. The availability of weekly forecasts to a store manager is essential to the efficient operation of a supermarket. Managers implement weekly promotional and advertising campaigns which coincide with the consumer's planning and/or budgeting period (Capps and Lambregts;Capps, 1989; Eastwood, Gray, and Brooker, 1991a; Eastwood, Gray, and Brooker, 1991b; Eastwood, Gray, and Brooker, 1991c). Another argument is that store managers use weekly predictions to schedule labor, keep track of inventory, and monitor individual departments' gross margins. The weekly time period would provide managers with insight into the relationship between unit sales and advertising and/or promotional campaigns. For example, weekly demand forecasts could provide store managers with information concerning consumer responsiveness to price (Thayer), advertising, seasonality, and holiday variations. This information could be useful in maintaining satisfactory inventory levels. Managers can identify the periods of high consumer traffic on a weekly basis and arrange labor schedules to increase efficiency, thus reducing the supermarket's labor cost (Thayer). These arguments support the importance of weekly information to supermarket managers.

The theoretical demand function for food group g is (Chapter II, equation 1).

(1) $Q_{gt}=F_g(P_{gt}, ADV_{gt}, SEA_{gt}, HOL_{gt}) + \epsilon_{gt}$, where

Q_{gt} = quantity sold of food group g in week t,

P_{gt} = vector of own- and substitute prices of food group g in week t,

A_{gt} = vector containing television, radio, and newspaper advertising of food group g in week t,

$SEAg_{gt}$ = vector of weekly binary variables to measure seasonality for food group g in week t,

HOL_{gt} = vector of weekly binary variables representing holiday periods for food group g in week t, and

ϵ_{gt}= random disturbance term for food group g in week t.

The theoretical demand function for brand b is as follows.

(2) $Q_{bt}=F_b(P_{bt}, P_{jt}, ADV_{bt}, POP_{bt}, SEA_{bt}, HOL_{bt}) + \epsilon_{bt}$, where

Q_{bt} = quantity sold of brand b in week t,

P_{bt} = own-price of brand b in week t,

P_{jt} = vector of substitute prices in week t ($j \neq b$),

A_{bt} = vector containing television, radio, and newspaper advertising of brand b in week t,

POP_{bt} = vector point of purchase advertising of brand b in week t,

SEA_{bt} = vector of weekly binary variables to measure seasonality for brand b in week t,

HOL_{bt} = vector of weekly binary variables representing holiday periods for brand b in week t, and

ϵ_{bt}= random disturbance term for brand b in week t.

Multicollinearity may be present among the subsitiute prices due to similar pricing strategies. The problem of multicollinearity is reduced by calculating a price index for the subsitute prices.

C. Statistical Model

An ARIMA statistical forecast, using the Box-Jenkins ARIMA methodology, is generated and used as a basis for comparison with the theoretical and composite forecasts. The procedure provides information for deciding whether an AR(p), MA(q), ARMA(p,q), or ARIMA(p,d,q) process is present.

The statistical model, ARIMA(p,d,q), (Chapter II, equation 32) for each of the selected food groups is as follows:

(3) $\varepsilon_t = \theta_q^{-1}(B)\phi_p(B)z_t$ where

ϕ = the AR(p) operator,

θ = the MA(q) operator,

z_t = the transformed stationary data series, and

B = the back shift operator.

D. Composite Model

The general transfer function (Chapter II, equation 42) is based on the econometric model described above, with appropriate modifications to accommodate the Box-Jenkins framework.

(4) $Q_{gt}=F_g(P_{gt},\ ADV_{gt},\ SEA_{gt},\ HOL_{gt},\ \theta_q^{-1}(B)\phi_p(B)z_t\ )$.

(5) $Q_{bt}=F_b(P_{bt},\ P_{jt},\ ADV_{bt},\ POP_{bt},\ SEA_{bt},\ HOL_{bt},\ \theta_q^{-1}(B)\phi_p(B)z_t\ )$.

The theoretical, statistical, and composite models are used estimate and compare forecasts described in objectives 2 and 3. The scan data base is divided into two subsets. One is to estimate the relationships and to generate historic record forecasts. The other is to generate two-week ahead ex post forecasts. The forecasts are updated weekly as additional information becomes available. The updated information is used to generate a new two-week ahead forecast. The two-week ahead forecast is considered to be appropriate because it corresponds to the amount of time required to provide local managers with a forecast. The two-week turn-around period occurs because of the elapsed time associated with forwarding the weekly scan data to corporate headquarters where they can be analyzed and used in generating weekly predictions. The weekly forecasts can then be forwarded to local managers (Eastwood, Gray, and

Brooker, 1991a).

E. Data

Supermarket management is concerned with increasing category sales as a method of increasing total sales revenue. Raju suggests that increased brand sales do not necessarily result in increased category sales. For example, a sales increase for one particular product in a category might lower the sales of competing products as consumers switch brands. The substitution between the promoted product and competing products may result in no significant change in total category sales. Product category sales are also important to supermarket managers because of the lack of brand specific products in certain departments (e.g. produce, fresh meats). These products are generally perishable which increase the need for accurate sales predictions to avoid losses from spoiling products.

Another reason for using product categories is because of the number of products available to consumers in supermarkets. A typical store offers 20 to 40 thousand products (Capps, 1989). For example, there are 1,700 bar codes representing meat products in the five stores (Eastwood, Gray, and Brooker, 1991b). This includes different package sizes of the same brand product and substitute brands.

Consumers, both domestically and globally, have changed their attitudes regarding brand and private label products. Consumer focus is changing from brand consciousness to price consciousness (Schiller). For example, in Britain, 32 percent of total consumer grocery expenditures consist of private label products. This trend is also expected to continue in the United States (Oster, Savery, and Templeman). In the United States this has led to increased sales of store brands which are cheaper than their competing national brands. This situation provides another motivation for using scan data to evaluate food demand at the product specific level. The problem of multicollinearity may arise if individual products were used in place of product categories. Multicollinearity emerges because of the large number of available substitutes and complements for brand b.

This dissertation forecasts the per customer consumption of two product categories and individual brands within the product group. Sales per thousand customers is used to account for the variation in sales between different stores. The variation in sales could be greater in the larger supermarkets compared to the smaller supermarkets because of consumer traffic. Sales variation could also be attributed to the nonreporting of stores in various weeks. Thus, a common unit of measurement is necessary to measure the impacts of the supermarkets advertising campaign.

The information for each of the selected product groups and individual products, as well as for substitute and complementary goods, was obtained from the scan data of a national supermarket chain operating five stores in the Knoxville metropolitan area. The data were collected over a 229 week period, May 28, 1988 through the last week in December of 1992. The supermarket chain represents a large share of the area's supermarket sales. The data are divided into two subgroups. The first is used to estimate the three forecasting models. The second is used to generate forecasts for use in comparing the three forecasting techniques.

Missing data can be the result of stockouts, the item being discontinued, a new product not available previously, a mechanical failure at the store, or technical difficulty at the headquarter's computer facility. A mechanical failure results in the data not being recorded and lost. Computer difficulty at the supermarket's headquarters prevents the weekly sales data from being transferred from the store accounts to the headquarter's computer. The inability to transfer the data means that new sales data are added to the previous week's and is called a "running total." This continues until the problem at the headquarter's computer facility is corrected. To adjust for this problem, the data for the combined weeks are divided by the number of weeks the running total was in effect. The interim weeks are not included in the data set to avoid entering incorrect data into the data set more than once. The problem of missing data has diminished over time.

Each product in the supermarket is assigned an individual code, by a national organization, which is called a universal product code or UPC code. The scanner system reads each product's UPC code as it is scanned at the register to process customer bills, and the number of times a bar code is read and the product's price are recorded by the management's software. Item movement refers to the number of times a bar code is read in a given period. The UPC code and its corresponding product description and package size are used to identify individual brand products. The weekly total item movements for each bar code for each supermarket are forwarded to the chain's corporate headquarters once a week. Good i's weekly item movement is calculated by aggregating (across customers, registers, and time) over a seven day period. Copies of the computer tape are sent to the University where the information is added to the historic record. The scan data set includes UPC codes and their corresponding item movements, product description, package sizes, and retail prices. These data also identify products on the basis of the chain's management structure. All products are grouped according to a three level system. The broadest group is the department followed by commodity and subcommodity.

Sorting the scan data by UPC cannot provide a useful list of the various package sizes marketed under brand b and substitute products because bar codes are not ordered by product (Eastwood, Gray, and Brooker, 1990). For example, the UPC bar codes for 18 ounce containers of Peter Pan, Jiffy, and Skippy creamy peanut butter are completely different (4530000040, 3700000407, and 4600127064, respectively).

The weekly data must be sorted using another procedure. Sorting the data by department, commodity group, and subcommodity groups provides a list of the various package sizes marketed under brand b as well as like products marketed under different brand names. A SAS program using the UPC codes of a few products, representing different brand products, was used to read six tapes to obtain the department, commodity, and subcommodity numbers for each brand. The department, commodity, and subcommodity numbers were entered into another SAS program which provided complete listing of brand products contained in the three specified categories. The list of brand b's various package sizes and product descriptions enabled brand groups to be compiled.

Once a list containing the various package sizes for brand b is compiled, it is possible to determine brand b's weekly quantity and weighted price by store. For fixed weight goods, quantity sold is item movement times the package size (ounces). Brand b's weekly quantity (e.g. equation 6) is determined by summing the weekly quantities of the individual products within a brand category.

$$(6) \qquad Q_{b,s,t} = \sum_{i=1}^{n_b} Q_{i,s,t},$$

where $Q_{i,s,t}$ = the quantity of good i sold in store s in week t,

$Q_{b,s,t}$ = the quanity of brand b sold in store s in week t, and

n_b = the number of food (bar codes) in brand b.

The price for the individual product brand is a weighted average price and is calculated using the price per package that is provided by the scan data for individual UPC codes. The price per ounce for good i is derived by dividing the price of the product by the number of ounces (standardized unit) contained in the package. The brand price (e.g. equation 7) is calculated by summing the price times quantity for each of the reporting stores and then dividing by the brand's total weekly quantity.

$$(7) \quad P_{b,s,t} = \frac{\sum_{i=1}^{n_b} (P_{i,s,t} * Q_{i,s,t})}{Q_{b,s,t}},$$

where $P_{i,s,t}$ = the price of good i sold in store s in week t, and

$P_{b,s,t}$ = the average price of brand b sold in store s in week t.

Substitutes and complements were identified from the lists of items in the selected subcommodities and included. Bar codes for close substitutes were used to define product categories. These product groups were updated quarterly to capture new products and update brand b's product mix. Thus, a general product category combines the various package sizes associated with a product as well as close substitutes. The department, commodity, and subcommodity lists provided the different package sizes associated with brand b as well as like products marketed under different names. This provided a list indicating products that are considered to compete with brand b. Brand b, and the competing products are combined into a single product group representing a general product category. For example, Skippy peanut butter and competing brands, including the various package sizes, are combined into a single product group called group g.

The quantity for product group g (e.g. equation 8) is determined by summing the quantities of the individual brands contained within product group g. That is, product group g's aggregate quantity is calculated by summing the quantities for each brand b contained within product group g. To obtain a weighted average price for product group g (e.g. equation 9), each brand's price per ounce is weighted by its quantity sold, and the sum is divided by the group's quantity sold.

$$(8)\quad Q_{g,s,t} = \sum_{b=1}^{B_g} Q_{b,s,t},$$

$$(9)\quad P_{g,s,t} = \frac{\sum_{b=1}^{B_g} (p_{b,s,t} * Q_{b,s,w})}{Q_{b,s,t}},$$

where $Q_{g,s,t}$ = the quantity of group g sold in store s in week t,
$P_{g,s,t}$ = the weighted average price of group g sold in store s in week t, and

B_g = the number of brands contained in group g.

Similar procedures are used for variable weight items. However, item movement is used as a proxy for pounds sold under the assumption that the distribution of package sizes per week does not change. The prices of variable weight products are expressed in prices per standard unit, \$/pound. Therefore, conversion into a price per standard unit is not necessary.

Aggregating weekly per store item movement to obtain product group g's total weekly item movement introduces two obvious problems. The first is that product group g's weekly item movement varies from one store to another. A store with a large weekly customer count, the number of individual sales receipts per store in week t, should logically have larger total sales compared to a store with a relatively small customer count. Second, nonreporting of stores influences the level of sales for product group g in week t. To resolve these two problems, product group g's weekly item movement per store is divided by the store's corresponding weekly customer count. Thus, the aggregated item movement across the reporting stores is independent of variation attributed to store size and/or nonreporting of stores.

Weekly customer counts are obtained by aggregating the number of individual sales receipts per store, over a week. This information was not contained in the weekly scan data provided by the supermarkets headquarters but was obtained from the regional technical marketing specialist.

The aggregated weekly price for product group g is calculated by summing all the reporting stores' weekly per unit prices times their respective weekly item movements, for product group g, and then dividing this total by the aggregated weekly item movement in standardized units. This provides a weighted average price index which represents the price of variable weight product group g.

The five supermarkets are all located in the same metropolitan area, so consumers are exposed to identical advertisements in all three of the advertising media. The advertising data represent advertising exposure. Weekly television and radio advertising data were provided by the supermarket's regional marketing manager and were expressed as gross rating points (GPR). Gross rating points is defined as the number of times an audience is exposed to an advertisement, television or radio, multiplied by the relative size of the audience (Eastwood, Gray, and Brooker, 1991 b). Gross rating points are used as a proxy to represent supermarket advertising levels (Tellis).

The second part of the advertising data, newspaper advertising, was obtained through a different source. Newspaper advertising data were collected from the Monday supplemental advertising section and occasional daily advertisements present in the Knoxville News Sentinel. The size of the advertisement in square inches was recorded for each newspaper advertisement. These square inches were aggregated to obtain measures for food group g and brand b in week t. This is consistent with Capps' 1989 research on retail meat demand in which newspaper advertising are recorded as squared centimeters. Squared inches of newspaper advertising are used as a proxy for newspaper advertising intensity. As square inches increase the number of locations increase. The use of different color advertisements also can be reflected in advertisement square inches because color ads tend to be larger.

Cross-advertising effects are not included in the demand function for the meat food category g. Funk, Meilke, and Huff, Marion and Walker found no significant relationship between meat sales and cross-advertising for meat products. The advertising for complimentary goods will be included in the demand function for Peanut butter. Cross-advertising will be incorporated into brand b's demand function to evaluate the impact of competing brands' advertising levels on the demand for brand b.

Before the promotional and advertising data can be incorporated into the demand model, it must be linked to the item movement of the promoted product. This is a time-consuming task because the promotional and advertising campaigns are not directly related to the promoted product's bar code. The information regarding the product's price is programmed into a store's computer price file using inventory codes and not related to the product's UPC code. Since the promotion does not provide the UPC code for the promoted product, it is manually matched using the product description that is provided along with the product's UPC code. The promotional measure in week t is then recorded (Eastwood, Gray, and Brooker, 1990).

The uniform code council has only assigned ranges of bar codes for variable wight items. Consequently, chains have some latitude in assigning bar codes to individual cuts of fresh beef, and the names of the cuts are not standardized. Occasionally, the bar code relating a specific cut could not be found for the promoted cut name. This problem is overcome by aggregating products into general product categories and relating the advertising level to a specific general category (Eastwood, 1990).

Point of purchase, P-O-P, advertising is an important promotional technique. However, P-O-P advertising data were not included in the demand function because at least one item within an aggregate was promoted nearly every week, this does not apply for brands. There is variation in the P-O-P variable at the brand level, and it is included in the brand's demand function.

Coupons are not included in the specified functions of the fixed and nongrocery products, although they are an important promotional technique. The chain's software does not capture the use of coupons to purchase specific products. Coupons are primarily manufacturer oriented and are not store-specific, although double and triple redemptions are store specific. Because this research is concerned with product/food groups and coupons have dates that cover multiple weeks, a dummy variable for coupons was not included. The omission of coupons is less problematic for the variable weight products because ground beef, roast, and steak are typically not promoted using coupons.

Scan data capture price and item movement information but not customer socioeconomic data. Therefore, matching a customer and his or her socioeconomic characteristics with the purchase of a product in this data base is impossible. The data used in this research cover a four and a half year period. Due to the relatively short time period, it is assumed that the socioeconomic composition of the metropolitan area has remained stable.

Two product categories are used. The logic behind selecting these products is that the forecasting literature revealed that different forecasting techniques provided superior forecasts for different products. Spriggs found the theoretical forecasting technique superior to the statistical and comparable to the composite (theoretical-statistical) techniques. Vere and Griffith reached the conclusion that no single technique (theoretical, statistical, or composite) provided a superior forecast. Hudson and Capps compared and evaluated the forecasting abilities of the theoretical and statistical techniques and suggested the superiority of the statistical method. Brandt and Bessler, 1981, evaluated several individual and composite forecasts and concluded that the poorest composite forecast was superior to the best individual forecasts. Thus, the forecasting abilities of the three techniques are compared and evaluated, using the MSE, RMSE, Chi-squared statistic, AIC, Theil's U, and directional accuracy criterion, to gain insight into the predictability of the three techniques given the different product categories.

Chapter V

Estimation and Analysis

A. Introduction

This chapter presents the results of the empirical work which are used to address the objectives of the study. It begins with descriptions of the dependent (item movement per thousand customers) and independent variables (own-price and advertising). The dependent variable for each of the three products is reported as standardized item movement per thousand customers, but for convenience it will be referred to as item movement. The advertising variables (television, radio, newspaper, and point of purchase) for each of the three products are described over time. The point of purchase (P-O-P) advertising variable is only present in the brand product's demand function.

B. Variable Description

Weekly item movement, price, advertising (television, radio, and newspaper), seasonal, holiday, and special event variables were measured over a 239 week period beginning with the first week of June of 1988 (06/04/88) and ending with the last week of December, 1992 (12/26/92). P-O-P advertising data did not become available until the last week of May, 1989 (05/27/89) and continued through last week of December, 1992 (12/26/92). Aggregating the P-O-P data for the group g and steak models revealed that the P-O-P variable was present in every time period so there was no reason to include this particular variable in these theoretical demand models. This provided a 239 week period for use in analyzing weekly item movement for group g and steak. On the other hand, the inclusion of own and cross P-O-P variables in brand b's theoretical forecasting model meant using a data set with fewer observations. The first year of data was excluded because of the lack of P-O-P advertising data. Thus, brand b's data set consisted of 187 observations starting with the last week of May 1988 and ending with the last week of December, 1992. Results in Table 1 include the description of the food categories used in the study.

Definitions of the independent variables used in the theoretical model for b are in Table 2. Brand b's price and the ten prices of competing peanut butter brands are weighted averages. Because three brands of peanut butter were not present during the entire study period, their prices were included in the theoretical forecasting model as binary variables (1=product was present in week t, 0=if product was not present in week t). Two of the peanut butter brands were present in the supermarket for less than 48 weeks out of the entire 161 week period while the other brand was present for slightly over one and a half years (72 weeks). One of the missing brands was present in the first quarter of the study while the remaining two were present during the middle of the study period. The prices of the three brands were used in calculating the weighted average price when the brands were present in the supermarkets.

Brand b's weekly item movement is the sum of the total ounces of brand b sold in week t. Descriptive statistics for brand b's historical and trial data are presented in Tables 3 and 4. Inspection of Table 3 reveals that brand b's mean price was the third highest of the seven entire period product prices, and its standard deviation was the second lowest. The relatively low standard deviation indicates that brand b's

Table 1. Category definitions for brand b, group g, and steak

Dep. Var.	Definition
Brand b	A specific brand of peanut butter aggregated across container size and product variations (creamy, crunchy, low salt. etc.). Brand b consists of nine different products (various package sizes and product variations) packaged under brand b's label.
Group g	The peanut butter category aggregated across brands. The individual brands contained within the group g are aggregated across container size and product variations (creamy, crunchy, low salt, etc.). Group g consists of ten individual brands of peanut butter.
Steak	The steak category represents an aggregate of all steak products across "cut" and size. The steak category consists of 60 different steak products.

Table 2. Brand b's theoretical model variable definitions

Variable	Definition
P_b	Brand b's weekly weighted average price.
P_{1-11}	Substitute brand's weekly weighted average price.
NWS_b	Brand b's weekly newspaper advertising, measured in square inches.
NWS_{1-11}	Substitute brand's weekly newspaper advertising, measured in square inches.
RD_b	Brand b's weekly television advertising, measured in gross rating points.
RD_{1-11}	Substitute brand b's weekly radio advertising, measured in square inches.
TV_b	Brand b's weekly television advertising, measured in gross rating points.
TV_{1-11}	Substitute brand's weekly television advertising, measured in gross rating points.
POP_b	Weekly point of purchase advertising, measured as being present or absent.
POP_{1-11}	Weekly point of purchase advertising, measured as being present or absent.
HOL_1	Represents the week containing Memorial day.
HOL_2	Represents the week containing the Fourth of July.
HOL_3	Represents the week containing Labor day.
HOL_4	Represents the week containing Thanksgiving day.
HOL_5	Represents the week containing Christmas day.
HOL_6	Represents the week containing January first.
SEA_1	Represents the spring season.
SEA_2	Represents the summer season.
SEA_3	Represents the fall season.
SCH	Represents the week containing the first day of school in Knox County.

Table 3. Descriptive statistics of the dependent and independent variables contained in brand b's theoretical forecasting model (historic period)

Variable	Mean	Std. Dev.	Minimum	Maximum
Q_b	1123.79	292.64	294.84	1844.24
P_b	0.12	0.01	0.10	0.13
P_1	0.09	0.01	0.06	0.11
P_2	0.20	0.39	0.00	1.00
P_3	0.15	0.36	0.00	1.00
P_4	0.46	0.50	0.00	1.00
P_5	0.12	0.02	0.07	0.18
P_6	0.11	0.01	0.10	0.12
P_7	0.12	0.01	0.10	0.13
P_8	0.14	0.01	0.12	0.17
P_9	0.10	0.01	0.07	0.14
P_{10}	0.76	0.42	0.00	1.00
NWS_b	0.42	2.16	0.00	16.31
NWS_1	1.76	4.45	0.00	22.50
NWS_5	0.07	0.85	0.00	10.50
NWS_6	0.31	1.79	0.00	13.51
NWS_7	0.02	0.25	0.00	3.25
NWS_9	0.02	0.22	0.00	2.81
TV_1	1.91	23.94	0.00	300.00
RD_b	3.82	33.75	0.00	300.00
RD_1	5.73	28.84	0.00	150.00
POP_b	0.61	0.48	0.00	1.00
POP_1	0.87	0.32	0.00	1.00
POP_6	0.74	0.43	0.00	1.00
POP_7	0.57	0.49	0.00	1.00
POP_8	0.15	0.36	0.00	1.00
POP_9	0.44	0.49	0.00	1.00
HOL^1	0.02	0.13	0.00	1.00
HOL_2	0.02	0.13	0.00	1.00
HOL_3	0.02	0.13	0.00	1.00
HOL_4	0.02	0.13	0.00	1.00
HOL_5	0.02	0.13	0.00	1.00
HOL_6	0.02	0.13	0.00	1.00
SEA^1	0.25	0.43	0.00	1.00
SEA_2	0.25	0.43	0.00	1.00
SEA_3	0.25	0.43	0.00	1.00
SCH	0.02	0.13	0.00	1.00

Table 4. Descriptive statistics of the dependent and independent variables contained in brand b's theoretical forecasting model (trial period)

Var.	Mean	Std Dev	Minimum	Maximum
Q_b	835.36	253.96	183.77	1214.10
P_b	0.11	0.00	0.11	0.12
P_1	0.08	0.00	0.08	0.09
P_2	1.00	0.00	1.00	1.00
P_3	1.00	0.00	1.00	1.00
P_4	1.00	0.00	1.00	1.00
P_5	0.10	0.01	0.09	0.11
P_6	0.11	0.00	0.10	0.12
P_7	0.10	0.00	0.09	0.10
P_8	0.14	0.00	0.14	0.15
P_9	0.06	0.02	0.01	0.10
P_{10}	1.00	0.00	1.00	1.00
NWS_b	0.11	0.58	0.00	3.00
NWS_1	1.91	3.23	0.00	10.75
NWS_5	0.00	0.00	0.00	0.00
NWS_6	0.00	0.00	0.00	0.00
NWS_7	0.00	0.00	0.00	0.00
NWS_9	0.00	0.00	0.00	0.00
TV_1	0.00	0.00	0.00	0.00
RD_b	0.00	0.00	0.00	0.00
RD_1	0.00	0.00	0.00	0.00
POP_b	1.00	0.00	1.00	1.00
POP_1	1.00	0.00	1.00	1.00
POP_6	0.96	0.19	0.00	1.00
POP_7	1.00	0.00	1.00	1.00
POP_8	0.38	0.49	0.00	1.00
POP_9	0.80	0.40	0.00	1.00
HOL_1	0.00	0.00	0.00	0.00
HOL_2	0.00	0.00	0.00	0.00
HOL_3	0.03	0.19	0.00	1.00
HOL_4	0.03	0.19	0.00	1.00
HOL_5	0.03	0.19	0.00	1.00
HOL_6	0.03	0.19	0.00	1.00
SEA_1	0.00	0.00	0.00	0.00
SEA_2	0.50	0.50	0.00	1.00
SEA_3	0.50	0.50	0.00	1.00
SCH	0.03	0.19	0.00	1.00

price varied somewhat less than did the prices of the other competing brands, with one exception.

Independent variable definitions for group g are found in Table 5. Group g's price was calculated using a weighted average procedure. Again, due to the absence of three of the ten peanut butter brands over the study period, group g's price consisted of a weighted average of the remaining seven brands. The three brands were included in group g's weighted average price for the time periods they were present. Group g's weekly item movement, in ounces per thousand customers, is the sum of all peanut butter brands sold in week t. Descriptive statistics for group g's historical and trial data are presented in Tables 6 and 7.

Comparison of brand b's and group g's mean weekly item movements disclosed that brand b contributed to roughly sixty percent (60%) of group g's weekly item movement, brand b's and group g's mean weekly item movements are 1123.79 and 1836.87 ounces, respectively. The remaining nine peanut butter brands comprise the remaining forty percent (40%) of group g's weekly item movement. The standard deviation for brand b's weekly item movement is smaller (292.64) than the standard deviation for group g's weekly item movement (600.10). Brand b's mean price ($0.1184/ounce) was higher than group g's mean price ($0.1099/ounce). The standard deviation for brand b's price (0.0075) is smaller than the standard deviation for group g's price (0.0108).

Steak variable definitions are found in Table 8. Descriptive statistics for steak's historical and trial data are presented in Tables 9 and 10. The price used to represent the steak category is a weighted average price. The mean weekly price of steak is $4.14 with a standard

Table 5. Group g's theoretical model variable definitions

Variable	Definition
P_g	Group g's weekly weighted average price.
NEW_g	Group g's weekly newspaper advertising, measured in square inches.
TV_g	Group g's weekly television advertising, measured in gross ratings points.
RD_g	Group g's weekly radio advertising, measured in square inches.
HOL_1	Represents the week containing January first.
HOL_2	Represents the week containing Memorial day.
HOL_3	Represents the week containing the Fourth of July.
HOL_4	Represents the week containing Labor day.
HOL_5	Represents the week containing Thanksgiving day.
HOL_6	Represents the week containing Christmas day.
SEA_1	Represents the spring season.
SEA_2	Represents the summer season.
SEA_3	Represents the fall season.
SCH	Represents the week containing the first day of school in Knox County.

Table 6. Descriptive statistics of the dependent and independent variables contained in group g's theoretical forecasting model (historical period)

Var.	Mean	Std Dev	Minimum	Maximum
Q_g	1836.87	600.10	247.41	3537.45
P_g	0.11	0.01	0.06	0.16
NEW_g	2.55	6.10	0.00	56.88
TV_g	1.40	20.55	0.00	300.00
RD_g	7.04	37.89	0.00	300.00
HOL_1	0.01	0.13	0.00	1.00
HOL_2	0.01	0.13	0.00	1.00
HOL_3	0.01	0.13	0.00	1.00
HOL_4	0.01	0.13	0.00	1.00
HOL_5	0.01	0.13	0.00	1.00
HOL_6	0.01	0.13	0.00	1.00
SEA_1	0.26	0.44	0.00	1.00
SEA_2	0.24	0.43	0.00	1.00
SEA_3	0.24	0.43	0.00	1.00
SCH	0.02	0.13	0.00	1.00

Table 7. Descriptive statistics of the dependent and independent variables contained in group g's theoretical forecasting model (trial period)

Var.	Mean	Std Dev	Minimum	Maximum
Q_g	1897.45	473.74	670.51	2527.49
P_g	0.11	0.00	0.09	0.11
NEW_g	1.51	2.70	0.00	9.00
TV_g	0.00	0.00	0.00	0.00
RD_g	0.00	0.00	0.00	0.00
HOL_2	0.00	0.00	0.00	0.00
HOL_3	0.03	0.19	0.00	1.00
HOL_4	0.03	0.19	0.00	1.00
HOL_5	0.03	0.19	0.00	1.00
HOL_6	0.03	0.19	0.00	1.00
HOL_1	0.00	0.00	0.00	0.00
SEA_1	0.00	0.00	0.00	0.00
SEA_2	0.50	0.50	0.00	1.00
SEA_3	0.50	0.50	0.00	1.00
SCH	0.03	0.19	0.00	1.00

Table 8. Steak's theoretical model variable definitions

Variable	Definition
P_S	Steak's weekly weighted average price.
$P_{(R,GB)}$	Roast's and ground beef's weekly weighted average price.
TV_S	Steak's weekly television advertising, measured in gross rating points.
$TV_{(R,GB)}$	Roast's and ground beef's weekly television advertising, measured in gross rating points.
RD_S	Steak's weekly radio advertising, measured in gross rating points.
$RD_{(R,GB)}$	Roast's and ground beef's weekly radio advertising, measured in gross rating points.
NWS_S	Steak's Weekly Newspaper advertising, measured in square inches.
$NWS_{(R,GB)}$	Roast's and ground beef's weekly Newspaper advertising, measured in square inches.
HOL_1	Represents the week containing Memorial day.
HOL_2	Represents the week containing the Fourth of July.
HOL_3	Represents the week containing Labor day.
HOL_4	Represents the week containing Thanksgiving day.
HOL_5	Represents the week containing Christmas day.
HOL_6	Represents the week containing January first.
SEA_1	Represents the spring season.
SEA_2	Represents the summer season.
SEA_2	Represents the fall season.
LAG	Lagged weekly Item movement.

Table 9. Descriptive statistics of the dependent and independent variables contained in steak's theoretical forecasting model (historical period)

Var.	Mean	Std Dev	Minimum	Maximum
Q_S	2423.73	858.00	287.20	6308.12
P_S	4.14	0.68	2.05	5.45
P_R	2.64	0.57	0.94	6.32
P_{GB}	2.57	0.47	0.99	2.84
NWS_S	21.99	33.86	0.00	210.00
NWS_R	15.78	22.76	0.00	98.25
NWS_{GB}	14.33	18.32	0.00	165.50
TV_S	54.59	139.69	0.00	684.00
TV_R	51.41	136.84	0.00	700.00
TV_{GB}	100.70	157.85	0.00	600.00
RD_S	7.28	36.10	0.00	300.00
RD_R	14.08	49.90	0.00	300.00
SEA_1	0.27	0.44	0.00	1.00
SEA_2	0.24	0.43	0.00	1.00
SEA_3	0.24	0.43	0.00	1.00
HOL_1	0.02	0.14	0.00	1.00
HOL_2	0.02	0.14	0.00	1.00
HOL_3	0.02	0.14	0.00	1.00
HOL_4	0.02	0.14	0.00	1.00
HOL_5	0.02	0.14	0.00	1.00
HOL_6	0.02	0.14	0.00	1.00
LAG	2423.81	857.99	287.20	6308.12

Table 10. Descriptive statistics of the dependent and independent variables contained in steak's theoretical forecasting model (trial period)

Var.	Mean	Std Dev	Minimum	Maximum
Q_S	2214.25	1053.45	554.09	4360.00
P_S	3.52	1.03	2.15	7.81
P_R	3.14	0.56	1.97	3.86
P_{GB}	2.44	1.18	1.86	7.60
NWS_S	54.64	73.18	0.00	299.44
NWS_R	25.48	33.41	0.00	94.74
NWS_{GB}	15.56	32.78	0.00	143.00
TV_S	109.54	205.58	0.00	512.00
TV_R	92.00	222.70	0.00	684.00
TV_{GB}	19.69	100.41	0.00	512.00
RD_S	25.00	62.05	0.00	200.00
RD_R	17.38	48.87	0.00	150.00
SEA_1	0.00	0.00	0.00	0.00
SEA_2	0.50	0.51	0.00	1.00
SEA_3	0.50	0.51	0.00	1.00
HOL_1	0.00	0.00	0.00	0.00
HOL_2	0.00	0.00	0.00	0.00
HOL_3	0.04	0.20	0.00	1.00
HOL_4	0.04	0.20	0.00	1.00
HOL_5	0.04	0.20	0.00	1.00
HOL_6	0.04	0.20	0.00	1.00
LAG	2209.51	1052.91	554.09	4360.00

deviation of $0.68. Steak's mean weekly item movement 2423.73 ounces and the corresponding standard deviation is 858.00.

In each of the three cases, the mean item movement, advertising, holiday, and seasonal values were significantly different between the historic and trial subperiods. The differences in mean values could be attributed to the number of time periods in each subgroup. For example, the seasonal and holiday variable means will be different because the trial forecast period covered the second half of 1992. This means that the spring, summer, and corresponding holidays were not included in the data series. The same logic could be applied to explain the differences in advertising means. The trial period only covered a six month period while the historic period consisted of 213 weekly observations. The larger time period would allow for a greater variation in advertising, seasonal, and holiday variable values. The difference in item movement means could be explained via seasonal purchasing patterns. For example, roast demand falls in the late fall and winter months. This is reflected in the steak data as the mean value is higher for the historic verses trial periods (Tables 9 and 10).

C. Data Description

The item movement for the brand b is shown in Figure 1. Inspection of the figure indicates the wide degree of variation present in the data set. A slight downward trend in the data is visible. Seasonal peaks were observed in the series for weeks corresponding to late summer or early fall and the winter holiday periods. After the late summer early fall peak periods, the item movement decreased until the winter holiday period in

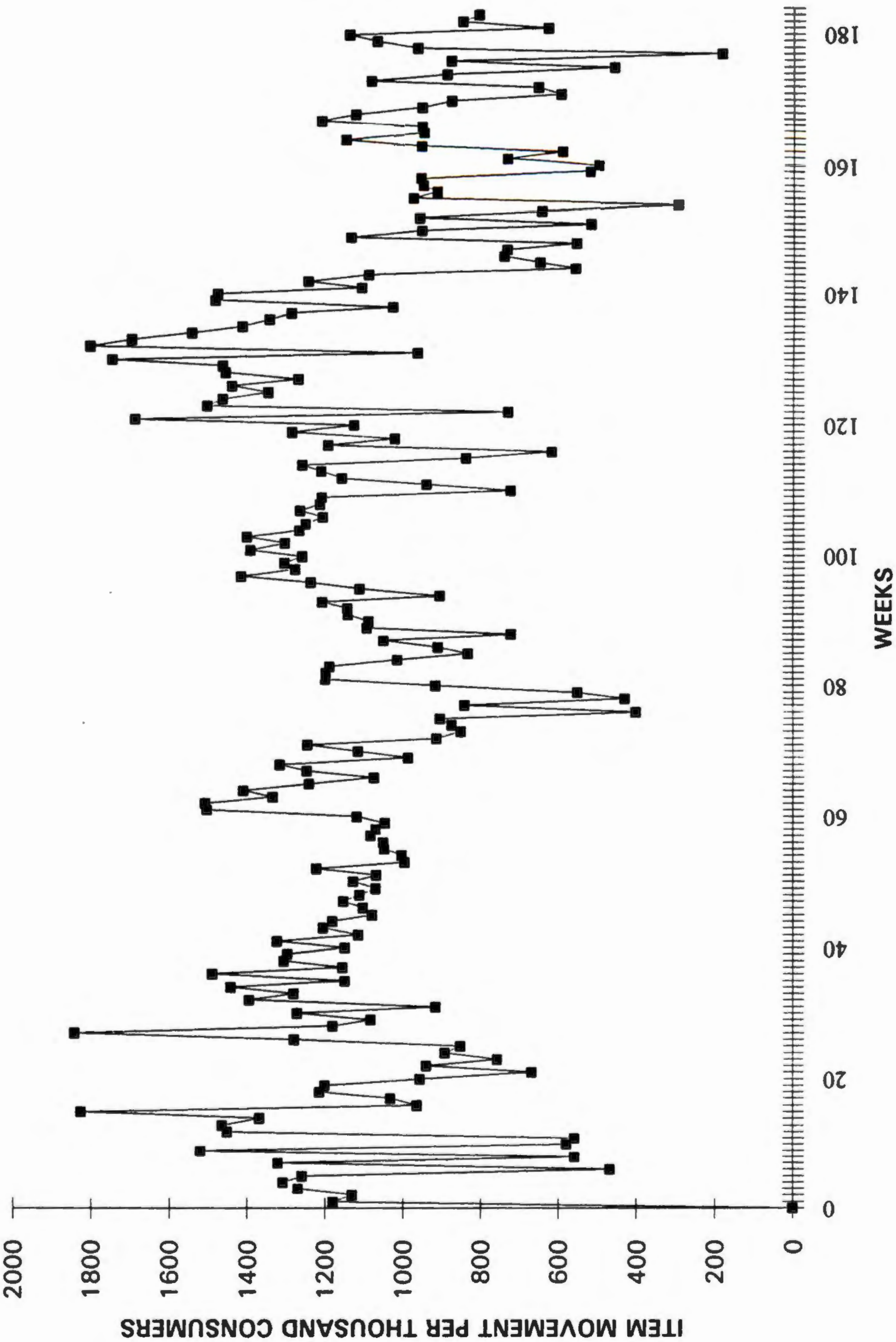

Figure 1. Brand b's weekly item movement

which item movement increased again. The remainder of the calendar year brand b exhibited a relatively stable item movement. Figures 2 and 3 display newspaper and GRP advertising. The former appears to decline during 1992. Electronic media advertising is too infrequent to identify any patterns.

Item movement for group g, shown in Figure 4, revealed more variation in product item movement than present in brand b's item movement. Decreases in demand occurred during the spring and early summer time periods. An increase in item movement is observed in the late summer/ early fall and late fall early winter time periods. For the remainder of the calendar year, no seasonality is noted. Figures 5 and 6 present group g's advertising.

The item movement for steak decreases during the fall, see Figure 7. Item movement increased gradually after the yearly low which occurred in the fall. There were peaks during the remainder of the calendar year which could not be attributed to specific holiday periods or seasons.

Inspection of the steak newspaper, radio, and television advertising revealed no apparent pattern. Newspaper advertising for steak occurred regularly throughout the study period (Figure 8). Radio promotion occurred on thirteen occasions during the two hundred and forty week time period (Figure 9). Television advertising occurred somewhat more frequently, but again inspection of the data over time revealed no apparent pattern (Figure 10). The various holiday periods were found to have no visual impact.

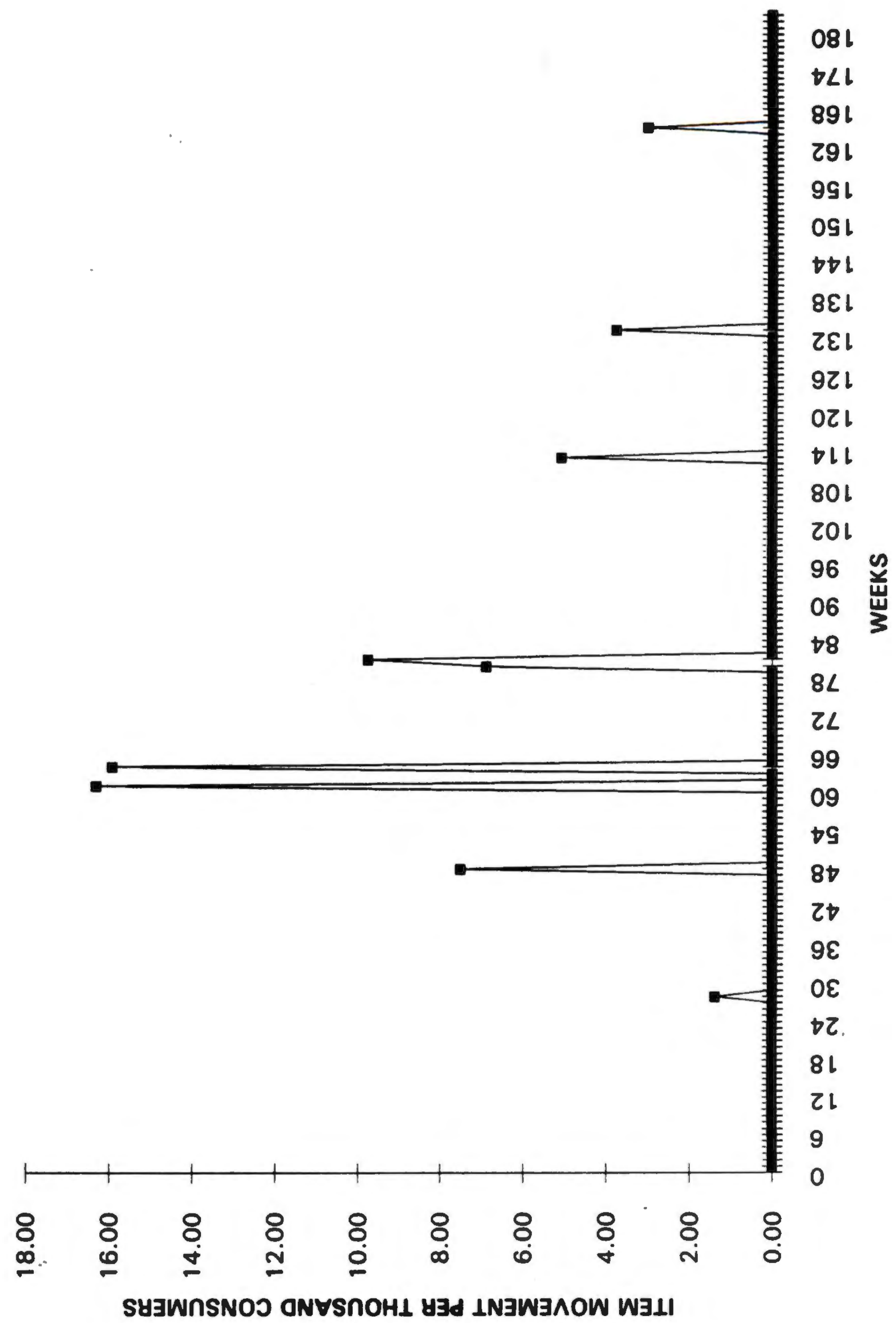

Figure 2. Brand b's weekly newspaper advertising

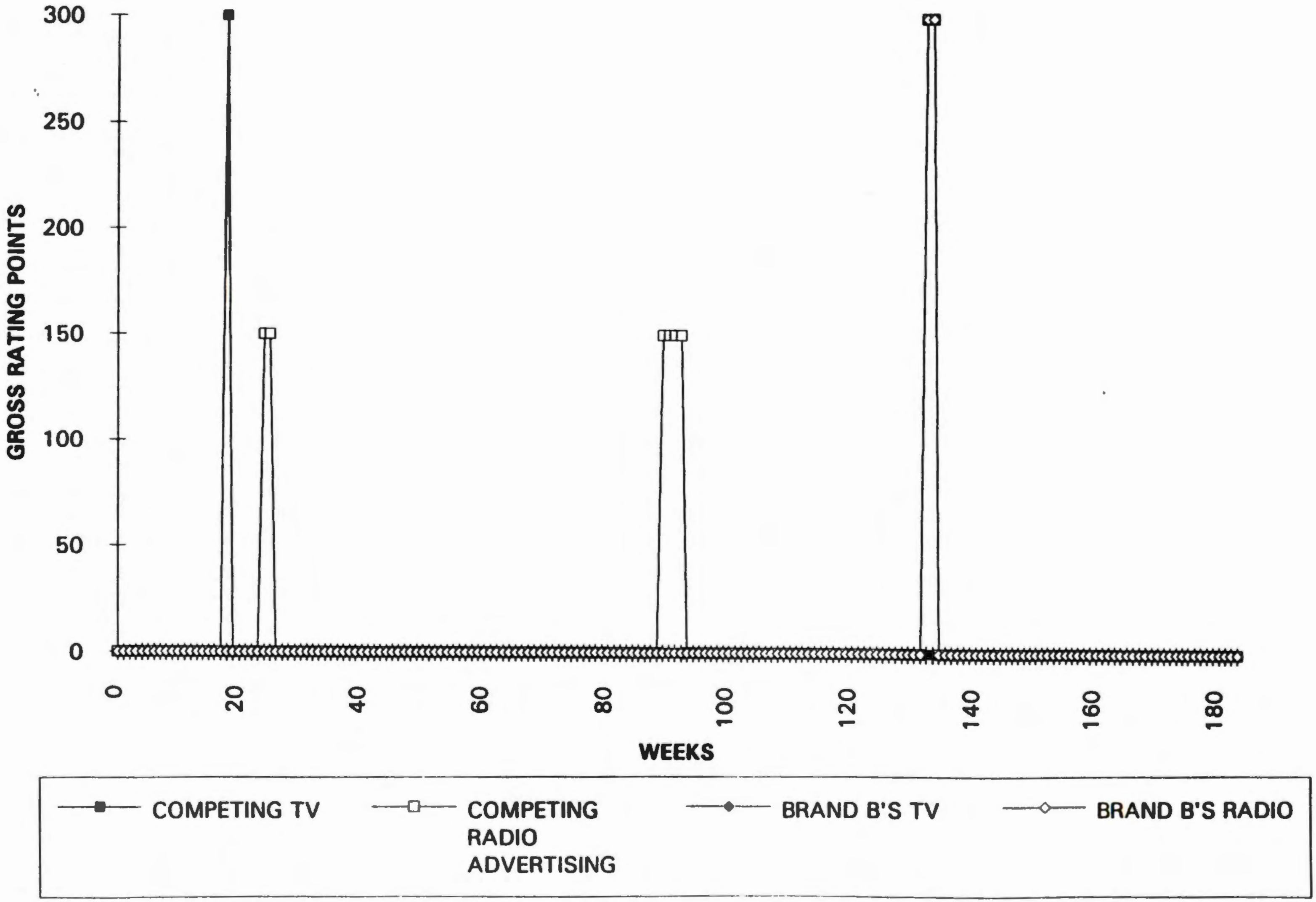

Figure 3. Brand b's weekly electronic media advertising

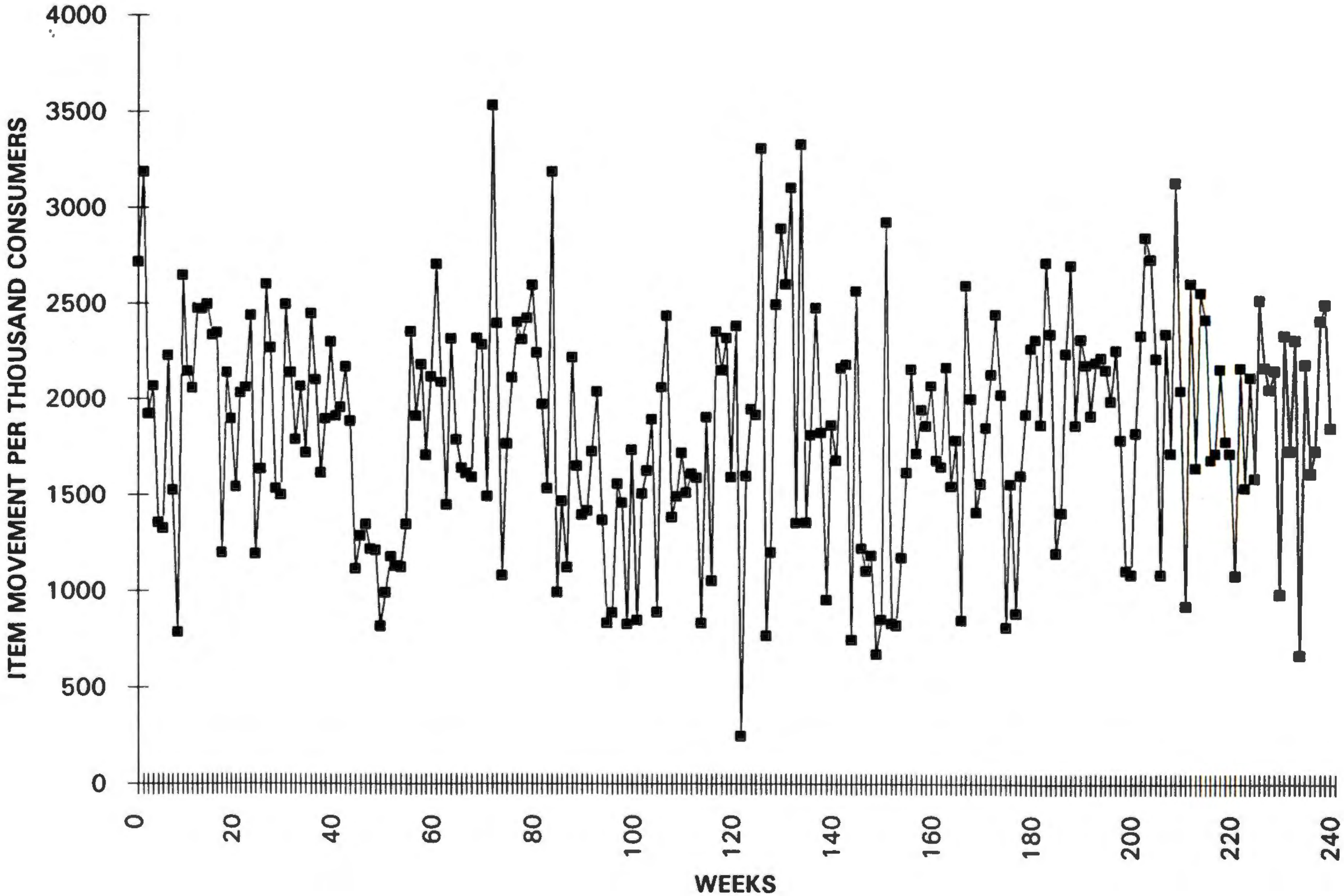

Figure 4. Group g's weekly item movement

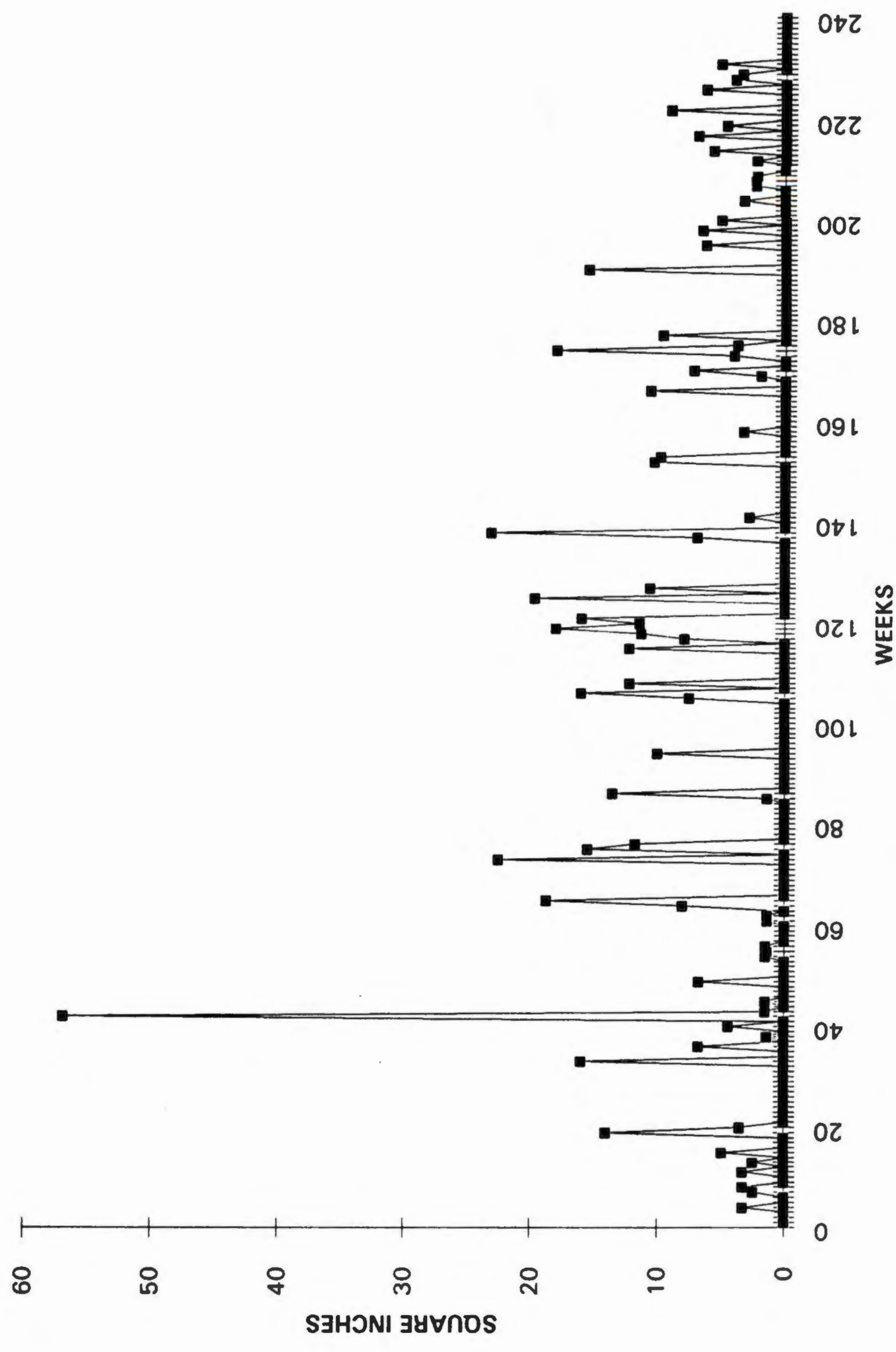

Figure 5. Group g's weekly newspaper advertising

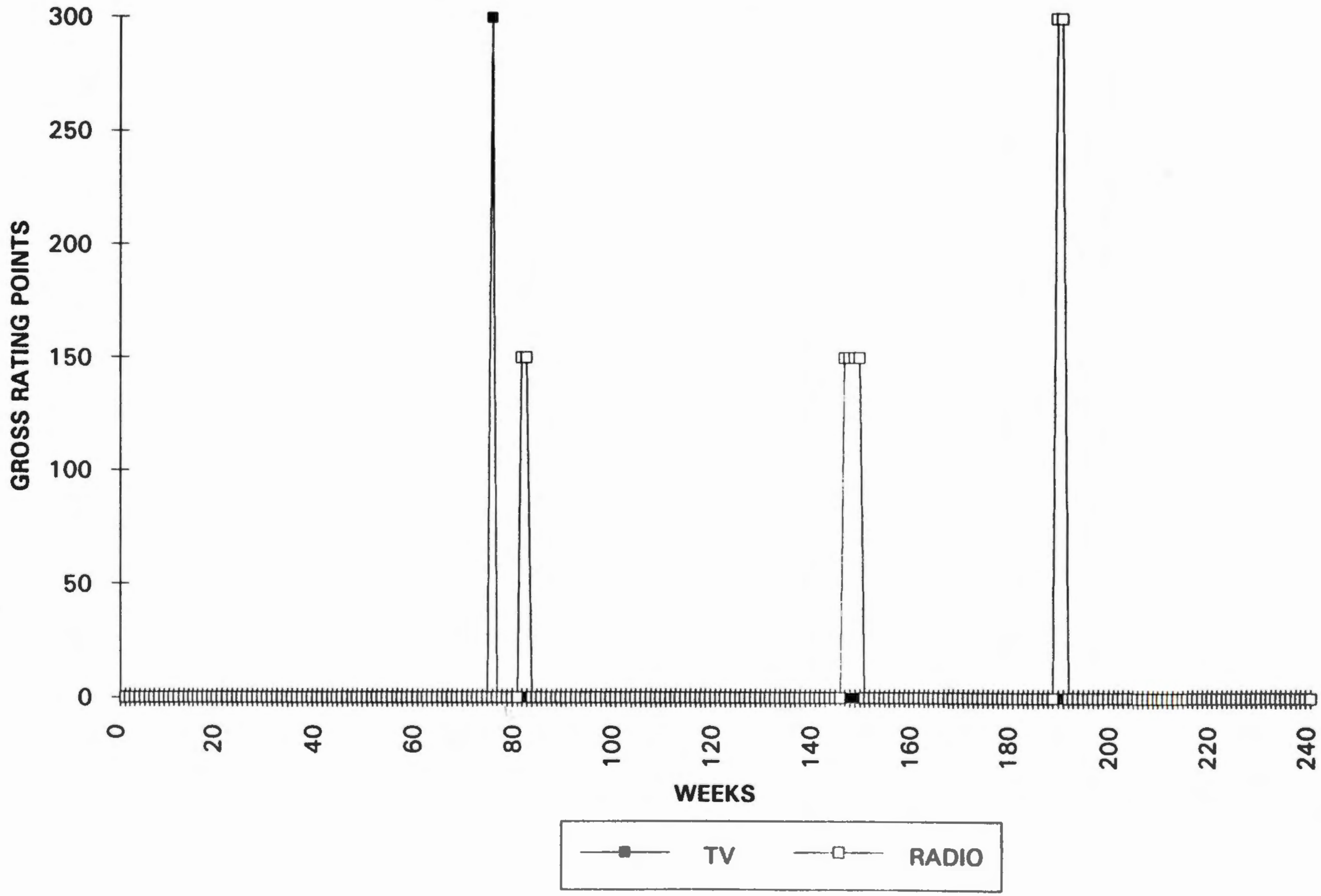

Figure 6. Group g's weekly electronic media advertising

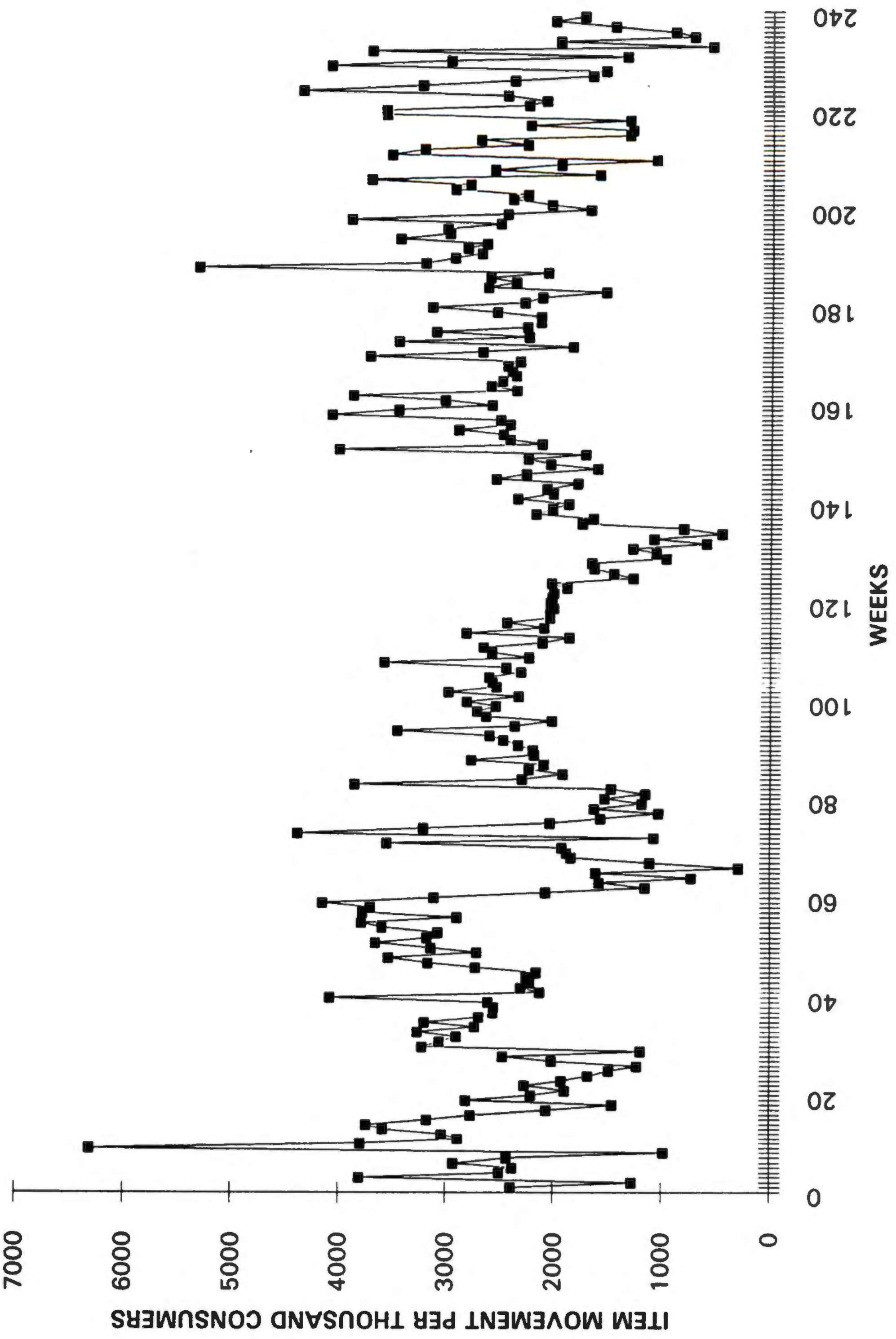

Figure 7. Steak's weekly item movement

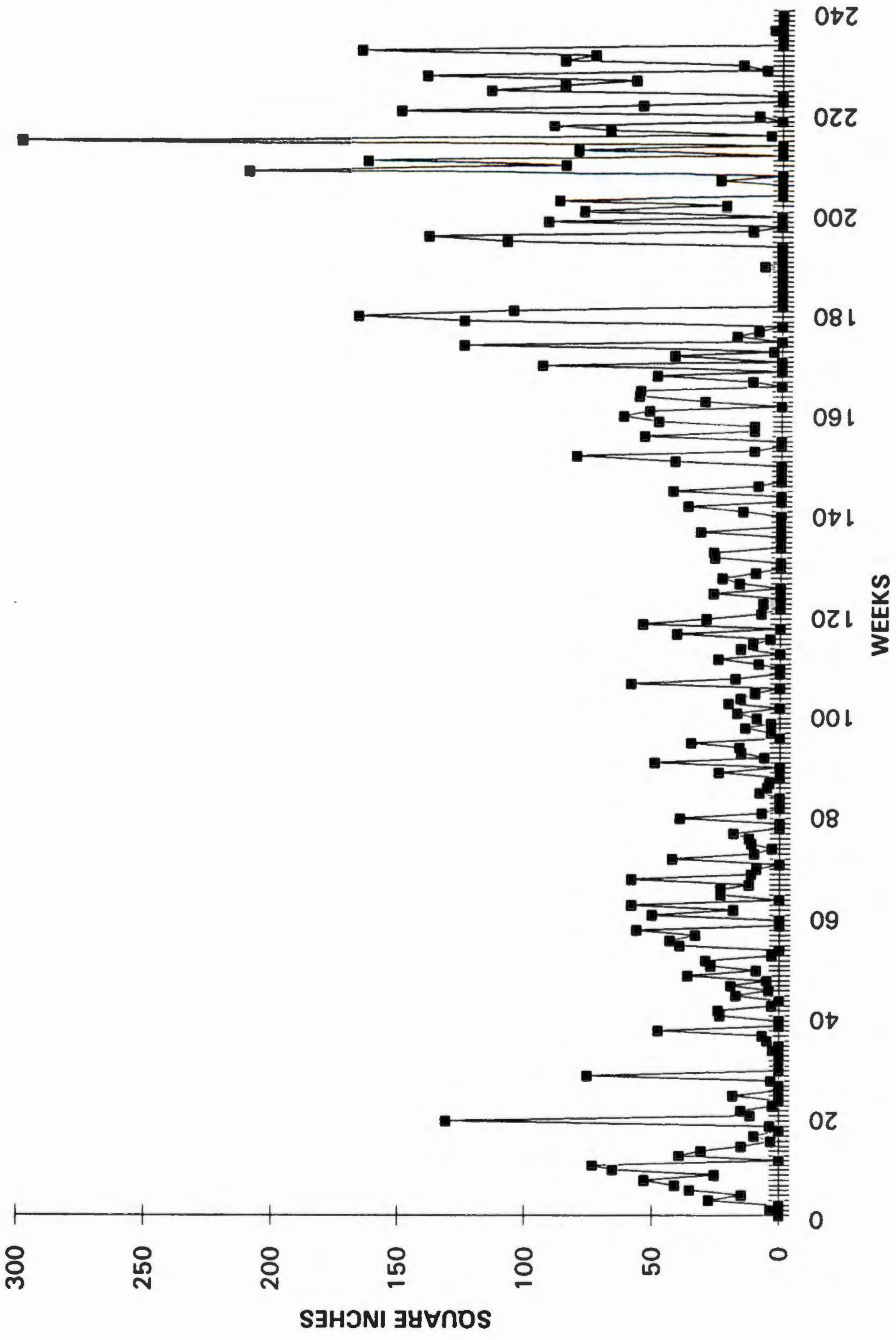

Figure 8. Steak's weekly newspaper advertising

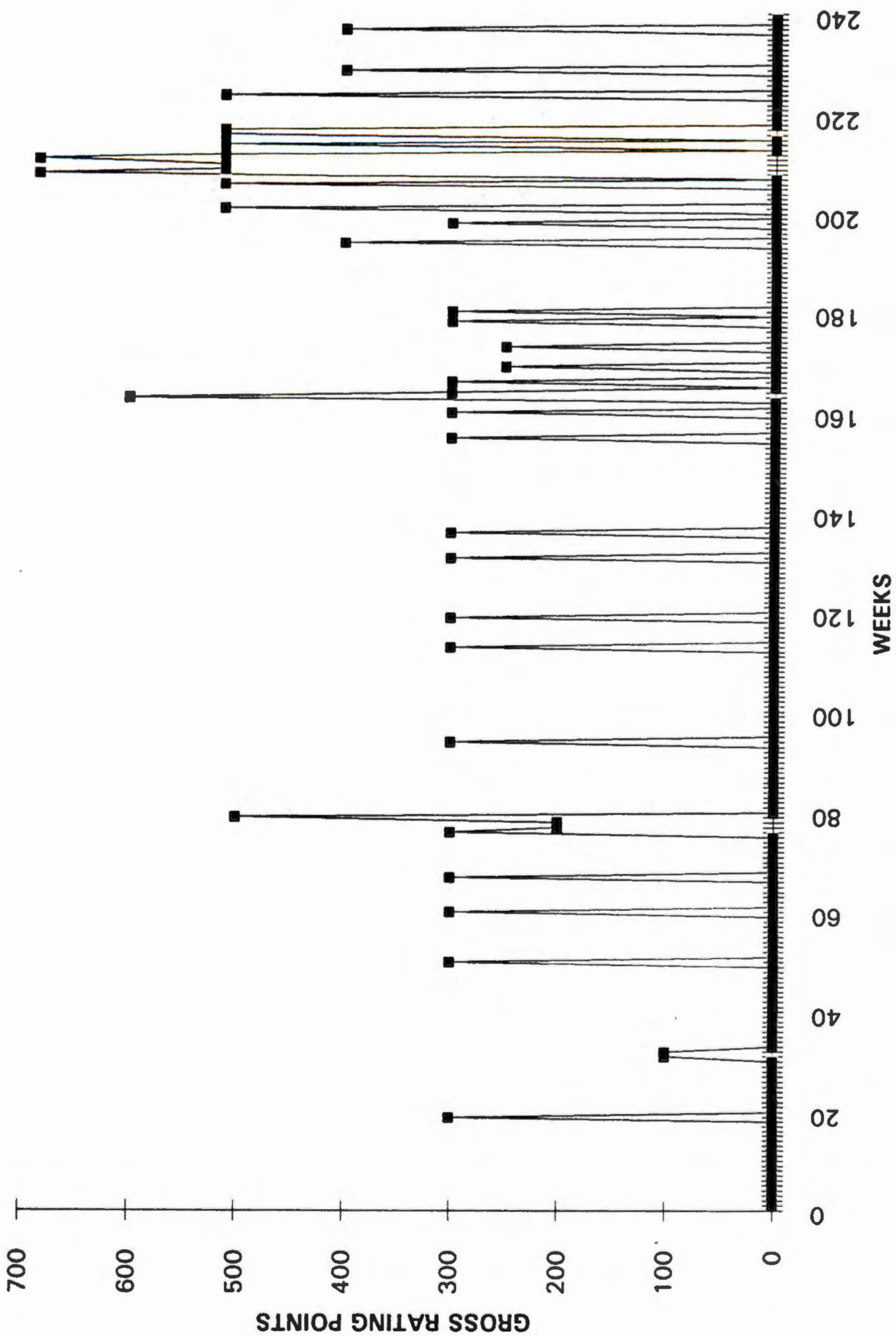

Figure 9. Steak's weekly radio advertising

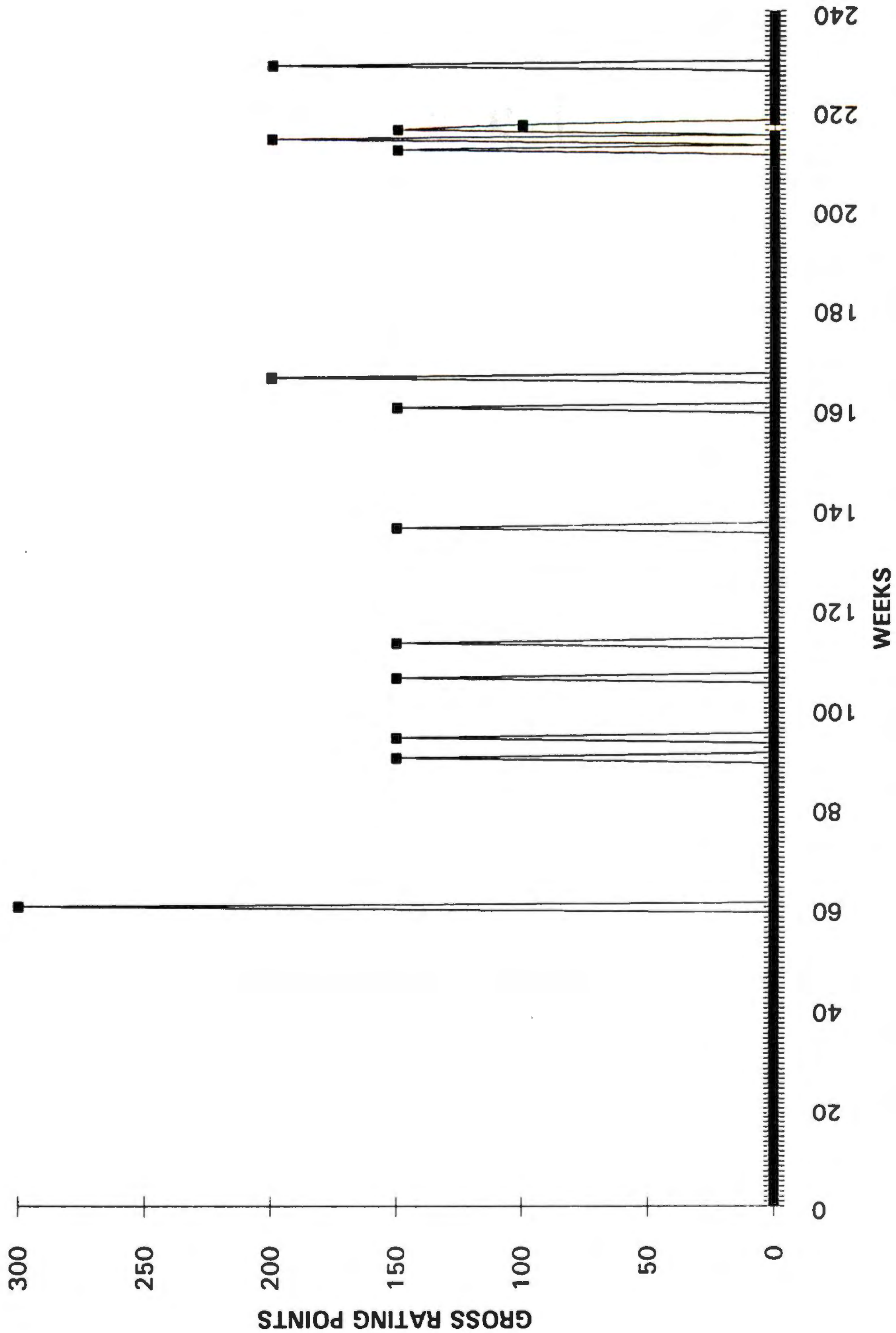

Figure 10. Steak's weekly television advertising

D. Estimates of the Theoretical, Box-Jenkins, and Transfer Function Models

1. Theoretical Models

The goodness of fit measures for each of the three estimated theoretical models are presented in Table 11. The R^2 coefficients ranged from 0.08 for group g's estimated theoretical model to 0.54 for brand b's estimated model. R^2 coefficients cannot be compared between models with different numbers of explanatory variable. The adjusted R^2 coefficient is better for comparison purposes because it takes into account the number of explanatory variables in a model.

Table 11. Theoretical models goodness of fit

Statistic	Brand b (161 weeks) Value	Group g (231 weeks) Value	Steak (231 weeks) Value
R-square	0.55	0.08	0.34
Adj R^2	0.43	0.02	0.27
F Value	4.64	1.28	4.96
Prob>F	0.0001	0.22	0.0001
RMSE	222.67	585.05	736.48
MSE	49582.83	342277.97	542397.91
DW	2.18	1.82	***

*** The Durbin-Watson D statistic and h statistic were not applicable. A modification of the D-W h statistic was used to evaluate the presence of serial correlation.

In specifying group g's theoretical model, dropping variables or including other variables resulted in inferior evaluation criteria values when compared to the evaluation criteria for the current model specification. This held true for various brand b and steak model specifications. In general, each of the evaluation criterion used individually led to the same model specifications presented below. The forecasting models with the highest R^2 generally had the highest adj. R^2, F value, and the lowest RMSE.

The purpose of specifying these theoretical models is to generate forecasts. The presence of multicollinearity is not considered a problem in forecasting if the model is able to predict accurately the dependent variable. However, the ability to interpret individual coefficients necessitates that some diagnostic checks be completed. A simple correlation matrix was evaluated for the independent variables used in each of the models as a partial check for the presence of multicollinearity. The correlation matrices for group g and steak indicated that there was no real pairwise collinearity, and thus, multicollinearity was not a problem. There was a slight degree of multicollinearity present between the price variables in brand b's model, but the degree of collinearity between these variables was not large enough to warrant the omission of any particular variable or variables.

Dropping one of two collinear variables was also performed as a second check for multicollinearity for all three of the theoretical demand functions. No significant changes were observed in the standard errors or estimated coefficients of the remaining independent variables further indicating that multicollinearity was not a problem.

The Durbin-Watson D statistic (D-W) was used to check for the presence of serial correlation in each of the three residual series. The D-W statistic for brand b's residual series was 2.18 which fell in the indeterminate range. A D-W that falls in the indeterminate or inconclusive range does not provide enough evidence to reject the null hypothesis of no serial correlation in the residual series. The D-W could not be used to check for the presence of serial correlation for the steak residuals because of the inclusion of a lagged dependent variable as an independent variable. A variation of Durbin's h statistic (Pindyck and Rubinfeld) was used to check for autocorrelation in the residual series. The test simply describes the residual series as being a function of the original model's explanatory variables and a lagged residual variable. A t-test was performed on the lagged residual term to test the null hypothesis that the lagged error term coefficient was not significantly different from zero. The null hypothesis of no serial correlation was not rejected. Even though group g's estimated model was not significant, the D-W statistic revealed that the residual series did not exhibit autocorrelation.

The first of the three theoretical forecasting models to be described is the model for brand b. The historic data consisted of 161 weekly observations starting with the week ending on June 4, 1989 and ending with the ending on December 27, 1992. Standard errors are below each estimated parameter in parenthesis.

$$
\begin{aligned}
(55)\quad Q_b = {} & \underset{(741.59)}{3526.26} - \underset{(4695.29)}{27942.00}(P_b) - \underset{(943.23)}{402.62}(P_3) + \underset{(101.85)}{(236.73}(P_5) \\
& + \underset{(51.22)}{116.70}(P\text{-}O\text{-}P_7) + \underset{(99.80)}{300.76}(P\text{-}O\text{-}P_1) - \underset{(74.70)}{409.79}(SEA_1) \\
& - \underset{(69.20)}{402.77}(SEA_2) + \underset{(132.10)}{283.39}(SCH) + e_t.
\end{aligned}
$$

The coefficient of determination or R^2 statistic for brand b's theoretical model was 0.55. An $R^2 = 0.55$ indicates that this model explained fifty-five percent of variation in brand b's weekly item movement. The corresponding F-statistic of 4.64 is statistically significant at the 95 percent level. The RMSE value of 222.67 does not appear to be large compared to the average weekly item movement (1123.79).

All of the variables in above model are significant at the ninety-five percent level. The negative sign for the own-price coefficient, P_b, is consistent with economic theory. Three cross-prices have the expected positive and significant coefficients leading to the inference that they are substitutes. The negative sign on the substitute price P_1 and the positive signs on the cross P-O-P advertising variables are not consistent with economic theory and thus, unanticipated. The prices for each of the individual brands, in general, tended to move in a similar direction. This might be attributed to similar pricing strategies by the various competitors for brand b. One reasonable explanation for the positive sign associated with the cross P-O-P advertising variables is that promoting peanut butter brands via P-O-P attracts the consumers attention. Once the consumer is reminded of the need for peanut butter, he/she inspects the promoted brand and compares the value of his/hers preferred brand with the promoted brand. The consumer then decides that the difference in price is not large enough to cause brand switching, and he/she purchases the preferred brand of peanut butter.

SEA_1 and SEA_2 estimated seasonal coefficients are negative and significant. This suggests that during these time periods band b's weekly item movement is significantly lower than during the winter. The SCH (school) variable represents the start of the Knox county school year. These estimated coefficients suggest that the first week of the school year results in an increase in the item movement for brand b peanut butter. This may be attributed to purchases of peanut butter by parents of school age children to make sandwiches for their children's school lunches.

One possible explanation for the insignificance of television, radio, and newspaper advertising may be simply that they are not effective. This may be a logical conclusion given how infrequently television, radio, and newspaper advertising were employed (Figures 2 and 3). There was only one week during the entire study period in which brand b was promoted via television. There was no television cross-advertising for competing brands. There were two radio advertisements for brand b and six radio advertisements for competing brands during the study period. Newspaper advertisements for brand b occurred ten times during the study period. Another possible explanation is that the advertising was for a specific jar of peanut butter not for ounces of peanut butter and thus was insignificant in explaining any variation in brand b's and group g's weekly item movement. None of the holiday variables was significant.

Figure 11 presents a graphical description of brand b's theoretical backcast. Visual inspection of the actual and backcast series implies that the theoretical model is capable of reproducing the historical data series. The RMSE and U^2 for brand b's backcast are 222.67 and 0.06, respectively. Decomposition of U^2 indicated a lack of bias in the error series and that the error series was the result of a random fluctuations. The theoretical model predicted 75 out of 161 directional changes.

Table 12. Brand b's theoretical backcast evaluation criteria

Criterion	Value
MSE	49582.83
RMSE	222.67
U^2	0.06
Directional change	75/161

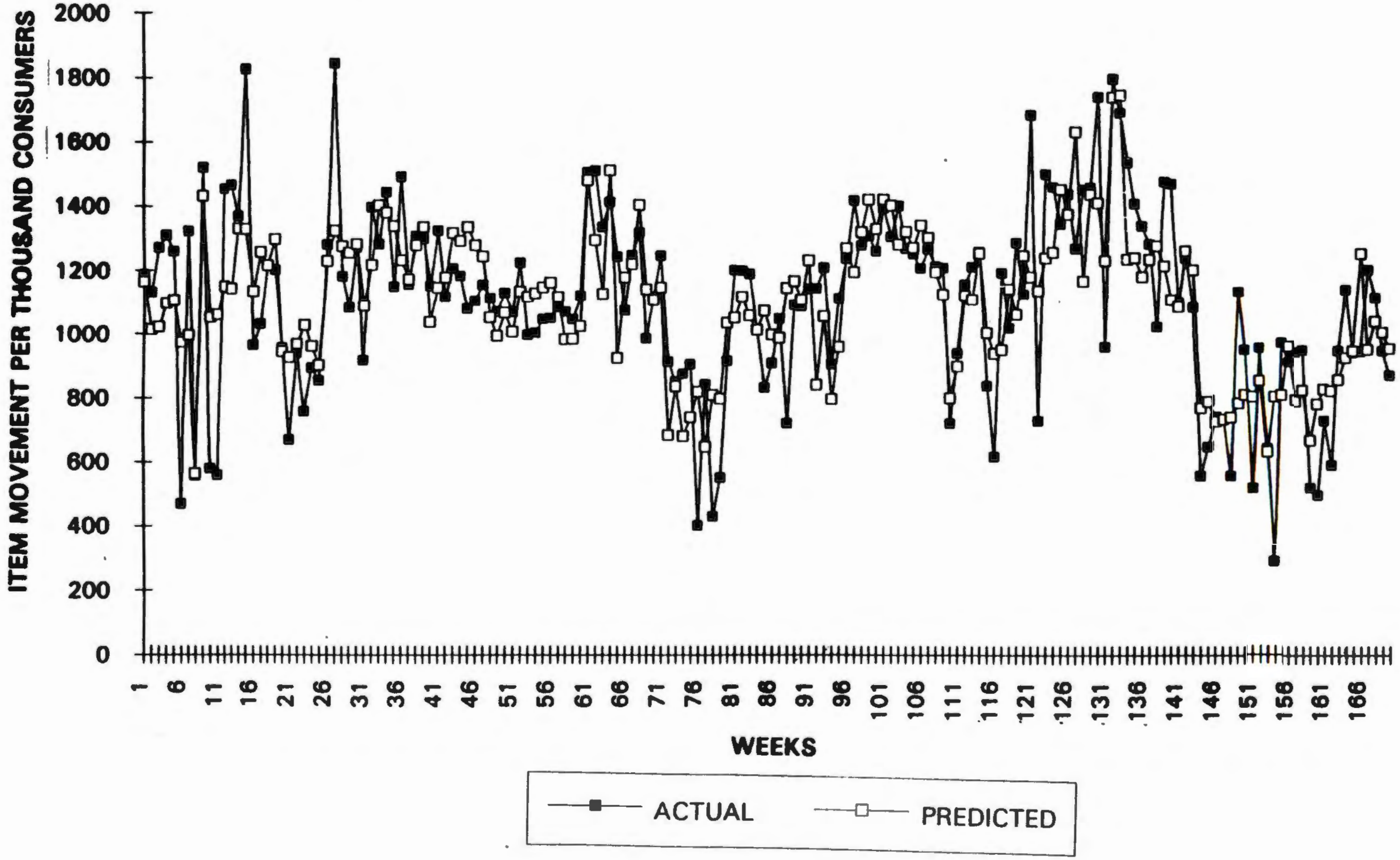

Figure 11. Brand b's theoretical backcast

The second model is the estimated theoretical model for group g. The historic data had 213 weekly observations starting with the week ending June 4, 1988 and ending June 27, 1992. The estimated model, with the standard errors below each estimated parameter in parenthesis, is as follows:

$$(56)\quad \underset{}{Q_g} = \underset{(776.88)}{1928.58} - \underset{(4698.83)}{707.23}(Hol_1) + e_t.$$

The R^2 statistic for group g's theoretical model was 0.08. $R^2 = 0.08$ indicates that this model explained eight percent of the variation in group g's weekly item movement. The RMSE value of 585.05 does not appear to be large compared to the average weekly item movement (1836.87). The corresponding F-statistic of 1.82 is not statistically significant at the 95 percent level, which leads to the inference that the set of estimated coefficients is not statistically different from zero.

The insignificance of the price and advertising variables in group g's theoretical demand model suggests that the linear theoretical model may be an inappropriate model specification for analyzing peanut butter group data or that peanut butter cannot be analyzed as a group. Inspection of other forecasting techniques, Box-Jenkins and transfer function model, suggested that the linear theoretical model was not the best modeling specification for analyzing peanut butter as a group.

Figure 12 presents a graphical description of group g's theoretical backcast. Visual inspection of the actual and backcast series revealed that the theoretical model is incapable of reproducing the historical data series. The backcast resembled a straight line over time with slight positive and negative variations The RMSE and U^2 for group g's backcast

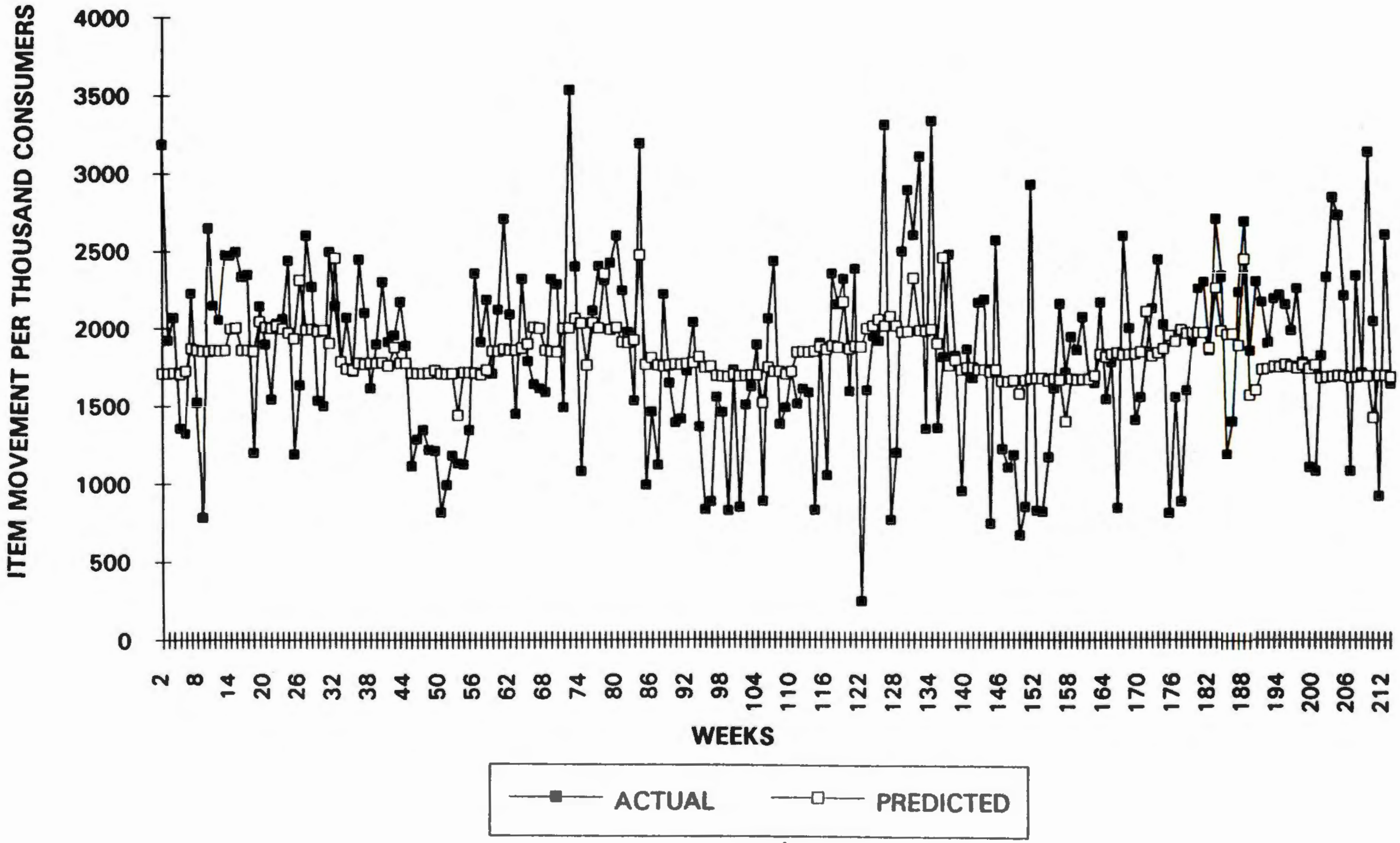

Figure 12. Group g's theoretical backcast

are 585.05 and 0.05, respectively. Decomposition of U^2 indicated a lack of bias in the error series and that the error series was the result of a random fluctuations. The theoretical model predicted 129 out of 213 directional changes, Table 13.

Table 13. Group g's theoretical backcast evaluation criteria

Criterion	Value
MSE	342277.97
RMSE	585.05
U^2	0.05
Directional change	129/213

The third model is the estimated steak item movement model. The historic data had 213 weekly observations starting with the week ending on June 4, 1988 and ending with the last week of June 27, 1992. The estimated model, with the standard error below each estimated parameter in parenthesis, is as follows:

$$(57)\quad \underset{}{Q_S} = \underset{(630.82)}{4862.81} - \underset{(83.27)}{384.79}(P_S) - \underset{(113.32)}{286.62}(P_{GBF}) - \underset{(116.19)}{541.60}(\text{Fall}) + \underset{(0.07)}{0.18}(\text{Lag}) + e_t.$$

An $R^2 = 0.34$ indicates that this model explained thirty-four percent of variation in the weekly item movement of steak. The corresponding F-statistic of 4.96 is statistically significant at the 95 percent level. The RMSE value of 736.48 does not appear to be large compared to the average weekly item movement (2423.73).

Each of the variables in the above theoretical model is significant at the ninety-five percent level. The sign on the own-price is negative as expected. The sign on the substitute price of ground beef is negative. A possible explanation could be that consumers enter a supermarket with a the intent on purchasing a specified amount of ground beef. If the price of ground beef is lower than expected, consumers may purchase their desired amounts and the remaining money is used to purchase steak.

Figure 13 presents a graphical description of steak's theoretical backcast. Visual inspection of the actual and backcast series revealed that the theoretical model was capable of reproducing the historical data series. The backcast, in general, resembled the historical data series with the exception of having a lower intercept and base line. The evaluation criteria results are in Table 14. The RMSE and U^2 for steak's backcast are 736.48 and 0.04, respectively. Decomposition of U^2 indicated a lack of bias in the error series and that the error series was consistent with a hypothesis of random fluctuations. The theoretical model predicted 102 out of 213 directional changes.

Table 14. Steak's theoretical backcast evaluation criteria

Criterion	Value
MSE	54397.91
RMSE	736.48
U^2	0.04
Directional change	102/213

2. Box-Jenkins Models

The Box-Jenkins technique was used to identify the model, obtain the appropriate estimates, and make a forecast for the last 26 weeks of 1992. Brand b's models were estimated using 161 weekly observations. Group g's and steak's models were estimated using 213 weekly observations. The best

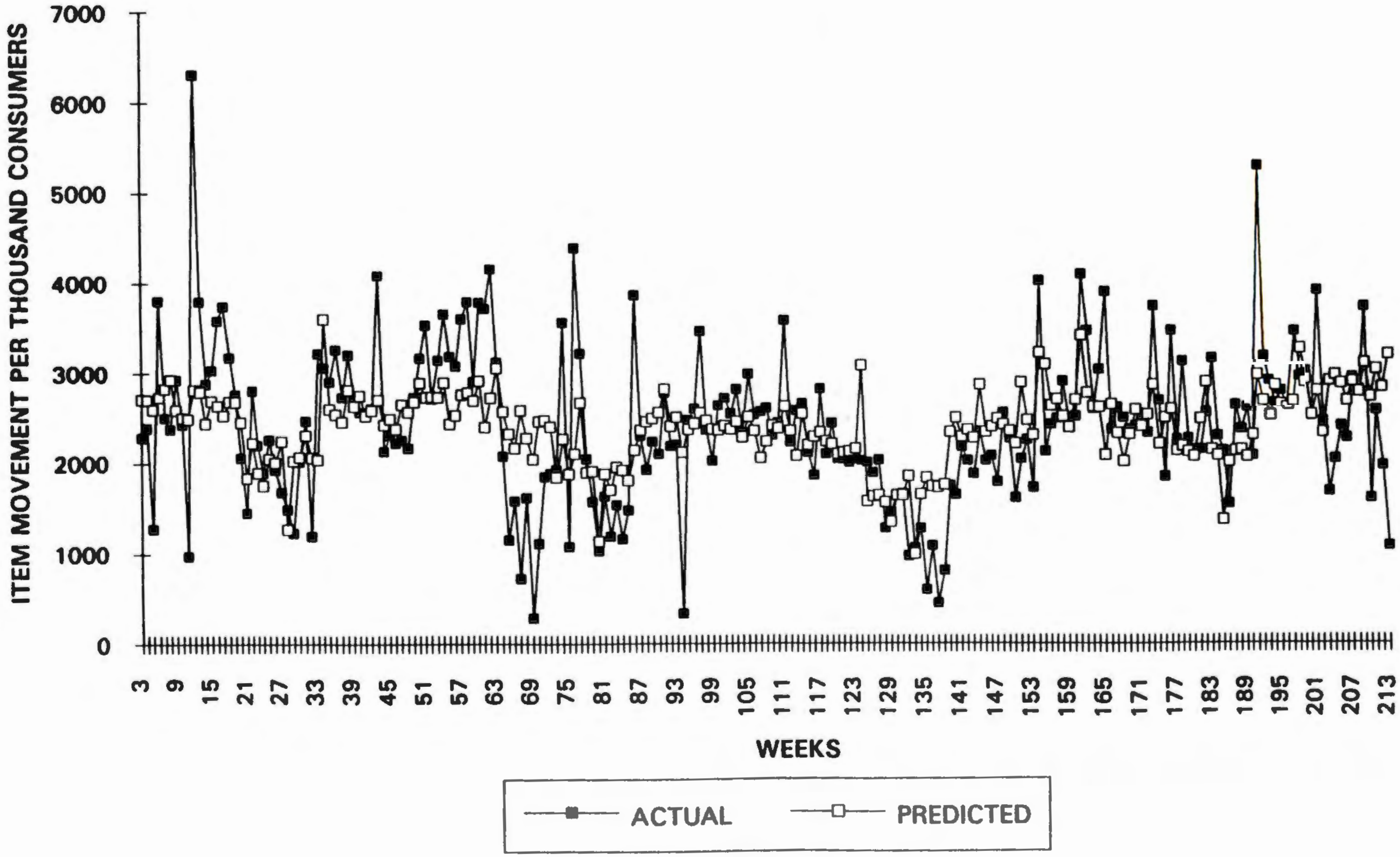

Figure 13. steak's theoretical backcast

model structures for forecasting weekly item movement were based on the AIC, computed chi squares of lagged autocorrelations, autocorrelation and partial autocorrelation plots, and significance of the autocorrelation coefficients. The ACF and PACF values were used to identify possible model specifications for forecasting brand b's, group g's, and steak's weekly item movement. Inspection of the autocorrelation and partial autocorrelation functions was also performed to reveal whether the data series was stationary. The ACF for each of the three products truncated quickly which is characteristic of a stationary data series. As a further test, the data series are differenced, and the ACF and PACF were again inspected to determine if the data series was stationary. The autocorrelation functions of the differenced data series revealed no more stationarity than the autocorrelation function for the undifferenced data series. Thus, there was not a need to difference the data.

The model used for forecasting brand b's weekly item movement was a second order autoregressive model with seasonal effects at 6 and 22 weeks. The U^2 coefficient for the historical series was 0.06 which indicated that the forecast was much better than a nochange forecast. Decomposition of U^2 indicated a lack of bias in the error series and that the error series was the result of a random fluctuations. The model is expressed as follows:

$$(58)\quad Q_{bt} = \underset{}{1110.18} + \underset{(0.065)}{0.29434(Q_{bt-2})} + \underset{(0.064)}{0.27588(Q_{bt-6})} \underset{(0.067)}{-0.14512(Q_{bt-22})}$$

The RMSE for the ARIMA $(2,0,0)_{6,22}$ model was 320.94. The second order autoregressive parameter estimate was 0.29434 with a t ratio of 4.50. The six period seasonal autoregressive parameter estimate was 0.27588 and had a t ratio of 4.34. The twenty-second period seasonal autoregressive parameter estimate was -0.14512 and had a t ratio of 2.17. The positive signs associated with second autoregressive and six period seasonal parameter estimates implies that high weekly steak item movement in periods t-2 and t-6 lead to a large item movement in period t. The negative sign associated with the twenty-second seasonal autoregressive parameter estimate implies that high weekly item movement in 22 periods earlier leads to a decrease in the item movement in week t.

The adequacy of the estimated model was checked by inspection of the chi square values of the residual autocorrelation function. The low chi square value at lag 6 (0.012 < 0.05) indicated the presence of autocorrelation between $residual_t$ and $residual_{t-6}$, Table 15. The model was respecified to incorporate a sixth order autoregressive term. The results of the alternative model indicated that overfitting the original model specification produced an inferior forecast. The inferiority of the alternative forecast was determined by comparing the evaluation criteria from the alternative and original forecasting models.

Table 15. Box-Pierce chi-square values for the brand b ARIMA(2,0,0) model

Lag	Chi Square	Prob						
6	8.86	0.012	0.126	0.001	0.089	0.002	0.111	0.005
12	13.41	0.099	0.051	-0.005	0.013	0.114	0.040	-0.027
18	14.79	0.393	0.024	0.008	0.030	0.021	-0.058	-0.000
24	16.27	0.700	-0.001	0.008	0.050	0.008	0.051	0.019
30	17.00	0.909	0.009	-0.010	-0.023	-0.006	0.003	0.044
36	20.36	0.945	0.075	-0.024	-0.029	-0.020	0.064	0.022
42	27.04	0.907	0.010	-0.037	-0.080	0.019	-0.020	-0.120

Chi square values for lags 12, 18, 24, 30, 36, and 42 indicated that there was no significant autocorrelation present between the residuals for the $AR(2,0,0)_{6,22}$ model at the ninety-five percent level of significance.

Figure 14 presents a graphical description of brand b's Box-Jenkins backcast. Visual inspection of the actual and backcast series revealed that the Box-Jenkins ARIMA theoretical model was very capable of reproducing the historical data series. The backcast, in general, resembled the historical data series except that it was not capable of predicting the large fluctuations in item movement. The RMSE and U^2 for brand b's backcast were 320.94 and 0.06, respectively. The model predicted 91 out of 161 directional changes, Table 16.

Table 16. Brand b's ARIMA backcast evaluation criteria

Criterion	Value
MSE	54397.91
RMSE	736.48
U^2	0.04
Directional change	91/161

The model used for forecasting group g's weekly item movement was a first order autoregressive first order moving average model with an autoregressive seasonal effect at t-18, $ARIMA(1,0,1)_{18}$. U^2 for the historical series was 0.05, which indicated that the forecast was superior a nochange forecast. Decomposition of U^2 indicated a lack of bias in the error series and that the error series was the result of a random fluctuations contained in the data series. The model is expressed as follows:

$$(59)\ Q_{gt} = \underset{(0.0004)}{0.99972(1)} - \underset{(0.065)}{0.26638(18)}/\underset{(0.03)}{0.89698(1)}$$

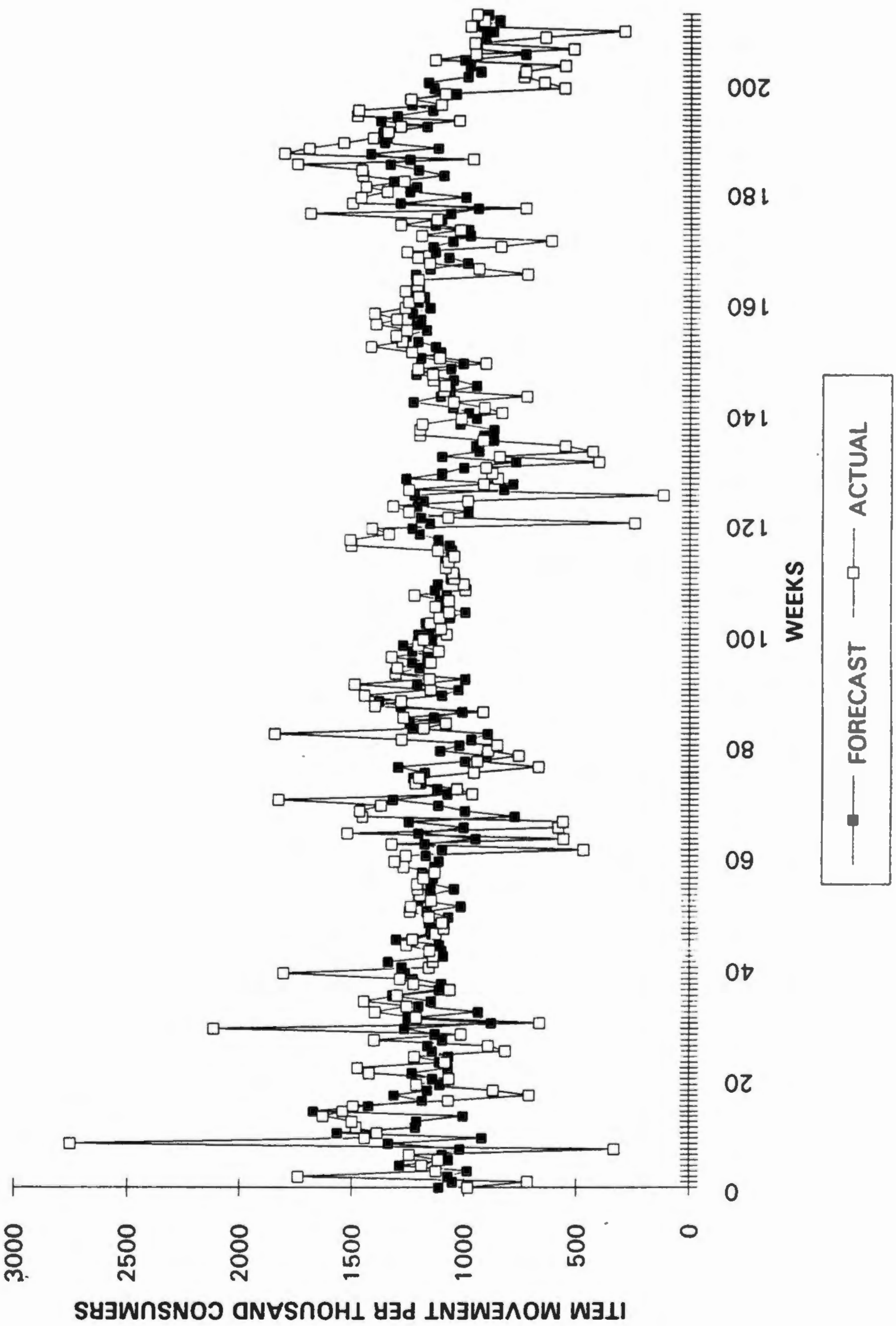

Figure 14. Brand b's statistical backcast

The RMSE for the ARIMA (1,0,1) model was 301.11. The first order autoregressive parameter estimate was 0.99972 with a t ratio of 30.29. The 18 period seasonal autoregressive parameter estimate was 0.26638 and had a t ratio of 240.66. The first order moving average parameter estimate was 0.89698 and had a t ratio of 4.13. The positive signs associated with first order autoregressive parameter estimates imply that high weekly item movement of group g in week t-1 lead to a high item movement in week t.

The adequacy of the estimated model was checked by inspection of the chi square values of the residual autocorrelation function, Table 17.

Table 17. Box-Pierce chi-square values for the group g ARIMA$(1,0,1)_{18}$ model

Lag	Chi Square	Prob						
6	2.01	0.365	0.037	0.013	0.006	0.037	0.069	0.023
12	12.10	0.147	-0.072	-0.039	0.017	-0.169	0.056	-0.037
18	17.06	0.253	0.054	-0.018	-0.107	-0.058	0.032	-0.004
24	18.84	0.533	0.000	-0.052	-0.007	0.030	-0.035	0.042
30	22.63	0.654	-0.013	0.004	-0.062	-0.039	-0.070	0.057
36	31.87	0.473	-0.037	-0.107	-0.072	-0.111	-0.051	-0.006
42	37.85	0.476	-0.019	-0.051	-0.011	-0.060	0.052	0.105

The chi square values for lags 6, 12, 18, 24, 30, 36, and 42, indicated that there was no significant autocorrelation present between the residuals for the AR$(1,0,1)_{18}$ model at the ninety-five percent level of significance.

Figure 15 presents a graphical description of group g's Box-Jenkins backcast. Visual inspection of the actual and backcast series revealed that the Box-Jenkins ARIMA model was very capable of reproducing the historical data series. The backcast, in general, resembled the historical data series except that it could was not capable of predicting the large

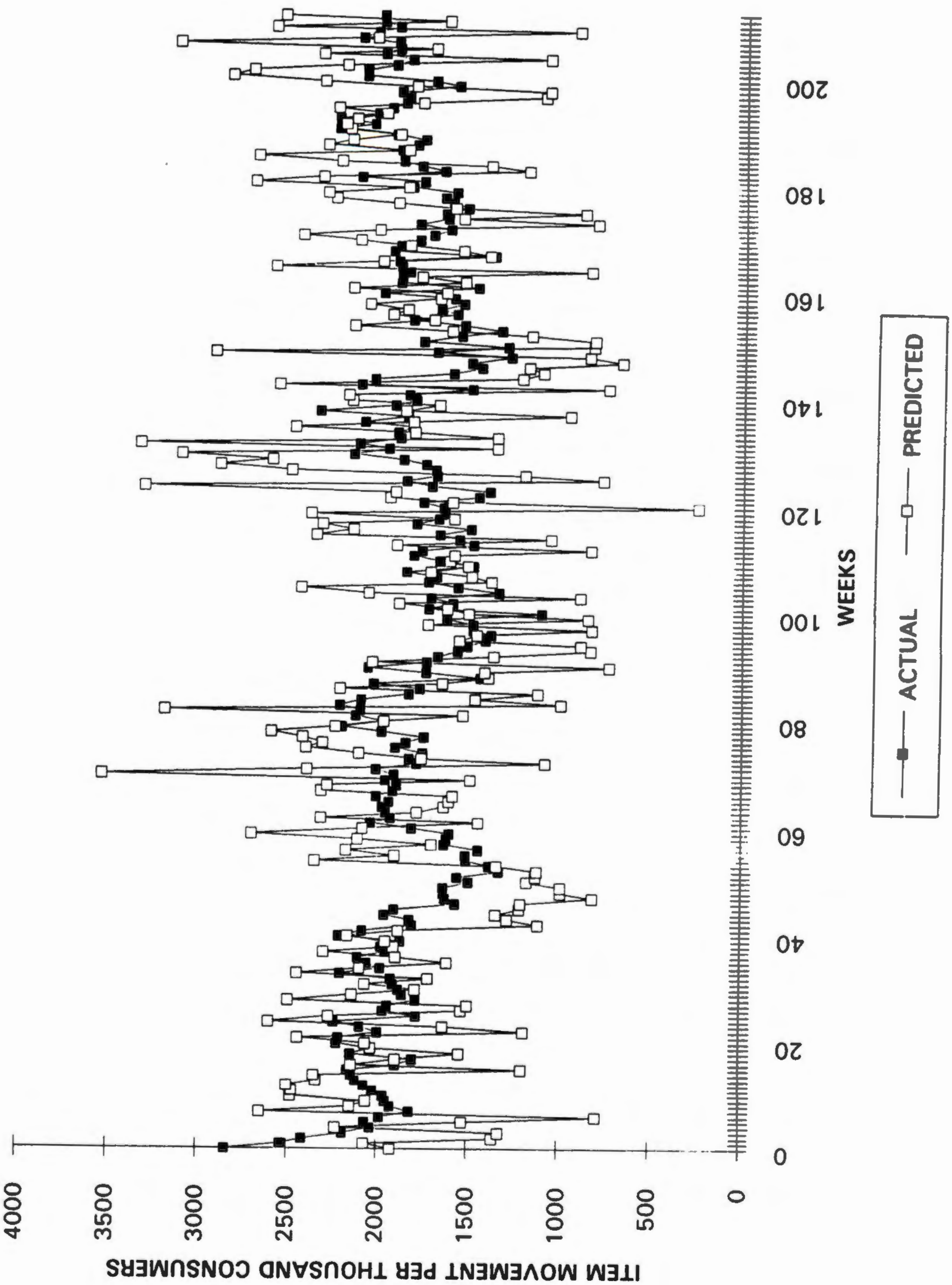

Figure 15. Group g's statistical backcast

fluctuations in item movement. The RMSE and U^2 for group g's backcast were 90667.23 and 0.056, respectively. The model predicted 101 out of 213 directional changes, Table 18.

Table 18. Group g's ARIMA backcast evaluation criteria

Criterion	Value
MSE	90667.23
RMSE	301.11
U^2	0.05
Directional change	101/213

The model used for forecasting steak's weekly item movement was a second order autoregressive model, ARIMA(2,0,0) with no seasonal effects. The mean value of the estimated subperiod was 2423.73. U^2 was 0.05, meaning that the steak Box-Jenkins forecast is superior to the nochange forecast. Decomposition of U^2 indicated a lack of bias in the error series, and the error series was the result of a random fluctuations contained in the data series. The model is expressed as follows:

$$(60)\ Steak_t = \underset{(100.39)}{2396.80} + \underset{(0.0633)}{0.25921(1)} + \underset{(0.0633)}{0.22246(2)}$$

The RMSE for the ARIMA (2,0,0) model was 812.32. The first autoregressive parameter estimate was 0.025921 with a t ratio of 4.09. The second autoregressive parameter estimate was 0.22246 and had a t ratio of 3.51. The positive signs associated with first and second autoregressive parameter estimates imply that high weekly steak item movement in period t-1 and t-2 will lead to a large item movement in the next period.

The adequacy of the estimated model was checked by inspection of the chi square values of the residual autocorrelation function, Table 19.

Table 19. Box-Pierce chi-square values for the steak ARIMA(2,0,0) model

Lag	Chi Square	Prob						
6	8.74	0.068	-0.032	-0.049	0.112	0.014	0.097	0.099
12	16.75	0.080	-0.093	0.015	-0.023	0.119	-0.076	-0.049
18	22.04	0.142	0.050	-0.010	0.046	-0.025	-0.059	0.107
24	25.73	0.264	0.014	-0.037	0.041	-0.098	-0.009	0.030
30	29.82	0.372	-0.014	0.021	-0.036	-0.060	-0.057	-0.078
36	35.37	0.403	0.050	-0.117	-0.053	0.003	0.025	-0.011
42	38.41	0.542	0.044	-0.009	-0.031	0.012	0.039	-0.076

The chi square values for lags 6, 12, 18, 24, 30, 36, and 42, indicated that there was no significant autocorrelation present between the residuals for the AR(2,0,0) model at the 95 percent significance.

Figure 16 presents a graphical description of steak's Box-Jenkins backcast. Visual inspection of the actual and backcast series revealed that the Box-Jenkins ARIMA model was very capable of reproducing the historical data series. The backcast, in general, resembled the historical data series except that it was not capable of predicting the large fluctuations in item movement. Table 20 contains the evaluation criteria. The RMSE and U^2 for steak's backcast were 659863.78 and 0.05, respectively. The model predicted 79 out of 213 directional changes.

Table 20. Steak's ARIMA backcast evaluation criteria

Criterion	Value
MSE	659863.78
RMSE	812.32
U^2	0.05
Directional change	79/213

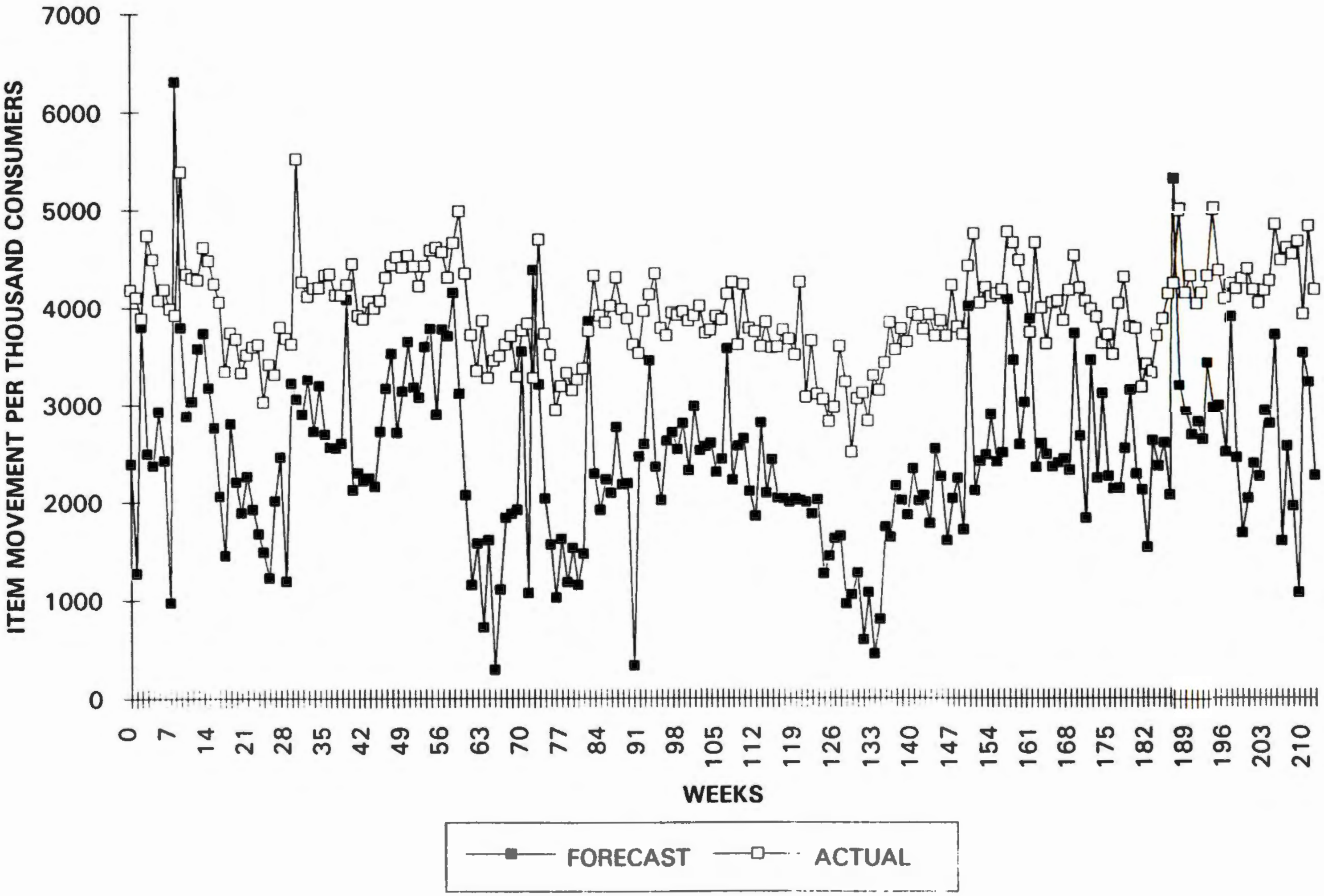

Figure 16. Steak's statistical backcast

3. Transfer Function Models

The estimated transfer functions forecasts were generated using a combination of the theoretical and Box-Jenkins modeling techniques. Again, brand b's model was estimated using 161 weekly observations. Group g's and steak's model was estimated using 213 weekly observations. The transfer function was estimated using the SAS analysis package. The proc ARIMA procedure allows the estimation of transfer function models by specifying the independent variables in the procedure statement. The SAS procedure estimates the transfer function model according to the nine steps outlined by Helmer and Johansson (Chapter 2 p.41).

The ARIMA part of the transfer function model used for forecasting brand b's weekly item movement is a third order autoregressive model, ARIMA(3,0,0). U^2 was 0.05 leading to the inference that the brand b Box-Jenkins forecast is superior to the nochange forecast. Decomposition of U^2 indicated a lack of bias in the error series and that the error series was the result of a random fluctuations contained in the data series. The model is expressed as follows:

$$(61)\quad \underset{(4.46)}{Q_{bt}=3327.8}\ \underset{(6.19)}{-27868.6(P_b)} + \underset{(2.88)}{152.53(POP_7)} + \underset{(2.46)}{243.2(POP_1)}$$

$$\underset{(5.81)}{-428.16(SEA_1)} - \underset{(6.12)}{405.62(SEA_2)} - \underset{(3.57)}{255.39(SEA_3)} - \underset{(2.08)}{0.18(3)}$$

The RMSE for the transfer function model was 206.50. The third period autoregressive parameter estimate was -0.18 with a t ratio of 2.08. The negative sign associated with third autoregressive parameter estimate implies a high weekly brand b item movement in period t-3 leads to a smaller item movement in period t. The own-price variable had a negative sign which was expected. Two cross-advertising POP variables were positive indicating that as shoppers were attracted to the POP advertising, they decided on purchasing brand b instead. Each of the three seasons (spring, summer, and fall) negatively impact brand b's item movement according to the estimated transfer function.

The adequacy of the estimated model was checked by inspection of the chi square values of the residual autocorrelation function.

Table 21. Box-Pierce chi-square values for the transfer function model of brand b's weekly item movement

Lag	Chi Square	Prob	Autocorrelations					
6	3.39	0.640	-0.046	0.005	0.003	-0.109	-0.056	0.027
12	14.06	0.230	-0.037	-0.148	0.049	0.109	0.076	0.107
18	19.41	0.306	-0.004	-0.093	0.064	-0.062	-0.097	-0.019
24	22.63	0.482	-0.101	-0.055	0.015	0.002	0.045	0.006
30	27.96	0.520	0.022	-0.032	0.056	-0.000	0.050	-0.130
36	29.36	0.737	0.025	0.022	-0.040	0.051	-0.025	0.016

Figure 17 presents a graphical description of brand b's transfer function backcast. Visual inspection of the actual and backcast series revealed that the transfer function model was very capable of reproducing the historical data series. The backcast, in general, resembled the historical data series and was capable of predicting the majority of the large fluctuations in item movement. Table 22 presents the evaluation criteria. The RMSE and U^2 for brand b's backcast were 42642.25 and 0.05, respectively. Decomposition of U^2 indicated a lack of bias in the error series and that the error series was the result of a random fluctuations. The model predicted 79 out of 161 directional changes.

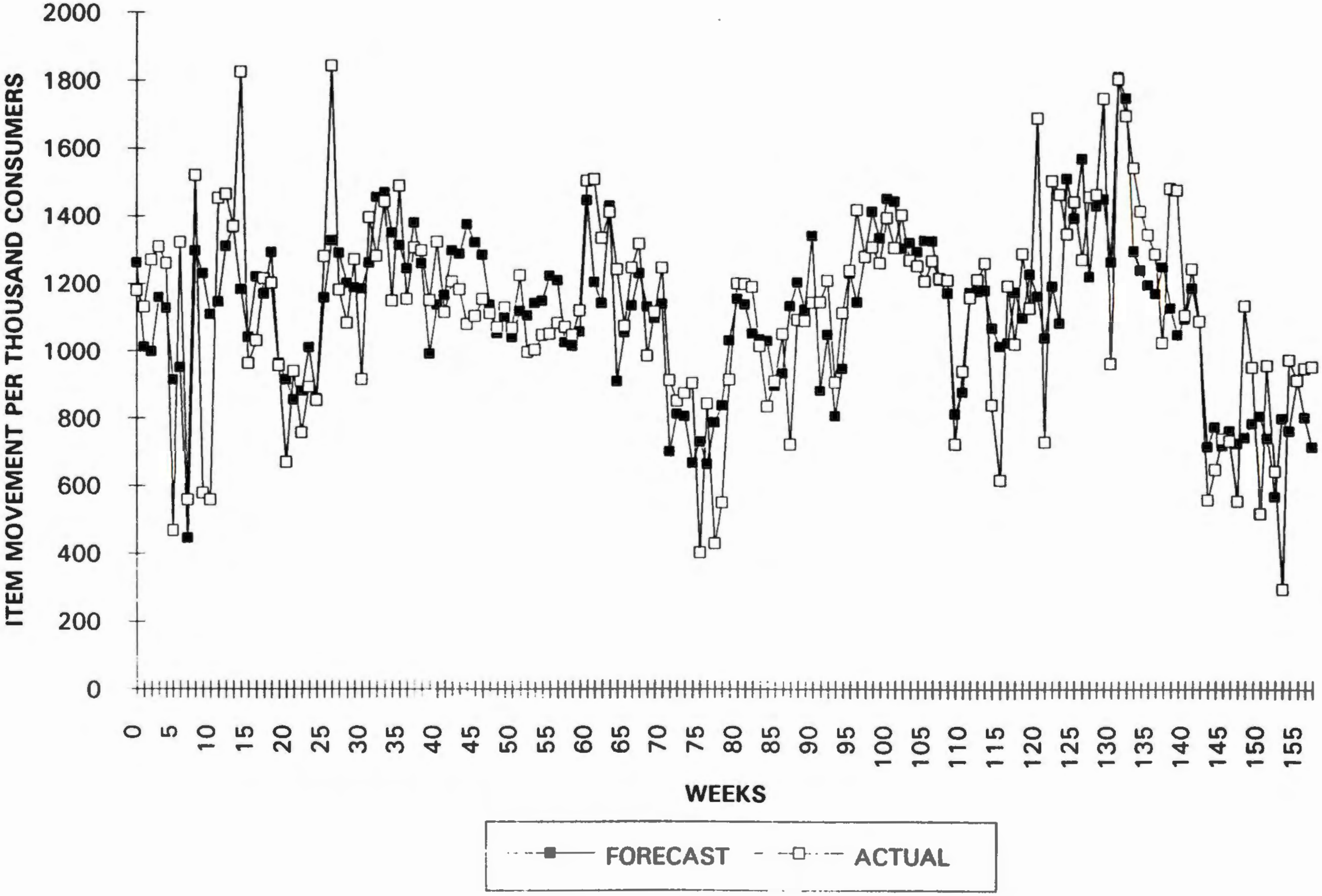

Figure 17. Brand b's transfer function backcast

Table 22. Brand b's transfer function backcast evaluation criteria

Criterion	Value
MSE	42642.25
RMSE	206.50
U^2	0.05
Directional change	79/161

The ARIMA part of the transfer function model used for forecasting group g's weekly item movement was an 18 order autoregressive model, ARIMA(18,0,0). U^2 was 0.07 leading to the inference that the steak Box-Jenkins forecast is superior to the nochange forecast. Decomposition of U^2 indicated a lack of bias in the error series and that the error series was consistent with a hypothesis of random fluctuations in the data series. The model is expressed as follows:

$$(62)\quad \underset{}{Q_{gt}}=\underset{(4.09)}{1620.20}\ \underset{(4.02)}{-0.27477(18)} + \underset{(2.06)}{589.03(HOL_1)}\ +e_t.$$

The RMSE for the transfer function model was 553.49. The negative sign associated with the autoregressive parameter estimate, t-18, implies a high group g weekly item movement in period t-18 leads to a smaller item movement in period t. The own-price variable was not significant at the 95 percent level. The HOL_1 variable, was positive indicating that item movement increased on January first.

The adequacy of the estimated model was checked by inspection of the chi square values of the residual autocorrelation function (Table 23).

Table 23. Box-Pierce chi-square values for the group g transfer function forecast of weekly item movement

Lag	Chi Square	Prob	Autocorrelations					
6	8.38	0.136	0.105	0.070	0.060	0.047	0.101	0.049
12	15.77	0.150	-0.055	0.057	0.021	-0.113	0.098	0.016
18	21.91	0.188	0.101	0.031	-0.082	-0.007	0.077	-0.002
24	23.70	0.420	0.039	0.003	0.026	0.042	0.019	0.049
30	26.00	0.625	0.047	0.046	-0.040	-0.019	-0.033	0.033
36	32.36	0.596	0.003	-0.092	-0.044	-0.085	-0.071	-0.004
42	35.80	0.700	-0.021	-0.029	-0.024	-0.077	0.010	0.062

Figure 18 presents a graphical description of group g's transfer function backcast. Visual inspection of the actual and backcast series revealed that the transfer function model was very capable of reproducing the historical data series. The backcast, in general, resembled the historical data series but was not capable of predicting the extremely large fluctuations in item movement. Table 24 presents the evaluation criteria. The RMSE and U^2 for group g's backcast were 306351.18 and 0.07, respectively. Decomposition of U^2 indicated a lack of bias in the error series and that the error series was consistent with the hypothesis of random fluctuations. The model predicted 105 out of 213 directional changes.

Table 24. Group g's transfer function backcast evaluation criteria

Criterion	Value
MSE	306351.18
RMSE	553.49
U^2	0.07
Directional change	105/213

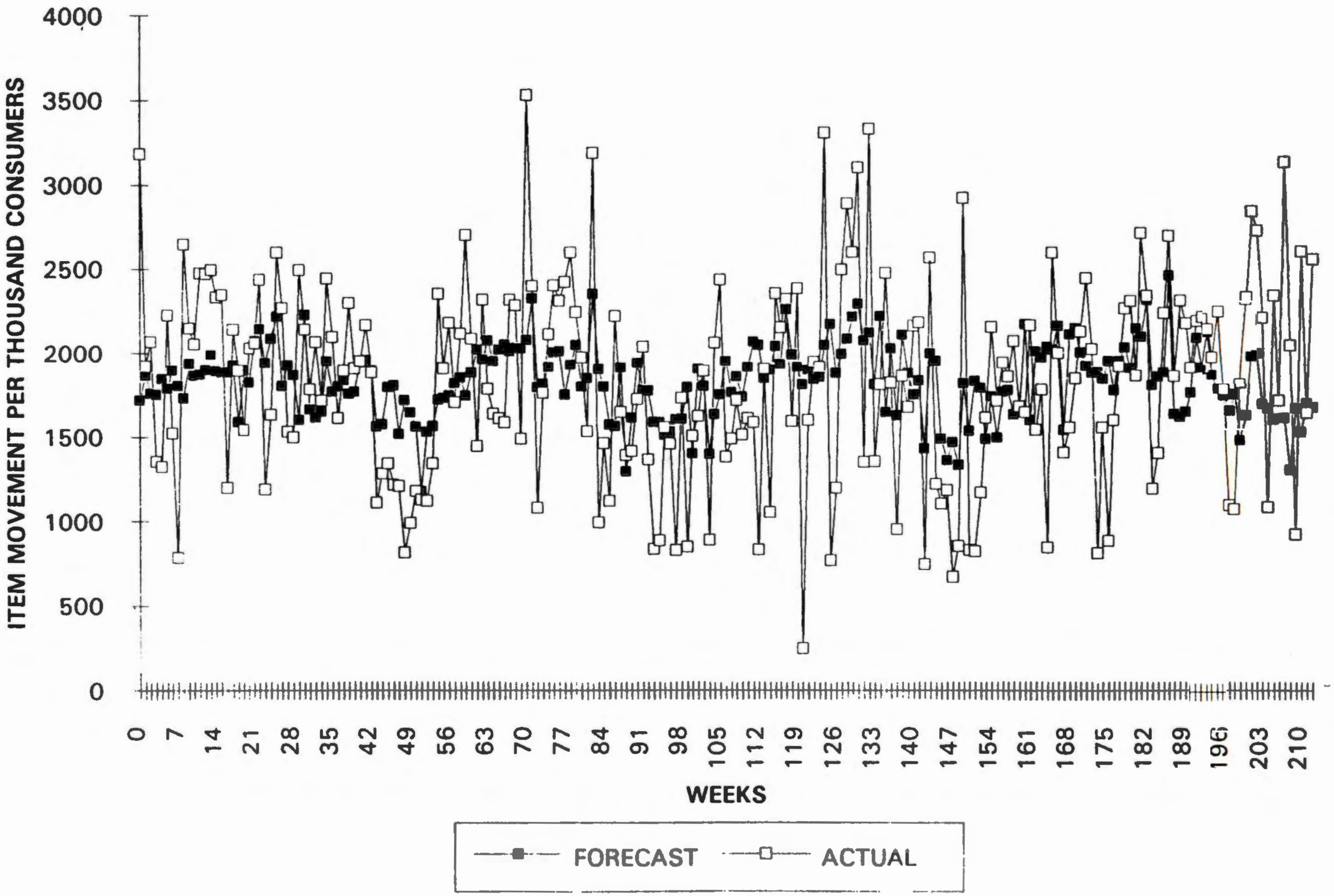

Figure 18. Group g's transfer function backcast

The ARIMA part of the transfer function model used for forecasting steak weekly item movement was a first order autoregressive model, ARIMA(1,0,0). U^2 was 0.04 leading to the inference that the steak Box-Jenkins forecast is superior to the nochange forecast. Decomposition of U^2 indicated lack of bias in the error series and that the error series was consistent with the hypothesis of random fluctuations contained in the data series. The model is expressed as follows:

$$(63)\ Q_{St}=\underset{(5.45)}{2625.20}\ \underset{(4.81)}{-0.37(1)} - \underset{(3.57)}{231.65(P_g)} + \underset{(2.06)}{259.15(SEA_1)} - \underset{(2.27)}{273.95(SEA_3)}$$

$$\underset{(2.12)}{-727.96(HOL_5)} + \underset{(8.33\ )}{0.546(LAG)} + e_t$$

The RMSE for the transfer function model was 680.57. The negative sign associated with the autoregressive parameter estimate, t-1, implies a high weekly item movement of steak in period t-1 leads to a smaller item movement in period t. The own-price variable was negative as expected and significant at the 95 percent level. SEA_1 and SEA_3 had their expected signs. The decrease in steak item movement in the fall time period could be attributed to a decrease in the frequency of consumers grilling outdoors because of the cooler weather. Steak weekly item movement increases in the spring when consumers start to grill-outdoors as the weather becomes more pleasant. The HOL_5 variable, was negative indicating that item movement decreased on the Thanksgiving holiday.

The adequacy of the estimated model was checked by inspection of the chi square values of the residual autocorrelation function.

Table 25. Box-Pierce chi-square values for the steak transfer function forecast of weekly item movement

Lag	Chi Square	Prob	Autocorrelations					
6	8.74	0.068	-0.032	-0.049	0.112	0.014	0.097	0.099
12	16.75	0.080	-0.093	0.015	-0.023	0.119	-0.076	-0.049
18	22.04	0.142	0.050	-0.010	0.046	-0.025	-0.059	0.107
24	25.73	0.264	0.014	-0.037	0.041	-0.098	-0.009	0.030
30	29.82	0.372	-0.014	0.021	-0.036	-0.060	-0.057	-0.078
36	35.37	0.403	0.050	-0.117	-0.053	0.003	0.025	-0.011
42	38.41	0.542	0.044	-0.009	-0.031	0.012	0.039	-0.076

Figure 19 provides a graphical description of steak's transfer function backcast. Visual inspection of the actual and backcast series revealed that the transfer function model was very capable of reproducing the historical data series. The backcast, in general, resembled the historical data series but was not capable of predicting the extremely large fluctuations in item movement. Table 26 presents the evaluation criteria. The RMSE and U^2 for steak's backcast were 463175.52 and 0.04, respectively. The model predicted 113 out of 213 directional changes.

Table 26. Steak's transfer function backcast evaluation criteria

Criterion	Value
MSE	463175.52
RMSE	680.57
U^2	0.04
Directional change	113/213

Comparing the RMSE, U^2, and directional accuracy evaluation criteria for each of the three estimated forecasting models using the historic subperiods revealed that forecasts of peanut butter brand products are more accurate than forecasts of the peanut butter category (Table 27).

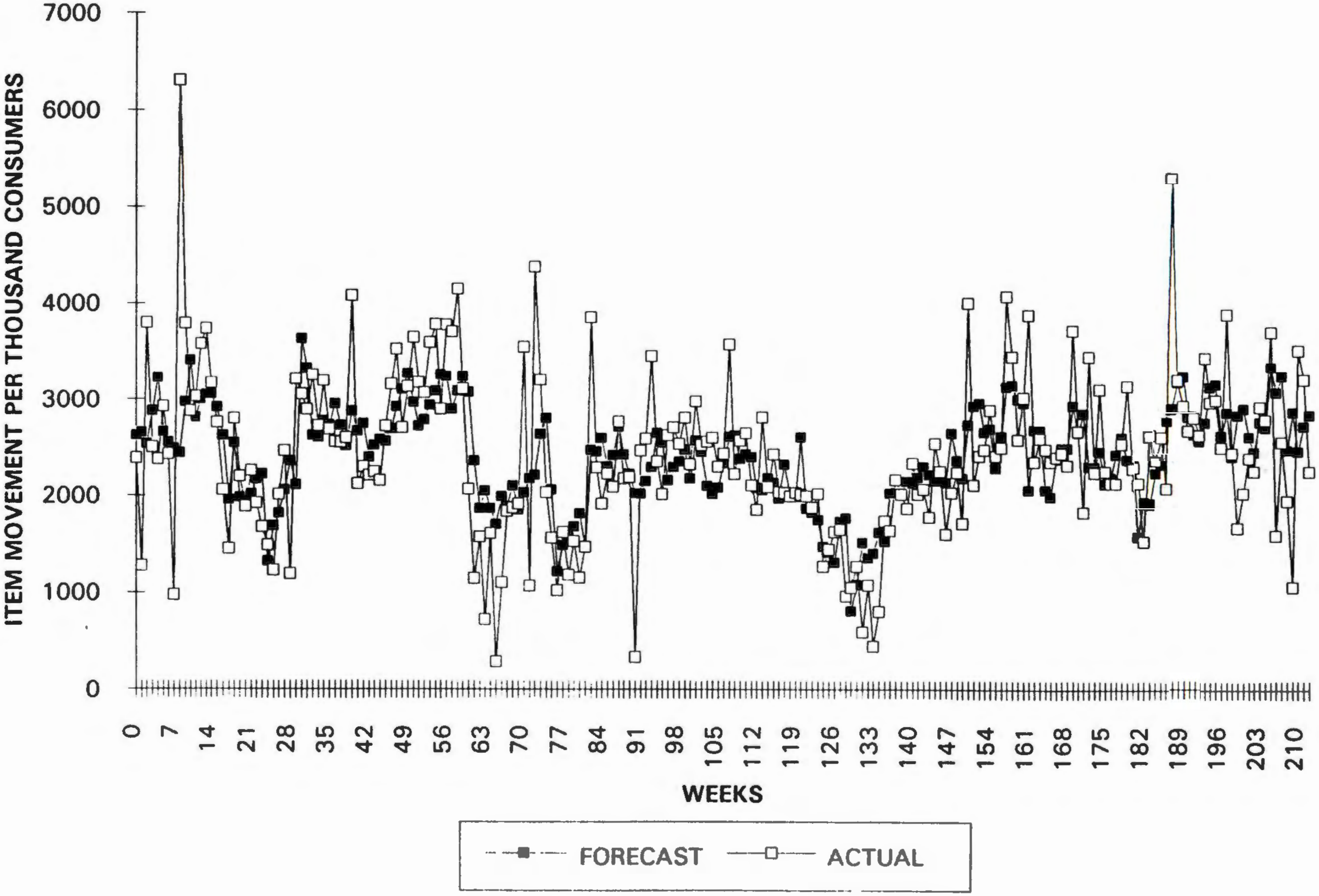

Figure 19. Steak's transfer function backcast

Table 27. Evaluation criteria summary for brand b, group g, and steak historic forecasts

Brand b	Theoretical	Box-Jenkins	Transfer
MSE	49582.83	103002.48	42642.25
RMSE	222.67	320.94	206.50
U^{2*}	0.06	0.06	0.05
Directional accuracy	75/161	91/161	102/161
Group g			
MSE	342277.97	90667.23	306351.18
RMSE	585.05	301.11	553.49
U^{2*}	0.05	0.05	0.07
Directional accuracy	129/213	101/213	105/213
Steak			
MSE	542397.91	659863.78	463175.52
RMSE	736.48	812.32	680.57
U^{2*}	0.04	0.05	0.04
Directional accuracy	102/213	79/213	113/213

* Decomposition of U^2 indicated a lack of bias in the error series and that the error series was the result of a random fluctuations.

The theoretical model, which was specified using economic theory, for brand b was clearly superior to the theoretical model specified for group g. Group g's model was statistically insignificant meaning that the parameter estimates were not statistically reliable. Again, inspection of brand b's and group g's weekly item movement suggested that there is more of a trend or pattern in the brand data relative to the category data. This would imply that the Box-Jenkins model would more accurately forecast brand item movement than category item movement. Brand b's transfer function model was clearly superior to group g's transfer function model, based on the given evaluation criteria. This may be attributed to consumer brand loyalty. Thus, if brand loyalty was present for a given product, these results would imply that supermarket management should focus on forecasting brand item movement as opposed to category item movement based on comparisons of the theoretical, Box-Jenkins, and Transfer function forecast model evaluation criteria.

Brand b's theoretical forecasting model was clearly superior to group g's theoretical demand model using the evaluation criteria present in Table 27. Brand b's estimated model goodness of fit values were consistently better than those associated with group g's estimated theoretical model. Brand b's estimated transfer function model was superior to group g's estimated transfer function model for brand b, but group g's estimated Box-Jenkins model was superior to brand b's estimated Box-Jenkins model.

The specified evaluation criteria suggest that brand b's and steak's estimated transfer function models were superior to the estimated Box-Jenkins and theoretical models. Group g's estimated Box-Jenkins model is superior to the estimated theoretical and transfer function models based on the evaluation criteria.

E. Trial Forecasts

Two week trial forecasts were generated to reflect the normal amount of time available to supermarket managers for forecasting weekly item movement. The two week period reflects the amount of time required to transmit the weekly store level data to corporate headquarters for analysis and generating forecasts. The forecast period started with the week ending on July 4, 1992 and ran through December 27, 1992. Twenty-six weekly observations were forecast for brand b, group g, and the steak category.

Each model was re-estimated by sequentially switching weeks from the trial to the historic subperiods. For example, after a forecast for the July 4, 1992, the actual value was included to update the model and generate the next two period ahead forecast. The sequential addition to the historic period, from the trial period, is appropriate to lessen the differences in the subperiods noted previously in tables 3-4, 6-7, and 9-10. The results of the theoretical, Box-Jenkins, and composite (transfer) forecasting models are provided in Table 28.

Table 28. Trial forecast evaluation criteria for brand b

	Theoretical	Box-Jenkins	Transfer
MSE	164536.55	103005.44	42640.76
RMSE	405.63	320.94	206.50
U^{2*}	0.14	0.06	0.02
Directional accuracy	5/25	13/25	13/25

* Decomposition of U^2 indicated a lack of bias in the error series and that the error series was the result of a random fluctuations.

Comparison of the three forecasting techniques revealed that the transfer function or composite forecasting technique was clearly superior to the theoretical and/or Box-Jenkins methods. The information contained in Tables 29-31 and inspection of the plots of the two week ahead forecasts over time (Figures 20-22) revealed that the transfer function forecast clearly outperformed the theoretical and Box-Jenkins forecasting techniques. The theoretical model, which was built using economic theory to describe the time series, yielded the poorest forecast. The theoretical model exhibited the highest MSE, RMSE, U^2 and predicted the lowest number

Table 29. Brand b theoretical trial forecast and residual series

	ACTUAL	TWO WEEK FORECAST	RESIDUAL^2
1	957.68	311.70	417295.07
2	522.87	342.44	32557.15
3	500.76	358.41	20264.06
4	734.17	305.32	183918.24
5	594.24	364.00	53010.64
6	956.07	428.26	278583.29
7	1150.43	403.11	558477.47
8	951.01	428.26	273260.77
9	955.08	475.76	229742.77
10	1214.10	475.76	545134.44
11	1126.41	738.49	150481.77
12	955.59	542.82	170375.85
13	879.15	489.74	151642.10
14	597.40	461.79	18388.55
15	655.26	478.56	31224.48
16	1085.69	852.71	54280.52
17	891.56	810.79	6523.73
18	458.79	863.88	164098.64
19	880.47	704.62	30923.36
20	183.77	665.50	232058.30
21	965.96	707.41	66846.14
22	1070.52	659.91	168604.02
23	1142.70	659.91	233084.64
24	629.54	718.59	7930.12
25	851.55	707.41	20775.79
26	808.68	690.64	13931.93

Table 30. Brand b Box-Jenkins ARIMA trial forecast and residual series

WEEK	ACTUAL	TWO WEEK FORECAST	RESIDUAL^2
1	957.68	909.17	2353.47
2	522.87	942.03	175694.27
3	500.76	965.50	215983.73
4	734.17	896.16	26239.24
5	594.24	840.49	60640.84
6	956.07	848.49	11573.99
7	1150.43	829.86	102764.03
8	951.01	886.29	4188.17
9	955.08	966.17	123.01
10	1214.10	982.64	53571.97
11	1126.41	991.97	18073.74
12	955.59	1048.87	8700.99
13	879.15	1072.68	37454.83
14	597.40	1054.05	208531.60
15	655.26	1024.66	136454.37
16	1085.69	945.60	19625.10
17	891.56	904.25	160.92
18	458.79	961.83	253045.52
19	880.47	967.34	7547.16
20	183.77	884.06	490403.14
21	965.96	905.91	3605.47
22	1070.52	789.31	79080.30
23	1142.70	856.32	82011.44
24	629.54	930.74	90724.45
25	851.55	995.59	20748.24
26	808.68	939.25	17049.10

Table 31. Brand b transfer function trial forecast and residual series

	ACTUAL	TWO WEEK FORECAST	RESIDUAL
1	957.68	717.81	57536.61
2	522.87	587.59	4188.10
3	500.76	749.64	61940.61
4	734.17	767.60	1117.38
5	594.24	821.12	51474.04
6	956.07	897.33	3450.52
7	1150.43	919.30	53423.39
8	951.01	974.30	542.33
9	955.08	943.76	128.20
10	1214.10	1062.39	23015.20
11	1126.41	1027.93	9698.84
12	955.59	1089.26	17866.33
13	879.15	924.15	2025.43
14	597.40	944.77	120663.76
15	655.26	937.45	79630.07
16	1085.69	968.85	13651.63
17	891.56	991.57	10002.30
18	458.79	822.67	132411.57
19	880.47	756.51	15366.85
20	183.77	661.87	228579.51
21	965.96	874.19	8421.07
22	1070.52	931.95	19201.89
23	1142.70	903.82	57063.46
24	629.54	776.29	21536.88
25	851.55	1000.48	22180.20
26	808.68	847.82	1532.10

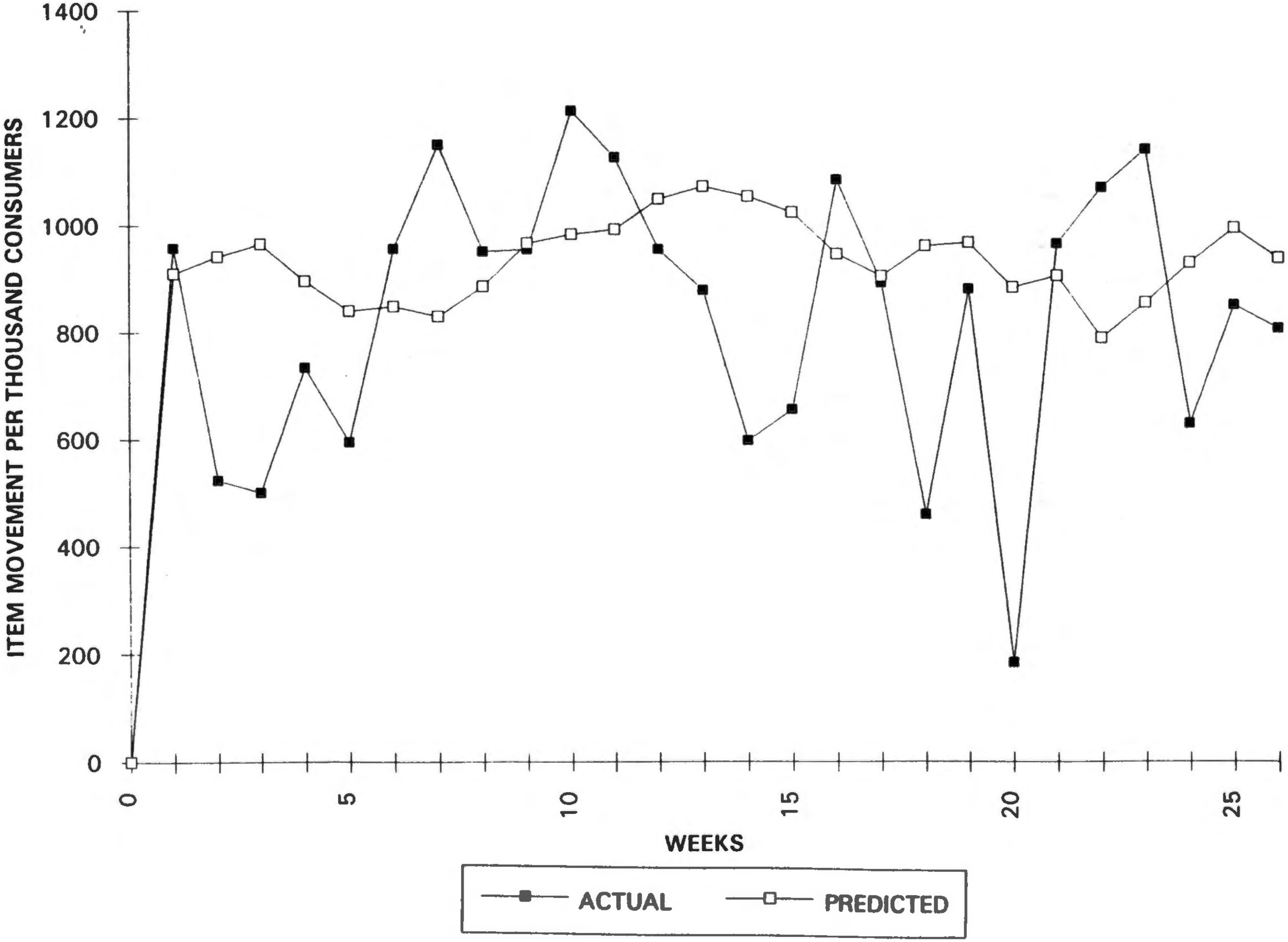

Figure 20. Brand b's theoretical trial forecast

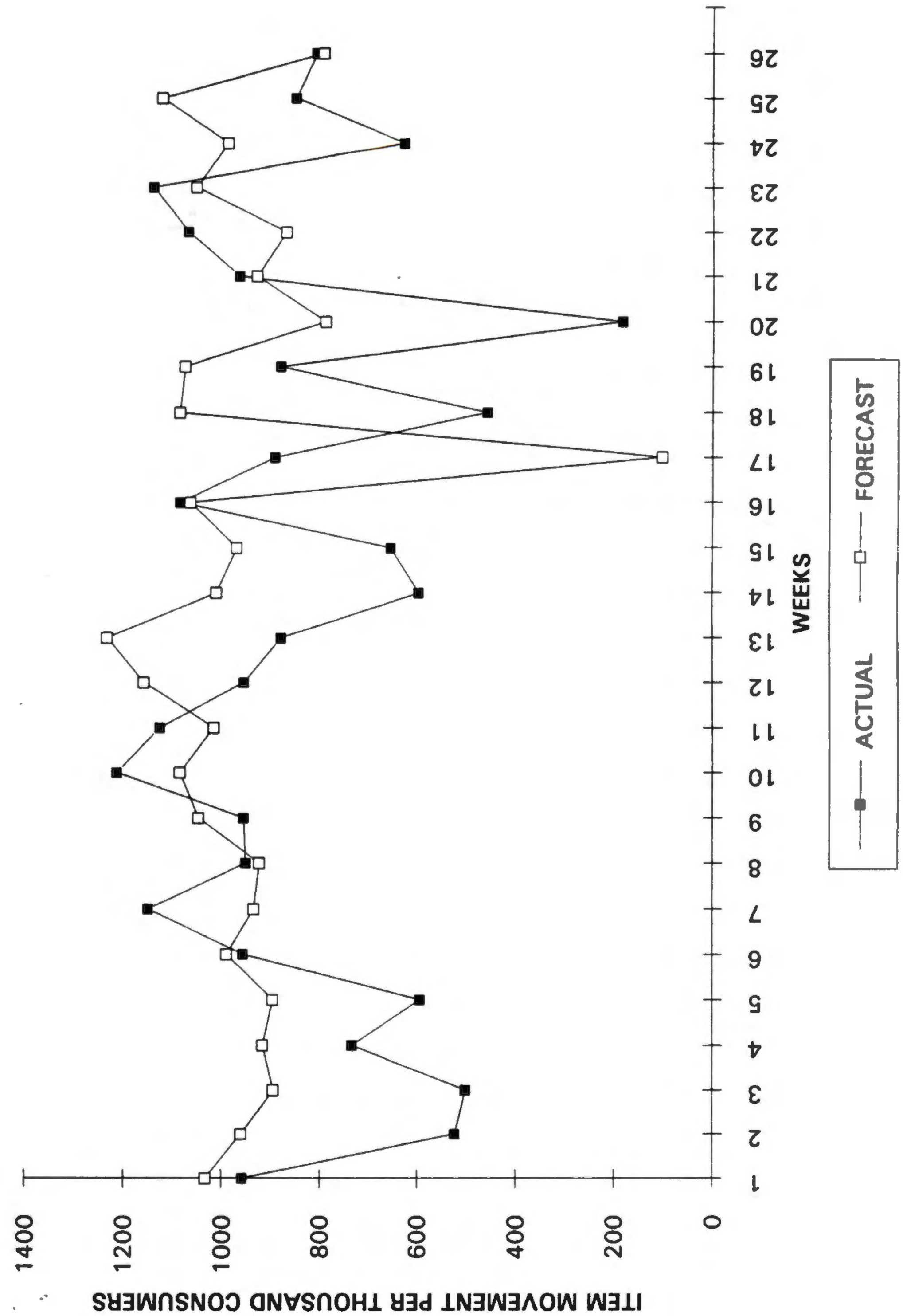

Figure 21. Brand b's statistical trial forecast

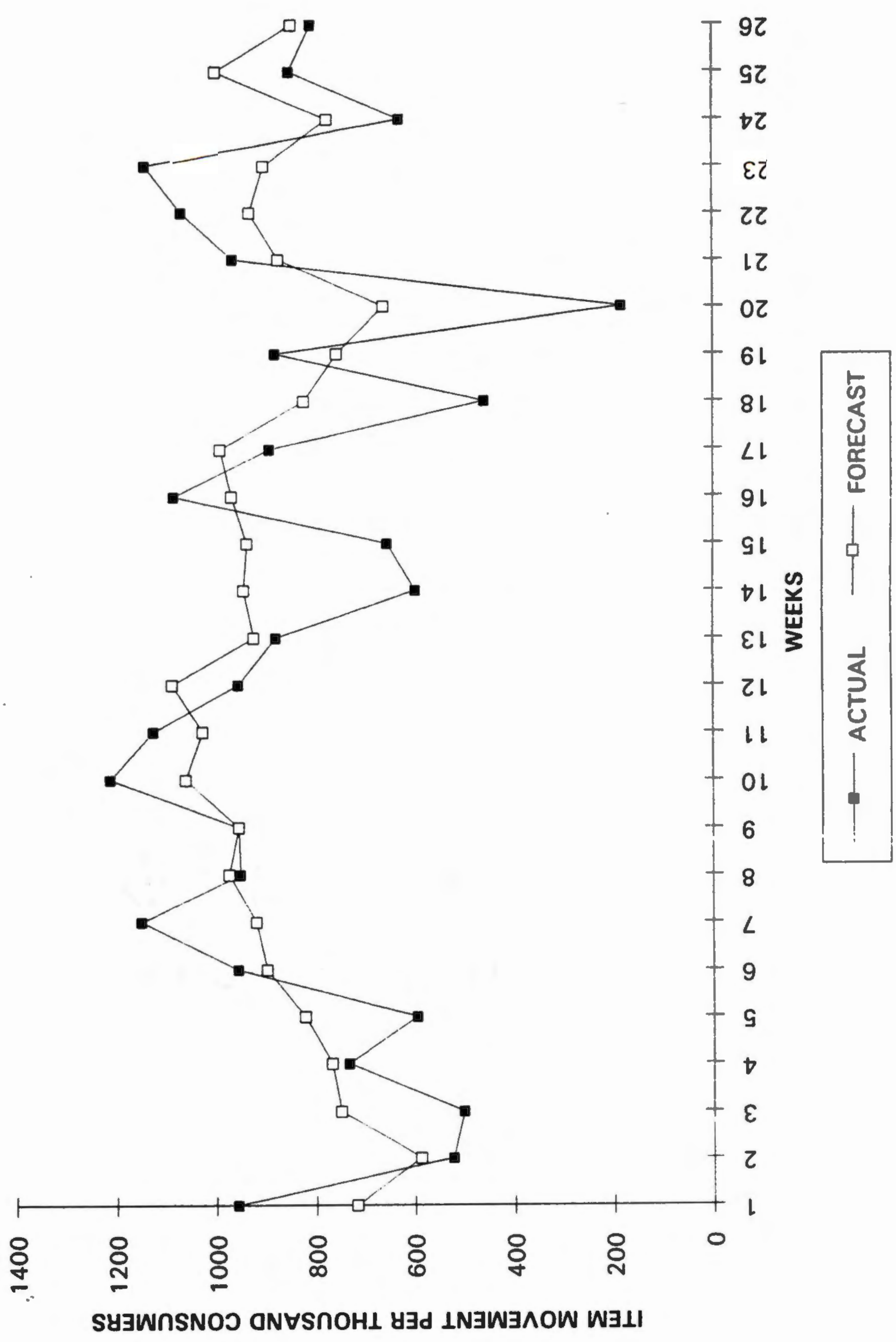

Figure 22. Brand b's transfer function trial forecast

of directional changes. The Box-Jenkins model exhibited the next highest MSE, RMSE, and U^2. Box-Jenkins was able to predict 13 out of 25 directional changes in brand b's weekly item movement. The transfer function weekly item movement forecasts generated the lowest MSE, RMSE, and Theil's U^2 coefficients. The transfer function and Box-Jenkins model each predicted 13 out of 25 turning points. Thus, based on the above criteria, the transfer function generated the superior forecast of brand b's weekly item movement.

The theoretical model for group g was determined to be statistically insignificant at the 95 percent level. Thus, it was not used to generate a forecast of group g's weekly item movement.

Table 32. Trial forecast evaluation criteria for group g

	Theoretical	Box-Jenkins	Transfer
MSE	----	228072.27	306348.35
RMSE	----	477.57	553.49
U^2	----	0.03	0.007
Directional accuracy	----	16/25	17/25

The Box-Jenkins forecast had the smallest MSE and RMSE evaluation criteria and predicted 16 out of the 25 directional changes in group g's weekly item movement. The transfer function forecast exhibited slightly larger MSE and RMSE's, but its U^2 was smaller, and it was able to predict accurately 17 out of the twenty five directional changes in group g's weekly item movement. Comparison of the forecast against the actual time series, Tables 33 and 34 and Figures 21 and 22, revealed a similar conclusion. The transfer function forecasts tended to be flatter than those for Box-Jenkins for weeks 20 through 24. Thus, using the above

Table 33. Group g Box-Jenkins ARIMA trial forecast and residual series

WEEK	ACTUAL	TWO WEEK FORECAST	RESIDUAL²
1	2424.05	1892.01	283066.56
2	1689.1	1947.97	67013.68
3	1722.91	1976.72	64419.52
4	2164.88	1929.74	55290.82
5	1785.81	1897.8	12541.76
6	1722.23	1915.11	37202.69
7	1084.12	1892.74	653866.30
8	2170.7	1870.09	90366.37
9	1542.18	1793.74	63282.43
10	2124.22	1837.94	81956.24
11	1592.05	1812.4	48554.12
12	2527.49	1847.54	462332.00
13	2174.84	1824.22	122934.38
14	2058.85	1893.15	27456.49
15	2157.57	1913.15	59741.14
16	988.41	1916.7	861722.32
17	2344.41	1928.55	172939.54
18	1737.04	1828.5	8364.93
19	2319.59	1880.07	193177.83
20	670.51	1862.02	1419696.08
21	2192.41	1902.8	83873.95
22	1618.85	1719.65	10160.64
23	1737.9	1829.43	8377.74
24	2421.87	1813.19	370491.34
25	2504.14	1812.14	478864.00
26	1857.67	1875.07	302.76

Table 34. Group g transfer function trial forecast and residual series

	TWO WEEK		
	ACTUAL	FORECAST	$RESIDUAL^2$
1	2424.05	1737.57	471254.79
2	1689.10	1830.82	20084.56
3	1722.91	2056.69	111409.09
4	2164.88	2040.74	15410.74
5	1785.81	1852.34	4426.24
6	1722.23	1675.64	2169.70
7	1084.12	1509.59	181016.21
8	2170.70	1592.80	333956.85
9	1542.18	1678.60	18610.42
10	2124.22	2071.88	2739.48
11	1592.05	1693.89	10371.39
12	2527.49	1868.29	434544.64
13	2174.84	1458.29	513429.57
14	2058.85	1726.22	110642.72
15	2157.57	2182.85	639.08
16	988.41	1734.39	556486.16
17	2344.41	2000.25	118446.11
18	1737.04	1737.99	0.90
19	2319.59	1804.12	265709.32
20	670.51	1987.51	1734462.66
21	2192.41	2020.53	29542.73
22	1618.85	1953.36	111896.94
23	1737.90	1975.40	56406.25
24	2421.87	1992.05	184745.23
25	2504.14	2172.76	109812.70
26	1857.67	1761.96	9162.32

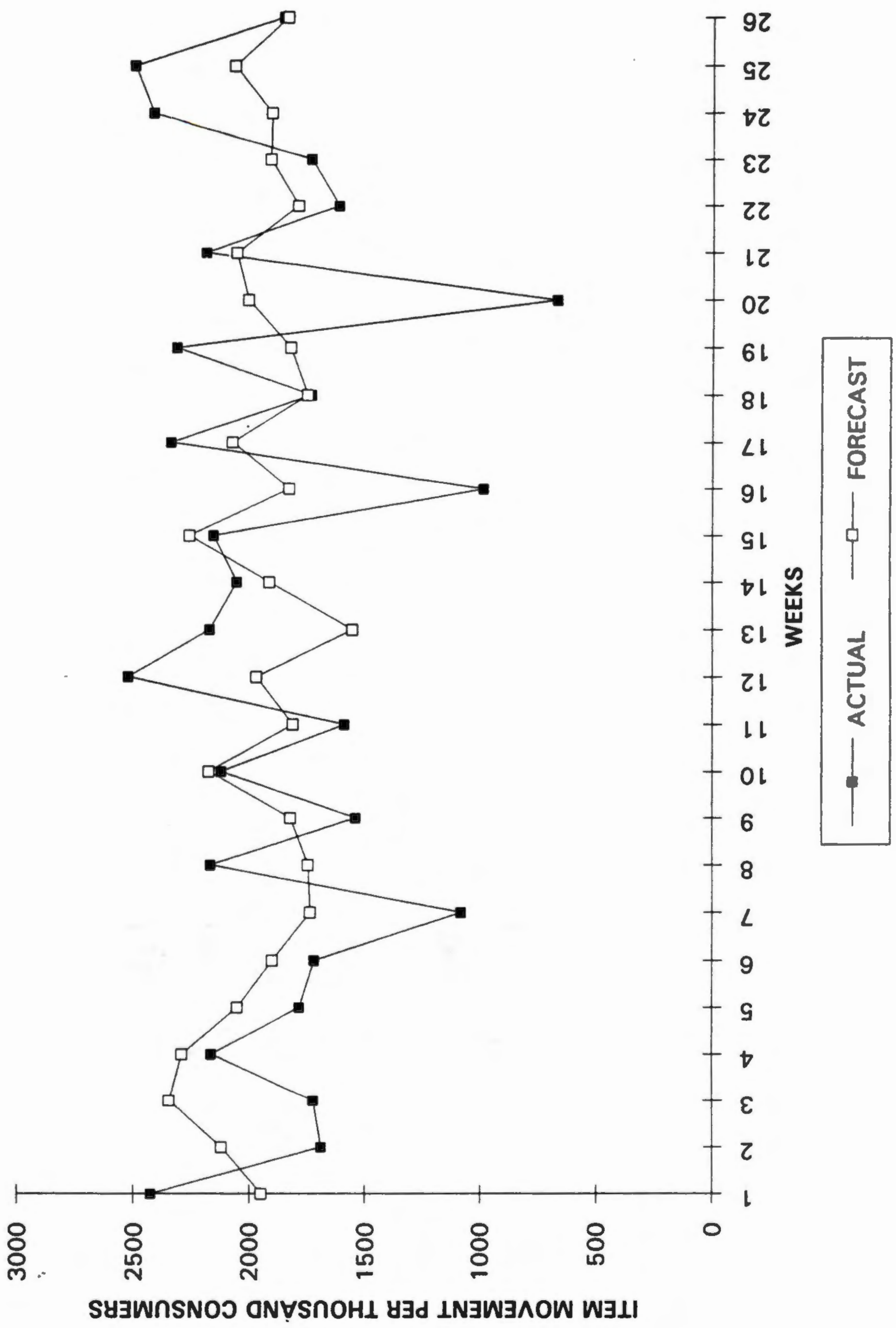

Figure 23. Group g's statisitical trial forecast

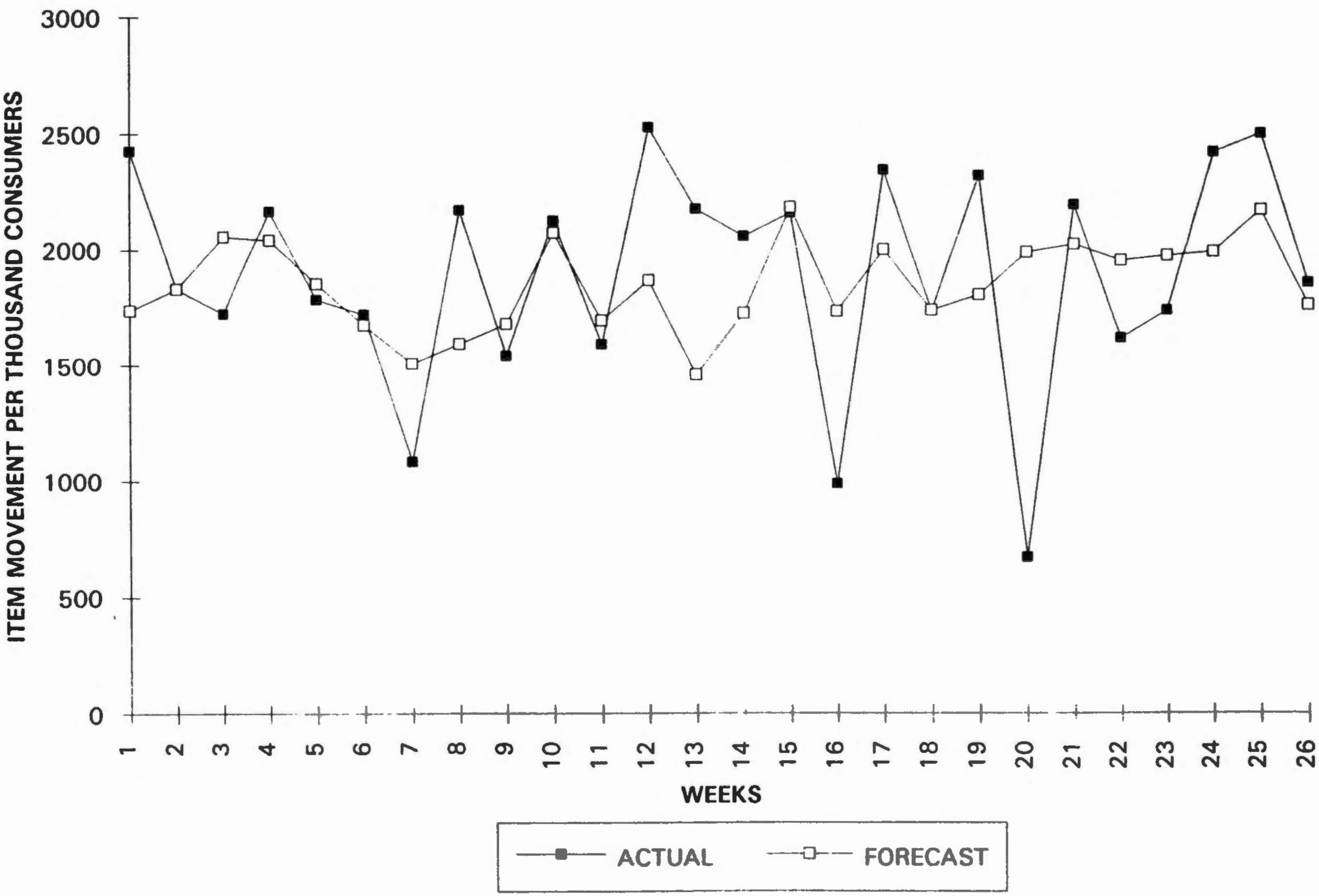

Figure 24. Group g's transfer function trial forecast

evaluation criteria it is inconclusive as to which forecasting technique provides the most accurate forecast of group g's weekly item movement.

Inspection of the evaluation criteria in Table 11, the actual and predicted values in Table 36, and the plots of the actual versus predicted values of steak weekly item movement, Figures 25-27, revealed that the transfer function forecast is clearly superior to both the theoretical and Box-Jenkins forecasts, Table 35.

Table 35. Trial Forecast evaluation criteria for steak

	Theoretical	Box-Jenkins	Transfer
MSE	2210624.29	1169265.32	306348.35
RMSE	1468.82	1081.33	680.58
U^2	0.14	0.09	0.04
Directional accuracy	9/25	9/25	17/25

The transfer function exhibits the smallest MSE, MSE, and U^2 coefficient of the three forecasting models. The RMSE of the transfer function forecast was approximately half the value of the theoretical forecast and approximately 60 percent of the value of the Box-Jenkins forecast. This indicates that the transfer function predicted values were closer, on average, to the actual values than the predicted values of the other two forecasting models. The transfer function was able to predict 17 out of the twenty five directional in the steak's weekly item movement which was almost twice as many as the theoretical or Box-Jenkins forecasts. The transfer function was able to predict high and low item movement levels more accurately than either the theoretical or Box-Jenkins forecasts.

The goodness of fit evaluation criteria suggest that the each of the

Table 36. Steak theoretical trial forecast and residual series

	ACTUAL	TWO WEEK FORECAST	RESIDUAL^2
1	2706.03	3093.17	149881.25
2	1319.58	3171.27	3428762.54
3	1295.41	3093.80	3234224.09
4	2240.66	3302.25	1126962.69
5	1321.06	3248.64	3715576.23
6	3580.48	3384.86	38269.21
7	3581.90	3568.23	186.63
8	2255.33	3490.16	1524808.53
9	2090.09	3219.49	1275544.74
10	2457.54	3233.61	602285.97
11	4359.99	3696.74	439907.81
12	3251.05	3762.54	261625.83
13	2390.18	3482.88	1194000.84
14	1668.14	2766.35	1206068.93
15	1545.65	2585.52	1081323.96
16	4103.65	2605.52	2244389.65
17	2983.36	3158.57	30699.00
18	1345.72	2916.92	2468684.95
19	3720.11	2797.05	852055.93
20	554.093	3187.60	6935405.52
21	1963.03	2927.20	929618.06
22	724.116	2750.16	4104856.02
23	897.526	2360.76	2141081.27
24	1458.87	-700.65	4663584.74
25	2014.9	2657.34	412719.02
26	2706.03	2670.47	1268.28

Table 37. Steak Box-Jenkins ARIMA trial forecast and residual series

WEEK	ACTUAL	TWO WEEK FORECAST	RESIDUAL²
1	2706.03	2706.03	0.0
2	1319.59	2407.10	1182684.53
3	1295.41	2478.87	1400572.84
4	2240.68	2102.60	19062.77
5	1321.06	2015.66	482467.77
6	3580.50	2288.06	1670370.14
7	3581.90	2076.20	2267132.49
8	2255.34	2677.62	178322.09
9	2090.10	2808.32	515845.71
10	2457.54	2424.16	1114.29
11	4359.99	2299.80	4244415.80
12	3251.05	2396.70	729920.76
13	2390.19	2968.94	334956.19
14	1668.15	2757.44	1186561.42
15	1545.66	2444.14	807269.90
16	4103.66	2185.36	3679859.54
17	2983.37	2108.24	765849.02
18	1345.72	2842.11	2239177.05
19	3720.12	2665.18	1112896.29
20	554.09	2126.17	2471425.78
21	1963.04	2719.48	572207.53
22	724.12	1936.36	1469534.06
23	897.53	2164.90	1606235.59
24	1458.88	1887.29	183536.84
25	2014.91	1866.07	22153.35
26	1741.92	2038.67	88061.16

Table 38. Steak transfer function trial forecast and residual series

Week	ACTUAL	TWO WEEK FORECAST	$RESIDUAL_2$
1	2706.03	2708.27	5.02
2	1319.59	1849.93	281292.34
3	1295.41	1758.59	214535.71
4	2240.67	1888.77	124538.41
5	1321.06	2452.19	1279455.08
6	3580.49	2362.21	1484206.16
7	3581.90	3339.16	58922.71
8	2255.34	3766.50	2283604.55
9	2090.10	2218.43	16468.59
10	2457.54	2499.97	1800.30
11	4359.99	2610.67	3060120.46
12	3251.05	3786.15	286332.01
13	2390.19	3070.70	463093.86
14	1668.15	3103.54	2060631.54
15	1545.66	1648.35	10539.08
16	4103.66	3128.50	950917.52
17	2983.37	2677.13	93782.94
18	1345.72	1852.57	256896.92
19	3720.12	3349.50	137359.18
20	554.09	1126.33	327458.62
21	1963.04	2446.40	233636.89
22	724.12	743.15	362.52
23	897.53	1642.36	554786.63
24	1458.88	135.35	1751731.66
25	2014.91	1945.44	4826.08
26	1741.92	1960.61	47825.32

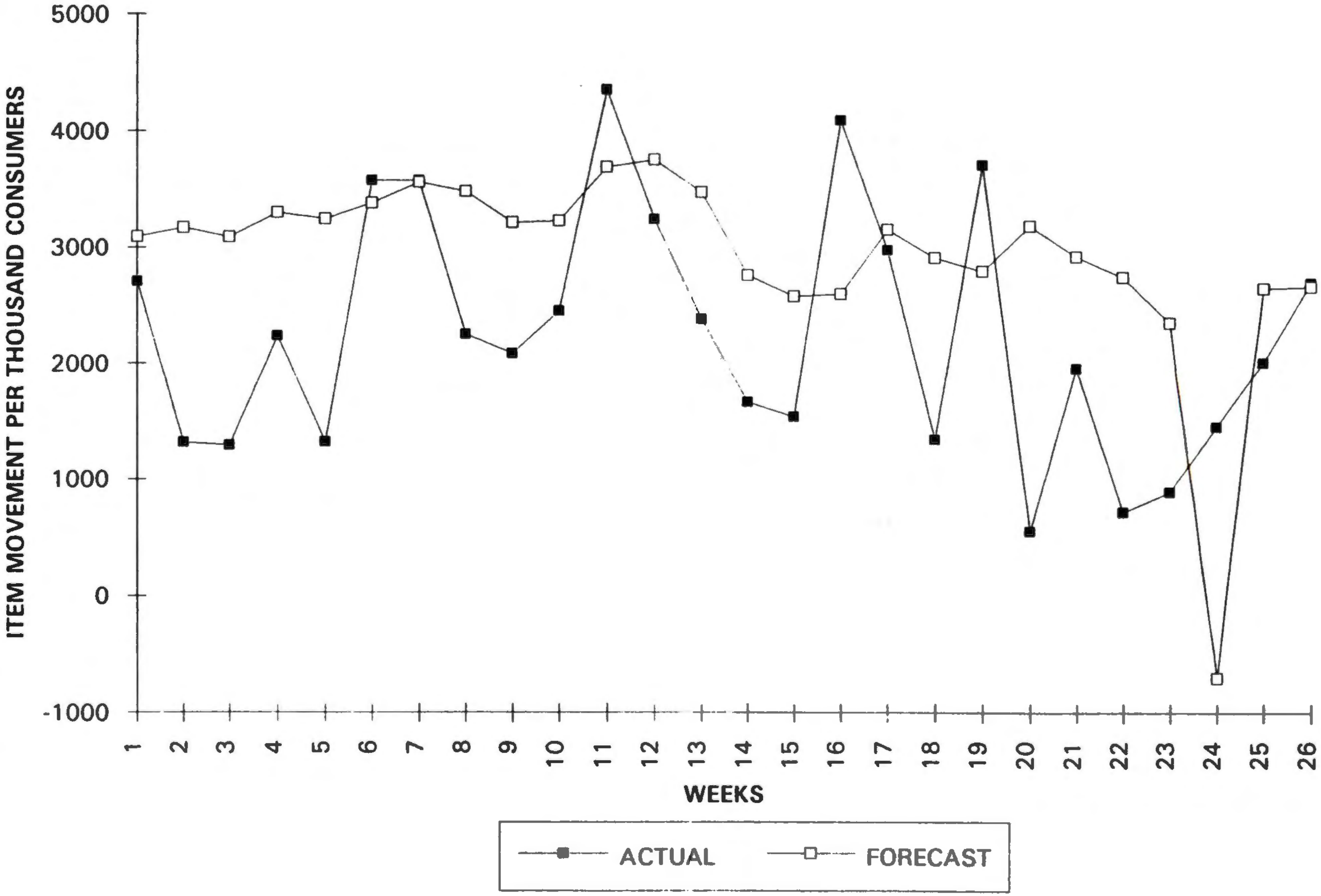

Figure 25. Steak's theoretical trial forecast

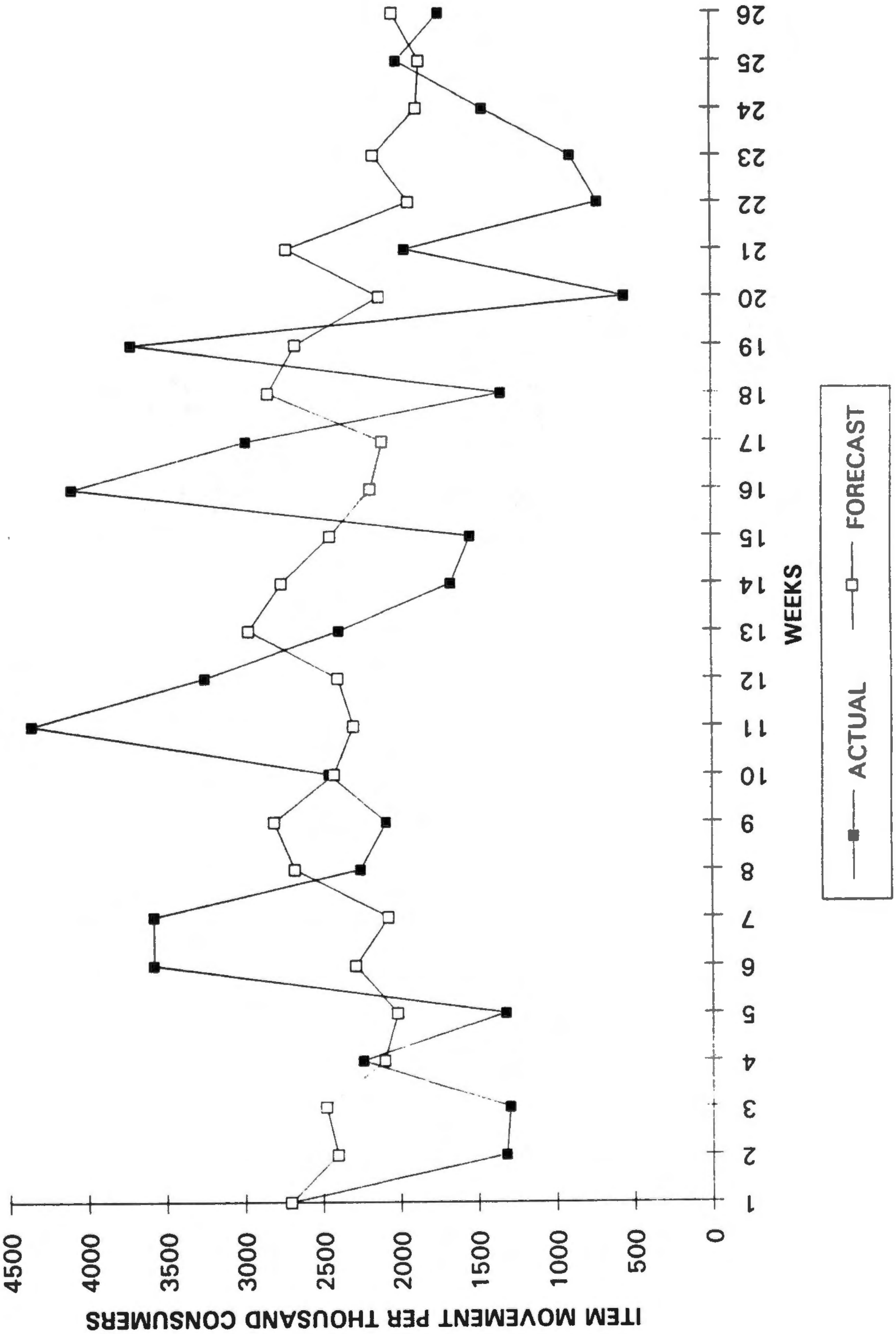

Figure 26. Steak's statistical trial forecast

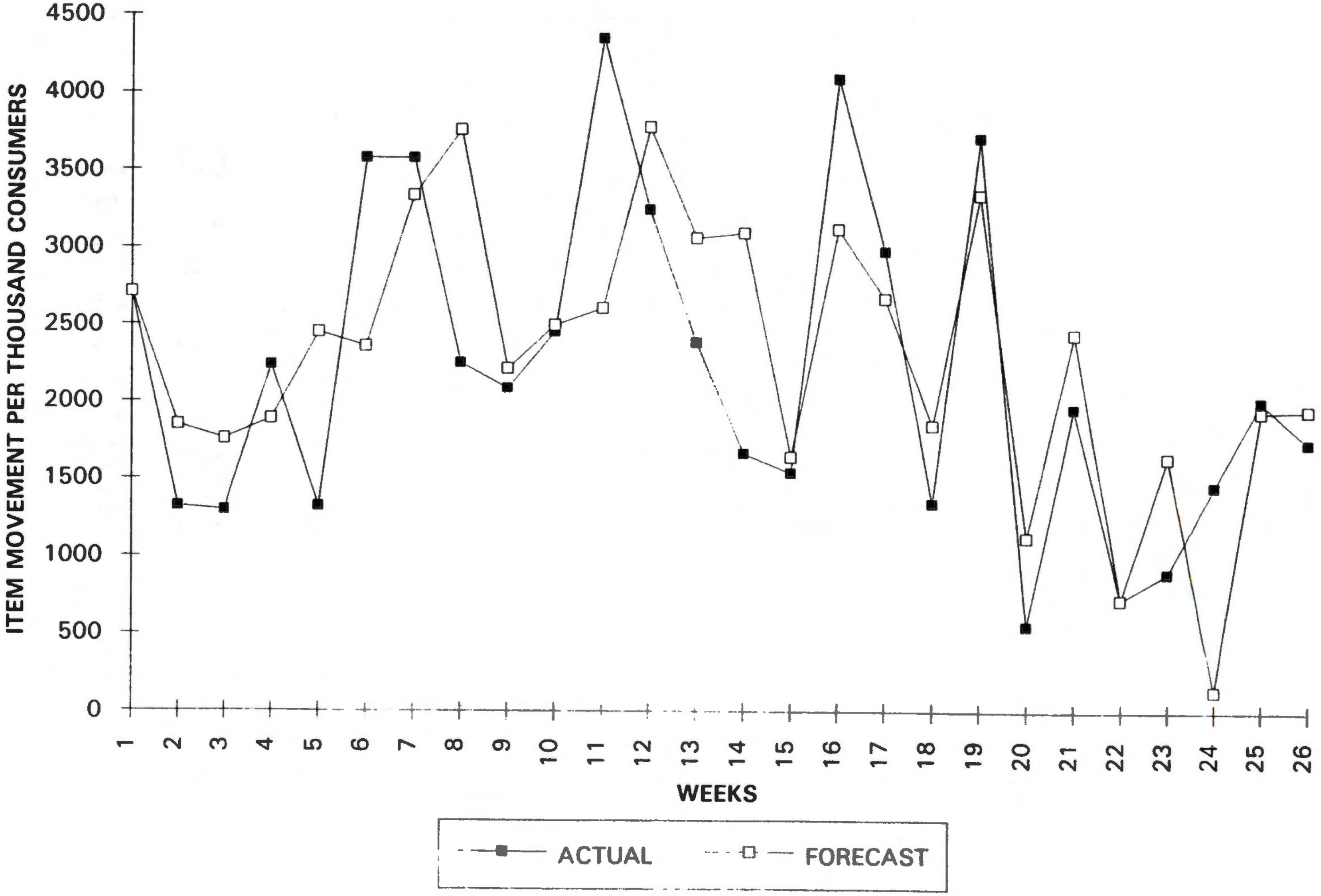

Figure 27. Steak's transfer function trial forecast

aforementioned forecasting techniques was relatively accurate in forecasting actual item movement while they all performed poorly at predicting directional change.

F. Conclusions

This chapter has presented an evaluation and comparison of three forecasting methods. The three foods are: a highly processed brand, the brand's corresponding category, and a fresh beef steak aggregate. The results of the research suggest that the transfer function forecasts, for each of the three foods, were superior to both the theoretical and Box-Jenkins methods. The superiority of the transfer function was demonstrated in both the backcast of the historical data series as well as the forecast of the trial data series. The forecasting techniques, in general, were better able to predict steak weekly item movement than the brand or group item movement.

Chapter VI

Summary, Conclusions, Implications, and Limitations

A. Summary

Scan data have provided a suitable data source for estimating consumer demand relationships at the retail level. These data capture actual quantity and price information which can be combined with additional explanatory variables, such as advertising, promotion and seasonality, to estimate the retail demand function for specific products. Despite the availability of scan data, relatively little research has focused on forecasting at the retail level.

The competitive nature of the supermarket industry, both the encroachment of warehouse food retailers and generic private label products, have lead to an increased interest in consumer demand analysis at the retail level. The increased competition from nontraditional retail outlets has eroded the traditional supermarket's market share. The nontraditional grocery outlets are perceived to be less expensive than their traditional grocery outlets. Thus, traditional grocery outlet managers have become increasingly interested in reducing operating costs. One method of reducing operating costs is to reduce inventory levels via implementation of an efficient consumer response (ECR) strategy, a version of just-in-time delivery. The ECR strategy has the potential to reduce inventory levels which can directly lower inventory costs. The rise of ECR has created a need for accurate product demand forecasts at the supermarket level to maintain adequate inventory levels. The ability to forecast weekly demand in response to changes in seasonality, holidays, and promotional and advertising campaigns is very important to retail managers for implementation of an ECR strategy.

The objectives of this study are to 1) develop alternative forecasting methods that are suitable for scan data, 2) estimate and compare the alternatives with respect to food groups and individual products in terms of their forecast accuracies using a scan data base, and 3) estimate and compare the alternatives with respect to food groups and individual products in terms of two week trial forecasts. The fist objective was met by a review of the literature. More specifically, the literature review suggested the theoretical, Box-Jenkins, and transfer function algorithms as well as forecast evaluation criteria. The second objective was achieved by selecting three types of foods from the available scan data base to be used for empirical work in evaluating the three methodologies. One product was a brand of peanut butter, another was peanut butter brand as a group, and the third was fresh beef steak which is a highly perishable product. The historical and trial forecast accuracies were then evaluated and compared, to reach objective three.

The theoretical forecasting model was developed utilizing economic theory and previous consumer demand research. The model described weekly product item movement as being a function of own- and cross-prices, own- and cross-advertising (television, radio, and newspaper), holidays, and seasonality. The theoretical model or brand b also included point of purchase and the start of the Knox County, Tennessee, school year.

The second forecasting model specification was developed using the Box-Jenkins methodology. This technique does not incorporate structural explanatory variables, but rather, identifies and replicates underlying patterns in the data series utilizing past item movement and disturbances in the series.

The third forecasting method combines the structural variables contained in the theoretical model with the pattern identification and replication ability of the Box-Jenkins model to produce a composite model known as a transfer function.

Scan data provide a relatively new and alternative data source that can be used to obtain new estimates of food demand relationships. They contains records that are capable of tracking individual products across time, and if pooled with data from other stores, may posses both time-series and cross-sectional data characteristics.

This study utilized weekly scan and advertising data (television, newspaper, radio, and point of purchase) which was supplied by a multi-regional supermarket chain. The data consisted of weekly UPC-level prices, item movement, and chain-initiated television, radio, and newspaper advertising. The data were pooled across five stores that catered to average to above average income food shoppers.

The data were divided into two subgroups. The first subgroup of data was used to estimate the alternative forecasting models and generate product backcasts for technique evaluation and comparison. The second subgroup of data, the last 26 weeks for each product, was used to generate a two week trial forecast. Again, the models and their forecasting abilities were evaluated and compared across alternative methods.

Group g's and steak's data were measured over a 239 week period starting with the first week of June 1988 and continuing through the last week of December, 1992. Brand b's was measured over 187 observations beginning with the last week of May 1989 and ending with the last week of December, 1992. The last 26 weeks of data, for each of the three food products, was set aside for the trial forecasting period.

B. Conclusions

1. Estimation and Comparison of Alternative Techniques

The three alternative forecasting models were estimated using the historic subgroup data. The alternative forecasting models were evaluated individually by the evaluation criteria to choose the "best model" to represent each technique. These model estimates were then used to generate backcasts of the data series for each of the three food products, brand b, group g, and steak. The alternative techniques were then evaluated and compared.

The results of the backcast forecast evaluation and comparison suggested that the transfer function forecast was superior to the Box-Jenkins and theoretical forecast in predicting weekly item movement for brand b and steak. Group g's weekly item movement was best forecast utilizing the Box-Jenkins methodology.

2. Trial Forecast Evaluation

Each of the alternative forecasting techniques was utilized to generate two week trial forecasts of weekly item movements for brand b, group g, and steak. The historic subperiod was sequentially updated to simulate data becoming available. This approach to trial forecasting has the advantage of accounting for the differences in dependant variable means and standard deviations between the historic and trial periods.

The results of the two week trial forecast evaluation suggested that the transfer function technique was superior to the theoretical and Box-Jenkins techniques in accurately forecasting weekly item movement for each of the three products, a highly process brand, its associated group, and steak a variable weight perishable product. This study has found that the transfer function is the best of the three techniques for use in forecasting weekly retail item movement for both brand and category peanut butter and the steak category. However, the results also indicate that each of the forecasting models was relatively accurate in forecasting actual item movement but performed poorly in predicting directional change.

C. Implications

The results of this study suggest that the transfer function forecasting technique is superior to the Box-Jenkins and theoretical forecasting techniques in predicting retail weekly item movement for peanut butter and steak food products. The poor results of the econometric model may suggest that additional research and analysis is needed to determine other structural variables and or model specifications which are more effective in explaining weekly item movement for retail food products.

Another implication is the difference in forecast accuracies observed between the brand and category level forecasts. This is an important finding given the increased interest in category management and popularity of private label products. Given that the brand forecasts were superior to the category forecasts, in this study, suggests that there are different factors involved in forecasting and managing brands and categories. The interest and implementation of ECR should focus on forecasting demand relationships at the brand level as opposed to the category level.

As pointed out in chapter III, different forecasting techniques were found to generate superior forecasts under different circumstances. Thus, there is a need for additional research in the area of forecasting demand at the retail level. For example, store managers will want to know which technique most accurately predicts weekly item movement for individual variable weight and nongrocery products.

D. Limitations

One limitation of this study was that actual steak quantity was not available and therefore item movement was used as a proxy. Another limitation is that the study did not include coupons, which could have a significant effect at the brand level, or special in-store advertising and/or promotional activity. The brand level forecasts were superior to the category level forecasts suggesting that the problem was not too severe. A third limitation is that only one market and three foods were used in this study. A fourth limitation is that seasonal price and advertising were not included. These variables may provide valuable information for making seasonal marketing decisions. A fifth limitation is the absence of competing supermarkets price and advertising levels. Further research should address the aforementioned limitations and should include more foods and markets. The forecasting methods were limited in their functional form. Specifically, the theoretical and transfer function forecasts were limited to linear relationships. These forecasting models could be compared and evaluated to the forecasting abilities of nonlinear model forecasts. A comparison of the linear and nonlinear forecasts (a modification of the theoretical approach) would provide additional insight into the feasibility of forecasting algorithms for supermarket food demand using scan data bases.

BIBLIOGRAPHY

BIBLIOGRAPHY

Allen, Geoffrey P. "A Note on Forecasting with Econometric Models." Northeastern Journal of Agricultural and Resource Economics 13(1984):264-267

Archabal, Dale P., Shelby H. McIntyre, Cherryl H. Bell, and Nancy Tucker. "The Effect of Nutrition P-O-P Signs on Consumer Attitudes and Behavior." Journal of Retailing, 63(1987):9-24.

Beilock, Richard P. and James W. Dunn. "An Application of the Box-Jenkins Approach to Forecasting the Cold Storage of French Fries." Journal of Northeast Economics Council, 9(1980):51-56

Bessler, David A. "Forecasting Multiple time Series with Little Prior Information." American Agricultural Economics Association 72(1990):788-792.

Bessler, David A. and Jon A. Brandt "Forecasting Livestock Prices with Individual and Composite Methods." Applied Economics 4(1981):513-522.

Brandt, Jon A. and David A. Bessler. "Composite Forecasting: An Application with U.S. Hog Prices." American Journal of Agricultural Economics. 63(1981):135-140.

Broadbent, Simon and Stephen Colman "Advertising effectiveness:Across Brands." Journal of Market Research Society 28(1985):15-24. *

Broadbent, Simon "Modeling with adstock." Journal of the Market Research Society 26(1984):295-312.

Canova, Fabio "An Alternative Approach to Modeling and Forecasting Seasonal Time Series." Journal of Business and Economic Statistics 10(1992):97-108.

Capps, Oral Jr., and Johannes Adrianus Lambregts "Assessing Effects of Price and Advertising on Purchases of Finfish and Shellfish in a Local Market in Texas." Southern Journal of Agricultural Economics 23(1991):181-190.

Capps, Oral C. Jr. "Utilizing Scanner Data to Estimate Retail Demand Functions for Meat Products." American Journal of Agricultural Economics, 71(1989):750-60.

Capps, Oral C. Jr. "The Ultimate Data Source for Demand Analysis?" Working Paper, Texas A & M University (1987).

Capps, Oral C. Jr. "Defining Management Responsibilities in the Retail Food Distribution Industry."Journal of Food Distribution Research, September, 1988:33-45.

Capps, Oral C. Jr. "MICRODATA BASES -- SCANNER DATA." Working Paper, Texas A & M University (1987).

Carman, H. F. and E. E. Figueroa. "An Analysis of Factors Associated with Weekly Food Store Sales Variation." Agribusiness, 2(1986):375-90.

Chang, Hui-Shung and Henry Kinnucan "Advertising and Structural Change in the Demand for Butter in Canada." Canadian Journal of Agricultural Economics 38(1990):295- 308.

Clover, Vernon F. "Relative Importance of Impulse Purchasing in Retail Stores." Journal of Marketing, 25(1950), 66-70.

Conway, Roger K., Charles B. Hallahan, Richard P. Stillman, and Paul T. Prentice."Forecasting Livestock Prices: Fixed and Stochastic Coefficients Estimation Comparisons." Agribusiness, Vol. 6, No. 1 (1990):15-32.

Craine, Roger and Arthur M. Havenner "Forecast Comparisons of Four Models of U.S. Interest Rates." Journal of Forecasting, 7(1988):21-29.

Culputa, Sunij. "Impact of Sales Promotions on When, What, and How Much to Buy." Journal of Marketing Research, 15(1988):342-55.

Curhan, Ronald C. "The Effects of Merchandising and Temporary Promotion Activities on Sales of Fresh Fruit and Vegetables in Supermarkets."Journal of Marketing Research, 11(1974):286-94.

Curhan, Ronald C. "The Relationship Between Shelf Space and Unit Sales in Supermarkets." Journal of Marketing Research, 9(1972):406-12.

Eastwood, David B.. "Potential Consumer Benefits of Scanner Data: The Case of Supermarkets." Proceedings of the 35th Annual Conference of the American Council on Consumer Interests (M. Carsky, ed.).: Columbus, MO: American Council on Consumer Interests, 1989:14-17.

Eastwood, David B., Morgan Gray, and John Brooker. "Forecasting Item Movement with Scan Data:Box-Jenkins Results." Northeastern Journal of Agricultural 20(1991):41-51.

Eastwood, David B., Morgan Gray, and John R. Brooker. "Advertising Food Products: Does a Supermarket Chain Coordinate Its Television, Radio, and Paper Ads?", presented at the annual Meeting of the Southern Agricultural Economics Association, Fort Worth, Texas, February, 1991.

Eastwood, David B., Morgan Gray, and John R. Brooker. "Box-Jenkins Forecasts of Perishable Foods: Is There Potential Consumer Benefit?",Proceedings of the 37th Conference of the American Council on Consumer Interest, M. Halderman (ed.). 1991.

Eastwood, David B., Morgan Gray, and John R. Brooker. "Using Scan Data to Evaluate Advertising: A Suggested Methodology and Preliminary Results." Presented at the Southern Agricultural Economics Meetings, Fort Worth, Texas, February, 1991.

Elam, Emmet M.and Shelby H. Holder. "An evaluation of the Rice Outlook and Situation Price Forecasts." Southern Journal of Agricultural Economics, 17(1985):155-161.

Fair, Ray C., and Robert J. Shiller."Comparing information In Forecasts from Econometric Models." The American Economic Review, 80(1990): 375-388.

Funk, T. F., K. D. Meilke, and H. B. Huff. "Effects of Retail Pricing and Advertising on Fresh Beef Sales." American Journal of Agricultural Economics, 59(1977):533-7.

Gagnon, Jean P., and Osterhaus, Jane T. "Research Note: Effectiveness of Floor Displays on the sales of Retail Products." Journal of Retailing, 61(1985):

Geurts, Michael D. and I.B. Ibrahim."Comparing the Box-Jenkins Approach with the Exponentially Smoothed Forecasting Model Application to Hawaii Tourists." Journal of Marketing Research, 12(1975):182-188.

Grubb, Howard "A Multivariate Time Series Analysis of Some Flour Price Data." Journal of Applied Statistics 41(1992):95-107.

Harris, Kim S., and Raymond M. Leuthold."A Comparison of Alternative Forecasting Techniques for Livestock Prices: A Case Study." North Central Journal of Agricultural Economics 7(1985):40-50.

Harvey, A.C. and P.H.J. Todd "Forecasting Econometric Time Series With Structural and Box-Jenkins Models: A Case Study." Journal of Business and Economic Statistics October, 1983:299-306.

Hayward, Russell A., George K. Criner, and Steven P. Skinner "Apple Price and Production Forecasts for Maine and the United States." Northeastern Journal of Agricultural and Resource Economics 13(1984):268-276

Holdren, B.R. The Structure of a Retail Market and the Market Behavior of Retail Units. Prentice-Hall, Inc., 1960.

Hooley G.J., and N. Wilson " Modeling the effects of advertising: Some methodological issues." Journal of the Marketing Research Society 30:45-58.

Hudson, Michael A., and Oral Capps, Jr. "Forecasting Ex-Vessel Prices for Hard Blue Crabs in the Chesapeake Bay Region: Individual and Composite methods." Journal of the Northeastern Agricultural Econometrics Council 13(1984):111-118.

Jensen, Helen H., and John R. Schroeter."Estimating Retail Beef Demand Using Household Panel Scanner Data." Contract paper for Beef Industry Council of the National Live Stock and Meat Board. 1989.

Jex, C.F. "Short-term modelling of advertising effectiveness: A confirmation of Broadbent's model." Journal of the Market Research Society 27:293-297

Jones, Eugene and Ronald W. Ward "Effectiveness of Generic and Brand Advertising on Fresh and Processed Potato Products." Agribusiness 5(1989):523-536.

Kaylen, Michael S. "A Note on the Forecasting of Turning Points. "North Central Journal of Agricultural Economics 8(19860: 156-158.

Kinnucan, Henry W. and Meenakshi Venkateswaran "Effects of Generic Advertising on Perceptions and Behavior: The Case of Catfish. "Southern Journal of Agricultural Economics 22(1990):1137-144.

Kluyver, Cornelis A., and Roderick J, Brodie. "Advertising- V e r s u s - Marketing mix Carryover Effects."Journal of Business Research, 15(1987):269-87.

Kumar, V. and R. P. Leone. "Measuring the Effect of Retail Store Promotions on Brand and Store Substitution. "Journal of Marketing Research, 25(1988):178-85.

Lattin, James A. and Randolph E. Bucklin. "Reference Effects of Price and Promotion on Brand Choice Behavior." Journal of Marketing Research, 26(1989):299-310.

Leitch, Gordon and J. Ernest Tanner. "Economic Forecast Evaluation: Profits Versus The Conventional Error Measures. "The American Economic Review, 81(1991).

Lesser, W. and J. Smith. "The Accuracy of Supermarket Scanning Data: An Initial Investigation." Journal of Food Distribution Research, 17(1986):69-74.

Levenbach, Hans and James P. Cleary The Modern Forecaster: The Forecasting Process through Data Analysis, Wadsworth, Inc. 1984.

Lin, Biing-Hwan, Mark Herrmann, Tsae-Yun Lin, and Ron C. Mittelhammer."Forecasting the Price of Farmed Atlantic Salmon: AN Integrated Econometric and Time Series Approach." Agribusiness, 5(1989):477-488.

Litterman, Robert B. "A statistical Approach to Economic Forecasting. "Journal of Business & Economic Statistics, 4(1986):1-4.

Liu, Lon-Mu "Sales Forecasting Using Multi-equation Transfer Function Models."Journal of Forecasting, 6(1987):223-238.

Marion, B. W. and F. E. Walker. "Short-Run Predictive Models for Retail Meat Sales." American Journal of Agriculture Economics, 60(1978):667-73.

McIntosh, Christopher S. and David A. Bessler "Forecasting Agricultural Prices Using A Bayesian Composite Approach. "Southern Journal of Agricultural Economics, 20(1988):173-80.

McLaughlin, E. and W. Lesser. "Experimental Price Variability and Consumer Response: Tracking Potato Sales with Scanners." Working Paper, Cornell University 1986.

McNees, Stephen K. "How Large are Economic Forecast Errors? "New England Economic Review, Federal Reserve Bank of Boston. July/ August 1992:25-42.

Menkhaus, D.J., and R.M. Adams."Forecasting Price Movements: An Application of Discriminant Analysis. "Western Journal of Agricultural Economics, 6(1981):229-238.

Moriarty, M. M. "Retail Promotional Effects on Intra- and Inter-brand Sales Performance. "Journal of Marketing, 61(1985):27-47.

Moriarty, M. M. "Feature Advertising-Price Interaction Effects in the Retail Environment." Journal of Marketing, 61(1985).

Nazem, Sufi M. Applied Time Series Analysis for Business and Economic Forecasting, Mackrel Dekker, Inc., 1988.

Olivieria, Ronald A., Carl W. O'Connor and Gary W. Smith. "Short-Run Forecasting Models of Beef Prices." Western Journal of Agricultural Economics, 4(1979):45-55.

Oster, Patrick, Gabrielle Savery, and John Templeman. "Brand Take A Beating." Business Week, July 1993, p. 22.

Park, Timothy "Forecast Evaluation for Multivariate Time-series Models: The U.S. Cattle Market" Western Journal of Agricultural Economics, 15(1990):133-143.

Paul,.P.B., N.L. Sar, and A. Bui-Lan."Forecasting Sawntimber Consumption: The Application of an End-Use Model. "Australian Journal of Agricultural Economics, 33(1989): 219-232.

Pindyck, Robert S., and Daniel L. Rubienfeld. Econometric models and Economic Forecasts, New York:McGraw-Hill, 1981.

Powers, Nicholas J."A Study of Demand Response to Grocery Advertising of Fresh California-Arizona Navel Oranges. "Agribusiness, 5(1989):423-435.

Raju, Jagmohan S. "The Effect of Price Promotions on Variability in Product Category Sales." Marketing Science, 11(1992):207-219.

Raunikar, Robert, and Huang, Chung-Liang. Food Demand Analysis. Iowa State University Press, 1987:14-15.

SAS/ETS User's Guide, Version 6 First edition. SAS Institute Incorporated, 1988.

Schrimper, Ronald A., "Cross-Sectional Data For Demand Analysis."The Economics of Meat Demand, October 1986:74-92.

Shiller, Zachery. "Procter and Gamble Hits Back. "Business Week, July 1993, pp. 20-22.

Spriggs, John "Forecasts of Indiana Monthly Farm Prices Using Univariate Box-Jenkins Analysis and Corn Futures Prices. "North Central Journal of Agricultural Economics, 3(1981):81-87.

Somers, Toni M., Yash P. Gupta and Scott R. Herroitt "Analysis of Cooperative Advertising Expenditures: A Transfer-function Modeling Approach." Journal of Advertising Research, Oct./Nov. 1990:35-47.

Stobaugh, Tobert B. and Phillip L. Townsend "Price Forecasting and Strategic Planning: The Case of Petrochemicals. "Journal of Marketing Research 12(75):19-29.

Supermarket Business. Supermarket Business, (September 1988).

Schwarz, Gideon "Estimating the Dimension of a Model." The Annuls of Statistics 6(1978):461-464.

Tellis, Gerald. "Advertising Exposure, Loyalty, and Brand Purchase: A Two Stage Model of Choice." Journal of Marketing Research, 25(1988): 134-44.

Thayer, Warren. "Measuring the Power of Promotion. "Progressive Grocer, 68(1989):38-46.

Tomek, W.G. "Limits on Price Analysis." American Journal of Agricultural Economics 67(1985):905-15.

Venkateswaran, Meenakshi and Henry W.Kinnucan. "Evaluating Fluid Milk Advertising In Ontario: The Importance of Functional Form." Canadian Journal of Agricultural Economics 38(1990):471-488

Vere, D.T., and G.R. Griffith. "Comparative Forecast Accuracy in the New South Wales Prime Lamb Market." Australian Journal of Agricultural Economics, 34(1990): 103-117

Watson, Mark W. and Lydia M. Pastuszek "Forecasting Commercial Electricity Sales. "Journal of Forecasting 6(1987):117-136.

Walters, Rockney G., and Heikke J. Rinne. "An Empirical Investigation into the Impact of Price Promotions on Retail Store Performance." 62(1986).

Weller, Barry R. "Predicting Small Region Sectoral Responses to Changes In Aggregate Economic Activity: A Time Series Approach." Journal of Forecasting, 9(1990):273-281

Wilkinson, J.B., J. B. Mason, and C. H. Paksoy. "Assessing the Impact of Short-Term Supermarket Strategy Variable."Journal of Marketing Research, 19(1982):72-86.

Wittink, D. R., M. J. Addona, W. J. Hawkes, and J. C. Porter. "The Estimation, Validation, and Use of Promotional Effects Based on Scanner Data." Working Paper, Cornell University (1988).

WU, L.S., N. Ravishanker, and J.R.M. Hosking. "Forecasting for Business Planning:A Case Study of IBM Product Sales." Journal of Forecasting, Vol. 10, 1991:579-595

Vita

Kent Lee Wolfe, the son of Dr. N. Lee and Janet K. Wolfe, was born in Las Vegas, New Mexico on March 10, 1966. Mr. Wolfe grew up in Athens, Georgia and graduated from Cedar Shoals High school in June of 1984. He then enrolled at The University of Georgia, Athens and graduated with a Bachelor of Science degree with a major in Agricultural Economics in August, 1988. He married Heidi Lynn Hunt of Powder Springs, Georgia in August of 1988. He then enrolled in the Department of Agricultural Economics at the University of Tennessee in the Fall of 1988. He had his first son, Jacob Wolfe in December of 1989 and graduated with a Masters of Science degree in Agriculture with a major in Agricultural Economics in August of 1990. He continued his studies in the Department of Agricultural Economics at the University of Tennessee in the Fall of 1990 to pursuing a Doctorate degree. In December of 1992 his second son, Jonathan Wolfe, was born. In May of 1994 he graduated with a Ph.D. in Agriculture with a major in Agricultural Economics. He is employed in Lincoln, Nebraska with the Gallup Organization as a Research Director.