To the Graduate Council:

I am submitting herewith a dissertation written by William Noel Blisard entitled "A Comparison of Alternative Price Forecasting Models for Slaughter Hogs." 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.

Dan L. McLemore

Dan L. McLemore, Major Professor

We have read this dissertation and recommend its acceptance:

Luther H. Keller

Glen D. Whipple

Hui-shyong Chang

Accepted for the Council:

C. W. Minkel

The Graduate School

t

A COMPARISON OF ALTERNATIVE PRICE FORECASTING MODELS FOR SLAUGHTER HOGS

A Dissertation

Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

William Noel Blisard

March 1984

ACKNOWLEDGEMENTS

I would like to thank the Agricultural Economics Department for making my study here at The University of Tennessee, Knoxville, a rewarding one, and also for the financial help which I have received. The members of my committee, Drs. Chang, Whipple, and Keller also deserve a special word of thanks for their cooperation and help on the final draft of this dissertation. A word of thanks also needs to go to Drs. Rawls and Klindt who helped in the formulation of this study.

A heartfelt thanks goes to my advisor Dr. Dan McLemore, who allowed me to pursue further coursework in econometrics, and formulate the research reported on in this dissertation. In addition, Dr. McLemore has given me more than my fair share of his time and patience. Without Dr. McLemore's competent guidance this work would have fallen far short of whatever merit it has.

Finally, I would like to thank my wife, Marcia, who has had to sacrifice these past five and one-half years so I could earn the doctorate degree. Words cannot express my gratitude.

Inevitably, there will be some errors in this work. Hopefully, they will be no worse than typos, and not lapses in the logic of the methodology. However, any error in this work is accepted as my own.

ABSTRACT

The purpose of this study was to develop a price forecasting model capable of accurately predicting the price of slaughter hogs one quarter in advance. Both time series and econometric models were hypothesized, estimated, and tested. In an attempt to improve forecasting accuracy, composite models were developed which utilized the price forecasts from both time series and econometric models as inputs.

Two types of time series models were estimated. The simplest was an ARIMA model, which assumes current price is related to past price plus a random error term. Also estimated was a transfer function which is similar to the ARIMA model, but which includes an input variable.

Three econometric models were hypothesized and estimated. The first was a recursive model which consisted of supply and demand equations separated by a change-in-storage equation. A reduced form model was derived from the recursive system. The final econometric model was a five-equation simultaneous system based upon equations and variables specified in a previous study. The simultaneous system was dropped from consideration due to poor estimation results.

Three composite models were estimated by regressing 30 forecasted prices from two individual models against actual price. The estimated beta coefficients became the weights assigned to each individual forecast.

Each model was evaluated over an eight-quarter out-of-sample period in terms of Theil's U_2 coefficient, root mean square error, and how well it could predict direction of price change. Results indicated that no model was able to predict better than a naive no-price-change model (calculated U_2 coefficient greater than one). Theil's U_2 coefficient ranged from a low of 1.11 for the transfer function to a high of 2.36 for the recursive system. Corresponding root mean square errors were \$6.43 and \$13.66, respectively. The best predictor of the direction of the change in price was the recursive model, which was correct 100 percent of the time.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The hog cycle has been a familiar phenomenon since the beginning of the current century. When hog prices are favorable producers tend to withhold gilts and sows from the slaughter market in order to increase the breeding herd. This action reenforces existing high prices for a short time. However, after a time lag the increased breeding herd leads to increased pork production with accompanying lower prices.

Lower prices then result in a liquidation of the breeding herd which further increases pork production and places additional downward pressure on prices. After liquidation has progressed for a time pork production diminishes and prices recover. The cycle then begins again.

This variation in price is a source of manmade uncertainty for producers (as well as consumers) and results in economic inefficiency. From the producers' standpoint price uncertainty can result in: (1) imperfect combinations of different products, (2) imperfect combinations of resources used to produce a product, and (3) imperfect levels of production [Tweeten].

In general, producers may attempt to cope with price uncertainty by seeking flexibility in product production and in liquidity. Thus, hog producers may forego investments in highly specialized and efficient production facilities, and farm equipment may be rented or custom hired. This can result in inefficient input ratios such as excessive labor relative to capital. This problem can then be aggravated by internal capital rationing. In order to survive in periods of low product prices, excessive cash reserves may be built up which leads to more inefficiency, since these reserves could be invested in productive and profitable inputs in the absence of uncertainty [Tweeten].

The effects of price uncertainty could be reduced if producers hedged in the appropriate futures market. However, farmers have tended to neglect this alternative. This forces the producer to become a speculator in the cash market. It can be said that unhedged producers are always "long" in the cash market.

In reality, even if the future price of slaughter hogs could be accurately predicted many of the above noted inefficiencies would still exist. But these inefficiencies are accentuated by the inability of producers to predict price with accuracy. However, if consumer demand for pork products were constant, then the future price of slaughter hogs would be primarily determined by the quantity of slaughter hogs coming to market.

In any given quarter the number of slaughter hogs on farms comes from two sources: (1) the inventories of slaughter hogs at the beginning of the current quarter, and (2) those hogs which reach some minimum slaughter weight during that quarter. Hogs are usually slaughtered at weights between 180 and 300 pounds. The minimum time for a hog to each 180 pounds is approximately 153 days while the minimum time for a hog to reach 300 pounds is approximately 247 days [Sullivan, et al.]. Hence, all hogs which have reached a minimum weight of 180 pounds at the beginning of a quarter will be slaughtered during that quarter unless they are retained for breeding purposes. Furthermore, if slaughter hogs are sold at an average weight of 220 pounds which requires approximately 180 days, then the number of sows farrowing two quarters ago largely determines the maximum number of hogs that can be slaughtered in the current quarter.

Therefore, given a relatively constant consumer demand for pork, current hog price is largely determined by the collective action of hog producers two quarters ago. Hence, it seems reasonable that some quantitative method could be used to predict the price that producers will receive for their product in the short run. A good starting point for attempting to predict the price of slaughter hogs would be a one-quarter-ahead forecast.

Such a forecast would benefit those hog producers who buy feeder pigs, since the time period required to fatten the pigs to market weight is approximately four months. Since the price of a feeder pig should tend to be equal to the price of a slaughter hog minus the cost of feeding the pig to market weight, feeder pig producers would also benefit from a one-quarter-ahead forecast. Finally, meat packers could also benefit by making short run adjustments if they knew what price they were going to pay for slaughter hogs in the next quarter.

I. PROBLEM OF THE STUDY

In general, prices can be forecast either by an econometric model or by time series analysis. Econometric models can be either a single equation or a system of equations. Time series analysis is usually conducted using univariate autoregressive integrated moving average (ARIMA) models in which the current observation is generated by a weighted average of past observation, plus a weighted average of random disturbances. Recent developments have expanded the scope of the time series models so that multi-input models (transfer functions) and simultaneous systems can be built as well.

Economists have tended to work with econometric models almost exclusively. The justification for this is that economic theory can be used to determine which variables should be included to the equation or system of equations. In this way, hypotheses can be derived from theory and tested. In addition, the hog-pork sector can be represented mathematically as a system of simultaneous equations. Modeling the pork industry with a system of statistical equations allows the investigator to explore the economic relationships that exist within the sector. In contrast, time series analysis has been called naive because the univariate ARIMA system utilizes only the past history of the variable in order to estimate the parameters of the model.

As noted, recent developments in time series analysis have enabled investigators to build multivariate time series models which can be based upon hypotheses derived from economic theory. In addition, it has been shown in the literature that time series models may be regarded as the "reduced forms" of simultaneous econometric models [Zellner and Palm]. In the macroeconomic area Sims has been particularly forceful in suggesting that much of macro theory which utilizes future expectations of endogenous and exogenous variables should be based upon multivariate ARIMA models.

However, alternative methods of forecasting are rarely in agreement, and it is possible that the most accurate method during one time period may not be the most accurate in other time periods. Naturally hog producers and meat packers would like to know which method of forecasting is best. But this may imply selecting one method to the exclusion of others. Yet alternative forecasts may contain valuable pieces of information not included in the method selected.

Therefore, hog producers and meat packers need to be concerned about the relative accuracy of forecasts that can be made from econometric and time series models, as well as the accuracy of price forecasts from single equation versus multiple equation systems. In addition, model builders and forecast users need to consider what, if any, benefits might be derived by combining time series and econometric models into composite forecasts.

II. REVIEW OF LITERATURE

This review of the literature concentrates on studies which have attempted to model hog prices at the national level in the United States over the last 15 years. Studies which were concerned with modeling quantity as the dependent variable are thus excluded, as are studies which were regional in nature, or which attempted to model both the United States and the Canadian markets.

In 1970 Hayenga and Hacklander build a five-behavioral-equation model of the live cattle and hog markets which was estimated by two-stage least squares. Even though price was the dependent variable in the demand equations for both cattle and hogs, neither was used for predictive purposes. Possibly due to model misspecification the cross-flexibilities of demand for beef and pork had unexpected opposite signs indicating complementarity.

During the same year Myers, Havlicek, and Henderson developed an eight-behavioral-equation system to describe the structure of the hog-pork industry in the U.S. covering the period 1949 through 1966. This model was also estimated by two-stage least squares. The reduced form equations from the first stage and the structural equations from the second stage were used to predict the monthly values of the ten endogenous variables for an 18-month period. Based on Theil's U_2 coefficient, the results indicated that in most cases a no-change extrapolation of the current price would have forecasted as well and in some cases better than the first stage or structural equations beyond a six-month period.

A rather well-known study also was completed in 1970 by Leuthold, MacCormick, Schmitz and Watts. Two models were built and compared which attempted to forecast daily hog prices in selected terminal markets. Due to the nature of the meat packing industry the econometric demand and supply equations were defined as a recursive model of the cobweb type. Thus, quantity supplied was first determined through the supply function, and then this quantity was sold in the market at a price determined by the demand function. In attempting to fit price and quantity ARIMA models, the price equation was found to be best represented by a random walk process, while the quantity supplied model had both moving average and autoregressive components.

Forecasts were made ex post meaning that true values of the exogenous variables were utilized in the econometric model to predict endogenous variables over the period used to fit the model. The justification for this is that if the exogenous variables are estimated, then the observed forecasting errors of the predicting model may be due to errors in the estimates of the predetermined variables as well as to incorrect specification of the prediction model. By using the true values of the exogenous variables, any errors in the forecast are attributable to the model itself. According to Theil's U_1 coefficient the econometric model predicted slightly better than the ARIMA model.

In 1972 Walter Myers expanded upon the work of Hayenga and Hacklander. Myers used the independent variables which they had found to have a high correlation with swine prices, and added to them variables to account for cycles of hog slaughter and price which were derived from a spectral analysis. A recursive model was then specified so that lagged independent variables could be used thereby resulting in a model capable of making ex ante forecasts (forecasts beyond the period of fit). By attempting to keep the equations within "manageable proportions" and yet produce useful information, severable desirable statistical properties were sacrificed or fell outside the bounds of conventional acceptability. In terms of forecast accuracy only the number of turning points missed was reported.

Ladd and Karg published a study in 1973 in which a complete econometric model was developed of the beef and pork marketing sector. Quadratic programming was used to determine quarterly and annual levels of farm marketings of hogs and cattle which would maximize the annual cash receipts of farmers. Price dependent demand functions were estimated for retail level prices but no forecasts were made since the main purpose of the study was to determine the effects of factor prices and labor productivity on marketing margins.

Also in 1973, Purcell and Elrod published a study that focused on trends and variations in prices received for hogs. Using a single equation, producer prices were hypothesized to be a function of the supply of pork, the wholesale price index, the unemployment rate, and population. Pork production was found to increase at an annual rate about equal to population growth. Thus, the increase in demand for pork on a per capita basis was reflected in an uptrend in prices. However, monthly hog price was found to be highly responsive to changes in hog slaughter.

In 1976, Roy, Foote, and Craven developed a simultaneous equation model which forecasted hog prices on a quarterly basis. The simultaneous system determined consumption, retail price of pork, live hog price, and the end-of-period storage level. Each quarter of the year was modeled separately via three-stage least squares. The authors attempted to use the same model on a monthly basis but obtained poor results. In addition Theil's U_2 coefficient was miscalculated so that the seemingly good predictive qualities of the quarterly model are questionable.

Other Relevant Literature

Three Ph.D. dissertations have dealt with the comparative aspects of price forecasting. The earliest of these was by Teigen in 1973. This study sought to compare the relative performance of four different methods of generating forecasts for cattle prices. Econometric, trend, and price difference models as well as futures market prices were compared as to their predictive powers. The models were evaluated over the period 1967-70 using monthly and quarterly data to generate one-period-ahead forecasts. The general conclusion was that the best one-period-ahead forecast consisted of projecting either the current cash price or the corresponding future price. This study therefore suggests that a simple naive forecasting method provides better results than more sophisticated methods.

A study by Edwards compared only econometric models for those commodities which were judged to be important to Georgia agriculture [1977]. Econometric models were gathered from a number of sources and re-estimated using current data. The emphasis in this study was to determine which econometric model predicted best for a particular commodity and how price predictions could be relayed on to farmers in an effective manner. The question of relative accuracy between naive, econometric and time series models did not arise.

A study by Bourke compared nine different models utilized for predicting the price of manufacturing grade beef [1978]. Included were a naive model, single and double moving averages, single and double exponential smoothing, an ARIMA model, single and simultaneous equation econometric models, and finally a method which combined forecasts. Except for the moving average and exponential models, the models were gathered from published sources and re-estimated using current data. Ex post forecasts were made over a one-year period.

Four tests were applied to determine each model's forecasting accuracy. Unfortunately each test resulted in a different ranking of the models, so the only conclusion drawn was that defining an acceptable forecast is difficult since acceptability depends on the purpose of the forecast. Bourke noted that the navie model came close to the sophisticated methods for quarterly and monthly forecasts. He also felt that the ARIMA model might prove equal to or better than the econometric model in ex ante forecasts.

In 1979 Hoffman completed a Ph.D. study which sought to improve the short-run livestock forecasting capabilities of the U.S.D.A. Unlike the other three dissertations discussed above, this study developed one comprehensive model for the livestock sector. A recursive system was specified because of the "short time unit and because of the existence of significant lags in the production-marketing process." The model consisted of 54 behavioral equations and the Cochrane-Orcutt procedure of estimation was used where serial correlation was evident in the ordinary least squares equations. Ex post forecasts indicated the absolute average percent errors on the 50 primary endogenous variables ranged from a high of about 34 percent for nonfed steer and heifer slaughter to a low of 1.6 percent for retail pork prices.

In a 1979 experiment station bulletin, Professors Bessler and Brandt attempted to improve forecasts by combining alternative forecast methods. In what the authors described as an attempt to keep the forecasting method simple, a single equation econometric model, an univariate ARIMA model, and expert opinion were utilized in order to predict hog, beef, and broiler prices one-period-ahead. Each model was evaluated by calculating the mean squared error, mean forecast error, and the mean absolute forecast error. A second group of performance indicators included tracking measures which indicated the number of turning points missed or falsely predicted compared to the number correctly forecast.

In an effort to improve the accuracy of the forecasts both the ARIMA and econometric models were combined to produce a composite forecast. The theory for combining forecasts has been developed in the literature and a rigorous statement can be found in Granger and Newbolds' Forecasting Economic Time Series. Bessler and Brandt analyzed three methods of combining forecasts: an adaptive process, a method based on minimum variance criterion, and a third method which gave equal weight to each model. Performance tests were similar to those discussed above for the individual forecasts and the composite forecasts were compared to those of the individual models.

The conclusions drawn by the authors were that for each commodity, composite forecasts were better than or almost as good as the best of the individual techniques. Furthermore, if one did not know which set of forecasts to use based on a priori information, then the risk of making a large error could be reduced by selecting a composite model.

III. OBJECTIVES OF THE STUDY

The specific goal of this study was to develop a model capable of accurately predicting, one quarter in advance, the price paid by packers for slaughter hogs in the United States. To reach this goal, both econometric and time series models were built. In an attempt to develop more accurate forecasts, composite models were built which utilized as inputs the forecasts from one econometric and one time series model. Comparisons between models were made by calculating the root mean square error (RMSE) and Theil's U_2 coefficient[1] and by

---

[1] $$\text{RMSE} = \sqrt{\frac{\Sigma(\hat{PH}_t - PH_t)^2}{N}}$$

$$\text{Theil's } U_2 = \sqrt{\frac{\Sigma(\hat{PH}_t - PH_t)^2}{\Sigma(PH_t - PH_{t-1})^2}}$$

where:

$\hat{PH}$ = predicted price of hogs.

PH = actual price of hogs.

N = number of forecasts.

determining how well each model predicted the direction of price change.

CHAPTER II

MODELS AND METHODOLOGY

I. THEORY UNDERLYING THE ECONOMETRIC MODELS

By definition econometric models are statistical equations containing variables suggested by economic theory. Since the objective of this study is to predict the price of slaughter hogs, and since price and quantity traded are determined by the intersection of a demand and a supply curve, a brief review of consumer demand and producer supply theory is appropriate.

Individual and Aggregate Demand

The quantity of any commodity that a consumer purchases depends upon a set of factors which include the individual's tastes and preferences, the price of the commodity in question, the prices of other commodities which are substitutes or complements, income, and perhaps other factors. If income and other prices are held constant, more of the commodity in question will be bought when price falls and less will be taken when price rises. This inverse or negative relationship is called the demand schedule for the consumer.

The individual consumer's demand schedule will shift with changes in his income, with changes in the prices of substitute and competing goods, and with changes in the consumer's tastes and preferences. Since each consumer may have particular tastes, and since incomes are not the same for all consumers, each individual may have a unique demand schedule for each commodity. Some individuals may greatly adjust their purchases in response to a price change, while others will make very little adjustment. Some consumers may buy large quantities at high prices although others may purchase little or none even at relatively low prices. Thus, each individual demand schedule may differ as to its shape, slope, and location from the origin.

If a consumer buys any of a given commodity, economic theory assumes that the quantity purchased will increase, at least slightly, as price decreases, all other things being equal. This assumption holds true for the sum or aggregate of the consumers in any given market. Thus, the aggregate or market demand schedule represents the sum of all the quantities that will be purchased by all consumers at each level of price. Very high prices will result in relatively small amounts taken by consumers, whereas relatively large amounts will be purchased at low prices. Assuming that there are many consumers in any given market, and that each consumer will take different amounts at different prices, the aggregate demand schedule can be expected to be a continuous curve from a relatively high price to a relatively low price.

Each commodity will have its characteristic demand schedule or curve in a particular market, and each curve will vary in shape, slope and location from the origin. Although it is assumed that the market demand for any commodity is a price-quantity relationship with incomes, population, other prices, and tastes held constant, the curve can shift with changes in any of these variables. Since these variables are changing through time, the demand schedules for all commodities are thus interrelated.

This interrelationship between commodities is often depicted for any two goods by an indifferent curve map. These curves show various combinations of the two goods which yield equal levels of satisfaction to the consumer. Since economists assume that utility declines as consumers get more of any particular good, these indifference curves are always negatively sloped and convex. Furthermore, economists assume that more of a good is preferred to less, so that indifference curves further away from the origin are preferred. Theoretically, an individual's indifference curve mapping will yield his demand schedule, and these demand schedules can then be summed to obtain an aggregate demand schedule [Bressler and King].

Individual and Aggregate Supply

The quantity supplied of a commodity is taken as a function of the price of the commodity in question, the price of inputs, the price of other outputs which are either complementary or competitive in production, technology, the number of other firms in the industry and perhaps other factors. From this starting point, production and supply theory parallels consumption and demand theory with a few important exceptions. Instead of a consumer's indifference map, an isoquant map shows the technology of factor substitution in the production of output. While economists assume that the consumer's utility cannot be measured, output usually can be measured in objective terms, so that isoquants have cardinal as well as ordinal values.

Constrained by the limits of technology the production problem may be thought of as twofold: first, with a given set of factor prices, how can inputs be combined to minimize the cost of a given output; and second, with a given product price, what level of output will maximize the positive difference between total revenue and total cost.

Under competetive market conditions, the final adjustment for the firm is conveniently summarized in terms of the marginal cost curve, with the profit maximizing output found where marginal cost equals product price. Like individual demand schedules, individual marginal cost curves may be different with respect to both shape and position from the origin. It is expected that these curves will increase with increases in output because of intensification on fixed factors and the principle of diminishing marginal productivity. In all cases marginal cost curves indicate the output response of individual firms to changes in product prices in the short run. By starting at a low price and increasing price gradually, it would be theoretically possible to sum together the output of all firms and obtain the aggregate market supply curve.

This curve would be positively inclined for two reasons: (1) in the short run, higher prices will bring greater output from firms with positively inclined marginal cost curves, and (2) in the long run new firms will enter the industry, attracting resources from other employments. The exact shape of the curve will depend upon factors noted earlier such as technology, factor prices, the prices of alternative products, and the number of firms in the industry [Bressler and King].

Price Equilibrium

In any market for a particular commodity, equilibrium is achieved when a price is established at which quantities offered for sale exactly equal the quantities demanded by purchasers. Any price higher than the equilibrium price will bring forth more quantity supplied from producers. Competition among sellers would make this an unstable situation and force price down to the equilibrium level. Conversely, any price below the equilibrium level will bring forth less output. Competition among buyers would bid price and quantity up to the equilibrium level.

The above is based upon the assumption that economic adjustments are made instantaneously, without time lags or frictions. This, of course, is not true of the real world. Consumers may take time to learn about and to adjust to changes in quantity and price situations. Lags are even more important on the supply side. Traditionally three supply curves exist through time. The first is the market curve which occurs with given supplies and where no change in production is possible. The second curve is the short run curve, where output may be changed by movement along the firm's marginal cost curve. The third supply curve is called the long run or planning curve, and this theoretically applies when all resources are variable.

II. PROPOSED RECURSIVE MODEL

The first model proposed to predict the price of slaughter hogs was a recursive system. Any model is recursive (rather than simultaneous) if each of the endogenous variables in the model can be determined sequentially. For biological reasons, it takes approximately six months from the time a sow is farrowed until a market hog is sold. Thus, current hog production is determined by the actions of producers at least six months in the past. This quantity produced then goes on the market and interacts with a current demand curve to determine price. Hence, the endogenous variables are determined sequentially.

In an attempt to keep the recursive model as simple as possible and yet include relevant variables, a three-equation system was specified. Economic theory hypothesizes both a demand and a supply curve, but study of the hog-pork industry indicated that cold storage holding was also a major factor in how much pork would be available in any given quarter. Hence, the recursive system consists of supply and demand equations separated by a cold storage equation, which can add to or subtract from current production to determine quantity supplied.

Quantity of Slaughter Hogs Supplied Equation

As noted previously, economic theory suggests that quantity supplied is a function of product price, the cost of inputs, technology, the price of competing products, and the number of firms in the industry. In hog production, technology has increased rapidly over the last few decades resulting in fewer producers supplying more hogs. In addition, hog production has become specialized to the extent that producers cannot readily switch from one livestock enterprise to another in the short run.

Thus, the quantity of slaughter hogs supplied (carcass weight in millions of pounds) was hypothesized to be a function of the number of sows farrowing two quarters previously. Furthermore, the number of sows farrowing has traditionally been taken as a function of the hog-corn price ratio. However, since soybean meal and corn are the predominate ingredients in feed, it was decided that a hog-feed price ratio would be more appropriate. The price of feed was assumed to be .88 times the price of corn (dollars per ton) plus .12 times the price of 44 percent of soybean meal (dollars per ton). The price of hogs and the price of feed were included as separate variables rather than as a single ratio since recent studies by Hobbs and Herndon and by Meilke have shown this to be a superior approach.

Due to the amount of specialization in the production of hogs, the price (or profitability) of competing enterprises was not considered in the equation. In addition, this study did not include technology as a supply equation variable although some studies have attempted to capture changes in technology by including a time variable in the equation. This approach was not used in this study because technological change normally results in changes in the structural parameters. But in regression analysis the structural parameters are assumed to be constant over the period of estimation. Including time as a variable in the equation cannot account for structural changes other than those that shift the supply curve.

Since there is a time delay between changes in the independent variables (price of hogs and price of feed) and the response of the dependent variable (quantity of hogs supplied) due to biological and planning reasons, it was hypothesized that a distributed lag was appropriate to relate variation in the dependent variable to variation in the independent variable. That is, in deciding upon how many sows to farrow, hog producers may look at hog prices and feed costs over some length of time in the past and present. A favorable response would then result in more sows farrowed, and thus a greater quantity of hogs produced and slaughtered approximately six months later.

A priori it would be difficult to hypothesize what weights should be assigned to past values of hog price and the cost of feed. However, a polynomial distributed lag can be estimated in which the weights assigned to each past value of the independent variable are determined by the data.

The polynomial lag model assumes that the lag weights can be specified by a continuous function, which in turn can be approximated by evaluating a polynomial function at the appropriate discrete points in time. The researcher specifies the degree of the polynomial and the number of lags [Pindyck and Rubinfeld].

Two recent studies have shown that good statistical results could be obtained by using this method to model the effect of the price of hogs and the price of feed [Hobbs and Herndon; Meilke]. In this study it was proposed to fit a second order polynomial lagged eight quarters in the past. A second degree or quadratic form would result in a concave or convex response function for both independent variables. The weights assigned to lags one through eight would be determined by the data. Since the hog cycle is typically assumed to be four years in length, it was hypothesized that two years of data (eight quarters) would completely capture the response of producers to changing hog and feed prices.

Since hog producers respond to the price of hogs relative to production and living costs, the price of hogs was deflated by the index of prices paid by farmers. In addition, the real cost of feed is its price relative to the price received for hogs, so the price of feed was divided by the price of hogs.

Quarterly dummy variables were included to account for seasonal factors in production. In each equation throughout this study the fourth quarter dummy variable was dropped so that the resulting matrix would be nonsingular. Traditionally, production has been greatest in the first and fourth quarters of the year, and least in the second and third quarters. Thus, inclusion of the seasonal dummy variables allowed shifts in the intercept term corresponding to shifts in seasonal supply.

Cold Storage Equation

The change in cold storage equation (thousands of pounds in 48 states) during any quarter was hypothesized to be a function of the current supply of fresh product (millions of pounds), the beginning level of cold storage (thousands of pounds in 48 states), and the current retail price of pork cuts. In this equation those variables pertaining to quantities supplied were converted to a per capita basis in order to account for population in the equation. Thus, the current production of fresh pork and the beginning level of stocks were specified on a per capita basis. Since inflation affects nominal prices, and since the intent of this model was to capture responses to real changes in the various variables, the retail price of pork cuts was deflated by the consumer price index.

In general, when the quarterly production of fresh product increases packers increase cold storage, especially if they have expectations of higher prices in the future. Hence, when quantity supplied has been large in the first and fourth quarters product has moved into storage. Thus, the sign of the current production of slaughter hogs should be positive. Conversely, if the beginning level of storage stocks were large one would expect a greater out-movement of stored product; hence its expected sign was negative.

The current retail price of pork was hypothesized to affect the change in cold storage holdings. However, since this model attempted to predict the price of slaughter hogs, the current period retail price of pork would not be available when needed. As a proxy, the retail price of pork cuts (weighted average price of retail cuts from pork carcasses) lagged one quarter was specified, since this variable should be highly correlated with the current retail price of pork cuts. It was hypothesized that if retail price were rising (consumer demand or producer supply shifts) then stored product would be expected to move out of storage. Hence, recently rising prices should result in reduced storage, so that the retail price of selected pork cuts lagged one quarter should have a negative sign.

Seasonal dummy variables were added to the cold storage equation. If higher prices are traditionally realized in the summer when fresh quantity is least, then stored product should move to market. Conversely, when fresh quantity is large and prices low, fresh product would be expected to move into storage. Thus, more fresh pork moves into storage during the first and fourth quarters with more frozen pork moving to market during the second and third quarters.

Price of Slaughter Hogs Equation

The price of hogs equation (all barrows and gilts in seven markets per cwt.) was hypothesized to be a function of the current quantity of fresh pork (carcass weight in millions of pounds per capita), change in frozen pork storage (in thousands of pounds per capita in 48 states), income (consumer disposable income in billions of dollars per capita), cattle on feed (quarterly estimates in seven states per capita), and the number of fully utilized work days per quarter. Since prices are affected by inflation and since this study is concerned with the real price offered by packers for slaughter hogs, the price of hogs was deflated by the producer price index. In order to account for the effects of changes in the population over time, income and quantity variables were divided by population.

The quantity of slaughter hogs coming to market should have a negative sign. Likewise, a positive increase in the change in cold storage variable would reduce the current quantity of pork on the market. Thus its expected sign was negative.

Since the demand for slaughter hogs is a derived demand, it was expected that variables which affect the retail demand for pork would also affect packer demand. Among these variables was per capita income. As per capita income increases, hog price would be expected to rise. Thus the expected sign of per capita income was positive.

The price that packers can pay for hogs should also be influenced by the quantity of substitutes and/or complements coming to market at the same time. Beef and chicken are the main substitutes that compete with pork for the consumer's dollars. Following Hayenga and Hacklander's finding that chicken had "little impact on hog or cattle monthly prices" this variable was not included in the model. However, beef competes directly with pork and was included in the model by using cattle-on-feed lagged one quarter as a proxy variable for beef production.

When the production of a competing product increases and its price declines, consumers substitute away from the relatively more expensive product to the relatively less expensive commodity. Thus, the expected sign of the lagged cattle-on-feed variable was negative.

Finally, the amount of slaughter capacity was hypothesized to have a positive affect on the price packers are willing to pay for a quantity of raw product. Since slaughter capacity is difficult to ascertain, it was hypothesized that the number of fully utilized work days each quarter would provide a barometer of the capacity pressure to which packers respond [Hayenga and Hacklander].

Again, seasonal dummy variables were specified for the demand equation, and the fourth quarter was dropped so that the resulting matrix would remain nonsingular. Traditionally, it would be expected that consumers would demand more pork products during the second and third quarters, and less during the first and fourth quarters of the year.

Thus, the proposed recursive model is:

$$HS_t = a_o + \sum_{i=1}^{8} \delta_1 \left(\frac{PH}{IPPF}\right)_{t-1} + \sum_{i=1}^{8} \Theta_i \left(\frac{PF}{PH}\right)_{t-1} + \phi_{11}D_1 + \phi_{12}D_2 + \phi_{13}D_3 + \varepsilon_{1t}$$

$$\left(\frac{\Delta S}{POP}\right)_t = \gamma_0 + \gamma_1 \left(\frac{HS}{POP}\right)_t + \gamma_2 \left(\frac{S}{POP}\right)_{t-1} + \gamma_3 \left(\frac{RPP}{CPI}\right)_{t-1} + \phi_{21}D_1$$

$$+ \phi_{22}D_2 + \phi_{23}D_3 + \varepsilon_{2t}$$

$$\left(\frac{PH}{PPI}\right)_t = \alpha_0 + \alpha_1 \left(\frac{HS}{POP}\right)_t + \alpha_2 \left(\frac{\Delta S}{POP}\right)_t + \alpha_3 \left(\frac{IN}{POP}\right)_t + \alpha_4 \left(\frac{COF}{POP}\right)_t + \alpha_5 \, WD_t$$

$$+ \phi_{31}D_1 + \phi_{32}D_2 + \phi_{33}D_3 + \varepsilon_{3t}$$

where:

HS_t = Commercial quantity supplied of slaughter hogs on a carcass weight basis in millions of pounds.

D_1, D_2, D_3 = Seasonal dummy variables (D_1 for the first quarter, etc.).

$\delta_i = \lambda_1 + \lambda_2 (i) + \lambda_3 (i^2)$; $\Theta_i = U_1 + U_2 (i) + U_3 (i^2)$ which are polynomial lags.

$\left(\frac{PH}{IPPF}\right)_{t-i}$ = Average quarterly price of all barrows and gilts per cwt. in seven markets, deflated by the index of prices paid by farmers.

$\left(\frac{PF}{PH}\right)_{t-i}$ = Price of feed (.88 times the price of corn per ton plus .12 times the price of 44 percent soybean meal per ton) divided by the average price of all barrows and gilts in seven markets.

$\left(\frac{\Delta S}{POP}\right)_t$ = Change in the quantity of pork in cold storage in 48 states in thousands of pounds per capita.

$\left(\frac{HS}{POP}\right)_t$ = Commercial quantity supplied of slaughter hogs, carcass weight in millions of pounds per capita.

$\left(\frac{S}{POP}\right)_{t-1}$ = Quantity of pork in cold storage at the beginning of the current quarter expressed on a per capita basis in 48 states in thousands of pounds per capita.

$\left(\frac{RPP}{CPI}\right)_{t-1}$ = Estimated weighted-average price of retail pork cuts per pound deflated by the consumer price index.

$\left(\frac{PH}{PPI}\right)_t$ = Average price of all barrows and gilts per cwt. in seven markets deflated by the producer price index.

$\left(\frac{IN}{POP}\right)_t$ = Disposable income in billions of dollars per capita.

$\left(\frac{COF}{POP}\right)_{t-1}$ = Estimated number of cattle on feed in seven states expressed on a per capita basis.

WD_t = Number of fully utilized work days per quarter.

t = Time period.

III. PROPOSED REDUCED FORM MODEL

The single equation model was deduced from the recursive system. By substituting (1) the per capita distributed lags from the quantity supplied equation for per capita quantity supplied in the price equation of the recursive model, and (2) the beginning-of-quarter level of per capita cold storage and the deflated retail price of pork cuts lagged one quarter for the change in per capita cold storage in the price equation, the single equation price model accounts for both per capita quantity supplied and the change in cold storage stocks. The other original variables from the recursive price equation were retained to form a complete single equation model.

Hence, the deflated price of slaughter hogs is expressed as a function of all exogenous variables of the recursive system. Since this equation is a combination of the supply, demand, and change in storage equations, the expected signs of the coefficients cannot be determined *a priori*. Stated another way, the coefficients are multipliers and thus measure the direct and indirect effects on the dependent variable. Thus, each individual coefficient on the right-hand-side represents the long run effect upon the deflated price variable. The reduced form model has the advantage of being somewhat simpler, but it is structurally less appealing.

Thus, the proposed model is:

$$\left(\frac{PH}{PPI}\right) = a_0 + \sum_{i=1}^{8} \delta_i \left(\frac{PH/IPPF}{POP}\right)_{t-i} + \sum_{i=1}^{8} \phi_1 \left(\frac{PF/PH}{POP}\right)_{t-i} + a_2 \left(\frac{COF}{POP}\right)_{t-1}$$

$$+ a_3 \left(\frac{S}{POP}\right)_{t-1} + a_4 \left(\frac{RPP}{CPI}\right)_{t-1} + a_5 \left(\frac{IN}{POP}\right)_t + a_6 WD_t$$

$$+ \phi_1 D_1 + \phi_2 D_2 + \phi_3 D_3 + \varepsilon_t$$

where all variables are defined in the recursive system except that:

$\left(\frac{PH/IPPF}{POP}\right)_{t-i}$ = The average price of all barrows and gilts per cwt. in seven markets deflated by the index of prices paid by farmers, all of which is expressed on a per capita basis.

$\left(\frac{PF/PH}{POP}\right)_{t-i}$ = The price of feed per ton deflated by the price of slaughter hogs, expressed on a per capita basis.

IV. SIMULTANEOUS EQUATION MODEL

The argument has been made that a recursive system is the appropriate model for the hog-pork sector since the current production of slaughter hogs is the end result of past actions and decisions of producers. Thus, quantity supplied is first determined, then this quantity interacts with a demand curve to determine price. Hence, the supply and demand equations are solved sequentially.

However, the case could be made that current price and quantity are determined simultaneously. For example, if price had been trending downward for several quarters and producers expected the trend to continue during the current quarter, they could begin liquidating their breeding herds, thereby increasing the quantity produced and simultaneously forcing price down further.

Since the objective of this study was to develop the best possible price forecasting model, a simultaneous equation system was specified and estimated. Based upon economic theory and a knowledge of the beef and pork industries, a five-equation system was specified. Specifically, the system consists of two supply equations, one for beef and the other for pork, a change in pork cold storage equation, and finally two demand equations to predict the price of cattle and hogs. Each equation in the system is overidentified except the demand equation for hogs which is just identified. This model is based upon an earlier study by Hayenga and Hacklander.

Quantity Supplied Equations

The supply equations for cattle and hogs relate quarterly commercial slaughter (carcass weight in millions of pounds) to quarterly U.S.D.A. estimates of feedlot inventories which were expected to be marketed during the quarter. Feedlot inventory for cattle (thousands of head) was based upon steers 700 to 1100 pounds and heifers 500 to 900 pounds in seven states. The hog inventory variable (thousands of head) was for market hogs 60 to 219 pounds in ten states. The expected sign of each inventory variable was positive.

Current commodity price deflated by the producer price index was also included in each supply equation. Normally the price variable would be deflated by the index of prices paid by farmers, but since the system is simultaneous, and since the predicted price of slaughter hogs (in the demand equation) would be deflated by the producer price index, that index was also used in this equation.

The inventory variable itself should be a function of past costs and returns for each respective enterprise. Thus, the current inventory variable was seen as the end result of producer adjustments to changes in past product prices and costs. Current product price will thus influence whether inventories are held back and fed or sent to market. Admittedly this assumption is more relevant for cattle than hog producers, since hogs reaching a minimum weight at the beginning of a quarter (approximately 180 pounds) would have to be marketed during that quarter unless retained for breeding purposes.

As in the recursive system it was hypothesized that livestock producers are specialized to such an extent that in the short run they cannot switch from one livestock enterprise to another. Thus, the price of the competing enterprise was not included in the supply equation. Likewise a proxy variable for technology was not included for the same reasons stated for the proposed recursive model.

Quarterly dummy variables were included to account for seasonal factors in production. As in all other equations, the fourth quarter dummy variable was dropped so that the resulting matrix would be nonsingular. In general, production would be expected to be greatest in the first and fourth quarters of the year for hogs, and during the second and third quarters for cattle. The inclusion of the seasonal dummy variables allows shifts in the intercept term corresponding to shifts in seasonal supply.

Cold Storage Equation

As in the recursive system, the change in the pork cold storage equation (48 states in millions of pounds) was hypothesized to be a function of the current price of retail pork cuts, the current production of fresh pork, and the beginning level of cold storage stocks. Since the simultaneous system would predict the price of slaughter hogs, and since this price and the retail price of pork cuts are highly correlated, current slaughter hog price was used as a proxy for the retail price. Thus, it was hypothesized that as current hog slaughter price (price of all barrows and gilts in seven markets) rises, more product would move out of storage. Hence, its expected sign was negative. In addition, this variable was deflated by the producer price index since this is an endogenous variable which is to be determined by the system.

Again it was hypothesized that if the current quantity of fresh production (carcass weight in millions of pounds) were large then fresh product would flow into storage, and thus its expected sign was positive. Conversely, if the beginning level of storage stocks (48 states in millions of pounds) were large then frozen pork would be expected to flow to market in order to reduce inventory. Hence, its expected sign was negative. Both variables were expressed on a per capita basis in order to account for population in this equation.

Seasonal dummy variables were included in the equation to account for changes in production and consumption. It was expected that fresh product would move into storage during the first and fourth quarters when supplies would be relatively great, and that frozen pork would move out of storage during the second and third quarters.

Demand Equations for Choice Steers and Slaughter Hogs

In the cattle demand equation the price of 900 to 1,100 pound choice steers at Omaha, deflated by the producer price index, was the dependent variable. Likewise, the quarterly average price of all barrows and gilts in seven markets deflated by the producer price index was the dependent variable in the hog demand equation.

For both demand equations, the price paid by packers was hypothesized to be influenced by the per capita quantity produced, per capita storage, as well as the supply of competing imported products. Study of various U.S.D.A. Livestock and Poultry Outlook and Situation Reports indicated that beef product inventories as well as beef import levels rarely exceeded 1 percent of total commercial production in any given quarter. Consequently these variables were not included in the model.

However, since changes in pork cold storage are quite large, and since it was included as an equation in the simultaneous system, it therefore became a variable in the pork demand equation. Its expected sign was positive.

As in the recursive system, slaughter capacity has been included in both demand equations by using the total number of working days in each quarter as a proxy variable. It was theorized that this variable would provide a barometer of the capacity pressure to which packers respond. Its expected sign was positive.

Beef and pork are considered to be major competitors for the consumer's expenditure. Thus, the per capita quantity produced of the competitive product was included as a variable in each demand equation. The expected sign was negative in both cases. Proposed by Hayenga and Hacklander and accepted in this study was the proposition that poultry would have a relatively small impact upon the consumer's decision to purchase beef or pork. Therefore, this variable was excluded from the model.

Finally, as suggested by economic theory, the level of income (disposable personal income in billions of dollars) was included on a per capita basis. Economic theory suggests that as income increases the price of all commodities should rise, other things being equal. Thus, the expected sign of this variable was positive.

Since the demand for both pork and beef may shift from season to season dummy variables were specified to allow the intercept term to shift over the course of a year. In general, more meat products are demanded by consumers during the warmer months of the year when outdoor activity is greatest. Hence, demand for meat would be expected to be greatest during the second and third quarters, and less the first and fourth quarters.

Thus, the proposed model is:

$$BS_t = a_0 + a_1 \left(\frac{PC}{PPI}\right)_t + a_2\, IC_{t-1} + \phi_{11}D_1 + \phi_{12}D_2 + \phi_{13}D_3 + \varepsilon_{1t}$$

$$HS_t = b_0 + b_1 \left(\frac{PH}{PPI}\right)_t + b_2\, IH_{t-1} + \phi_{21}D_1 + \phi_{22}D_2 + \phi_{23}D_3 + \varepsilon_{2t}$$

$$\left(\frac{\Delta S}{POP}\right)_t = c_0 + c_1 \left(\frac{PH}{PPI}\right)_t + c_2 \left(\frac{HS}{POP}\right)_t + c_3 \left(\frac{S}{POP}\right)_{t-1} + \phi_{31}D_1 + \phi_{32}D_2 + \phi_{33}D_3 + \varepsilon_{3t}$$

$$\left(\frac{PC}{PPI}\right)_t = \alpha_0 + \alpha_1 \left(\frac{BS}{POP}\right)_t + \alpha_2 \left(\frac{HS}{POP}\right)_t + \alpha_3 \left(\frac{IN}{POP}\right)_t + \alpha_4\, WD_t + \phi_{41}D_1 + \phi_{42}D_2 + \phi_{43}D_3 + \varepsilon_{4t}$$

$$\left(\frac{PH}{PPI}\right)_t = \delta_0 + \delta_1 \left(\frac{HS}{POP}\right)_t + \delta_2 \left(\frac{BS}{POP}\right)_t + \delta_3 \left(\frac{\Delta S}{POP}\right)_t + \delta_4\, WD_t + \delta_5 \left(\frac{IN}{POP}\right)_t + \phi_{51}D_1 + \phi_{52}D_2 + \phi_{53}D_3 + \varepsilon_{5t}$$

where:

BS_t = Quarterly commercial slaughter of cattle, carcass weight in millions of pounds.

$\left(\frac{PC}{PPI}\right)_t$ = Average quarterly price per cwt. of choice steers, 900 to 1,100 pounds, Omaha, deflated by the producer price index.

IC_{t-1} = Quarterly estimates of cattle on feed in seven states (steers 700 to 1,100 pounds and heifers 500 to 900 pounds).

D_1, D_2, and D_3 = Seasonal dummy variables.

HS_t = Quarterly commercial slaughter of hogs, carcass weight in millions of pounds.

$\left(\frac{PH}{PPI}\right)_t$ = Average quarterly price per cwt. of all barrows and gilts in seven markets deflated by the producer price index.

IH_{t-1} = Quarterly estimates of the inventory of slaughter hogs, 60 to 210 pounds in ten states.

$\left(\frac{\Delta S}{POP}\right)_t$ = Per capita change in pork cold storage stocks in 48 states in millions of pounds.

$\left(\frac{BS}{POP}\right)_t$ = Per capita supply of beef in millions of pounds carcass weight.

$\left(\frac{S}{POP}\right)_{t-1}$ = Per capita level of pork cold storage at the beginning of the quarter in millions of pounds.

$\left(\frac{HS}{POP}\right)_t$ = Per capita supply of slaughter hogs in millions of pounds carcass weight.

$\left(\frac{IN}{POP}\right)_t$ = Disposable income in billions of dollars per capita.

WD_t = Number of fully utilized work days per quarter.

t = Time period.

V. RATIONALE AND DESCRIPTION OF TIME SERIES MODELS

The models proposed in this section all make the assumption that the time series to be forecasted has been generated by a stochastic process. This is to say, that each value of the variable in question is drawn randomly from a probability distribution. Thus, a time series model provides a description of the random nature of the stochastic process that generated the sample of observations under study. The description is not given in terms of a cause-and-effect relationship as in the case of an econometric model, but rather in terms of the way that the randomness is embodied in the process.

If the probability distribution function for a particular time series could be numerically specified, it would be possible to determine the probability of one or more future outcomes. However, it is almost impossible to completely specify such a probability distribution, but it is possible to construct a simplified model of the time series which can explain its randomness in a manner that is useful for forecasting purposes.

For example, the researcher might believe that the values of a time series are normally distributed, and are correlated with one another according to a simple first order process. The actual distribution may be much more complex, but the model may be a good approximation. To be useful, the model merely needs to capture the main characteristics of the series' randomness; it is not expected to match the actual past behavior since the series and the model itself are stochastic [Pindyck and Rubinfeld]. Two time series models are proposed in this section: an univariate integrated moving average (ARIMA) model and a transfer function model.

ARIMA Time Series Model

If the time series to be modeled is stationary, that is, it is in statistical equilibrium about a constant mean, say C, then it can be represented by a class of models called autoregressive moving average (ARMA) models. That is:

$$Z_t - C = \phi_1 (Z_{t-1} - C) - \ldots - \phi_p (Z_{t-p} - C) + a_t - \theta_1 a_{t-1} \cdots - \theta_q a_{t-q}$$

Thus, the model represents the current value of the series as a linear function of: (a) the past values of the series (Z_{t-i}), and (b) the current and past values of the residuals (a_{t-i}).

Introducing the backward shift operator B, we can write:

$BZ_t = Z_{t-1}$ and $B^j Z_t = Z_{t-j}$

$Ba_t = a_{t-1}$ and $B^j a_t = a_{t-j}$

Then the ARMA model may be written in operator form as:

$$Z_t - C = \frac{\theta B}{\phi B} a_t = \frac{1 - \theta B_1 - \ldots - \theta_q B^q}{1 - \phi_1 B_1 - \ldots - \phi_p B^p} a_t .$$

The model above represents the series Z_t as the output from a linear filter whose input is a random series a_t with zero mean and constant variance, and whose filter function is the ratio of two polynomials in the backward shift operator B. Writing the model in the above form and factoring the polynomials in B helps in understanding the generating process. If, for example, the polynomial consists of a complex factor then some periodicity underlies the process.

If the time series under question, Z_t, is not stationary in the statistical sense, then the series may be transformed by differending one or more times. Let $W_t = \nabla^d Z_t$. The stationary series W_t can be represented by an ARIMA model:

$$W_t - C = \frac{(B)}{(B)} a_t = \frac{1 - {}_1B_L - \ldots - {}_qB^q}{1 - {}_1B_1 - \ldots - {}_pB^p} a_t$$

This model is called an autoregressive integrated moving average model or ARIMA. These models are capable of describing a broad class of stochastic trends whose coefficients adapt as each observation comes to hand. Models involving a single differencing can be used to describe a series whose level is continuously updated by random shocks; models involving double differencing can describe a series whose level and slope are updated by random shocks.

To describe a series containing seasonal patterns with period S, and which may be evolving in a nonstationary manner, a seasonal model may be defined as:

$$W_t = \nabla^d \nabla^D s Z_t$$

so that:

$$W_t - C = \frac{\theta(B)\ \theta(B^s)}{\phi(B)\ \phi(B^s)} a_t$$

where $\phi(B)$, $\theta(B)$ are nonseasonal autoregressive and moving average operators defined above, and

$$\nabla_s^D Z_t = Z_t - Z_{t-s}$$

is the seasonal differencing operator, and

$$\phi(B^s) = 1 - \phi_1 B^s - \phi_2 B^{2s} - \dots - \phi_p B^{ps}$$

$$\theta(B^s) = 1 - \theta_1 B^s - \theta_2 B^{2s} - \dots - \theta_q B^{qs}$$

are the seasonal autoregressive and moving average components [Jenkins].

Thus, the anticipated ARIMA hog price model can be represented as:

$$\nabla^d \nabla_s^D PH_t - C = \frac{\theta(B)\ \theta(B^s)}{\phi(B)\ \phi(B^s)} a_t$$

Transfer Function Models

Assume that there is one variable Y_t which is to be forecasted and one input variable X_t which is to be related to the variable to be forecast. Y_t may be split into two components

$$Y_t = U_t + N_t$$

where U_t contains that part of Y_t which can be explained exactly in terms of the input variable X_t, and N_t is an error or noise term which cannot be explained in terms of the input variable; it represents all the missing input variables. A general way of representing a linear dynamic relationship of this kind is:

$$U_t - \delta_1 U_{t-1} - \dots - \delta_r U_{t-r} = W_0 X_{t-b-1} - \dots - W_s X_{t-b-s}$$

that is:

$$U_t = \frac{W_0 - W_1B - \ldots - W_sB^w}{1 - \delta_1B - \ldots - \delta_rB^r} X_{t-b} = \frac{W(B)}{\delta(B)} X_{t-b} = V(B)\ X_{t-b}$$

The transfer function V(B) contains (1) a moving average operator W(B), (2) an autoregressive operator $\delta(B)$, and (3) a pure delay parameter b, which represents the number of complete time intervals after which a change in X_t begins to have an effect on Y_t.

To derive the parameters for U_t the output series Y_t must be differenced one or more times to produce stationarity. For the input series X_t a univariate model may be written:

$$a(B)X_t = b(B)a_t$$

From this it is easy to see that:

$$\frac{a(B)}{b(B)} X_t = a_t$$

That is, the ratio of the two polynomials a(B) and b(B) times the input variable is approximately white noise (a random variable). This enables the model builder to find the estimates of the transfer function. To see this, multiply the original equation by the ratio of the above two polynomials. Thus:

$$Y_t \frac{a(B)}{b(B)} = V(B)\ X_{t-b} \frac{a(B)}{b(B)} + N_t \frac{a(B)}{b(B)}$$

Then substituting:

$$Y_t \frac{a(B)}{b(B)} = V(B)\, a_t + N_t \frac{a(B)}{b(B)}$$

The cross correlations between $Y_t \frac{a(B)}{b(B)}$ and $\hat{a}_{t-k}$ (k=0, 1, 2, ...) are calculated, where $\hat{a}_{t-k}$ are the estimated residuals from the univariate model of the input series X_t. The cross correlation function is then used to suggest the appropriate parameters for:

$$V(B) = \frac{W(B)}{\delta(B)} X_{t-b}$$

It is not necessary that the input and output series be differenced the same number of times [Jenkins]. Each is differenced the number of times necessary to produce stationarity. Thus the model may be written

$$\nabla^d Y_t = V(B)\, \nabla^d X_{t-b} + N_t$$

In general the noise term will be nonstationary, and hence in many practical situations could be represented by an ARIMA model of the form:

$$\nabla^{dn} N_t = C + \frac{\theta(B)}{\phi(B)} a_t$$

Upon substituting this into the model above, the complete model is obtained:

$$\nabla^d Y_t = C + \frac{W(B)}{\delta(B)} \nabla^d X_{t-b} + \frac{\theta(B)}{\phi(B)} a_t$$

Box and Jenkins have recommended fitting the noise or error term by assuming it is white noise. That is, the model

$$\nabla^d Y_t = \frac{W(B)}{\delta(B)} \nabla^d X_{t-b} + e_t$$

is estimated. The autocorrelation function of the residuals $\hat{e}_t$ is calculated from the fitted model, which will then suggest an error structure

$$\phi(B)e_t = \theta(B) N_t$$

This error structure is inserted into the model giving

$$\nabla^d Y_t = \frac{W(B)}{\delta(B)} \nabla^d X_{t-b} + \frac{\theta(B)}{\phi(B)} N_t$$

The coefficients are then estimated in the usual way.

If several input variables $X_{1t}, X_{2t}, \ldots, X_{ky}$ are to be related to the output Y_t, then the model may be generalized to:

$$\nabla^d Y_t = C + \sum_{i=1}^{k} \frac{W_j\ (B)}{\delta_j\ (B)} \nabla^d X_{j_{t-b}} + \frac{\theta(B)}{\phi(B)} a_t$$

Each X-variable has its own transfer function with its own moving average operator $W_j(B)$, autoregressive operator $\delta_j(B)$, and pure delay parameter b_j. Finally, we can include seasonal variations

$$\nabla^d\ \nabla_s^D\ Y_t = \sum_{j=1}^{L} \frac{W_j\ (B)}{\delta_j\ (B)} \nabla^{dj}\ \nabla_s^{Dj}\ X_{jt} + \nabla^{dn}\ \nabla_s^{DN}\ N_t$$

where:

$$\nabla^{dN} \nabla_s^{DN} N_t = \frac{\theta(B)\ \theta(B^s)}{\phi(B)\ \phi(B^s)} a_t$$

Thus, the anticipated transfer function model for the price of slaughter hogs is:

$$\nabla^d \nabla_s^D Ph = \sum_{j=1}^{L} \frac{W_j(B)}{\delta_j(B)} \nabla^{dj} \nabla_s^{Dj} X_{jt} + \frac{\theta(B)\ \theta(B^s)}{\phi(B)\ \phi(B^s)} a_t$$

VI. COMPOSITE FORECASTING

Rather than choose either an econometric or time series forecast the analyst might consider the benefits of a composite forecast. As suggested by Bates and Granger, most if not all forecasts contain some information which is independent of that contained in other forecasts. Theoretically, the gain in information achieved by combining forecasts should result in a model which will outperform the individual model forecasts.

If the assumption is made that each individual forecast is consistent over the forecast period, then the forecast error variance can be denoted as σ^2. A composite of n forecasts will then have a variance σ_c^2, given as

$$\sigma_c^2 = \sum_{i=1}^{n} a_i^2 + 2 \sum_{i<j} a_i a_j P_{ij} \sigma_i$$

where P_{ij} is the correlation coefficient between the errors of forecast i and j, and a_i represents the weight assigned to forecast i. If the problem is to combine a set of unbiased forecasts, the minimum composite variance will be given by the set of n weights which minimize the following Lagrangian function:

$$L = \sum_{i=1}^{n} a_i\sigma_i^2 + 2 \sum_{i<j} a_i a_j P_{ij}\sigma_{ij} + \lambda\left(\sum_{i=1}^{n} a_i - 1\right) = 0$$

By solving this function for the n + 1 equations, the forecast weights for each individual forecast can be determined. In the general case where n = 2 and assuming uncorrelated forecast errors, the optimum weights reduce to:

$$a_1 = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} \qquad\qquad a_2 = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}$$

Thus, the higher the error variance for forecast method 2, the higher the weight placed on forecast method 1.

An easy way of calculating the minimum variance weights has been suggested by Nelson. He suggests the regression of the actual series of P_t on the n forecasts Y_t, Y_t, ... , Y_t^n, so that weights a_0, a_1, ... , a_n can be found that are unbiased and have minimum variance. That is,

$$P_t = a_1 P_t^1 + a_2 P_t^2 + \ldots + a_n P_t^n + U_t$$

Here the intercept is hypothesized to be zero and the coefficients sum to one. Clearly use of this method assumes that the forecast values are consistent relative to each other over the forecasting period. Put another way the relative performance of the two forecasting methods must be stationary.

Thus, the composite minimum variance forecasts in this study will be:

$$P_t = a_1PF_1 + a_2PF_2 + \varepsilon_t$$

where P_t is the actual price of slaughter hogs and PF_1 and PF_2 are the price forecasts of the chosen forecasting models.

Note that if the PF_1 and the PF_2 predictions are individually unbiased, then the above equation may be rewritten as

$$P_t = a_1PF_1 + (1 - a_1)PF_2 + \varepsilon_t$$

The least-squares estimate of a_1 is then given by

$$\hat{a}_1 = \frac{\Sigma(PF_1 - PF_2)(P_t - PF_2)}{\Sigma(PF_1 - PF_2)^2}$$

which is seen to be the coefficient of the regression of the PF_2 prediction errors $(P_t - PF_2)$ on the difference between the two predictions. Thus, the greater the ability of the difference between the two predictions to account for errors committed by PF_2 the larger the weight given to PF_1.

Consider now the case where the PF_1 prediction subsumes the PF_2 prediction and contains additional information, say PF_1' so that $PF_1 = PF_2 + PF_1'$. Then

$$\hat{a}_1 = \frac{\Sigma(PF_1')[P_t - PF_1 + PF_1']}{\Sigma(PF_1')^2}$$

and it can be shown that in large samples $\hat{a}_1$ will approach unity or that the probability limit of $\hat{a}_1$ is

$$\text{plim } \hat{a}_1 = 1$$

since the PF_1 predictions are assumed to be uncorrelated with their associated errors. Thus, if the PF_1 predictions incorporate all the information provided by PF_2 predictions, then the estimates of $\hat{a}_1$ and $\hat{a}_2$ above should be approximately unity and zero, respectively [Nelson].

VII. DATA COLLECTED FOR THE STUDY

In general, data on 16 variables were collected on a quarterly basis covering the years 1963 through 1983. Each of the proposed models was estimated using the first 19 years of data, giving 76 quarterly observations. The final two years of data were not used in estimation so that price predictions from the estimated models could be compared to actual values.

Over the eight-quarter forecasting period, all exogenous variables were known at the time of the forecasts, except for population and income. Rather than build models to predict values for these two variables, actual values were used where needed. The rationale for this was that a one-period-ahead forecast of income or population should be fairly easy to obtain by either regression or time series analysis, or by extrapolation. Thus, all other variables were known when the one-period-ahead forecasts were made.

Price of Cattle and Slaughter Hogs

The monthly average prices per cwt. of all barrows and gilts in seven markets were averaged to obtain quarterly prices for slaughter hogs. Likewise, the monthly average prices per cwt. of Choice steers, 900 to 1,100 pounds at Omaha, were averaged to obtain the quarterly price paid for beef cattle. Both series of prices were taken from the supplements to the U.S.D.A. Livestock and Meat Statistics.

Quantity Produced

Both the quantity of hogs slaughtered and the quantity of cattle slaughtered during each quarter were specified by carcass weight in millions of pounds. Figures obtained were for commercial slaughter. Monthly values were used to calculate quarterly totals, and were taken from the supplements to the U.S.D.A. Livestock and Meat Statistics.

Pork Cold Storage

Statistics for the amount of cured and frozen pork in storage is reported on a monthly basis in the U.S.D.A. Cold Storage bulletin. Values for the end-of-month cold storage level of pork are given in thousands of pounds for 48 states. This study utilized the end-of-month level of cold storage for the third month of each quarter.

Income and Population

Statistics on population and income are reported on a monthly basis by the U.S. Department of Commerce. Figures for consumer disposable income in billions of dollars were taken from the Survey of Current Business. Population figures were collected from Current Population Reports, Series p. 25 and were given in millions of persons. Both series were converted to a quarterly basis.

Number of Fully Utilized Work Days Per Quarter

Values on the number of fully utilized work days each quarter were calculated on a quarterly basis by omitting national holidays and counting Christmas Eve day and Good Friday as half a work day. No account was taken of possible overtime due to seasonal fluctuations in product supply or demand.

Price of Feed

The price of feed was assumed to be .88 times the price of corn in dollars per ton plus .12 times the price of 44 percent soybean meal in dollars per ton. The price of corn and soybean meal were taken from the U.S.D.A. Agricultural Prices Annual Summary, 1963 to 1983. The price of corn was converted from a per bushel basis to a per ton basis. Likewise, the price of soybean meal was converted from a one hundred pound basis to a per ton basis. Monthly prices were averaged to obtain quarterly prices. Both price series were on a national basis.

Retail Price of Pork Cuts

The values for the retail price of pork cuts were taken from the U.S.D.A. Livestock and Meat Statistics, supplemental bulletins. Monthly figures were averaged to obtain quarterly values on a cents per pound basis. The figures reported in the bulletins were based upon the estimated weighted average price of retail cuts from pork carcasses.

Cattle on Feed and Hog Inventory

Cattle-on-feed and the hog inventory values were quarterly estimates taken from the U.S.D.A. Livestock and Meat Statistics supplemental bulletins. Cattle-on-feed were reported in thousands of head for steers 700 to 1,100 pounds and heifers 500 to 900 pounds. Reported estimates were for seven states.

The hog inventory estimates were from the same source, and were also reported in thousands of head. The inventory variable included hogs that weighed 60 to 219 pounds, and the estimates were based upon ten states.

Number of Sows Farrowing

Quarterly values for the number of sows farrowing were obtained from the supplemental bulletins for the U.S.D.A. Livestock and Meat Statistics, and from the U.S.D.A. Hogs and Pigs report. Figures were reported in thousands of head, and were for 14 states from 1963 through 1981. Thereafter, the figures were reported for ten states. Thus, for the last two years of data, the 14 state figure was obtained by adding the appropriate numbers for the deleted four states reported in the Hogs and Pigs report, to the estimated totals of the ten state report.

Price Indices

Three price indices were utilized in this study. The index of prices paid by farmers, the consumer price index, and the producer price index. The index of prices paid by farmers was reported in the U.S.D.A. Agricultural Prices Annual Summary, 1963 to 1983. Data for both the consumer price index and the producer price index were taken from the U.S. Department of Labor, News bulletin.

CHAPTER III

ESTIMATED MODELS

I. TIME SERIES MODELS

The estimated time series models are reported in this section. Both the univariate ARIMA and the transfer function were estimated using the most recent SAS ARIMA package. Since time series models are nonlinear, the SAS estimation procedure uses an iterative linearization method in which the nonlinear function is linearized (by a Taylor series expansion) around an initial set of coefficients. Oridnary least squares is then applied to estimate a new set of coefficients. This process is repeated until convergence of the estimated coefficients is attained.

ARIMA Model

The first model fitted to the data was the univariate autoregressive integrated moving average model, or ARIMA. Inspection of the hog price series indicated that the series was not stationary, and that differencing would be necessary to achieve stationarity. First differencing was applied, and the estimated autocorrelation function (ACF) and the estimated partial autocorrelation function (PACF) were obtained. Inspection of these two functions indicated that first differences produced a constant mean in the series.

Further study of the ACF and the PACF suggested that a fifth order autoregressive model would be adequate to model the differenced price series. For the price of all barrows and gilts in seven markets the estimated model is:

$$\underset{(.101)}{(1 + .583B^5)}\nabla PH = a_t$$

where B is the backshift operator, and a_t is the normally and independently distributed error term, and ∇PH refers to the first difference of hog price series. The standard error is in parentheses. The estimated coefficient was significant at the 5 percent level. This equation was further evaluated statistically using the Ljung and Box Q statistic.[1] For the above equation Q = 5.76 through lag 12. This may be compared with the Chi-square statistic of 19.7 with 11 degrees of freedom at the 5 percent level of significance. Hence, we can conclude that the residual autocorrelations are not significantly different from zero as a set, and that the fitted model is an adequate representation of the behavior of the hog price time series. The estimated autocorrelations of the residuals for lags 6 through 24 are shown in Table 1.

---

[1] $Q = n(n + 2)\ (n - k)^{-1}\ r_k^2\ (a)$

where:

$$r_k^2 = \sum_{T=1}^{n-k} \frac{(z_t - \bar{z})(z_{t+k} - \bar{z})}{\sum_{T=1}^{n} (z_t - \bar{z})^2}$$

Table 1. Estimated Autocorrelation of Residuals from ARIMA and Transfer Function Models

To Lag	Ljung and Box Statistics	Chi-Square (.05)	Degrees of Freedom	Autocorrelations					
				ARIMA Model					
6	2.92	11.10	5	.03	-.18	-.02	.07	.01	.01
12	5.76	19.70	11	.05	.11	-.01	-.03	-.02	.13
18	12.32	27.60	17	.07	.13	.00	.11	.00	.18
24	24.79	35.20	23	.21	.09	.16	.07	.04	.18
				Transfer Function					
6	.86	5.99	2	.01	.00	-.06	.03	.07	.04
12	3.69	15.50	8	.03	-.10	-.11	-.02	-.09	.05
18	11.21	23.70	14	-.18	.13	-.14	.09	.06	-.03
24	16.13	31.40	20	-.05	.02	.13	.02	.02	.16

Transfer Function Model

The second model fitted to the data was the transfer function. This model differs from the ARIMA model in that it utilizes an exogenous input variable. Only one input variable, the number of sows farrowing, was used in this study. Examination of the ACF and PACF of the price series indicated that first differencing induced stationarity. A study of the same two functions for the input series indicated that fourth differences were necessary to achieve stationarity.

An univariate model for the input series was built, and the residuals from this model and the hog price series were cross-correlated. This crosscorrelation function suggested that a second order autoregressive component could be fitted to the input variable. Diagnostic checking of the residuals from this model then indicated that the error term could be represented by a fifth order auto-regressive term and a second order moving average term.

Thus, the estimated transfer function for the price of all barrows and gilts in seven markets is:

$$\nabla PH = \frac{\overset{(.002)}{-.010}}{1 + \underset{(.163)}{.319}B + \underset{(.149)}{.793}B^2} \nabla^4 SF_{t-1} + \frac{1 - \overset{(.126)}{.286}B - \overset{(.122)}{.307}B^2}{\underset{(.113)}{.456} + \underset{(.112)}{.594}B^5} a_t$$

where B is the backshift operator, a_t is the normally and independently distributed error term, and ∇PH and $\nabla^4 SF_{t-1}$ refer to the first and fourth differences for the price of hogs and the number of sows farrowing, respectively. Standard errors are in parentheses. All coefficients were significant at the 5 percent level.

The above equation was also statistically evaluated using the Ljung and Box Q statistic. Through lag 12, this equation had a Q statistic of 3.69, which may be compared with a chi-squared statistic of 15.5 for 8 degrees of freedom at the 5 percent level of significance. In addition to this, the crosscorrelation check of the residuals with the input variable indicated independence. This model was also accepted as an adequate representation of the behavior of the hog price time series. The estimated autocorrelations of the residuals are shown in Table 1.

II. ECONOMETRIC MODELS

The results of estimating the three econometric models are reported in this section. For the recursive model the SAS AUTOREG procedure was used where first order autocorrelation was initially found in the residuals. The reduced form equation was estimated by ordinary least squares (OLS). Finally, for the simultaneous equation model two-stage least squares (2SLS) was used to derive the estimated coefficients.

Recursive Model

The estimated coefficients of the three-equation recursive model are shown in Table 2. In the first equation, the quantity of slaughter hogs, carcass weight in millions of pounds, was hypothesized to be a function of a polynomial distributed lag for both the deflated

Table 2. Estimated Coefficients and Associated Standard Errors of the Recursive Model

Independent Variable	Supply Equation (HS_t) Coeff.	(Std. Error)	Cold Storage Equation ($\Delta S_t/POP_t$) Coeff.	(Std. Error)	Demand Equation (PH/PPI_t) Coeff.	(Std. Error)
Constant	6837.640	(986.210)	-0.761	(0.413)	62.100	(13.820)
D_1	-216.960	(44.350)	0.024	(0.063)	-1.710	(0.360)
D_2	-288.300	(50.650)	-0.059	(0.073)	-2.600	(0.530)
D_3	-480.080	(45.280)	-0.498	(0.079)	-1.270	(0.650)
$(PH/IPPF)_{t-1}$	-25.180	(11.000)				
$(PH/IPPF)_{t-2}$	-13.290	(5.820)				
$(PH/IPPF)_{t-3}$	-4.490	(5.300)				
$(PH/IPPF)_{t-4}$	-0.150	(6.370)				
$(PH/IPPF)_{t-5}$	1.080	(6.360)				
$(PH/IPPF)_{t-6}$	-1.230	(5.330)				
$(PH/IPPF)_{t-7}$	-7.100	(6.080)				
$(PH/IPPF)_{t-8}$	-16.520	(11.440)				
$(PF/PH)_{t-1}$	-97.780	(87.090)				
$(PF/PH)_{t-2}$	-74.239	(50.020)				
$(PF/PH)_{t-3}$	-61.600	(43.330)				
$(PF/PH)_{t-4}$	-59.880	(48.570)				
$(PF/PH)_{t-5}$	-69.700	(47.650)				
$(PF/PH)_{t-6}$	-89.180	(40.120)				
$(PF/PH)_{t-7}$	-120.180	(45.250)				
$(PF/PH)_{t-8}$	-162.130	(83.240)				
WD_t				(0.015)	0.026	(0.206)
$(HS/POP)_t$			0.061	(14.510)	-2.680	(0.220)
$(\Delta S/POP)_t$					4.000	(1.000)
$(S/POP)_{t-1}$			-0.320	(0.08)		
$(IN/POP)_t$					0.200	(0.420)
$(RPP/CPI)_{t-1}$			0.482	(0.301)		
$(COF/POP)_{t-1}$					-0.002	(0.006)
R^2	.71		.81		.77	

price of hogs and the real price of feed. Since the hog cycle is believed to be approximately four years in length, it was hypothesized that eight lags would be adequate to capture the adjustments which producers make to changes in the price of the output and the price of their main input. A second order polynomial was estimated over the eight lags. Only the constant term for the price of hogs was statistically significant at the 5 percent level.

The Durbin-Watson test for first order correlation of the residuals from the OLS estimation was significant. Since no other variable in the data set was capable of removing this autocorrelation, the model was estimated by the SAS AUTOREG procedure. This procedure is equivalent to generalized least squares. The results of the AUTOREG procedure are shown in Table 2. The individual beta coefficients of the second degree polynomials have been calculated[2] from the estimates obtained from the AUTOREG procedure. All variables for the recursive model are defined on pages 27 and 28. From Table 2 only the first, second, and sixth lags of the deflated price of hogs were significant at the 5 percent level, and only lags six and seven were significant on the real price of feed. The three seasonal dummy variables were significant at the 5 percent level as was the intercept term. Since the fourth quarter dummy variable is embedded in the

---

[2] $B_i = \hat{a}_0 + \hat{a}_1 i + \hat{a}_2(i^2)$

$$\mathrm{Var}(\hat{B}_i) = \sum_{j=0}^{m} i^{2j}\,\mathrm{Var}(\hat{a}_j) + \sum_{j<p} i(j + p)\,\mathrm{Cov}(\hat{a}_j\hat{a}_p).$$

intercept term, and since the first quarter dummy variable has the smallest negative value, the conclusion can be drawn that more fresh product comes to market during the first and fourth quarters as expected.

The supply equation in Table 2 indicates that if the price of slaughter hogs increased by one dollar relative to production and living costs, then the amount of slaughter hogs supplied would decline by 25.180 million pounds or approximately .70 percent the first quarter.[3] During the second, third, and fourth quarters supply would continue to decrease by 13.290, 4.940, and 0.150 million pounds, respectively. In the fifth quarter the quantity of slaughter hogs supplied would increase by 1.080 million pounds. Quantity produced would then again decrease the next three quarters by 1.230, 7.100 and 16.520 million pounds, respectively. Thus, the estimated response function is an inverted U shape, which is typical of second order polynomial distributed lags [Pindyck and Rubinfeld].

Relative to the real price of feed the individual beta values again indicate an inverted U shape, but all the estimated values are negative. Looking at selected lags, if the real cost of feed increased by one dollar, supply would decrease by 97.780 million pounds or by 2.9 percent the first quarter, 59.880 million pounds or by 1.8 percent the fourth quarter, and by 162.130 million pounds or by 4.8 percent the eighth quarter.

---

[3]Percentage declines are based upon the mean value for each variable.

The second equation in Table 2 is the estimated per capita change in pork cold storage. This equation was estimated by OLS since the Durbin H statistic did not indicate the presence of first order autocorrelation (H = .8263). Of the three continuous variables in the model, per capita hog slaughter and the beginning level of per capita pork cold storage were significant at the 5 percent level, and their respective signs were positive and negative as expected. The deflated retail price of pork cuts lagged one quarter had a positive sign but was not significant. Of the intercept and the three seasonal dummy variables, only the third quarter dummy variable was significant at the 5 percent level.

The coefficients of the three continuous variables in the per capita change in the pork cold storage equation may be interpreted in the following manner. An increase in the per capita production of slaughter hogs by one pound per person will increase the change in per capita storage by .061 of a pound per person. Likewise, if the beginning level of pork cold storage increased by one pound per person, the current change in per capita cold pork storage would decline by .320 of a pound per person. Finally, if the deflated price of retail pork cuts rose by one dollar, .482 of a pound of pork per person would move into storage.

The third equation of the recursive system was the price of hogs deflated by the producer price index. Of the five continuous explanatory variables only two, per capita hog slaughter and per capita change in pork cold storage, were significant at the 5 percent level. In addition, the sign of each, negative for quantity supplied and positive for the change in storage was as expected. Of the three other explanatory variables, per capita income and number of fully utilized work days per quarter were positive in sign but insignificant. Per capita cattle-on-feed lagged one quarter was negative as expected, but not significant at the 5 percent level. The intercept coefficient was positive in sign, and the first and second quarter dummy variables were negative, but all three variables were significant at the 5 percent level.

The estimated coefficients indicated that if the number of fully utilized work days increased by one, then the deflated price of slaughter hogs would increase by approximately $.03. A one pound increase in per capita hog supply (approximately 209 million pounds based upon the mean value) would decrease deflated hog price by approximately $2.68. If the change in per capita storage increased by one pound per person, the deflated price of slaughter hogs would rise by $4.00. Likewise, a $1000.00 increase in per capita disposable income would increase the deflated price of slaughter hogs by $.20. Finally, if the per capita number of cattle-on-feed increased by one animal per person, the deflated price of slaughter hogs would decrease by $.002.

Reduced Form Model

The reduced form econometric model was derived from the recursive model. Estimated coefficients are shown in Table 3. All variables are defined on pages 27, 28 and 29. In this model the

Table 3. Estimated Coefficients and Associated Standard Errors of the Reduced Form Model

Independent Variable	Price Prediction Equation (PH_t/PPI_t) Coefficient	(Std. Error)
Constant	-33.490	(33.050)
D_1	2.990	(1.020)
D_2	1.76	(1.450)
D_3	5.24	(1.340)
$((PH/IPPF)/POP)_{t-1}$	59.130	(50.930)
$((PH/IPPF)/POP)_{t-2}$	9.31	(25.950)
$((PH/IPPF)/POP)_{t-3}$	-23.640	(12.430)
$((PH/IPPF)/POP)_{t-4}$	-39.700	(12.370)
$((PH/IPPF)/POP)_{t-5}$	-38.890	(12.890)
$((PH/IPPF)/POP)_{t-6}$	-21.210	(8.310)
$((PH/IPPF)/POP)_{t-7}$	13.36	(10.840)
$((PH/IPPF)/POP)_{t-8}$	64.81	(30.740)
$((PF/PH)/POP)_{t-1}$	584.230	(168.250)
$((PF/PH)/POP)_{t-2}$	320.000	(74.240)
$((PF/PH)/POP)_{t-3}$	129.180	(70.510)
$((PF/PH)/POP)_{t-4}$	11.770	(93.650)
$((PF/PH)/POP)_{t-5}$	-32.220	(91.030)
$((PF/PH)/POP)_{t-6}$	-2.790	(64.560)
$((PF/PH)/POP)_{t-7}$	100.04	(79.540)
$((PF/PH)/POP)_{t-8}$	276.300	(184.240)
WD_t	0.190	(0.500)
$(S/POP)_{t-1}$	-7.00	(2.000)
$(IN/POP)_t$	-0.551	(0.349)
$(RPP/CPI)_{t-1}$	30.448	(17.579)
$(COF/POP)_{t-1}$	0.345	(0.090)
R^2	.81	

third term of the per capita deflated price of hogs in the polynomial distributed lag was statistically significant. Likewise, the second and third terms of the per capita real cost of feed in the polynomial distributed lag were significant at the 5 percent level. Individual coefficients are again shown for each lag. For the per capita price of hogs, lags four, five, six, and eight were statistically significant at the 5 percent level. For the per capita price of feed, lags one and two were found to be significant.

Two other continuous variables were statistically significant at the 5 percent level: per capita cattle-on-feed lagged one quarter and per capita cold pork storage lagged one quarter. The number of fully utilized work days each quarter, per capita income, and the deflated retail price of pork cuts lagged one quarter were not significant at the 5 percent level. Since the AUTOREG procedure calculated the first order autocorrelation of residuals at .11, this model was not corrected for autocorrelation.

The individual beta values of both polynomial distributed lags indicate that both response functions are U shaped. Looking at selected lags, if the deflated price of hogs increased by one dollar <u>on a per capita basis</u> the deflated price of slaughter hogs would increase by $59.130 the first quarter, decline by $39.700 the fourth quarter, and increase by $64.810 the eighth quarter. Likewise, if the price of feed relative to the price of hogs increased by one dollar <u>on a per capita basis</u>, the price of hogs would rise by $584.23 the first quarter, $11.77 the fourth quarter, and by $276.30 the eighth quarter.

The coefficients of the other continuous variables indicated that if the number of fully utilized work days increased by one, the deflated price of hogs would increase by $.19 per hundredweight. If the beginning-of-quarter per capita level of pork cold storage were increased by one pound per person, the deflated price of hogs would decline by $7.00 per hundredweight. If the deflated per pound price of retail pork cuts lagged one quarter rose by one dollar, the deflated price of slaughter hogs would increase by $30.448 per hundredweight. If the level of per capita cattle-on-feed lagged one quarter were increased by one animal per person, the deflated price of hogs would increase by $.35 per hundredweight. Finally, if per capita income increased by $1000.00 per person, the deflated price of hogs would decline by $.55 per hundredweight.

Simultaneous Equation Model

The final individual econometric model was a simultaneous equation system estimated by 2SLS. These estimations with asymptotic standard errors are shown in Table 4 and all variables are defined on pages 35 and 36. The system consists of two supply equations (one for pork and the other for cattle), a change in per capita pork cold storage equation, and two equations which predict the price of cattle and the price of slaughter hogs. Seasonal dummy variables were included in each equation.

In the first equation the quantity of beef supplied in millions of pounds was hypothesized to be a function of the number of cattle-on-feed lagged one quarter and the price of cattle per hundred pounds

Table 4. Estimated Coefficients and Associated Standard Errors of the Simultaneous System

Independent Variables	Dependent Variables: BS_t Coeff.	BS_t Std. Error	HS_t Coeff.	HS_t Std. Error	$(\Delta S/POP)_t$ Coeff.	$(\Delta S/POP)_t$ Std. Error	$(PC/PPI)_t$ Coeff.	$(PC/PPI)_t$ Std. Error	$(PH/PPI)_t$ Coeff.	$(PH/PPI)_t$ Std. Error
Constant	6302.876	(563.313)	191.817	(710.926)	42.254	(560.232)	2.859	(75.864)	-14.042	(102.597)
D_1	-245.781	(116.520)	-90.622	(104.784)	-37.600	(65.218)	-0.082	(1.744)	0.403	(2.408)
D_2	-245.794	(109.510)	(113.627)	(121.103)	-105.867	(72.729)	2.894	(2.182)	0.714	(3.225)
D_3	-5.684	(109.389)	132.356	(131.607)	-564.152	(79.369)	3.189	(1.975)	1.166	(3.180)
$(S/POP)_{t-1}$					-0.351	(0.085)				
$(COF)_{t-1}$	0.298	(0.030)								
$(MH)_{t-1}$			0.131	(0.020)						
$(HS/POP)_t$					42.085	(23.047)	1.143	(0.632)	-0.737	(0.971)
$(\Delta S/POP)_t$									1.000	(5.000)
$(IN/POP)_t$							-1.246	(0.429)	-0.911	(0.577)
$(BS/POP)_t$							1.829	(0.843)	3.345	(1.142)
$(PH/PPI)_t$			6.457	(13.560)	-0.004	(0.009)				
$(PC/PPI)_t$	-114.877	(23.684)								
WD_t							-0.608	(1.053)	-0.560	(1.321)
R^{2} [a]	.60		.59		.80		.14		.29	

[a] R^2 as reported by SAS for the individual equations.

deflated by the producer price index. The deflated price variable was significant at the 5 percent level, but the sign was negative. The wrong sign on the deflated price variable may reflect model specification error. Specifically, the current supply of cattle produced should probably be a function of lagged price as in the recursive model. Thus, the negative sign on current price reflects the law of demand: when quantity increases price declines. The cattle-on-feed variable was also significant and the sign positive as expected. The intercept term and the first and second quarter dummy variables were also significant at the 5 percent level.

The estimated coefficients indicate that if per capita cattle-on-feed increased by 1000 head, cattle supply would increase by 298,000 pounds the following quarter. An increase in the deflated price of cattle by one dollar, would be associated with a decline of 114.877 million pounds in the quantity of cattle slaughtered during that quarter.

In the second equation of Table 4 the quantity supplied of hogs in millions of pounds was hypothesized to be a function of the deflated price of hogs and the inventory of market hogs lagged one quarter. The inventory variable was comprised of hogs weighing 60 to 219 pounds. In addition, three seasonal dummy variables were specified. Only one variable, the inventory of market hogs, was statistically significant at the 5 percent level. Both signs were positive as expected, although this is a moot point for the insignificant price variable.

Hence, in the second equation of the simultaneous model, the estimated coefficients suggest that a 1000 head increase in the market hog inventory will result in 131,000 additional pounds of hog slaughter during the following quarter. Likewise, a one dollar increase in the deflated price of hogs would result in 6.457 million pounds of fresh pork coming to market.

The third equation in Table 4 is the percapita change in frozen pork storage. This was hypothesized to be a function of the deflated price of hogs, the per capita supply of pork, the biginning-of-quarter per capita level of pork cold storage, and seasonal dummy variables. The deflated hog price variable was negative but insignificant at the 5 percent level. Per capita hog slaughter was positive in sign but insignificant. The beginning level of pork cold storage stocks was negative in sign and significant. Of the intercept term and the three seasonal dummy variables only the third quarter dummy was significant at the 5 percent level.

Thus, the estimated coefficients suggest that if the beginning-of-quarter level of per capita pork cold storage were increased by one pound per person, the per capita change in pork storage would decline by .351 pounds per person. If the per capita supply of slaughter hogs increased by one pound per person, the per capita change in storage would increase by .042 pounds per person. Finally, if the deflated price of hogs rose by one dollar, .0046 pounds of pork cold storage per person would flow to market.

The fourth equation in Table 4 is the deflated price of cattle. This variable was hypothesized to be a function of per capita income, the number of fully utilized work days, the per capita quantity produced of beef and pork, and three seasonal dummy variables. At the 5 percent level of significance, the quantity of beef supplied was significant, but its sign was positive rather than negative. In the demand equation, price and quantity produced should have opposite signs. The other significant variable was per capita income, but its sign was negative rather than positive. As per capita income increases, the price of beef would be expected to rise since the demand curve would shift outward. The per capita quantity of pork supplied was positive in sign but insignificant. Finally, all three dummy variables were insignificant.

Looking at the estimated coefficients, a one day increase in the number of fully utilized work days would result in a $.61 decrease in the deflated price of cattle. A one pound increase in the per capita supply of slaughter hogs would result in a $1.143 increase in the deflated price of cattle. Likewise, a $1000.00 increase in per capita income would result in a $1.246 decline in cattle price. Finally, if the per capita supply of beef increased by one pound per person, the deflated price of cattle would rise by $1.829.

The final equation in Table 4 is the slaughter hog demand equation. The deflated price of slaughter hogs was hypothesized to include all the variables in the deflated price of cattle equation, plus the change in per capita frozen pork storage variable. Out of all the variables, only the per capita quantity of beef supplied was significant at the 5 percent level, but its sign was unexpectedly positive. If the supply of a competitive product increased, its price would fall and consumers would substitute the relatively less expensive good. Hence, its sign should theoretically be negative.

Of the insignificant variables, the per capita quantity of hogs supplied was negative in sign and the change in frozen pork storage was positive, as expected. Both per capita income and the number of fully utilized work days had negative signs, contrary to expectations. All three seasonal dummy variables as well as the intercept term were insignificant.

The estimated coefficients may be interpreted in the following manner. A one day increase in the number of fully utilized work days would decrease the deflated price of slaughter hogs by $.56. If the per capita quantity of slaughter hogs were increased by one pound per person, the deflated price of slaughter hogs would decline by $.737 per hundredweight. A one pound increase in per capita pork cold storage would result in the deflated price of slaughter hogs rising by $1.00. If per capita income increased by $1000.00, the price of slaughter hogs would decline by $.911. Finally, if the per capita quantity of slaughter cattle increased by one pound per person, the deflated price of slaughter hogs would increase by $3.345.

In general, the signs and the magnitude of the estimated coefficients of this model lack credibility. It is apparent that the model has been misspecified to such an extent that the coefficients are unreliable. In reality, one would not attempt to predict the price of slaughter hogs without reformulating the simultaneous model. However, since several types of models had already been estimated, and since, econometrically, the hog market is probably best represented on a quarterly basis by a recursive model, the simultaneous system was not used to forecast the price of slaughter hogs in succeeding sections.

III. COMPOSITE MODELS

Results of the composite models are shown in Table 5. Each individual time series and econometric model was used to predict price over the period of fit (1963-81). Three composite forecast models were developed. For Composite Model 1, the reduced form model and the ARIMA were combined since they represented the simplest models from each group. For Composite Model 2, the recursive model and the transfer function were combined, since the recursive model appeared from the foregoing discussion to be superior to the simultaneous system, and since the transfer function was the most sophisticated time series model. Composite Model 3 was developed by combining the reduced form with the transfer function. The reason was that the reduced form model appeared to predict nearly as well as the recursive model based upon Theil's U_2 coefficient (which will be discussed later). As stated in the model proposal section, each individual model was used to predict slaughter hog price over the period of fit.

Table 5. Estimated Coefficients (Weights) for the Individual Model Forecasts Comprising the Composite Models

Individual Model	Weights		
	Composite Model 1	Composite Model 2	Composite Model 3
ARIMA Model	.44		
Recursive Model		.51	
Transfer Function Model		.48	.77
Reduced Form Model	.56		.23

The predictions from two individual models were then regressed against actual price to obtain the composite model weights.

In Composite Model 1, .56 of the predicted hog price from the reduced form model, and .44 of the predicted hog price from the ARIMA model were used to predict slaughter hog price. In Composite Model 2, .51 of the price prediction from the recursive system and .48 of the price prediction of the transfer function were used to form a composite price prediction. Finally, for Composite Model 3, .23 and .77 of the price predictions from the reduced form and the transfer function, respectively, were used to formulate a composite hog price forecast.

CHAPTER IV

FORECASTING RESULTS

I. INTRODUCTION

Forecasting results will be reported in two parts. First, ex post forecasts over the last eight observations (1980-81) of the fitted data will be reported in terms of the root mean square error (RMSE) and Theil's U_2 coefficient. Researchers have often limited themselves to this type of analysis in the past. The justification for this is that all variables are known, hence any forecasting error is due to the model and not to the forecasts of exogenous variables within the model. Since the models of this study relied upon lagged independent variables (except for the variables of population and income) this problem was largely overcome. However, the reporting of ex post forecasts will give the reader a better idea of how well the models fit the data.

In the second part of the results actual or ex ante forecasts will be made. These forecasts are for the out-of-sample period covering the years 1982 and 1983. The criterion for comparison will again include the RMSE and Theil's U_2 coefficient. In addition, the ability of each model to track actual price movements will be assessed. This is added to the analysis since a model might have a U_2 coefficient larger than one (the model forecasts worse than a naive no-price-change model) and yet be valuable if it could forecast an increase, decrease, or no change in price. In addition, the reader may be interested in knowing whether those models which are the best *ex post* forecasters are the best *ex ante* forecasters as well.

II. INDIVIDUAL MODEL *EX POST* FORECASTS

ARIMA Model

Results of the ARIMA model forecasts for 1980-81 are shown in Table 6. This model forecasted below actual price five times and above actual price three times. The most accurate forecast showed an \$.80 error while the least accurate missed by \$9.84. The RMSE over the eight forecasts was \$4.79 and Theil's U_2 coefficient was .66. Hence, over the period this model forecasted better than a naive no-price-change model.

Transfer Function Model

Also shown in Table 6 are the results from the eight forecasts for the transfer function. This model forecasted above actual price seven times and below once. The most accurate forecast missed actual price by \$.21 and the least accurate forecasted under actual price by \$8.60. The RMSE was \$3.99 and Theil's U_2 coefficient was .55. As with the ARIMA model the transfer function provided a better forecast than a naive no-price-change model.

Recursive Model

The forecasts of the recursive econometric model are also shown in Table 6. This model forecasted better than a naive no-price-change

Table 6. Comparison of Ex Post Forecasts of Quarterly Hog Prices from Estimated Individual Models for the 1980-81 Period, in Dollars Per Hundredweight

Time	Actual	ARIMA		Transfer Function		Recursive Model		Reduced Form	
Period	Price	Forecast	Error	Forecast	Error	Forecast	Error	Forecast	Error
8001	36.31	35.51	-.80	37.34	1.03	35.92	-0.39	37.28	.97
8002	31.18	35.18	4.00	35.76	4.58	29.67	-1.51	40.46	9.28
8003	46.23	36.69	-9.85	37.63	-8.60	47.34	1.11	44.36	-1.87
8004	46.44	48.87	2.43	49.07	2.63	46.29	-0.15	44.57	-1.87
8101	41.13	47.68	6.55	45.31	4.18	43.52	2.39	47.18	6.05
8102	43.63	41.18	-2.45	45.71	2.08	48.37	4.74	44.38	.75
8103	50.42	46.62	-3.80	52.05	1.63	56.57	6.15	47.51	-2.91
8104	42.63	41.64	-.99	42.84	.21	50.84	8.21	41.74	-.89
Root Mean Square Error		4.79		3.99		4.14		4.19	
Theil's U_2 Coefficient		.66		.55		.59		.60	

model as indicated by a Theil's U_2 coefficient of .59. In addition, five forecasts were above actual price while three were below, with the most accurate missing by $.15 and the least accurate missing by $8.21. The RMSE for the eight forecasts was $4.14.

Reduced Form Model

The eight forecasts from the reduced form model are also shown in Table 6. This model forecasted slightly worse than the recursive or transfer function but better than the ARIMA model. Over the last eight observations Theil's U_2 coefficient was .60, and the RMSE was $4.19. The model forecasted above and below actual price four times. The most accurate forecast was above actual price by $.75 while the least accurate forecast was above actual price by $9.28.

III. COMPOSITE MODEL *EX POST* FORECASTS

Composite Model 1: Reduced Form and ARIMA Model

The forecasts from the first composite model are shown in Table 7. It was hypothesized that the composite models would forecast better than any individual model from which they were composed. In the case of the reduced form and ARIMA composite model this was true. Theil's U_2 coefficient was calculated to be .58 and the RMSE was $4.06. Thus, the composite forecasting model was better than either individual model used alone. The composite model forecasted above and below actual price four times. The most accurate forecast was below actual price by $.31 while the least accurate forecast was above actual price by $7.27.

Table 7. Comparison of Ex Post Forecasts of Quarterly Hog Prices from Estimated Composite Models for the 1980-81 Period, in Dollars Per Hundredweight

Time Period	Actual Price	Reduced Form and ARIMA Model		Recursive Model and Transfer Function		Reduced Form and Transfer Function	
		Forecast	Error	Forecast	Error	Forecast	Error
8001	36.31	36.80	.49	36.36	.05	37.36	1.05
8002	31.18	38.45	7.27	32.41	1.23	36.89	5.71
8003	46.23	41.18	-5.05	42.32	-3.91	39.24	-6.99
8004	46.44	46.84	.40	47.32	.88	48.06	1.62
8101	41.13	47.79	6.66	44.09	2.96	45.79	4.66
8102	43.63	43.32	-.31	46.75	3.12	45.44	1.81
8103	50.42	47.50	-2.92	53.99	3.52	51.03	.61
8104	42.63	42.04	-.59	46.62	3.99	42.65	-.01
Root Mean Square Error		4.06		2.84		3.72	
Theil's U_2 Coefficient		.58		.41		.53	

Composite Model 2: Recursive Model and the Transfer Function

Ex post forecast results of the second composite model are also shown in Table 7. Like Composite Model 1, this composite model forecasted better than either individual model. The RMSE for the composite model was \$2.84 and Theil's U_2 coefficient was .41. This can be compared to the forecasts of the transfer function with a RMSE of 3.99 and a Theil's U_2 coefficient of .55, or the forecasts of the recursive model with a RMSE of 4.14 and a U_2 coefficient of .59. The composite model forecasted above actual price seven times and below actual price once. The most accurate forecast was below actual price by \$.05 while the least accurate was above by \$3.99.

Composite Model 3: Reduced Form Model and the Transfer Function

The last composite model shown in Table 7 also forecasted better than either individual model. Theil's U_2 coefficient was calculated to be .53 and the RMSE over the eight forecasts was 3.72. Accuracy of the forecasted values ranged from a low of \$.01 below actual price to a high of \$6.99 below actual price. The composite model forecasted above actual price six times and below twice.

IV. INDIVIDUAL MODEL *EX ANTE* FORECASTS

The results for the out-of-sample forecasts for the individual models are reported below. These forecasts and the forecasts of the composite models that will follow are the ones that the pork industry would primarily be concerned about. If any of the models below could forecast better than a naive no-price-change model, then hog producers and others should be able to utilize the one-quarter-ahead forecasts for planning purposes.

ARIMA Model

Forecasts over the eight quarters of 1982 and 1983 for the ARIMA model are shown in Table 8. Theil's U_2 coefficient of 1.75 indicates that a naive no-price-change model would have forecasted more accurately than this model. The most accurate forecast was above actual price by \$1.67 while the least accurate forecast was below actual price by \$17.84. Table 9 shows how well the model could track actual price movements. Diagonal elements, left to right, in the table indicate when the model is tracking actual price movements accurately. Thus, summation of the diagonal elements divided by the number of forecasts gives an indication of how well the model mimics real price movement. For the ARIMA model, the direction of three out of eight movements in price were predicted, thus the model was accurate 38 percent of the time. A comparison of actual versus forecasted price is shown in Figure 1.

Transfer Function Model

Statistics on the price predictions of the transfer function are also shown in Table 8. As with the ARIMA model, the transfer function failed to predict any better than a naive no-price-change model, as indicated by Theil's U_2 coefficient of 1.11. The RMSE over the eight-period forecast was \$6.43, an improvement over that of the ARIMA model. Errors in the forecasted prices ranged from \$2.04 to

Table 8. Comparison of *Ex Ante* Forecasts of Quarterly Hog Prices from Estimated Individual Models for the 1982-83 Period, in Dollars Per Hundredweight

Time Period	Actual Price	ARIMA		Transfer Function		Recursive Model		Reduced Form	
		Forecast	Error	Forecast	Error	Forecast	Error	Forecast	Error
8201	48.17	42.51	-5.66	46.13	-2.04	60.64	12.47	38.60	-9.57
8202	56.46	45.60	-10.86	51.18	-5.28	62.58	6.12	35.24	-21.22
8203	61.99	44.15	-17.84	50.43	-11.56	74.76	12.77	47.86	-14.13
8204	55.12	40.19	-14.93	48.16	-6.96	70.44	15.32	42.63	-12.49
8201	55.00	44.73	10.27	53.14	-1.86	70.25	15.25	44.20	-10.80
8302	46.74	44.80	-1.94	52.08	5.34	61.13	14.39	44.90	-1.84
8303	46.90	42.99	-3.91	50.13	3.23	63.53	16.63	41.62	-5.28
8304	42.18	43.85	1.67	50.82	8.64	55.85	13.67	30.89	-11.29
Root Mean Square Error		10.11		6.43		13.66		12.11	
Theil's U_2 Coefficient		1.75		1.11		2.36		2.09	

Table 9. Ability of Estimated Individual Models to Predict the Direction of Quarterly Change in Hog Prices, 1982-83

Actual Price Movement	ARIMA		Transfer Function		Recursive Model		Reduced Form	
	Rising	Falling	Rising	Falling	Rising	Falling	Rising	Falling
Rising	2	2	2	2	4	0	1	3
Falling	3	1	2	2	0	4	2	2
Percentage of Correct Tracking	38%		50%		100%		38%	

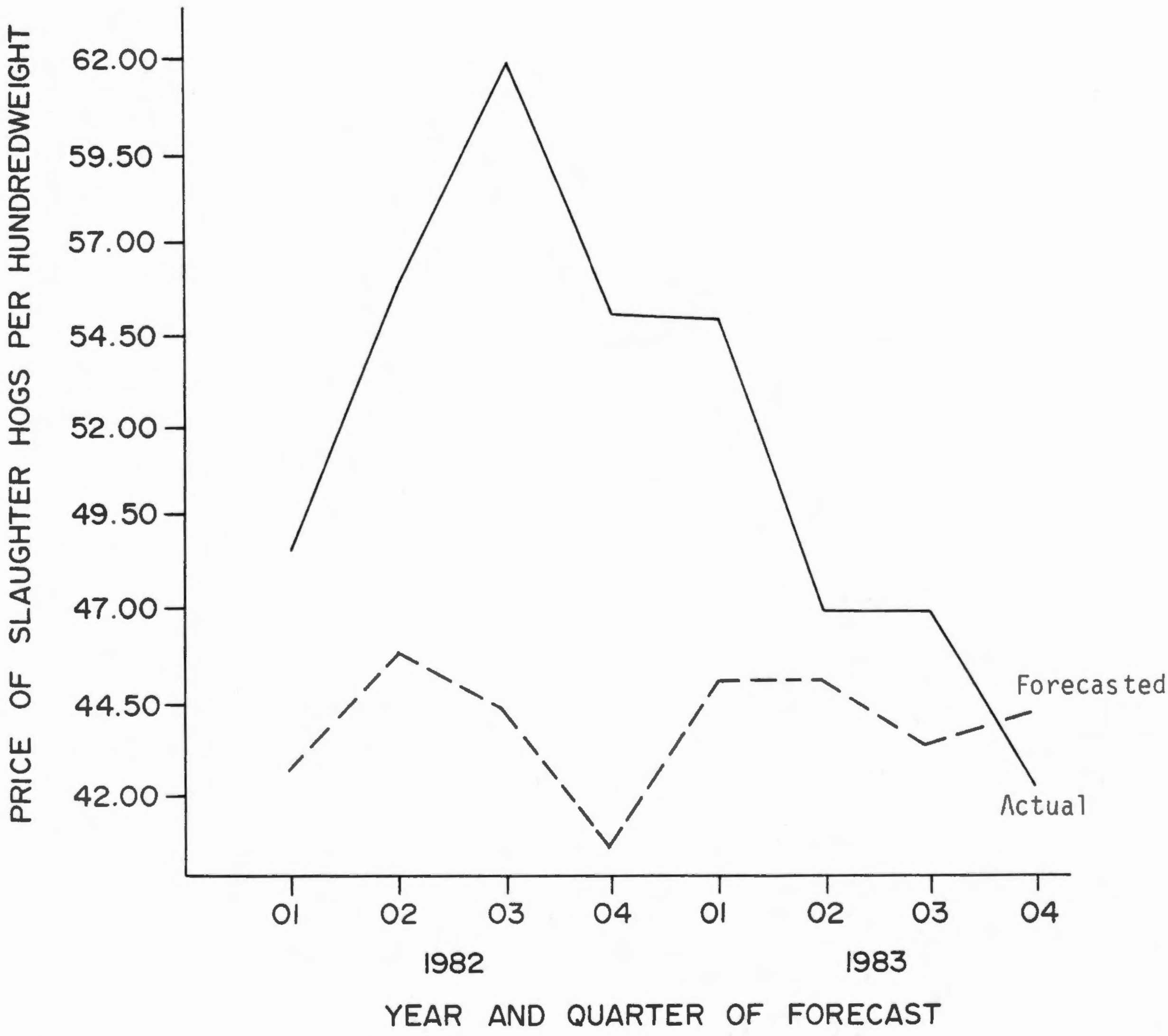

Figure 1. Actual Versus Predicted Prices of the ARIMA Model.

\$11.56. The price tracking measures of Table 9 indicate that the transfer function was able to predict the direction of price movement 50 percent of the time. While this model performed better than the ARIMA model it was still less than satisfactory. Actual versus forecasted prices are shown in Figure 2.

Recursive Model

The recursive model forecasted well over the fitted data, but Table 8 indicates that Theil's U_2 coefficient was 2.36 and the RMSE was \$12.66. Thus a naive no-price-change model would have been more accurate than this model. All eight forecasts from this model were above actual price. The most accurate forecast missed actual price by \$6.12 while the least accurate was above actual price by \$16.63. A redeeming feature of this model is indicated in Table 9. The recursive model was able to track the direction of price change 100 percent of the time. Actual versus predicted prices are shown in Figure 3.

Reduced Form Model

From Table 8 the calculated U_2 coefficient for the reduced form model is 2.09. The RMSE over the eight-quarter forecast was \$12.11, and the model forecasted below actual price each quarter. Thus, while this model can forecast better than the recursive system it is still worse than either time series model. The most and least accurate forecasts were \$1.84 and \$21.22, respectively, below actual price. Table 9 indicates that the reduced form model could track the direction of price movement 38 percent of the time, which was the same

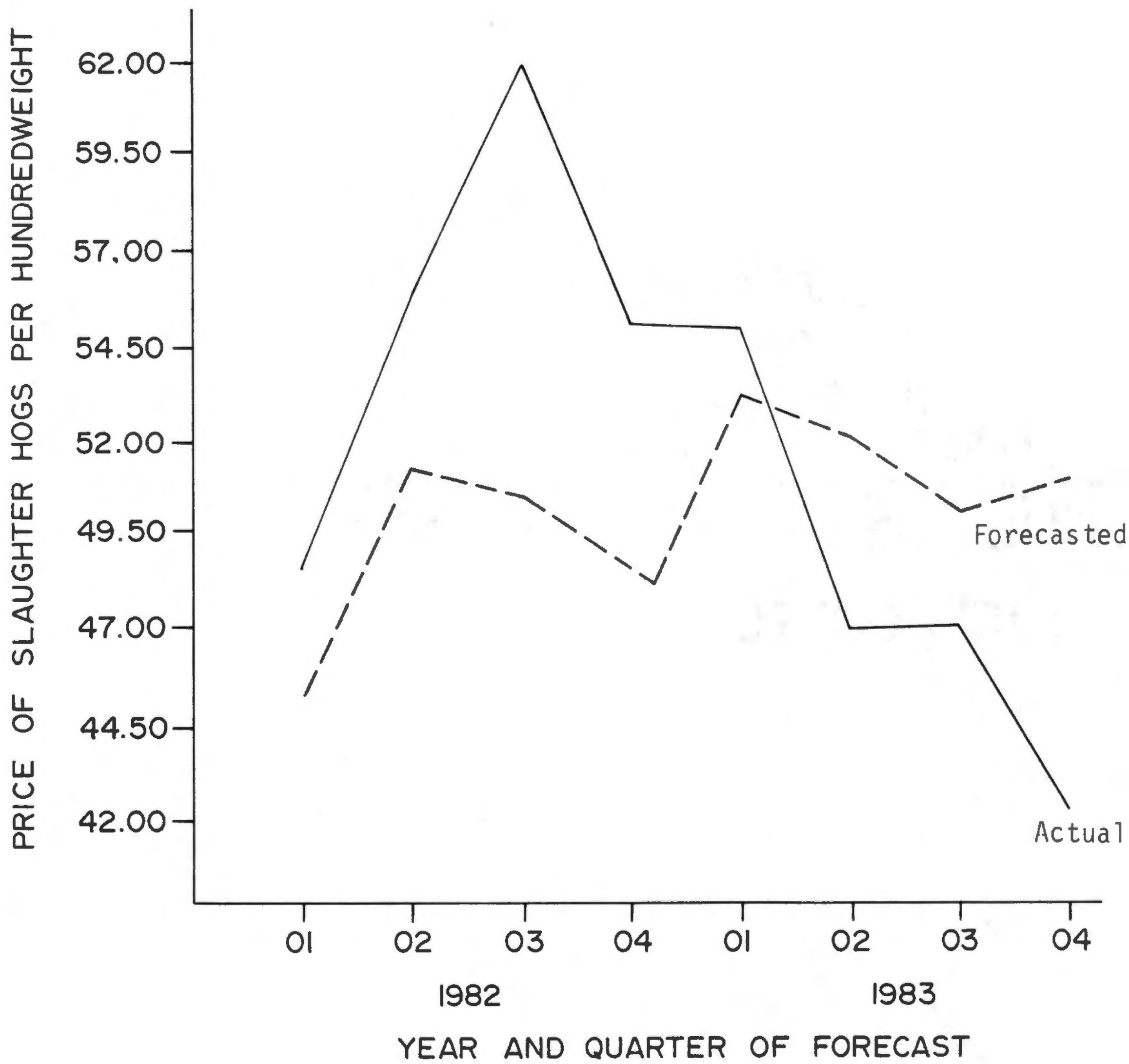

Figure 2. Actual Versus Predicted Prices of the Transfer Function Model.

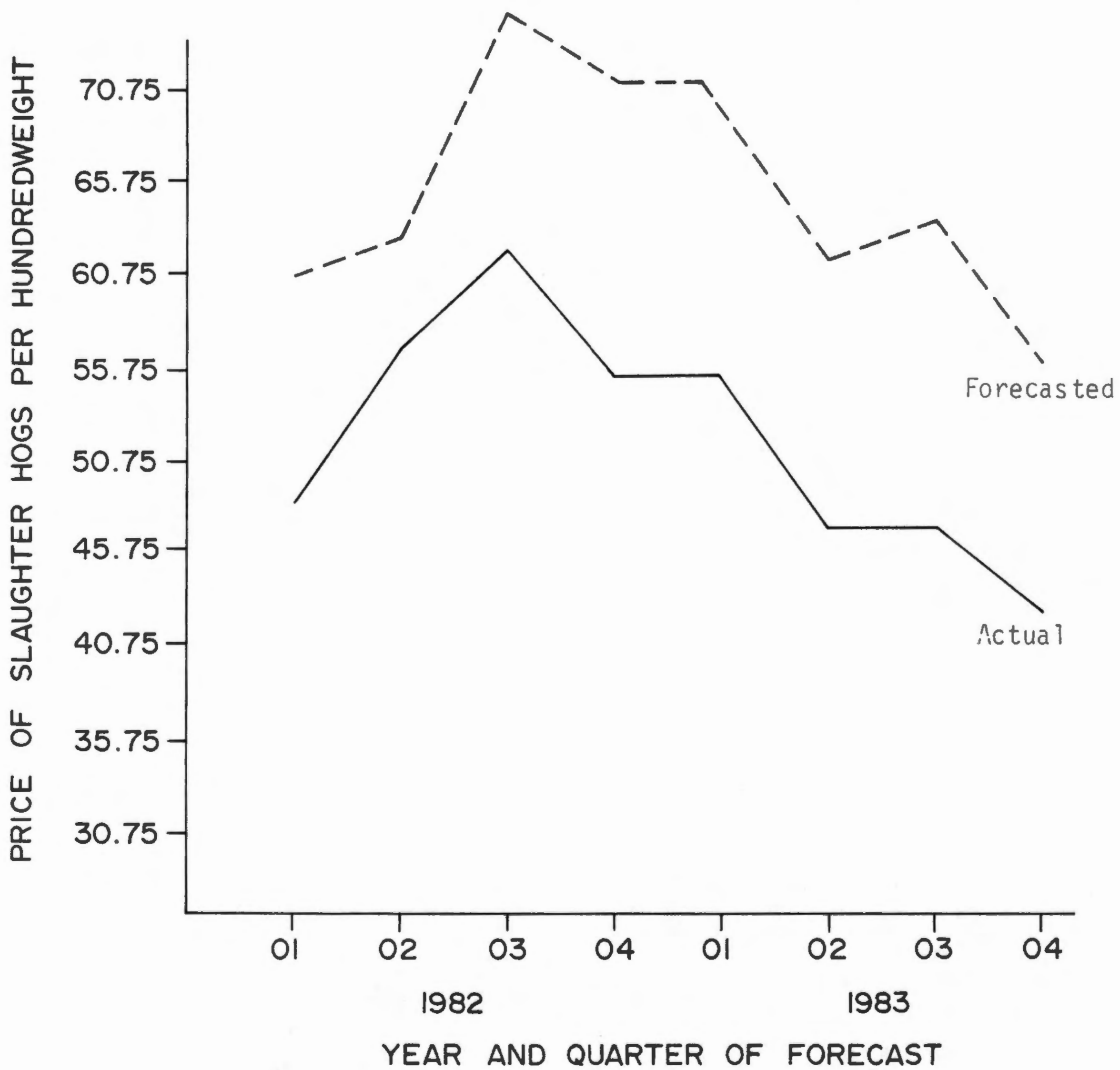

Figure 3. Actual Versus Predicted Prices of the Recursive Model.

as the ARIMA model, but worse than the transfer function or the recursive system. Actual versus forecasted prices are shown in Figure 4.

V. COMPOSITE MODEL *EX ANTE* FORECASTS

Results of the out-of-sample forecasts of the composite models are presented below. Since a composite model was the best *ex post* forecaster it was hoped that at least one composite model would be able to forecast *ex ante* better than a naive no-price-change model.

Composite Model 1: Reduced Form Model and the ARIMA Model

Table 10 shows that the composite model made up of forecasts from the reduced form and ARIMA model was able to forecast better than the reduced form model used alone, but not better than the ARIMA model over the 1982-83 out-of-sample period. Theil's U_2 coefficient was calculated to be 1.82 which again indicates that a naive no-price-change model would have given more accurate price forecasts. This U_2 coefficient can be compared to the ARIMA U_2 coefficient of 1.75 and the reduced form model's U_2 coefficient of 2.09. The RMSE for the composite model was $10.54. Each forecast was below actual price, and ranged from a low of $1.52 to a high of $16.33. Table 11 indicates that this composite model was able to track the direction of price movement 38 percent of the time, which was the same as both individual models. Actual versus forecasted prices are shown in Figure 5.

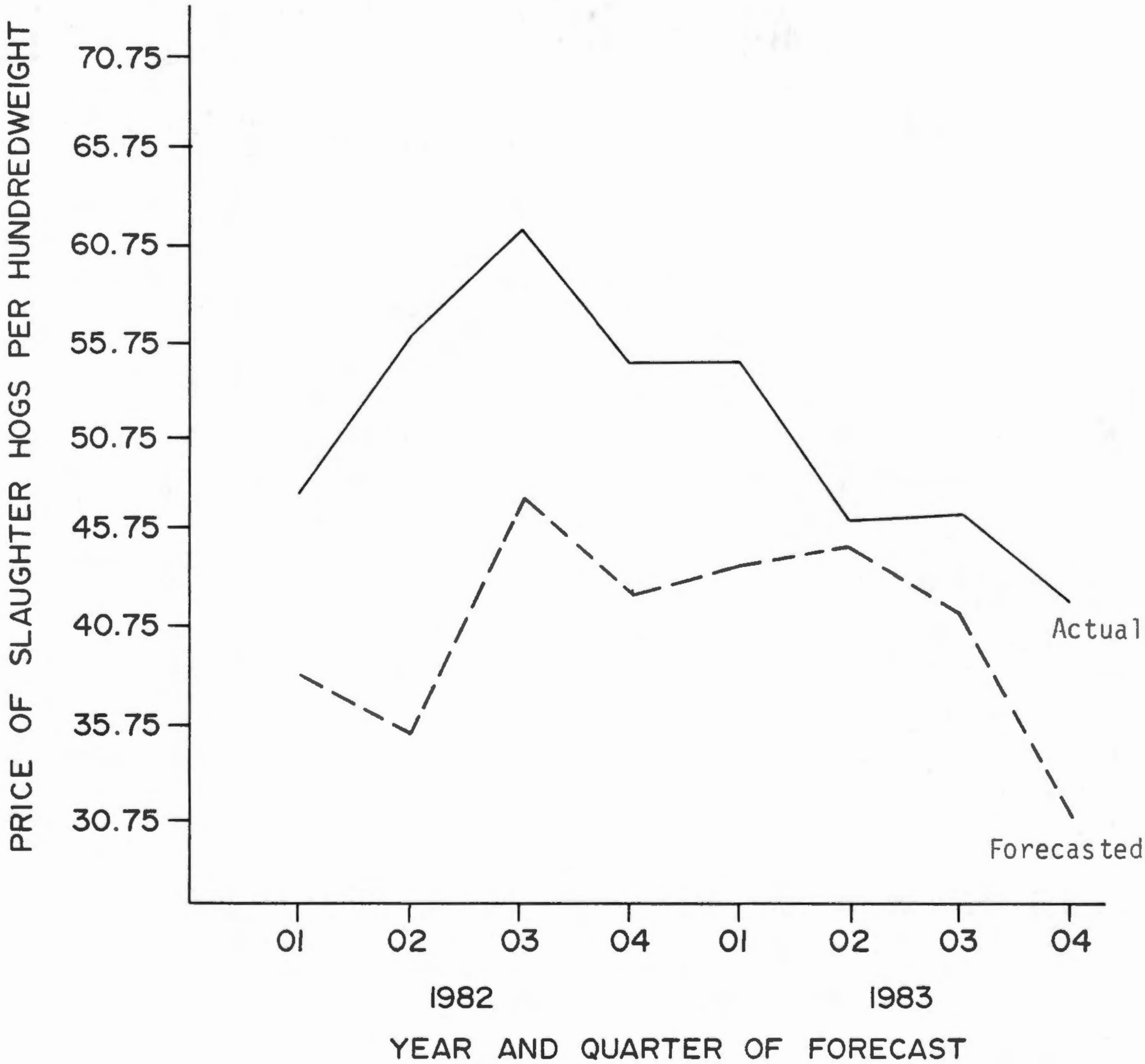

Figure 4. Actual Versus Predicted Prices of the Reduced Form Model.

Table 10. Comparison of *Ex Ante* Forecasts of Quarterly Hog Prices from Estimated Composite Models for the 1982-83 Period, in Dollars Per Hundredweight

Time Period	Actual Price	Reduced Form and ARIMA Model: Forecast	Reduced Form and ARIMA Model: Error	Recursive Model and Transfer Function: Forecast	Recursive Model and Transfer Function: Error	Reduced Form and Transfer Function: Forecast	Reduced Form and Transfer Function: Error
8201	48.17	40.65	-7.20	53.20	5.03	44.41	-3.76
8202	56.46	40.13	-16.33	56.63	.17	47.50	-8.96
8203	61.99	46.60	-15.39	62.47	.48	49.87	-12.12
8204	55.12	41.89	-13.23	59.18	4.06	46.91	-8.21
8301	55.00	44.80	-10.20	61.49	6.49	51.10	-3.90
8302	46.74	45.22	-1.52	56.33	9.59	50.45	3.71
8303	46.90	42.57	-4.33	56.61	9.71	48.18	1.28
8304	42.18	36.89	-5.29	53.03	10.85	46.20	4.02
Root Mean Square Error		10.54		6.97		6.67	
Theil's U_2 Coefficient		1.82		1.20		1.15	

Table 11. Ability of Estimated Composite Models to Predict the Direction of Quarterly Change in Hog Prices, 1982-83

Actual Price Movement	Reduced Form and ARIMA Model		Recursive Model and Transfer Function		Reduced Form and Transfer Function	
	Rising	Falling	Rising	Falling	Rising	Falling
Rising	1	3	4	0	3	1
Falling	2	2	1	3	1	3
Percentage of Correct Tracking	38%		88%		75%	

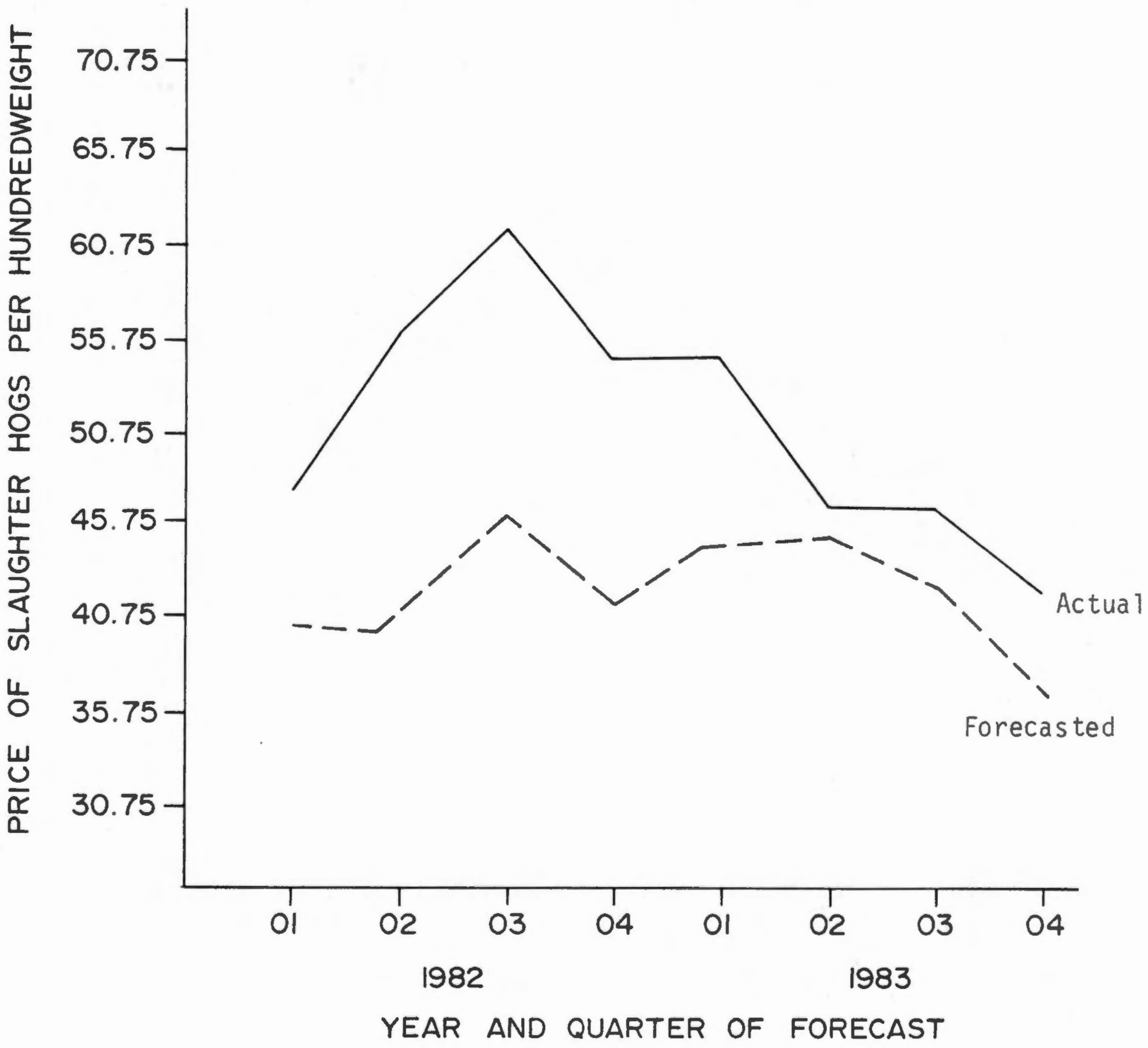

Figure 5. Actual Versus Predicted Prices of Composite Model 1.

Composite Model 2: Recursive Model and the Transfer Function

Table 10 also shows the results for the composite model made up of the transfer function and the recursive model. Unfortunately, Theil's U_2 coefficient of 1.20 again indicates that a naive no-price-change model would have been more accurate over the eight-quarter forecast period. Thus, the composite model's U_2 was worse than the transfer function's U_2, but better than the 2.36 calculated for the recursive model. The RMSE for the composite model was $6.97 which was $.54 larger than that of the transfer function forecasting alone. All eight forecasts were above actual price. Forecast errors ranged from a low of $.17 to a high of $10.85. Table 11 indicates that this composite model was able to track the direction of price movement 88 percent of the time. This was better than the transfer function forecasting alone, but worse than the 100 percent accuracy rate of the recursive system when used alone. Actual versus forecasted prices are shown in Figure 6.

Composite Model 3: Reduced Form Model and the Transfer Function

The last composite model of Table 10 is the reduced form and the transfer function. It was more accurate than either composite model 1 or 2. Theil's U_2 coefficient was calculated to be 1.15 which again indicates that a naive no-price-change model would have forecasted better. The RMSE was $6.67 which was the second lowest of all models. However, the transfer function used alone had both a lower RMSE and Theil's U_2 coefficient.

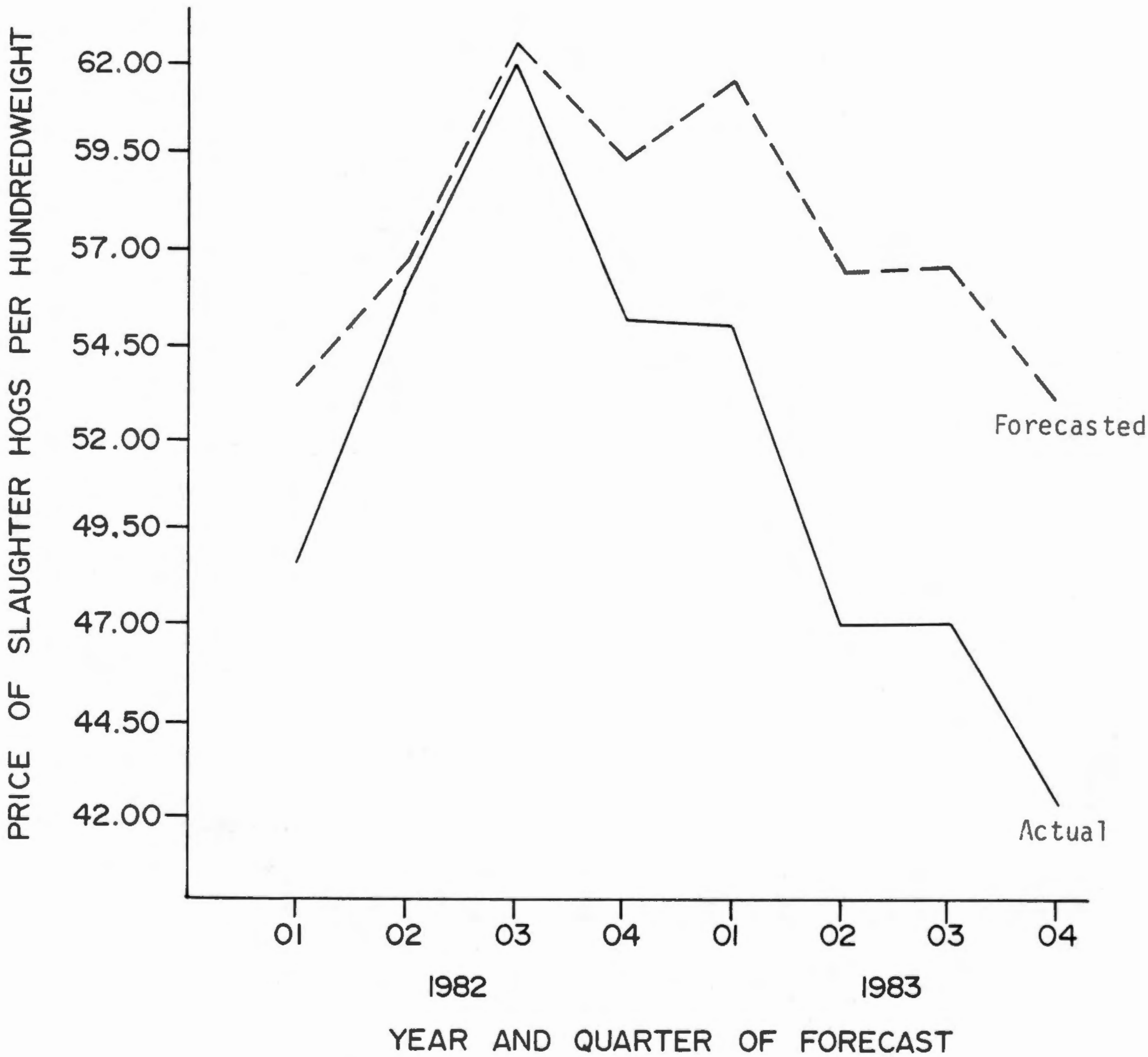

Figure 6. Actual Versus Predicted Prices of Composite Model 2.

The errors in the forecasts ranged from a low of \$1.28 to a high of \$12.12. Forecasted values were below actual price five times and above the remaining three times. Table 11 indicates that this composite model was able to track the direction of price change 75 percent of the time which was third best of all the models. Actual versus forecasted prices are shown in Figure 7.

VI. SUMMARY OF FORECASTING RESULTS

In general, no individual or composite model was able to provide *ex ante* forecasts which were more accurate than a naive no-price-change model. However, the recursive model and composite models two and three were able to predict the direction of price change more than half the time (100, 88, and 75 percent, respectively) during the out-of-sample period. Thus, some utility might be gained by hog producers who would like to know whether the direction of price change would be up or down. More analysis over a longer forecasting period would be needed before the reliability of the model to predict the direction of price change could be ascertained.

Comparing the *ex post* forecasts with the *ex ante* forecasts indicates that a model which forecasted well over the latter part of the period of fit does not necessarily forecast well outside the period of fit. The calculated Theil's U_2 coefficient for the *ex post* forecasts indicated that all models forecasted more accurately than a naive no-price-change model. Yet, all *ex ante* forecasts were less accurate than a naive no-price-change model.

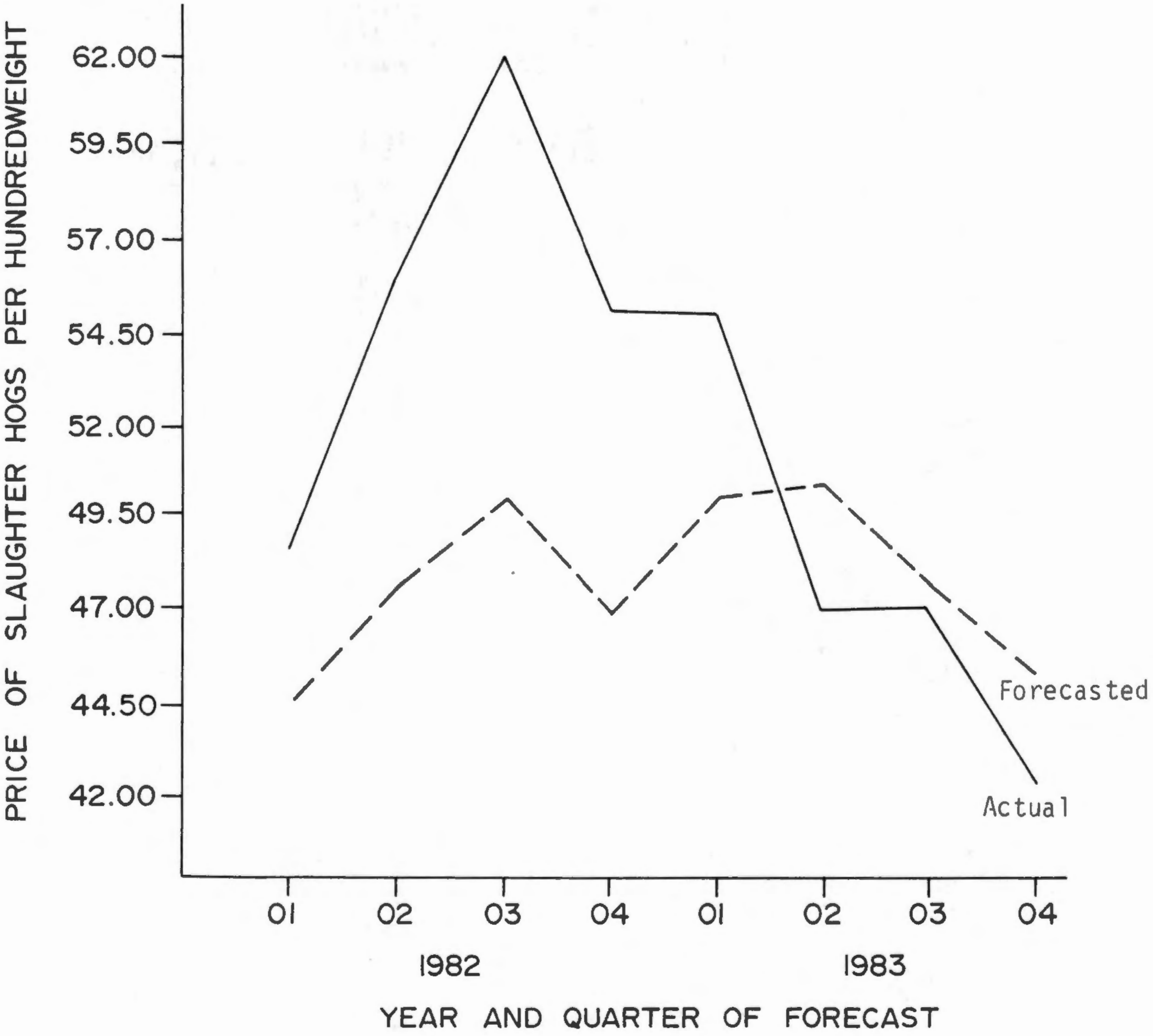

Figure 7. Actual Versus Predicted Prices of Composite Model 3.

In addition, in this study the ranking of ex post forecasts by Theil's U_2 coefficient of the individual and composite models was not the same as the ex ante rankings. Over the last eight quarters of the period of fit the recursive and transfer function composite model forecasted best followed by the reduced form and transfer function composite model, the transfer function, the reduced form and ARIMA composite, the recursive model, the reduced form model, and the ARIMA model. However, over the ex ante forecasting period, the best forecasting model was the transfer function followed by the reduced form and transfer function composite, the reduced form and ARIMA composite, the ARIMA model, the reduced form and ARIMA composite, the reduced form model, and the recursive model. The simultaneous model was not tested by forecasting since its estimation results were highly suspect.

From these results it might be loosely stated that either time series models or composite models would provide better out-of-sample forecasts than econometric models used alone, but that neither class of models is likely to forecast better than assuming that next quarter's price will be the same as this quarter's price.

CHAPTER V

GENERAL SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

I. GENERAL SUMMARY

This study attempted to develop price forecasting models which could be utilized by hog producers to predict the price of slaughter hogs one quarter in advance. Several models were constructed which included both time series and econometric models, as well as composite models which consisted of one model of each type. Econometric models included a recursive system, a single equation model which was the reduced form of the recursive system, and a simultaneous model whose variables were suggested by Hayenga and Hacklander.

Time series models consisted of an univariate autoregressive integrated moving average model (ARIMA), and a transfer function model which utilized an exogeneous input variable. Three composite models were also developed. The first composite model combined the reduced form and the ARIMA model. The second composite consisted of the recursive system and the transfer function. And the third composite model was comprised of the reduced form and the transfer function models.

Each model was fitted to quarterly data over the period from 1963 through 1981. Quarterly forecasts were made over 1982 and 1983 (*ex ante* forecasts). Over this forecast period the root mean square error (RMSE) of each model was calculated as well as Theil's U_2

coefficient. In addition, each model was ranked as to how well it could predict the direction of price change from quarter to quarter. Also included in the analysis were the root mean square error and Theil's U_2 coefficient for the last eight observations of the fitted data (*ex post* forecasts).

Composite models were developed by combining forecasts from two models. This was accomplished by forecasting with each model over the last thirty observations. The forecasts from two individual models were then regressed against actual price in order to determine the weight given to each individual forecasting model.

The recursive model consisted of three equations: hog slaughter in millions of pounds (carcass basis), per capita change in pork cold storage, and the price of hogs deflated by the producer price index. The hog supply equation was hypothesized to be a function of the price of dogs deflated by the index of prices paid by farmers, and the ratio of the price of feed to the price of hogs. Each variable was fitted as a second order polynomial distributed lag eight quarters in length. In addition, seasonal dummy variables were added to allow for changes in the intercept term.

The per capita change in frozen pork storage was hypothesized to be a function of per capita hog slaughter, the per capita level of frozen pork storage at the beginning of the current quarter, and the retail price of pork cuts lagged one quarter deflated by the consumer price index. This last variable was used as a proxy for the current price of pork cuts.

The final equation of the system, the demand or price equation, hypothesized that the price of hogs deflated by the producer price index is a function of, among other variables, per capita quantity of hogs slaughtered and the change in per capita pork cold storage. Since packers have fixed capacity in the short run, and since a minimum work week is guaranteed, it was hypothesized that this capacity pressure would affect the price packers were willing to pay for slaughter hogs. A proxy variable for this pressure, the number of fully utilized work days per quarter, was introduced into the equation.

Since the demand from packers is derived demand, several variables from the retail demand curve have been included in the packer demand equation. Foremost among these are population and income. However, since quantity variables were divided by population in order to account for that variable in each appropriate equation, income was entered on a per capita basis also. Since beef is a substitute which competes directly with pork for the consumer's dollar, a per capita cattle-on-feed variable lagged one quarter was entered to capture this effect. As in the other two equations of the recursive model, seasonal dummy variables were added to allow for shifts in the intercept term.

Since the above system was recursive it was easy to substitute terms from the first two equations into the third or price prediction equation, and thus collapse the three equation system into one. Specifically, terms for per capita hog slaughter and per capita change in pork cold storage were substituted into the third equation to create the single equation reduced form model. This became the second price prediction model.

The third econometric model developed was a simultaneous equation system whose variables were suggested by Hayenga and Hacklander. This model consisted of five equations: beef and hog production, a change in pork cold storage equation, and two demand equations to predict the price of beef and hogs.

Beef production was hypothesized to be a function of current beef price deflated by the producer price index and the inventory of cattle-on-feed lagged one quarter. This last variable consisted of steers 700 to 1100 lbs. and heifers 500 to 900 lbs. Likewise, hog production was hypothesized to be a function of the current hog price deflated by the producer price index and the inventory of market hogs 60 to 219 lbs. lagged one quarter. Both supply equations contained seasonal dummy variables.

The change in per capita pork cold storage equation was much like the corresponding equation in the recursive system. This equation was hypothesized to be a function of per capita hog production and the beginning-of-quarter level of per capita pork cold storage. However, instead of the retail price of pork cuts deflated by the consumer price index, the current hog price deflated by the producer price index was included in this equation. The rationale was that the model would determine the current price paid by packers, and that this price would be highly correlated with the retail price of pork, which in turn would influence the amount of pork moving into or out of storage.

The price of cattle deflated by the producer price index was hypothesized to be a function of the quantity of beef and pork supplied (substitutes), per capita income, and the number of fully utilized work days per quarter. This last variable was again a proxy variable to represent the capacity pressure to which packers might react. The price prediction equation for hogs was exactly the same as that for beef, except that the per capita change in frozen pork storage was added to take into account additions to or substractions from the current quantity supplied.

By definition, the ARIMA model was hypothesized to be an autoregressive function of past hog prices and a moving average function of the random error term in the price series. The transfer function was hypothesized to contain the same two components plus an input variable. For the transfer function the input variable chosen was the number of sows farrowing. Composite models were developed from the forecasts generated by the above five models.

Results of the fitted models were somewhat disappointing. For the recursive model the polynomial distributed lags were not significant except for the constant term of the second degree polynomial of the deflated price of hogs variable. Each polynomial was converted to a coefficient for each lag (for a total of eight lags) and of these, lags one, two, and six of the deflated price of hogs and lags six and seven for the price of feed variable were significant at the 5 percent level. Since this equation exhibited a high degree of autocorrelation among the residuals it was estimated by the SAS AUTOREG procedure which is equivalent to generalized least squares. In addition, the intercept term and all three dummy variables were significant.

The change in storage equation had two significant continuous explanatory variables. These were per capita hog supply and the beginning-of-quarter frozen pork storage. In addition, each variable had the expected sign. Only the third quarter dummy variable was found to be significant. No significant first order autocorrelation was present in the residuals.

For the demand equation, the intercept and the first and second quarter dummy variables were significant at the 5 percent level. In addition, per capita hog slaughter and the change in storage variable were significant and had the expected signs. Per capita income, cattle-on-feed lagged one quarter, and the number of fully utilized work days were not statistically significant. Since the residuals from ordinary least squares were autocorrelated this equation was estimated by the AUTOREG procedure.

For the reduced form equation the third term of the per capita deflated price of hogs and the second and third terms for the price of feed were significant in the polynomial distributed lags. Of the individual coefficients, lags four, five, six, and eight on the price of hogs, and lags one and two on the price of feed were significant at the 5 percent level. In addition, per capita cattle on feed lagged one quarter, and the first and third seasonal dummy variables were significant at the 5 percent level. Thus, the intercept term, the second quarter dummy variable, the deflated retail price of pork lagged one quarter, and the first and third seasonal dummy variables were significant at the 5 percent level. Thus, the intercept term, the second quarter dummy variable, the deflated retail price of pork lagged one quarter, per capita income and the number of fully utilized work days were not statistically significant.

The least satisfying of the estimated models was the simultaneous equation system. For the quantity of beef supplied equation the cattle-on-feed variable was significant and had the expected sign, but the deflated price variable had a negative sign and was significant. The intercept and the first and second seasonal dummy variable were also significant.

The hog production equation had only one significant variable at the 5 percent level. That was the inventory of market hogs lagged one quarter.

The change in per capita pork cold storage had one continuous variable which was significant. Beginning-of-quarter pork cold storage was significant and had the expected sign. Both the deflated hog price and per capita hog slaughter were insignificant. Of the intercept term and the three seasonal dummy variables only the third quarter dummy variable was significant.

The beef demand equation also was less than satisfactory. While the quantity of beef supplied was significant its sign was positive, and per capita income was significant but its sign was negative. All other variables were not significant. However, the unexpected signs on the two variables cast doubt on the equation.

For the hog demand equation only the quantity of beef supplied was significant but the sign was unexpectedly positive. Overall the simultaneous equation model estimated by two-stage least squares was the least appealing model, probably due to model misspecification, and was not used to make forecasts.

Over the fitted data for 1980-81, eight forecasts were made and compared to the actual prices. For these ex post forecasts, the best forecaster based upon Theil's U_2 coefficient was the composite model made up of the recursive model and the transfer function model. For this model the root mean square error (RMSE) was \$2.84 and Theil's U_2 coefficient was .41. A close second was Composite Model 3 made up of the reduced form and the transfer function model. The RMSE was 3.72 and the U_2 coefficient was .53. Following this was the transfer function with a RMSE of 3.99 and a Theil's U_2 coefficient of .55. The next best forecasting model was a composite forecast made up of the reduced form and ARIMA models. This composite model had a RMSE of 4.06 and a Theil's U_2 of .58. The fifth best forecasting model was the recursive system which had a RMSE of 4.14 and a Theil's U_2 of .59. This was closely followed by the reduced form model with a RMSE of 4.19 and a Theil's U_2 of .60. The seventh best forecasting model was the ARIMA model. The calculated RMSE and Theil's U_2 coefficient were 4.79 and .66, respectively.

Over the period of ex ante (out-of-sample) forecasts the best predicting model was the transfer function with a RMSE of \$6.43 and a Theil's U_2 coefficient of 1.11. A close second was the composite model made up of forecasts from the reduced form and the transfer function. The RMSE of this model was \$6.67 and Theil's U_2 coefficient was 1.15. Hence, the transfer function forecasting alone was slightly better than the best composite model.

The third best model was a composite forecast made up of the recursive model and the transfer function model. The RMSE was \$6.97 and Theil's U_2 coefficient was 1.20. The fourth best model was the ARIMA function. Its RMSE was \$10.11 and its U_2 coefficient was 1.75. The fifth best forecaster was the composite model made up of the reduced form equation and the ARIMA model. The RMSE for the eight forecasts was \$10.54 and the U_2 coefficient was 1.82.

The two econometric models forecasting alone made the poorest price predictions for the *ex ante* forecasts for 1982-83. The reduced form model had a RMSE of \$12.11 and a Theil's U_2 coefficient of 2.09. The recursive system did slightly worse, with a RMSE of \$13.66 and a U_2 coefficient of 2.36.

II. CONCLUSIONS AND RECOMMENDATIONS

In general no individual or composite model was able to provide *ex ante* forecasts better than a naive no-price-change model. Hence, more accurate price forecasts could have been obtained if one had assumed that next period's price will be the same as this period's price. However, three models could track the direction of price change better than half the time. These models were the recursive system, the composite model made up of the recursive and transfer function models, and the composite model made up of the reduced form and transfer function.

Thus, no model, individual or composite, utilized in this study could have helped hog producers predict the price of slaughter hogs one quarter in advance. However, several conclusions can be noted. In general the time series models and the composite models forecasted better than the econometric models used alone. Thus, if one were attempting to forecast for planning purposes, one would probably make smaller errors if a composite forecast were used as opposed to an individual forecast. Hence, model builders might be well-advised to expand their efforts in building both time series and econometric models.

In this study the transfer function was a better forecaster of hog price than the ARIMA model. Hence, researchers might concentrate on fitting this type of model, perhaps with another input variable, or with an additional input variable added to the model.

Further work might be carried out in reformulating the recursive system. Specifically the hog supply equation was not adequately represented by a polynomial distributed lag model. Previous research in this area had obtained models with a high R^2 statistic by deleting the intercept term. Hence, careful thought should go into the specification of this equation. Since the current quantity supplied of market hogs is roughly a function of sows farrowing two quarters ago, it seems reasonable that an econometric or time series model could be specified to predict this variable.

In addition to the above, it would probably be worthwhile if all or most of the variables suggested by economic theory were included in the models. Specifically, in retrospect the demand equation of the recursive system probably should have included a variable representing poultry. In this way one could attempt to include the major variables which might explain the variation in the dependent variable.

In addition, in the change in storage equation one might include the price of the futures frozen pork belly contract since packers may use this contract at certain times of the year in order to "lock in" a price for their product. The inclusion of this variable would necessitate fitting the model to data from a shorter time period. However, the argument can be made that pork production has undergone structural change due to technological advance over the past few years. Hence, from this point of view a shorter time period for estimation could be justified.

Finally, one might wish to try some nontraditional models. For example in the recursive system one might try fitting a transfer function for both hog production and the change in storage equation. These predictions could then be used in a traditional econometric model to predict hog price. Also it could be worthwhile to update each forecasting model after the data become available. Hence, before each quarter's price prediction, the data for the latest quarter could be added to the data set and the model reestimated.

LIST OF REFERENCES

LIST OF REFERENCES

Bates, J. M. and C. W. J. Granger. "The Combination of Forecasts." Operations Research Quarterly, 20(1969), pp. 451-468.

Bessler, D. A. and J. A. Brandt. Composite Forecasting of Livestock Prices: An Analysis of Combining Alternative Forecasting Methods. Purdue University Agricultural Experiment Station Bulletin 265, 1979.

Bourke, I. J. "A Comparison of Price Forecasting Models for Manufacturing Beef in the U.S.A." Ph.D. Dissertation, Texas A & M University, 1978.

Bressler, R. G. and R. A. King. Markets, Prices, and Interregional Trade. Raleigh: Norman-Weathers, 1978.

Brown, R. A. and W. D. Purcell. Price Prediction Models and Related Hedging Programs for Feeder Cattle. Oklahoma State University Agricultural Experiment Station Report B-734, 1978.

Edwards, R. A. "An Evaluation of Econometric Commodity Models and Their Roles in a Proposed Price Reporting System." Ph.D. Dissertation, University of Georgia, 1977.

Granger, C. W. J. and P. Newbold. Forecasting Economic Time Series. New York: Academic Press, 1977.

Hayenga, M. and D. Hacklander. "Monthly Supply-Demand Relationships for Fed Cattle and Hogs." American Journal of Agricultural Economics, 52(1970), pp. 535-544.

_______. Short-Run Livestock Price Prediction Models. Michigan State University Agricultural Research Bulletin 25, 1970.

Hobbs, J. C. and C. W. Herndon. "Hog-Corn Price Ratio Vs. Separate Hog and Feed Prices in Estimating Pork Production." Oklahoma State University Agricultural Experiment Station Current Farm Economics 56-3, 1983.

Hoffman, G. H. "Monthly Forecasting of the U.S. Livestock Industry." Ph.D. Dissertation, University of Minnesota, 1979.

Jenkins, G. M. "Practical Experiences with Modelling and Forecasting Time Series." Forecasting. Ed. O. D. Anderson. New York: North-Holland, 1977.

Ladd, G. W. and G. Karg. Econometric and Programming Analysis of the Beef-Pork Marketing Sector. Iowa State University Agricultural Experiment Station Research Bulletin 577, 1973.

Leuthold, R. M., A. J. A. MacCormick, A. Schmitz, and D. G. Watts. "Forecasting Daily Hog Prices and Quantities: A Study of Alternative Forecasting Techniques." Journal of the American Statistical Association, March 1970, pp. 90-107.

Meilke, K. D. "Another Look at the Hog-Corn Ratio." American Journal of Agricultural Economics, 59(1977), pp. 216-219.

Myers, L. H., J. Havlicek, and P. L. Henderson. Short-term Price Structure of the Hog-Pork Sector of the United States. Purdue University Agricultural Experiment Station Research Bulletin 885, 1970.

Myers, Walter. "Combining Statistical Techniques with Economic Theory for Commodity Forecasting." American Journal of Agricultural Economics, 54(1972), pp. 785-790.

Nelson, C. R. Applied Time Series Analysis for Managerial Forecasting. San Francisco: Holden-Day, 1977.

Pindyck, R. S. and D. L. Rubenfield. Economic Models and Economic Forecasts. New York: McGraw-Hill, 1976.

Purcell, J. C. and J. C. Elrod. Price Related Risk in the Swine-Pork Sector. University of Georgia Agricultural Experiment Station Research Report 154, 1973.

Roy, S. K., R. I. Foote, and G. Sadler. Econometric Models of Quarterly and Monthly Live Hog Prices. Texas Technical University College of Agricultural Science Publication T-1-149, 1976.

Sullivan, J. D., C. Y. Liu, and W. Vincent. A Systems Analysis of the Hog-Pork Subsector. USDA, ERS, Technical Bulletin 1535, 1976.

Teigen, L. "Costs, Loss and Forecasting Error: An Evaluation Models for Beef Prices." Ph.D. Dissertation, Michigan State University, 1973.

Tweeten, L. G. Foundations of Farm Policy. 2nd ed. Lincoln: University of Nebraska Press, 1979.

United States Department of Agriculture. Agriculture Prices Annual Summary, ESCS Crop Reporting Board. Washington, D.C., 1963-1983.

_______. Cold Storage, ESCS, Crop Reporting Board. Washington, D.C. Selected issues, 1963-1983.

______. *Hogs and Prices*. ESCS, Crop Reporting Board. Washington, D.C. Selected issues, 1963-1983.

______. *Livestock and Meat Statistics*. Statistical Bulletin 333 and 522, ESCS, Washington, D.C., 1963-1983.

______. *Livestock and Poultry Outlook and Situation Report*. ERS. Washington, D.C. Selected issues, 1963-1983.

Zellner, A. and F. Palm. "Time Series Analyses and Simultaneous Equation Econometric Models." *Journal of Econometrics*, 2(1974), pp. 17-54.

VITA

William Noel Blisard was born in Oakland, California, on November 9, 1949. He attended schools in that city, and graduated from Fremont High School in June 1967. Four years later he received a B.A. in English from Humboldt State University in Arcata, California.

Since that time he has worked at a variety of jobs. From May 1974 to April 1976 he served in the United States Peace Corps at Udorn Teacher's College in Udorn, Thailand. Later he served as an English instructor for the Royal Iranian Army. It was in Iran where he met and married Marcia Clark.

In June 1979 he enrolled in the graduate program in Agricultural Economics at The University of Tennessee. The M.S. degree was awarded in December 1981, and the Ph.D. degree in March 1985.