A PRICE DISCOVERY STUDY OF CASH SLAUGHTER HOG PRICES, WHOLESALE PORK PRICES, AND LIVE HOG AND PORK BELLY FUTURES MARKET PRICES

By

BOO–YONG SONG

Bachelor of Science
Dongguk University
Seoul, Korea
1982

Master of Science
Dongguk University
Seoul, Korea
1984

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
July, 1993

A PRICE DISCOVERY STUDY OF CASH SLAUGHTER HOG PRICES, WHOLESALE PORK PRICES, AND LIVE HOG AND PORK BELLY FUTURES MARKET PRICES

Thesis Approved:

Clement E. Ward

Thesis Adviser

Joseph E. Williams

Stephen R. Koontz

W. Ward

Thomas C. Collins

Dean of the Graduate College

ACKNOWLEDGEMENTS

I wish to express a thank you to each of the members of my dissertation committee. I extend cordial appreciation to my research advisor, Dr. Clement Ward for providing the ideas, and his helpful comments and encourgement. I will never be able to repay his generosity and kindness. My sincerest appreciation goes to my advisor, Dr. Joseph Williams, for encouragement and advisement throughout the OSU Ag. Econ. program. I also wish to express my gratitude to Dr. Stephen Koontz for his guidance and detailed comments. I am also thankful to Dr. William Warde in the department of Statistics for his helpful advisement. I would like to thank all faculty, my fellow graduate students and staff in the Ag. Econ. department.

I cannot forget saying thanks to the many individuals, while not directly involved with my thesis, who encouraged me. I am particularly grateful to Dr. Mike Woods for the opportunity to work with him about rural regional development. I have enjoyed many interesting discussions for rural development with Him. His support is greatly appreciated. I always will be especially thankful for the instruction and guidance of Drs, Chonghwan Chu, Byeongdong Lee, Hongchul Kim, and Byeongchan Cho at Dongguk University in Seoul, who convinced me to study Agricultural Economics.

With my feeling of both happiness and sadness resulting from this thesis, I am also grateful to a random friend of mine, Jinsoo Yoo, who has encouraged me throughout the Ph. D program. Moreover, my special thanks go to my wife, Sunkyung Lee, who is devoted to me and our lovely children, Junhyuk and Jinha.

Without their love, patience, and cheerfulness, the old man will not reach the culmination of his graduate experience.

My heartfelt appreciation is expressed to my brother, Chanyong, and sister–in–law. They have a larger part in this than they know. My final warmest thanks are reserved for my dear mother and father. Their endless love, support, sacrifice, encouragement and understanding have alleviated the burden I could have had in my life and motivated me to go further.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

LIST OF FIGURES

Page

CHAPTER I

INTRODUCTION

Background

The study of price discovery, including the lead–lag or causal relationships among farm level, wholesale and futures prices of a livestock commodity is of importance to producers, wholesalers, packers, and retailers. Such a study provides information on pricing efficiency to them. The importance has increased since livestock and meat contracts have been traded on futures markets, although futures market trading in livestock and meat has a short history compared to futures trading market in grains. The introduction in 1961 of a pork belly contract, which was the first meat complex contract, and in 1966 of a live hog contract by the Chicago Mercantile Exchange (CME) increased the importance of price discovery.

The development of meat futures contracts encountered the difficulty of defining a contract, because meat products cannot be easily defined as commodities due to the heterogeneity of the animals from which they are made. Live hogs cannot be stored, but they can be defined as commodities if their production techniques reach some standardized quality such as optimum slaughter weight, for slaughter–ready hogs. Another difficulty in defining a contract for live hogs was setting delivery terms because the major cash markets are dispersed. This problem was resolved but with reservations about physical delivery. Increasing uniformity in the age and weight of animals at slaughter increased the possiblility of viewing a range of animals and meat products as commodities. Pork bellies, the first meat complex contract by the CME, are treated as a commodity with defined quality standards.

Contracts on live hogs provide hedging opportunity for farmers more than contracts on meat products. Futures contracts have been developed because of the volatility of commodity prices in the cash or spot market. The volatility results from variations in the level of hog production, changes in the corn–hog price ratio, inventories, and other factors.

Despite being related, the futures prices of pork bellies and live hogs are determined differently. Hogs are not storable, and so the prices of the various months of the live hog futures contract are not related by storage costs. However, prices for the various months of the pork belly futures contract are related by carriage, since pork bellies are stored in refrigerated warehouses.

Live hogs are slaughter–ready animals. Hence the price of these in the spot market is the price paid by abattoirs for live hogs for immediate slaughter. But meat pricing at the wholesale level has been less well understood by producers. Wholesale prices are determined in part by retailers and reflects consumers' demand. Since pork bellies, one of several wholesale pork products from each hog are the underlying commodity for a futures contract, these wholesale pork prices are expected to be affected by futures market prices.

There are some interesting things about the performance of futures markets. It is known that futures markets reduce cash market price volatility. Futures markets are also used to forecast cash prices. Ironically, futures trading reduces cash market volatility, but futures prices are also highly volatile (Atkin, 1989).

In general, futures prices are known to have an effect on pricing efficiency (Brorsen, Bailey and Richardson, 1984). That is because pricing efficiency involves the availablility of information in markets. According to Fama (1970), efficient prices can be divided into three categories with different types of available information; weak, semi–strong, and strong efficient prices. Weakly efficient prices are formed with the information set consisting of past prices, and semi–strong efficient prices reflect all publicly available information. Strongly efficient prices includes all relevant information. Accordingly, the study of price discovery in hog and pork markets are performed in the weak efficiency sense. That is, price discovery is investigated by using past prices.

There is little research to describe and document the role of live hog futures or pork belly futures trading in the pricing of cash slaughter hogs and wholesale loins, hams, and bellies. But relatively similar research for cattle has been conducted (Oellerman and Farris, 1985; Koontz, Garcia, and Hudson, 1990). Previous research uses different approaches and tests the role of futures markets in risk transfer and price discovery. Previous livestock futures research on risk transfer and price discovery used five methods; cross–spectral analysis, a univariate residual cross–correlation approach, simultaneously dynamic analysis, vector autoregressive (VAR) analysis, out–of–sample performance and cointegration between price series.

Rausser and Cargill (1970) used spectral analysis to determine whether any significant lead–lag relationships exist among the time–series data that normally are employed to illustrate cycles in broilers. They found no evidence of a meaningful lead–lag pattern among various time series. Barksdale, Hilliard and Ahlund (1975) studied beef prices at different market levels. They reported the lead–lag relations among beef prices at four different market levels — feeder, live cattle, wholesale and retail. Spectral analysis revealed the direction of influence and that prices at the feeder, live animal and wholesale levels move together without any time lag.

Miller's two studies are representative of research on price discovery using cross–correlation analysis in the livestock market. Miller (1979) examined alternative price discovery mechanisms for beef. He concluded that farm prices lead wholesale prices by about one week, and in turn, wholesale prices lead retail prices by about three weeks. Miller (1980) also applied univariate residual cross–correlation analysis to pork prices at the retail, wholesale, and farm levels. In the pork markets, farm level prices lead wholesale prices by up to 2–3 weeks and wholesale prices lead retail prices by 2–3 weeks. Faminow (1981) also tested the evidence of a lead–lag relationship between two wholesale beef price quotations using residual cross correlation analysis.

Results from dynamic analysis on price discovery are usually compared to those from other methods. Oellerman and Farris (1985) investigated the lead–lag relationship between changes in futures and cash prices for live cattle using dynamic analysis. Their results confirm the role of futures markets in that futures prices lead cash prices. Garbade and Silber (1983) specified and estimated a simultaneous dynamic model which describes the interrelationship between cash prices and futures prices for storable commodities.

Price discovery analyses using VAR are somewhat varied due to the application of different test methods. The Sims' test has been used but some drawbacks have been identified (Jacob et al., 1979). Bessler and Brandt (1982), and Hudson (1984) used Geweke's causality test to analyze lead–lag relationships in price discovery. Bessler and Brandt provide specific evidence on leads and lags for several variables in the hog and cattle markets. They did not consider futures prices but livestock prices and various causal variables to examine the lead–lag relationships using related topics of exogeneity and Granger causality. A specific feature of their research is that they test causal relationships using the out–of–sample forcasting method.

The above empirical methods mentioned can be applied to the role of futures markets in providing price information. Some are related to the lead–lag relationships among prices and markets. Another popular method is tests for cointegration among price series in order to identify the relationship of futures markets to the cash market. Bessler and Covey (1991) used cointegration as well as out–of–sample causality tests. Schroeder and Goodwin (1991) analyzed the price discovery role of live hog cash and futures markets, and examined the longer–run stability of live hog futures prices to cash market prices using the concept of cointegration. The authors found that causality exists from the futures to the cash market, that the two futures price series operate independently, and that the long–term basis is nonstationary.

Problem Statement

It is important that one knows the conceptual nature of the markets, and the interrelationship between prices in the different markets. The major part of this study includes examining the price discovery process among: live hog and pork belly futures prices, cash slaughter hog prices, and wholesale pork prices, for loins, hams, and bellies. A schematic diagram in Figure 1 shows the important roles and interrelationships between each market agent and price series. And Figure 2 shows interrelationships among the different prices. In Figure 2, causal direction for each arrow will be determined in this study.

There exists limited information on the exact nature of the interrelationships among each price series in the hog and pork markets. Most previous price discovery studies conclude there is strong evidence of the relationship between live hog futures and cash slaughter hogs.

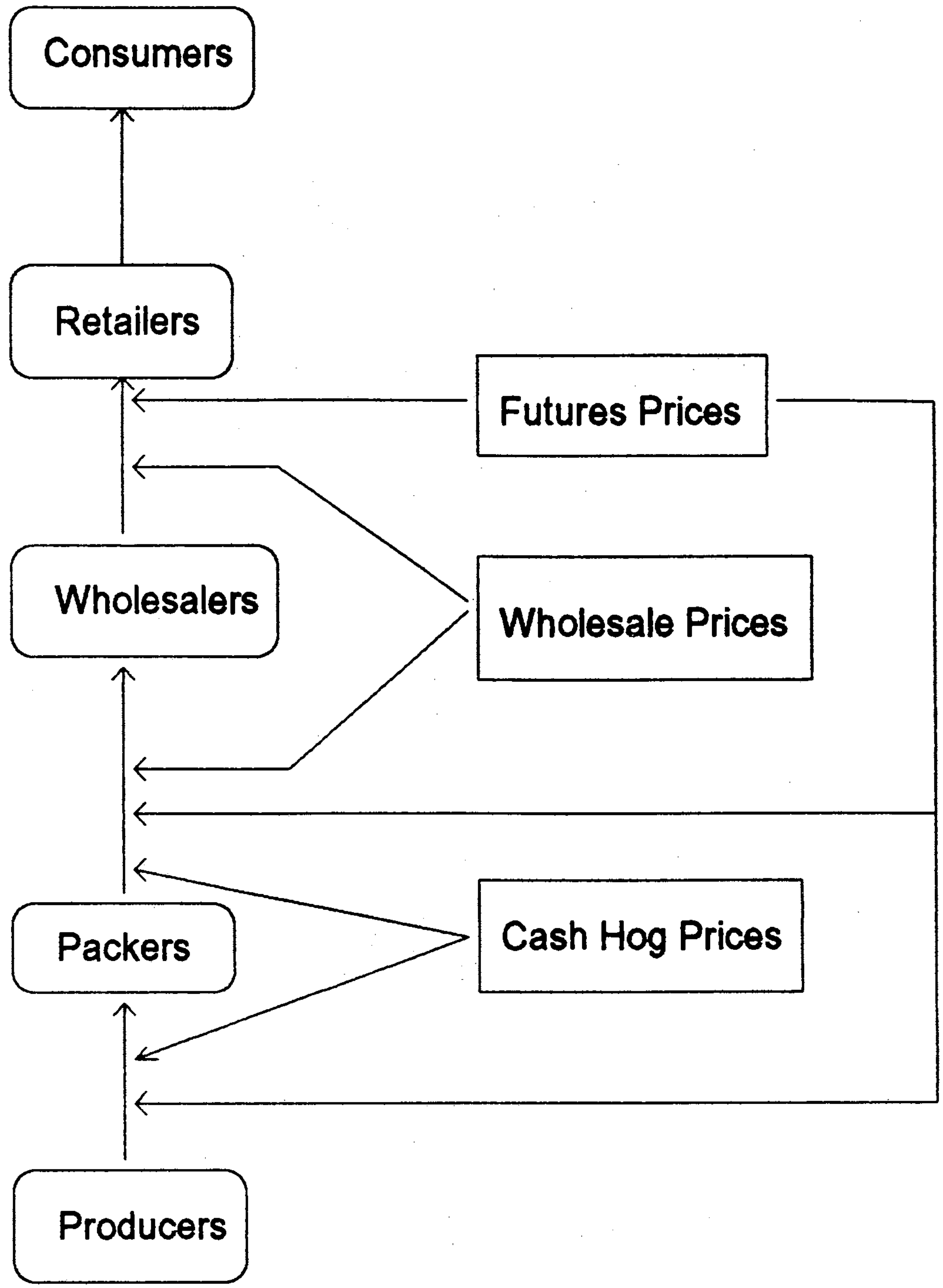

Figure 1. A Schematic Diagram of Interrelationships between Each Market Agent and Price Series

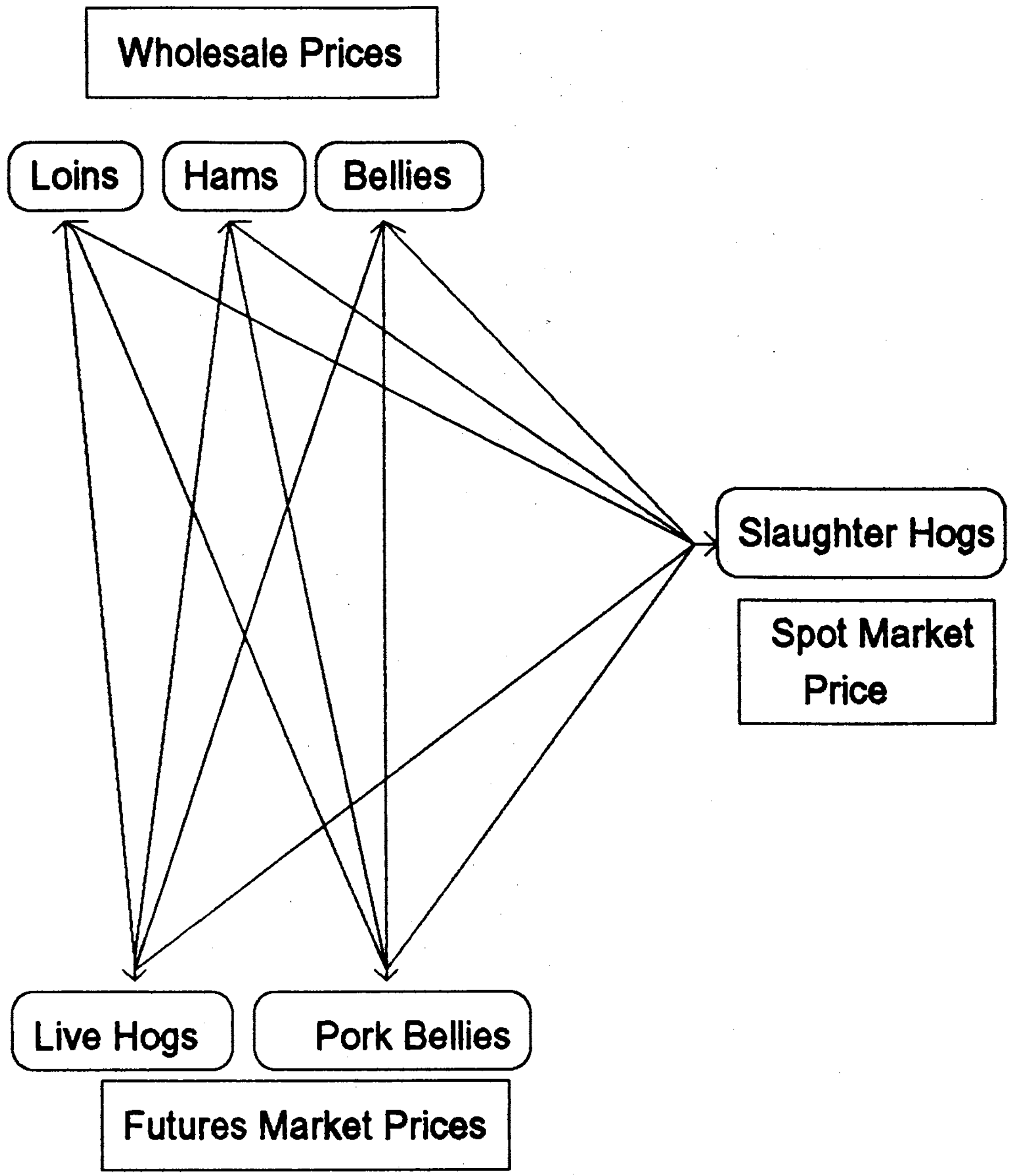

Figure 2. Interrelationships among Different Market Prices

Due to limited information, a number of problems are suggested. Among three markets — cash, wholesale and futures — which market is contributing most to price discovery of slaughter hogs? Similarly, which market plays the most important role in price discovery of each wholesale pork product? Is it possible to observe feedback relationships among the three price series? If futures market prices strongly lead cash or wholesale prices, then which futures market — pork belly or live hog — is most important for price discovery? For the above questions, has price discovery changed over time?

Another question is related to short–run and long–run price relationships between daily cash, wholesale and futures prices using cointegration. A discussion follows on how cointegration reflects on the economic interrelationships of the price discovery process.

The short or long run relationships between prices need to be investigated because a market with leading price quotations must be a more efficient mechanism for rational price formation than another market with following price quotations. Since an efficient market is defined as one in which the current price fully reflects all relevant available information (Stein, 1981), this study will provide information in which market prices are more efficient.

Information is costly. One concern then is which markets have the lowest costs for gathering new supply and demand information? Markets which have the lowest costs for gathering new market information will be price discovery leaders. Newly discovered prices then become public goods because prices are publicly reported and are reflected in other markets. So, other markets will be price discovery followers. Sometimes, different markets gather different information efficiently, which implies feedback. This efficient market theory discussion underlies causality testing methods.

Based on the literature, this price discovery study could be done by using several different methods. When these methods are applied to the above questions, which method is best to analyze price discovery? Is it possible to determine one preferred method? If each method has independent properties, one could argue about the statistical problems and identify advantages and disadvantages with each method.

In addition, there may exist problems due to seasonality in some price series. If seasonality exists in futures or cash price series, can we conclude that the answers to the above questions are the same as in the absence of seasonality?

The questions mentioned above can be summarized and categorized into two major problems: (1) Are futures market prices centers of pricing in cash and wholesale markets, as is generally expected? Are futures market prices more efficient than cash hog market prices or wholesale pork prices? And, (2) do futures market prices forecast cash or wholesale prices? A broad base of information results from answering the above questions, for hog and pork producers, marketing agents, and economists.

Objectives

Futures markets are considered to perform two functions: price setting through speculative activity (Tomek and Gray, 1970), and hedging (Working, 1962). Any person who owns a commodity automatically assumes risks. Hog producers have seen potential profits vanish as market prices decline. A packer who has a commitment to supply meat to a wholesaler or retailer for a stipulated price may see his potential profits vanish as market prices move up. Hedging in the futures market is the transfer of price risk to others, i.e. speculators who are willing to assume price risk. Therefore, the documentation and description of the interrelationships between prices in cash, wholesale, and futures markets are important. The overall objective is to analyze the price discovery process and determine which markets provide more information to each marketing agent in cash hog and pork markets.

To accomplish the overall objective, specific objectives are:

(1) to examine the lead–lag relationship between live hog futures market prices and cash slaughter hog prices;

(2) to examine the lead–lag relationship between pork belly futures market prices and wholesale pork belly prices;

(3) to examine the lead–lag relationships between wholesale pork prices — loins, hams, and pork bellies, and cash slaughter hog prices;

(4) to examine the lead–lag relationships between futures market prices — live hog futures and pork bellies futures market prices, and wholesale pork prices — loins, hams, and pork bellies;

(5) to check for long–run equilibrium relationships between pairs of price series under consideration;

(6) to identify the statistical problems of each method;

(7) to compare results to previous research;

(8) to draw conclusions from the above research findings.

Procedure

This price discovery study could find different results when different methods are applied to the same time–series price data.

If two price time series are jointly stationary, the lead or lag structure can be determined using the phase diagram drawn from cross–spectral functions. Spectral methods do not require specification of the model. The estimation procedure is independent of the form of the model (i.e. a non–parametric procedure). However, spectral analysis, initially applied in the physical sciences, brings unfamilar concepts to some economists and presents difficulties in interpretation. Increasing familarity with time series models such as distributed lag and autoregressive integrated moving average (ARIMA) models (time domain methods) makes it easy to interpret results without the use of complicated mathematics. The time domain and the frequency domain (spectral) are two different ways of looking at the same model. Thus, spectral analysis will not be utilized in this study.

The application of univariate residual cross–correlation analysis in assessing economic lead–lag relationships between price series is not as complicated. But the procedure is parametric, that is, it allows us to specify the form of the model from two jointly covariance–stationary time series. Then the residual cross–correlation of two series is used to assess linear causal or lead–lag relationships between two prices series.

It is exactly comparable to the choice between parametric and non–parametric statistical estimation methods. If parameterization is correct, the parametric procedure is much more efficient than the non–parametric procedure. If not, however, only the non–parametric method will give valid results.

Another general process of assessing the economic lead–lag relationships among multiple time series is a VAR process. In the estimation of the dynamic structure of each variable, the lagged dependent or lagged independent variables can be included in the regressors. A reduced form of the system of equations is called VAR. If the regressor of the parameters cannot be consistently estimated with a single regression, the Granger causality test using VAR is used to resolve this problem. To estimate the VAR process, the stationarity property of each series must be estimated, and an appropriate number of lags of the lagged dependent variables which are used as independent variables must be pre–determined. Then testing the estimated parameters could denote the direct causal relations between variables. However, there are several views on test methods and the interpretation of VAR results which will be dissused in later chapters.

The concept of cointegration can be applied to price discovery studies. Cointegration tests cannot be used to directly analyze price discovery, but they provide a basis of forecasting performance and longer–run stability of the relationship between two prices series. In general, two time–series are called cointegrated if both series are integrated of order 1, but a linear function of the two series is integrated of order 0. By using Dickey and Fuller's unit root test, integration of each time series can be tested. Then a cointegrating regression function can be estimated from both integrated series. Using Engle and Granger's tests for cointegration or a cointegrating regression function will determine whether each market is efficient. Efficient markets for the same asset are cointegrated. Evidence of cointegrating relationships between the two price series suggest some dependency between series. Thus a Granger–type causal relationship would be induced. As a result of developments in cointegration theory, the error correction model was developed, which is a kind of restricted vector autoregressive model.

Cointegration thoery is closely related to the error correction model. When pairs of economic time series are cointegrated, they will be integrated as in a long run equilibrium state. Cointegration, therefore, relates to the long run relationship that potentially exists. The characteristics of the short run dynamics which are not expressed through cointegration will be adjusted by the corrected error resulting from the cointegrating regression model. When the vector error correction term approaches the VAR representation, the model may be an 'error correction model (ECM).[1] Using this model, the causal relationship between pairs of series can be determined.

The VAR representation is a special case of the ECM. It incorporates a restriction, sometimes implied by the data. Restrictions, if they are correct, increase efficiency and reduce bias in estimated models. However, if pairs of series are not cointegrated, the vector error correction term cannot be utilized in an ECM. As an alternative, the augmented unrestricted vector autoregressive (AUVAR) model can be used. The AUVAR model is based on an augmented or higher order verson of vector autoregression in levels which are not restricted to satisfy the cointegration constraints (Engle and Granger, 1987).

The methods of price discovery analysis are varied. With each method, it is important to check for stationarity of the data series under consideration. To do that, the Dickey–Fuller test or the augmented Dickey–Fuller test will be applied (Dickey and Fuller, 1979; Fuller, 1976). These two tests are simply to check the nonstationarity of price series using the property of autoregressiveness of the data series.

Hypotheses

This study is based on several hypotheses, which arise from the problem statement and objectives. Each hypothesis is closely related to causality and cointegration tests. The specific hypotheses of the dissertation are as follows:

(1) A statistically significant level of causal influence running from the CME futures prices for live hog or pork belly contracts to wholesale pork prices in Omaha will exist.

(2) A statistically significant level of causal influence running from the CME futures prices for live hog or pork belly contracts to cash slaughter hog

prices in Iowa–Southern Minnesota will exist.

(3) A statistically significant level of causal influence running from wholesale pork prices in Omaha to cash slaughter hog prices in Iowa–Southern Minnesota will exist.

(4) A statistically significant level of cointegration between CME futures prices for live hog or pork belly contracts and wholesale pork prices in Omaha will exist.

(5) A statistically significant level of cointegration between CME futures prices for live hog or pork belly contracts and cash slaughter hog prices in Iowa–Southern Minnesota will exist.

(6) A statistically significant level of cointegration between wholesale pork prices in Omaha and cash slaughter hog prices in Iowa–Southern Minnesota will exist.

(7) There exists a statistically significant level of co–movement between any two of the price series.

Overview

The overall structure of the dissertation is as follows.

In chapter 1, the objectives of the study are specified with a brief description of price discovery research, previous research methods and the direction of the present study.

The theory and concept of the relationships between cash and futures prices, and between price discovery and each price series are developed and described in chapter 2. Moreover, the concept of causality is discussed, comparing it as an economic term to one of a philosophical principle. The relationships between price discovery and causality are also documented in this chapter.

In chapter 3, several approaches are explained, providing some advantages and disadvantages of each method, and providing general step by step procedures.

In chapter 4, data description, stationarity tests, and test methods are carefully demonstrated. Cointegration and error correction models, and also out–of–sample causality tests will be applied to several pairs of series. All test for Granger–type causality. Findings will be discussed for the whole data period and for subperiods.

Finally, a comparison of the efficiency of each method will be described in chapter 5. Conclusions are drawn regarding the objectives pursued at the beginning of the study. Findings of the research are summarized and some implications will be suggested.

CHAPTER II

THEORY AND CONCEPTUAL DEVELOPMENT

Introduction

Little can be said about price discovery and causality unless there is a theoretical connection between them. The purpose of this chapter is to explain the theory of price discovery and related concepts, thus providing the rationale for this study.

There are several definitions of price discovery in the literature. This chapter begins with a comparison among them, and defines price discovery for this study. Generally, price discovery deals with relationships between cash and futures market prices. In the next section, a rationale for the price discovery process will be discussed. Most price discovery studies do not provide the theoretical background. Providing the theory will form the basis for this study. Price discovery analyses are frequently conducted using Granger's causality concept. However, there are some arguments against the concept of Granger–type causality and its usage in economics. For the use of causality in economics, the philosophical basis must be documented. This will be discussed in the fourth section. The fifth part of this chapter describes the relationship between causality defined as an economics concept and price discovery. Finally theories and developed concepts are summarized.

Definition of Price Discovery

Economics has a long history of arguing the desirability of price stabilization. And impacts of futures market on efficiency has often been described and demonstrated in research for agricultural commodities. Much attention has been given to agricultural products because their prices fluctuate along with demand and supply. Economic theory also argues for maintaining quality differentials. Nonstorable agricultural commodities differ from storable commodities. Much economics research in livestock has emphasized price stability and quality differences. Those economic objectives brought about a change in marketing mechanism, introducing trade in futures contracts.

Livestock futures trading has been the subject of continual debate regarding whether or not futures trading can play a role in price stability with quality differences. Differences between cash and futures market prices depend in part on differences in quality, location, or delivery time of the commodity. Those differences are usually called 'basis'. The placing of a hedge against a purchase of a cash commodity is one part of a double transaction which is arbitrage both in fact as well as in form because it depends on judgements regarding the relationship between two prices (Working, 1977). The effectiveness of arbitrage in price discovery can be found between two sets of prices determined in a single commodity market when both sets of prices are not treated as determined in two separate commodity markets. If determined in separate markets, however, the effectiveness of arbitrage must be considered differently from price discovery which will be described and demonstrated here, because in such a case, a degree of independence of each market must be implied.

The major roles of futures markets are their contributions to risk transfer and price discovery (Working, 1962 and 1970; Garbade and Silber, 1983; Evans, 1978; Schroeder and Goodwin, 1991). It is widely accepted that futures markets' role in risk transfer is closely related to hedgers who are using futures contracts in order to transfer price risk to others. Moreover, Working (1948) and Garbade and Silber (1983) mentioned that price discovery refers to the use of futures prices for pricing cash market transactions. That is, if prices discovered in futures markets are used to price cash market transactions, futures markets may improve pricing efficiency in the cash commodity markets. According to Ward (1988), price discovery begins with a general price level and concludes with a transaction price.

The market price which is a price discovery leader has the lowest costs for gathering supply and demand information for a commodity. A discovered price will have an effect on pricing in other markets for the same commodity. Then prices in other markets are said to be followers. Thus, if futures market prices lead wholesale and cash hog prices, it can be concluded that futures market prices have lower transaction costs and more market agents -- buyers and sellers, and also are more efficient for pricing live hogs and pork products.

In general, price discovery has the same function as price determination in terms of pricing efficiency because both use relevant information to achieve the prices which equalize supply and demand for the commodity. However, the latter refers to the theory of pricing under the whole range of relevant economic factors and market structures whereas the former deals with a process to achieve at any satisfactory price for both buyers and sellers. Hudson (1984), in research on the cattle markets, distinguishes the differences between them as: "Price determination focuses on the factors which affect live cattle prices and the net impact of these factors in generating market clearance prices. Price determination is largely irrespective of which market is examined. Price discovery, on the other hand, is concerned with the relative efficiency of the process in cash, futures, and carcass markets in assembling the price related information and arriving at the market clearing prices (p. 13)." Therefore, price discovery is concerned with the relative efficiency of the processes in cash, wholesale, and futures markets in arriving at the market clearing price using relevant information.

Price determination includes all information, so the level of price achieved is very important because it depends on whether the price satisfies the market clearing condition. But, price discovery does not focus on the clearing condition even though it uses all possible relevant information. Price discovery is concerned with the relative efficiency and the process of generating prices.

Based on the above price discovery definition, the efficient market hypothesis requires that the future spot and current futures prices of a commodity are closely related. That means the difference between the current futures price and the future spot price must rely only on new unanticipated information because price discovery processes are using all relevant information. Then, a futures market price is said to be relatively efficient.

Therefore, if the prices discovered in futures markets are used to determine cash market prices, futures markets may contribute to increased efficiency of price adjustments in that commodity. Then futures prices will become a set of information to increase pricing efficiency in cash markets.

Most previous price discovery studies analyze the relationship between cash spot market prices and futures market prices for various commodities. However, this study is concerned with different prices; cash slaughter hog prices, wholesale pork prices and futures market prices generated in a specific hog and hog–related meat market. Therefore, the price discovery process must be defined differently than previous studies. Considerations in this study include the relationships between cash slaughter hog prices and live hog futures market prices, between wholesale pork prices and pork belly futures market prices, and between cash hog prices and wholesale pork prices. Because all price series generated in different markets are based on similar information, e.g., similar characteristics of the commodities, a price discovery study for hogs must consider prices in all three separate markets. According to Brorsen et al. (1984), prices actually discovered in one market lead those discovered in the other. Generally, wholesale or the futures markets are believed to be the center of price discovery for hog and pork markets.

Given that three market prices are considered in this study, price discovery is defined as the process of finding prices which are generated in a specific market based on the available price–related information. Ultimately, price discovery focuses on finding the market clearing price with the available information set. Then, it is said that any market which discovers price leads other markets with regard to the formation of price and subsequent stream of prices. Any market price discovered, of course, cannot be said to be the market clearing price because no one can obtain and use all relevant information in the price discovery process. The discovered price would be said to be relatively more efficient than other price(s) for the commodity. Therefore, price discovery can also be defined as the process of finding relevant information in order to reach the true price which could exist under market clearing conditions.

Rationale for Price Leadership by Futures Markets

Information about the lead and lag relationships among the farm, wholesale and retail level prices of a livestock commodity are general factors in determining producers', packers' and retailers' margins for that commodity. The role of price discovery and risk transfer by futures markets is mathematically analyzed here, which will form the rationale for this study.

It is well–known that uncertainty has a decisive influence on economic behavior. However, if all economic agents were not averse to risk, risk resulting from uncertainty would have no influence on economic behavior. The futures market is one of the most important institutions that facilitate adaption to risk aversion. The papers by McKinnon (1967), Sandmo (1971), Danthine (1978), Feder, Just and Schmitz (1980), Benninga, Eldor and Zilcha (1983) and Antonovitz and Roe (1986), etc. examined producers' decision–making under price and output uncertainty using futures markets.

In general, risk in agriculture arises from the variability of both prices and production. The hog and pork sector considered in this study is no exception. First of all, price risk can be classified into three categories: cash slaughter hog prices at the farm level, wholesale pork prices at an intermediate level, and finally retail prices. All these prices are variable, and related to futures market prices for live hogs and pork bellies.

The following is an example for a producer's decision–making process. It will provide evidence that the predetermined futures market prices are more important than cash prices formed in the future spot market. Assume that at the firm level, a hog producer encounters a band of prices which includes all three prices. For convenience, assume no production uncertainty and zero basis. So, the desirable output level is determined at the beginning of production by Q. Also, a producer should determine the amount to contract or the amount sold forward, F, with known futures prices of live hogs denoted by P_f. The producer does not know the market price of live hogs, say P, at the end of production. Based on these assumptions, the profit function of the producer will be:

$$\Pi = PQ - C(Q) + F(P_f - P) + \Pi_0, \tag{2.3.1}$$

where Π is profit of the producer, Π_0 is the level of initial wealth which the producer has, and C(Q) is total cost of production. Assume that transportation cost is not considered, and all input markets are fully informed without uncertainty. Also, assume that interest rates are not considered.

Actually, a hog producer has numerous uncertain factors, focusing here on price. If he/she prices all output with the futures market, then he/she can eliminate uncertainty for his/her output prices. However, since the producer's goal is to achieve maximum profits from production, under the situation that prices are random variables, the ultimate goal is to maximize the expected utility from production. Therefore, the hog producer's objective function would be to maximize expected utility, E[W], which is a function of choice variables Q and F under unknown spot prices in the future, P. That is:

$$Z = \underset{Q\,,\,F}{\mathrm{MAX}} \; \{\underset{P}{E}[W(Q,F)]\}, \tag{2.3.2}$$

where E[.] is the expected value operator conditional on information currently available to the decision maker. Then the expected utility function can be rewritten as:

$$Z = \underset{Q\,,\,F}{\mathrm{MAX}} \; \{\underset{P}{E}[W(\Pi)]\}, \tag{2.3.3}$$

because the producer's expected utility is composed of the profit shown in equation (2.3.1). Choosing the optimal amounts of Q and F determine the maximum level of the expected utility to the producer. Therefore, substituting equation (2.3.1) into equation (2.3.3), and differentiating Z with regard to Q and F, respectively, leaves

$$\frac{\partial Z}{\partial Q} = \underset{P}{E}[(\frac{dW}{d\Pi})\{P - C'(Q)\}] = 0, \tag{2.3.4}$$

and

$$\frac{\partial Z}{\partial F} = \underset{P}{E}[P_f - P] = (\frac{dW}{d\Pi})[P_f - E(P)] = 0 \tag{2.3.5}$$

These equations (2.3.4) and (2.3.5) are known as first–order conditions[1)] of the objective function Z in order to choose the optimum levels of choice variables, Q and F. Solving (2.3.4) and (2.3.5) will generate the optimum amount of Q which is the amount determined to be sold in cash markets after production, and F which is the amount of future contracts sold at the beginning of production.

The problem arising out of the first–order condition, equation (2.3.5), is whether $\partial Z/\partial F$ is equal to zero or not. This problem can explain the behavior of the producer who is willing to hedge. Usually, the amount of output to be produced must be positive, but the quantity F may not be greater than zero. Thus, the relation between P_f and E(P) yielded from equation (2.3.5) allows the following interpretation. If $P_f > E(P)$, then a producer tries to sell products with the futures markets, and F can go to positive infinity. If $P_f < E(P)$, then the producer will sell his/her products in the future cash market, and then F can be negative infinity. In the case that $P_f < E(P)$, he/she will purchase some amount of products with the futures market in order to sell them at a future cash market price as a speculator. If $P_f = E(P)$, then all the output that a producer produces would be fully hedged. That is, all outputs may be sold with the futures markets. In this case, profit in equation (2.3.1) is not a random variable anymore because

1) Remember that $W'(\Pi) > 0$ is a property of the utility function. The second–order conditions, $W''(\Pi) < 0$ and $C''(Q) \geq 0$, respectively. If a producer is risk–neutral, then $W'''(\Pi) = 0$ in equation (2.3.5).

price uncertainty has disappeared. This provides proof for the general belief that future prices have an effect on production decisions rather than cash prices.

When the futures and cash market prices for a commodity exist simultaneously, production decisions would be made based only on futures market prices. That is, one thing can be concluded from the first–order conditions. If two first–order conditions are added up, then the new relation would yield

$$E_P[(\frac{dW}{d\Pi})\{P - C'(Q) + P_f - E(P)\}] = 0. \qquad (2.3.6)$$

Equation (2.3.6) can be reformulated as

$$E_P(\frac{dW}{d\Pi})[E(P) - C'(Q) + P_f - E(P)] = 0. \qquad (2.3.7)$$

Equation (2.3.7) in reduced form is

$$P_f - C'(Q) = 0. \qquad (2.3.8)$$

Therefore, $P_f = C'(Q)$ and solving it with regard to Q, $Q^* = g(P_f)$, where g(.) is notation of some function. That means the level of output which a producer must establish will be determined only by the forward or contract price, P_f. This is important in that when both cash and futures markets exist for a commodity, production decisions are determined only by the predetermined futures prices, but not future cash price. This is because of the role of hedging in removing price risk in the future cash market.

The result that futures price determines production will be extended to wholesalers. So far, models presented assume maximization of only a producer's expected utility. If models described above are modified to include a wholesaler's expected utility maximization, then a wholesaler's decision on the amount of hedging and speculation would also depend only on futures prices because futures prices provide a risk–free situation. Therefore, where two market prices exist, a future cash market and futures market prices for one commodity, only futures prices play an important role in determining hedging or speculating. That provides the theoretical belief that futures prices lead cash prices.

Another rationale for this study would be to prove that futures prices lead cash prices by comparing the expected profits between one who has been informed, one who has been partially informed, and one who has not been informed.

Assume that prices are random variables as before. There are three hog production firms, I, II, and III. Without any information on prices, all producers assume the output price is the same as the level of expected price, $E(P) = \overline{P}$, based on experience in the past several periods. There are assumed to be two kinds of information: e.g. "α" denotes quantifiable price information that futures prices cause wholesale meat prices, and "β" is additional quantifible price information that futures prices cause cash hog prices. Thus, the output price function at the end of production which producers consider at the beginning of production will be formulated as:

$$P = E(P) + \alpha + \beta, \tag{2.3.9}$$

where α and β are random variables with $E(\alpha) = E(\beta) = E(\alpha\beta) = 0$ and $E(\alpha + \beta) = 0$.

The first firm I is assumed not to be informed. So, this firm is assumed to know only E(P) and the distribution of $\alpha + \beta$ when the firm tries to determine the level of output at the beginning of production. Another firm II has partial information on price. That is, when this firm chooses the production level in the first instance, firm II knows the E(P), α, and distribution of β. Meanwhile, the third firm III is assumed to have all relevant information on output prices. This firm knows E(P), α and β exactly before the optimal level of output is determined. Then assume that all three firms have the same profit function, such as

$$\Pi = (\overline{P} + \alpha + \beta)Q - C(Q). \tag{2.3.10}$$

Since firms I and II have some distributions on $\alpha + \beta$ or β, the profit function must include expectational notation as

$$E(\Pi) = \underset{\alpha,\beta}{E}[(\overline{P} + \alpha + \beta)Q - C(Q)]. \tag{2.3.11}$$

That is, firms I and II's objective is to maximize their expected profit.

Case A: For the uninformed firm I, the first–order condition on equation (2.3.11) is as follows:

$$\frac{\partial E(\Pi)}{\partial Q} = \overline{P} - C'(Q) = 0, \tag{2.3.12}$$

because $E[E(P)] = E(P) = \overline{P}$ and $E(\alpha + \beta) = 0$. Firm I can determine its output level by $Q^I = g(\overline{P})$ from equation (2.3.12). Again, note that g(.) indicates 'some function of'. Then the profit which firm I receives under variation of α and β would be Π^I in equation (2.3.13), such as:

$$\Pi^{I} = (\overline{P} + \alpha + \beta)\, g(\overline{P}) - C[g(\overline{P})]. \tag{2.3.13}$$

Case B: Now, for the well–informed firm III, the first–order condition cannot be derived from equation (2.3.11) because this firm has at least some knowledge about $\alpha + \beta$. So, equation (2.3.10) instead of (2.3.11) can be applied to determine the optimum level of output of this firm. The first–order condition on equation (2.3.10) is somewhat different from equation (2.3.12) because the firm knows $\alpha + \beta$. The first–order condition is

$$\frac{\partial \Pi}{\partial Q} = \overline{P} + \alpha + \beta - C'(Q) = 0. \tag{2.3.14}$$

Firm III determines the output decision from equation (2.3.14) as $Q^{III} = g(\overline{P} + \alpha + \beta)$. Thus, the profit level for firm III can be derived as:

$$\Pi^{III} = (\overline{P} + \alpha + \beta) g(\overline{P} + \alpha + \beta) - C[g(\overline{P} + \alpha + \beta)]. \tag{2.3.15}$$

Case C: Consider a firm II that has partial information. Taking the first–order condition on equation (2.3.11) yields the optimum level of output. However, before taking it, the objective function (2.3.11) should be reconsidered because even though $E(\beta) = 0$, $E(\beta | \alpha)$ could not be zero. That is, the objective function (2.3.11) can be rewritten as

$$\begin{aligned} E(\Pi) &= \underset{Q}{\text{MAX}}\ \underset{\beta}{E}[\overline{P} + \alpha + \beta)Q - C(Q)] \\ &= [\overline{P} + \alpha + E(\beta | \alpha)]Q - C(Q). \end{aligned} \tag{2.3.16}$$

Assume that $E(\beta|\alpha) = 0$ which means there is no covariance correlation between α and β. Then the first–order condition for firm II can be derived from equation (2.3.11), such as

$$\frac{\partial E(\Pi)}{\partial Q} = (\overline{P} + \alpha) - C'(Q) = 0, \tag{2.3.17}$$

because $E(\beta) = 0$. Then the desirable level of output can be solved from equation (2.3.17) by some function of $\overline{P}$ and α, denoted by $Q^{II} = g(\overline{P} + \alpha)$. Substituting Q^{II} into the objective function will generate the maximum level of profit which firm II can achieve,

$$\Pi^{II} = (\overline{P} + \alpha + \beta)g(\overline{P} + \alpha) - C[g(\overline{P} + \alpha)]. \tag{2.3.18}$$

Random variables described so far, α and β, represent information on actual prices which has resulted from previous price discovery. A rationale for this price discovery study can be provided by comparing the level of profit for each firm.

Consider the following example. Assume that each firm has the same cost function, given by the following quadratic functional form,

$$C(Q) = \kappa Q + \frac{\varphi Q^2}{2}, \tag{2.3.19}$$

where κ and φ are numeric. Differentiating it with regard to Q will give the marginal cost structure of the firms.

$$\frac{\partial C(Q)}{\partial Q} = MC(Q) = \kappa + \varphi Q. \tag{2.3.20}$$

The marginal cost function, MC(Q), is assumed to be linear. Firms' profit function is as follows:

$$\Pi = (\overline{P} + \alpha + \beta)Q - \kappa Q - \frac{\varphi Q^2}{2}, \text{ or} \tag{2.3.21}$$

$$E(\Pi) = [\overline{P} + E(\alpha) + E(\beta) - \kappa]Q - \frac{\varphi Q^2}{2}. \tag{2.3.22}$$

Based on the same procedures which generated optimum outputs and related maximized profits of each firm used above, estimated output levels and profits are as follows:

$$Q^{I} = (\overline{P} - \kappa)/\varphi \tag{2.3.23}$$

$$Q^{II} = (\overline{P} + \alpha - \kappa)/\varphi$$

$$Q^{III} = (\overline{P} + \alpha + \beta - \kappa)/\varphi,$$

and

$$\Pi^{I} = (\overline{P} + \alpha + \beta)Q^{I} - \frac{\varphi}{2}(Q^{I})^2 \tag{2.3.24}$$

$$\Pi^{II} = (\overline{P} + \alpha + \beta)Q^{II} - \frac{\varphi}{2}(Q^{II})^2$$

$$\Pi^{III} = (\overline{P} + \alpha + \beta)Q^{III} - \frac{\varphi}{2}(Q^{III})^2$$

From equations in (2.3.24), the price discovery effect is found when the magnitude of each firm's profit is compared with each other.

Comparison One: Comparison of profits between the fully–informed firm III and the partially–informed firm II: The difference is

$$\Pi^{III} - \Pi^{II} = \frac{\beta^2}{2\varphi}. \tag{2.3.25}$$

The sign of the result in equation (2.3.25) depends only on φ, which is the slope of the marginal cost curve represented in expression (2.3.20). Thus, when the slope of the marginal cost curve is positive, the magnitude of the difference in (2.3.25) will always be positive. Since the value of β is not known to firm II, equation (2.3.25) must be an expectation as in expression (2.3.26). That is,

$$\underset{\beta}{E}(\Pi^{III} - \Pi^{II}) = \frac{\sigma_\beta^2}{2\varphi}. \tag{2.3.26}$$

When the slope of the marginal cost curve is positive, the value of information from the price discovery process would depend positively on the variance of random variable, β.

Comparison Two: Between the partially–informed firm II and the uninformed firm II: The difference is

$$\Pi^{II} - \Pi^{I} = \frac{\alpha}{\varphi}(\frac{\alpha}{2} + \beta), \tag{2.3.27}$$

and its expected value is

$$\underset{\alpha,\beta}{E}(\Pi^{II} - \Pi^{I}) = \frac{1}{2\varphi}[E(\alpha^2) + E(2\alpha\beta)] = \frac{\sigma_\alpha^2}{2\varphi}, \tag{2.3.28}$$

because $E(\alpha\beta) = 0$ was assumed. Thus, the magnitude of $\Pi^{II} - \Pi^{I}$ depends on the signs of α and β. If α and β have the same sign, the magnitude will be greater than zero. Otherwise, it will be negative. Equation (2.3.28) shows that the expected value of the difference depends only on the variance of α if the marginal cost curve is positively sloped. Note that the signs of α and β depend on the accuracy of price discovery. That is, if the price discovery process correctly provides some evidence that futures market prices cause cash prices, the sign of α is positive, otherwise it will be negative. The magnitude of α and β will depend on the degree of accuracy.

Comparison Three: Between the fully–informed firm III and the uninformed firm I: The difference is

$$\Pi^{III} - \Pi^{I} = \frac{(\alpha + \beta)^2}{2\varphi}, \tag{2.3.29}$$

and its expected value is

$$\underset{\alpha,\beta}{E}(\Pi^{III} - \Pi^{I}) = \frac{\sigma_\alpha^2 + \sigma_\beta^2}{2\varphi}. \tag{2.3.30}$$

Equation (2.3.29) shows that the magnitude of the difference will always be positive if and only if the slope of the marginal cost curve is positive. And then the expected difference of profits in expression (2.3.30) depends on the variances of α and β.

In conclusion, for firms having positively sloped supply functions, price discovery provides benefits to the one who has more information. Since the marginal cost curve represents the supply curve of each firm, the greater the elasticity of supply, the larger the effect price discovery could have. On the positively sloped, but not vertically sloped supply curve, the smaller value of φ will have a greater effect from price discovery.

Since newly discovered prices become public goods, producers obtain knowledge of α and β at no cost. If producers are charged for getting α and β, the above results may be influenced. Impacts from such extra costs can be analyzed in another study.

Also in this section, regarding lead–lag relationships, it was found that when both cash market and futures market prices exist for a commodity, futures market prices leads cash market prices.

Causality and Its Janus

Some portion of economic benefits to be realized in the future may already have been determined by plans or commitments by the individual decision–making units of the economy. Such plans and commitments will be based on factors concerned with the future and the present, as well as the past. The relations between the past, the present and the future could be explained or defined by the concept of causality.

Many economists have contributed to defining causality conceptually and operationally. Two representative economists are Hicks (1979) who defined it conceptually or in principle, and Granger (1969) who defined it for an operational purpose to be applied in economics.

Besides them, Sims (1972) demonstrated his own causality concept, a two–sided distributed lag method, which modified Granger's operational definition of causality and reached a new conclusion. Sims demonstrated explicitly the existence of an econometric test for exogeneity. Sims concluded that Granger–type causality is equivalent to econometric exogeneity so that unidirectional causality from the independent to the dependent variable is a necessary condition for the consistent estimation of distributed lag models involving other variables rather than lagged dependent variables. Pierce (1977) and Haugh (1976) also suggested a concept close to Granger causality. Geweke (1978) developed two tests of exogeneity suggested by Sims and showed that his methods are more powerful than Sims' even without stationarity conditions for the variables. Caines, Keng and Sethi (1981) and Geweke, Meese, and Dent (1983) also proposed similar concepts to Granger causality. As Sargent (1976) formalized a one–sided distributed lag approach implied by Granger, these other definitions since Sims, differing from Hicks and Granger, were constructed and elaborated for testing causality. Therefore, a comparison of causality is made between Hicks and Granger.

Hicks' Causality: Hicks defined causality as the relationship between cause and effect, which is thought to be the business of philosophers. Following his opinion, though economists often talk about effects and causes, they are usually content to leave the question of the meanings of these terms to others.

Hicks' classification of causality is as follows. He defined causality as weak, strong, and sequential causality. Define X as an event occuring at some time, t_X, and Y as some event occuring at time t_Y. If X was one of the causes of Y, then the relation is defined as "weak causation". Strong causation implies a relationship between X and Y when an event X was the sole cause of Y. Again, weak causality is divided into two kinds, separable and non–separable causation. Separable causation was defined by a statement by a philosopher, Hume (1978), that cause has priority in time before the effect. Non–separable causation follows the recognition by another philosopher, Kant (1943).

Separable causation assumed that there are several events, e.g. X_1 and X_2. If X is to be a separable cause of Y, such a hypothetical situation in which X did not happen must be conceived. There are five kinds of separable causation defined: (i) additive causes, (ii) sole causes, (iii) overlapping causes, (iv) negative causes, and (v) ultimate causes. Additive causation means that the effect Y will not appear unless both causes X_1 and X_2 are present. Sole causation implies that either X_1 and X_2 causes Y. But in the case that X_1 causes Y solely, X_2 must be assumed not to present, vice versa. Overlapping causation is when Y occurs if either X_1 or X_2 is present. To explain negative causation and ultimate causation as separable causations, another cause, say X_3, other than X_1 or X_2 is introduced. Assume a case that the effect Y would have happened if neither X_1 nor X_2 was present, but X_3 was present, and also happened when X_1, X_2 and X_3 were all present. Under this situation, if X_1 alone was not present, Y would not have happened. Then the X_2 is called negative causation or preventive measure. That is, if X_2 acted alone, it would offset the effect of X_3. In this case, X_3 is called ultimate causation.

Hicks' explanation of Hume's philosophy that cause precedes effect led to the theory of separable causation. On the other hand, he explains non–separable causation based on Kant's critique of Hume's principle. Kant did not deny that an event X precedes another event Y, and also event Y precedes event X. This kind of causal relation is usually called contemporaneous causation or mutual causality. In reality, such a contemporaneous causal relationship is found frequently. Keynes' explanation of the relation between the money supply and interest rate is often interpreted in terms of contemporaneous causality. Also the relationship between stocks and flows in the economy is contemporaneous in the sense of Keynes.

Another important explanation of causality by Hicks is sequential causality which means the relationship between lags and reserves. Sometimes it is called a causal chain. Sequential causality implies that X was a cause of X_1, X_1 was a cause of X_2, X_2 of X_3 and so on, and finally X_n was a cause of Y. Economics is concerned with decisions, and decision—makers sometimes must decide something in the intermediate stage of such a causality chain. Price discovery is closely related to sequential causality. This price discovery study may pose an intermediate stage of variously changeable economic circumstances.

Granger Causality: While the definition of causality by Hicks was somewhat of a conceptual or philosophical one, Granger defines it as an operational or practical one in an economic time series context. The general time series analysis concentrates on the use of alternative lag structures in modeling. The process of modeling includes some present and past information, and the analysis of predictability of behavior of the economic time series.

An operational definition of causality which Granger (1969, 1980) derived is more practical than Hicks' definition, because it is testable with economic time series. Unidirectional causal or feedback relationship between pairs of random variables in the economy must be identified before market efficiency is tested, and after some spurious regression problems are addressed. Thus, his definition provides a rationale for identifying the direction of causality from alternative models. And also, the tests following his operational definition will provide empirical support for model building and the assertion of lead—lag relationships.

Granger presents two perspectives to defining causality. One is a correlation between a pair of variables (1969), the one is a Bayesian view (1980). Causality is generally a difficult concept to model in the analysis of real economic time series variables. One way Granger explained it through the use of bivariate variables was to consider the statistical interpretation and prediction for the potential implication of causal relationships after removing behavioral properties. Let's say X and Y are two random variables in the universe. Granger's view of causality by means of the Bayesian principle is that one would change one's initial belief or causal relationship using test results when he/she believed X causes Y initially. The person's initial belief is the prior probability. However, the existence of a correlation between pairs of random variables provides more clear causal symptoms because generally the prior density function in a Bayesian framework cannot frequently be found. Assume one knows X does not cause Y. Since one still has a question about whether Y can cause X, he found that one could use an observed significant correlation between two variables in order to interpret causal relationships (see Granger and Newbold, 1977, p.224, for more discussion).

Time series analysis usually focuses on the use of different lags in modeling the behavior of economic time series variables. For bivariate variables, both have their own past and present values to be modeled with each other. Then three rules must be assumed when causality is defined: that those variables must result from stochastic processes but not deterministic processes; that the variables are stationary, and that the future cannot predict the past. With these assumptions, one can find one or more of the following four general relationships: (i) X causes Y, (ii) Y causes X, (iii) feedback or instantaneous causality exists between X and Y, and (iv) no relationship between X and Y exists. Causality is defined such that X causes Y, with respect to the information set including at least X and Y, if current values of variable Y can be better predicted using past information of another variable X, but not vice versa. On the other hand, if current values of X can be better predicted using past information of Y, not vice versa, then Y causes X. This definition may not be accepted by philosophers like Hume, because cause is too strong. When the definition of causality is interpreted as being 'temporally related[1], the above notions are acceptable for the temporal analysis of economic time series. The third relationship is feedback or instantaneous causality. If both possibilities of X causing Y and Y causing X are proven, there is evidence of an instantaneous causal relationship between X and Y. Finally, if past values of X are not useful in the sense of the predictions of current Y and also past values of Y are not also useful in the prediction of X, it is said there is no relationship.

The four definitions are too general to be testable. Also, in the case of the third definition about feedback or instantaneous causality, the above rule described such that the future cannot cause the past, may not be admitted. Of course, instantaneous causality would be differentiated from feedback. More accurate definitions proven by Granger (1969) which are applicable to test causality are as follows:

Definition 1: Causality

If $\sigma^2(X|\Omega) < \sigma^2(X|\overline{\Omega - Y})$, we say that Y_t is causing X_t, denoted by $Y_t \longrightarrow X_t$. We say that Y_t is causing X_t if we are better able to predict X_t using all available information than if the information apart from Y_t had been used.

Definition 2: Feedback

If $\sigma^2(X|\overline{\Omega}) < \sigma^2(X|\overline{\Omega - Y})$, $\sigma^2(Y|\overline{\Omega}) < \sigma^2(Y|\overline{\Omega - X})$, we say that feedback is occurring, which is denoted $Y_t \longleftrightarrow X_t$, i.e., feedback is said to occur when X is causing Y_t and also Y_t is causing X_t.

Definition 3: Instantaneous Causality

If $\sigma^2(X|\overline{\Omega}, \overline{\overline{Y}}) < \sigma^2(X|\overline{\Omega})$, we say that instantaneous causality $Y_t \longrightarrow X_t$ is occurring. In other words, the current value of X_t is better predicted if the present value of Y_t is included in the prediction than if it is not.

Definition 4: Causality Lag

If $Y_t \longrightarrow X_t$, we define the (integer) causality lag m to be the least value of k such that $\sigma^2(X|\Omega - Y(k)) < \sigma^2(X|\Omega - Y(k+1))$.

> Thus, knowing values of Y_t, j = 0, 1, 2, ..., m–1, will be of no help in improving the prediction of X_t (pp. 428–429),

where Ω_t represents all the information in the universe at time t, and $\Omega_t - Y_t$ is all the information apart from Y_t. The bar notation above a variable represents the set of past values of the variable and the the double bar notation above a variable represents the set of past and present values. The lag number, k, in parentheses implies the set of all past values having fewer than k of the variable. These definitions are more testable on economic time series. The three rules described before are still legitimate to the above four definitions.

Of Granger's definition of causality, feedback and instantaneous causality are frequently used as a synthesized notation in practice. The so–called "instantaneous feedback" also implies the mutual causal relationship between two variables such as pure feedback. Instantaneous feedback is distinguished differently from feedback because most procedures for modeling and testing causality use the past and present values of each variable.

Comparison: Granger and Newbold (1977) said that it is doubtful philosophers would completly accept Granger's definition of causality. Also, Zellner (1979) and Granger (1980) discussed whether Granger–type causality tests are truly causal in nature. Nevertheless, Hicks tried to explain the relationship between stocks and flows using lags and reserves, and all the liquidity preference relationships for the Keynesian macro–constructions by using his definition of causality, which stemmed from philosophers Hume and Kant. All three definitions are related to one another.

Table 1 is constructed by using the definitions of each scholar, and shows the relationships among those definitions. Granger's definitions are narrower compared to Hicks' definitions. Except for definitions based on Kant's principle, Granger's definitions of causality seem a subset of Hicks', whereas Hicks'

TABLE 1

CATEGORIES OF CAUSALITY

Philosopher	Hicks	Granger
Hume*	1. Weak causation	
	i) Separable causation	
	a. Additive	Instantaneous Causality
	b. Sole	
	c. Negative	
	d. Ultimate	
	e. Overlapping	Causality Lag
	ii) Sequential causation	
	2. Strong Causation	Causality
Kant**	1. Weak Causation	
	i) Non–separable causation	Feedback

*: Cause precedes Effect. Cause has priority in time before the effect.
**: When an event (X) causes another event (Y), it is possible that Y also causes X.

definitions also look like a subset of Hume and Kant in a broad sense. The acceptability of Granger–type definitions of causality occurs when it allows testing true causality in nature. Since the study of price discovery includes identifying the temporal interactions between prices in alternative markets, the Granger–type definition of causality is an effective method used in modeling the price discovery process for agricultural commodities.

Therefore, the definition of Granger causality is used as a theoretical basis for the analysis of the price discovery process of hog and pork markets.

Price Discovery and Causality

Price discovery, as defined earlier, refers to the process of discovering a price in a specific market using all available information for the relevant commodity. Also, price discovery was defined as the process of finding relevant information to reach the market clearing price. All the relevant information is assumed to involve the economic factors that influence the true price. Since newly discovered prices include more information, the study of price discovery involves efficiency in pricing a commodity. On the other hand, causality, especially Granger causality, provides a convenient technique which can be used to analyze the lead or lag relationships between any two price series.

To discuss the relationship between price discovery and causality, several possibilities are considered.

The use of Granger causality in analyzing the price discovery process can indicate the presence of lead and lag relationships, feedback, or no relationships between any two economic time series. First, assuming that hog futures market prices lead cash slaughter hog prices by testing for Granger causality, futures market prices are said to have more relevant information to arrive at a market clearing price for live slaughter hog than do cash prices. It means that live hog futures market prices are more efficient than cash slaughter hog prices. Second, if both pork belly futures prices and wholesale pork belly prices have feedback resulting from testing for Granger causality, both prices are said to have relevant information influencing the market clearing price of pork bellies.

In the above first case assumed, cash slaughter hog prices follow live hog futures prices with some time lag to reach a market clearing price when the leading price, i.e., the futures market price, is assumed to be a market clearing price. There is no assurance that live hog futures prices are market clearing prices even if they lead cash live hog prices. In the second case, a question arises about which market price is more efficient to provide insight in arriving at a market clearing price. Only one thing that both market prices have in common is that similar information is known.

Third, let's assume that wholesale ham prices lead cash slaughter hog prices. Since both ham and live hogs are not exactly the same commodity, in this special case it is hard to say that ham prices have more available information to find the market clearing price. The finding of lead/lag relationships between them will mean that wholesale ham prices lead cash slaughter hog prices with some time lags, whereas cash hog prices follow wholesale ham prices with the same time lags. As a result, wholesale ham prices are said to have forecasting power for the future spot market price of cash slaughter hogs because they have more information on economic factors influencing the market prices for hams and live hogs due to tests for Granger causality.

Fourth, another assumption which may result from Granger causality tests between two prices is that there is no specific lead/lag relationships between them. This means that neither price has any relevant information to achieve the market clearing condition nor that one price leads or follows the other price. It is doubtful that both prices in two markets for a commodity do not have any information to satisfy the true market price level. That is, price discovery in that commodity takes place in some other commodity market or that the price of the commodity does not respond to market information.

Fifth, the existence of no relationship between wholesale pork and cash slaughter hog prices or between wholesale pork and live hog futures market prices could occur because they are different commodities.

Finally, consider that cash prices lead futures prices from the tests for Granger causality. This case violates one of the basic rules that the future cannot predict the past, and would be difficult to interpret. However, even though it violates the rule, it is possible that cash prices contain more information than futures prices in order to achieve the market clearing price. That is, cash prices would be more efficient in pricing hogs and pork than futures prices. But, it violates the theoretical argument that the future leads the past. Compared to the theoretical base, it might be thought that the cash spot price in the future was formed by persons who are not risk averse but risk loving. Such a conclusion is based on the belief that risk averse persons do not prefer risky prices, and that they prefer risk–free futures market prices.

As discussed so far with six different possibilities, identifying the lead and lag, feedback, or no relationship between two prices from alternative markets using tests of Granger causality can be useful to find the market which discovers price. That is, Granger causality is a very useful technique to analyze price discovery processes in agricultural commodities.

Summary

Price discovery has been defined as the process of tracing the role of futures markets for pricing in cash markets. But, when there exist related markets for a commodity, e.g., the wholesale meat market in this study, price discovery would involve the use of one market price, even if it is not a futures market price, for determining other market prices.

Theoretically, futures market prices have been proven to lead producers' decision making. Thus, the existence of futures markets will produce a stream of information resulting from the price discovery process. Another finding is that price discovery provides some profit gain to producers, packers and wholesalers, etc..

Granger–type causality is often used in price discovery studies. Some disagreement exists concerning the use of Granger–type causality in economics. Since the analysis of causation by Hicks permits the establishment of the relationship between causal phenomena in economics and philosophical terms, Granger causality will be utilized in this study.

CHAPTER III

PROCEDURES FOR ANALYZING PRICE DISCOVERY —BIVARIATE MODEL APPROACHES—

Introduction

Various procedures are used in applying Granger–type causality criterion to two economic time series. The procedures indicate the progress in developing theories and techniques. Economic causality was studied originally by spectral analysis. But developing a univariate Box–Jenkins type model provides some alternatives such as univariate residual cross–correlation analysis and vector autoregressive model analysis. Both alternatives are substitutes for spectral analysis if an ARIMA model is well identified.

The following section discusses the univariate residual cross–correlation method, giving advantages and disadvantages of using it. Then, out–of–sample performance using ARIMA techniques developed to test Granger causality is discussed.

The fourth section describes VAR analysis used in price discovery research. Testing for cointegration of the series estimates the long run equilibrium relationship between the series. Cointegration with an error correction model, the current most popular technique, is discussed in the fifth section.

Out–of–sample causality testing is discussed again. Out–of–sample performance was originally devised for reducing some of the disadvantages with the ARIMA technique. However, the original out–of–sample tests did not consider the properties of autoregressiveness for the data series. Out–of–sample performance using the VAR analysis enables causality testing, considering the autoregressive properties of the data.

Univariate Residual Cross–Correlation Analysis

Univariate residual cross–correlation analysis, Box–Jenkins (1976) techniques to pre–whiten the original series of interest, is a method to empirically assess lead–lag relationships between economic time series. The interrelationships existing between two time series were discovered first by system identification, and second, by checking the independence of the two series before the univariate residual cross–correlation approach was developed. This approach was convenient and easy to check for the independence of two time series. Assessing lead–lag relationships between two time–ordered variables using univariate residual cross–correlation analysis resulted from Granger's causality theory (1969).

Basically, if the two economic time series are transformed so as to satisfy their joint covariance stationarity, then their interrelationships would be determined by the cross–correlation function. The analysis is a two–stage method proposed for investigating the independence of two covariance–stationary time series (Haugh, 1976). A two–step method implies that first when univariate models for each series are fitted to their residual series, and second, a cross–correlation function for the two residual series is obtained, the independence of the two series will be tested with the residual cross–correlations.

Let a time–ordered variable X be represented by autoregressive and/or integrated moving average forms, and also another variable Y be represented by autoregressive and/or integrated moving average forms (ARIMA) as

$$H(L)X_t = v_t \qquad (3.2.1)$$
$$\longrightarrow \; X_t = K(L)v_t, \text{ and}$$

$$M(L)Y_t = u_t \qquad (3.2.2)$$
$$\longrightarrow \; Y_t = Q(L)u_t, \text{ respectively,}$$

where H(L), K(L), M(L) and Q(L) are infinite polynomials, L is a lag operator, u_t and v_t are white noise process at time t with zero mean and constant variances σ_u^2 and σ_v^2, respectively. Note that $H(L) = \sum_{j=0}^{\infty} h_j L^j$, $K(L) = \sum_{n=0}^{\infty} k_n L^n$, $M(L) = \sum_{g=0}^{\infty} m_g L^g$, $Q(L) = \sum_{i=0}^{\infty} q_i L^i$. To satisfy the stationarity of X_t and Y_t processes, H(L) and M(L) in equations (3.2.1) and (3.2.2) must be invertible. Then we can express X_t and Y_t processes as the above expressions. That is, to apply the univariate residual cross–correlation model to two raw time series, each series must be stationary and an invertible univariate model as shown in equation (3.2.1) and (3.2.2).

If those two processes, X_t and Y_t, are stationary and invertible, so are the univariate models, then the cross–correlation between the white noise residual processes of the u's and v's, would be obtained. The cross–correlation function at some lag k is defined as (Haugh and Box, 1977)

$$\sigma_{uv}(k) = \frac{E\{(u_{t-k}, \; v_t)\}}{\sigma_u \sigma_v} \qquad (3.2.3)$$

where E{.} is a general expression of expectation. In practice, a set of cross–correlation estimates $(r_{\hat{u}\hat{v}}(k))$ are proven to be constant estimators for $\rho_{uv}(k)$ and to be asymptotically distributed under the reasonable assumptions of joint covariance stationarity and a normal distribution of each series (Hannan, 1970). The estimated cross–correlation function may be used to describe the interrelationship, lead–lag, or causal relationship. Pierce (1977) and Miller (1979) defined linear causal relationship implied by various cross–correlations as in Table 2.

TABLE 2

CAUSALITY PATTERNS REPRESENTED BY RESIDUAL CROSS–CORRELATIONS

Lead–Lag	Cross–Correlations at lag k
1. X leads Y	$\rho_{uv}(k) \neq 0$ for some $k > 0$
2. Y leads X	$\rho_{uv}(k) \neq 0$ for some $k < 0$
3. Instantaneous causal relation	$\rho_{uv}(0) \neq 0$
4. Feedback between X and Y	$\rho_{uv}(k) \neq 0$ for some $k > 0$ and for some $k < 0$
5. Y does not lead X	$\rho_{uv}(k) = 0$ for all $k < 0$
6. X does not lead Y	$\rho_{uv}(k) = 0$ for all $k > 0$
7. X leads Y, no feedback from Y to X	$\rho_{uv}(k) \neq 0$ for some $k > 0$, and $\rho_{uv}(k) = 0$ for all $k < 0$
8. Y leads X, no feedback from X to Y	$\rho_{uv}(k) \neq 0$ for some $k < 0$, and $\rho_{uv}(k) = 0$ for all $k > 0$
9. X and Y are related instantaneously but in no other way	$\rho_{uv}(k) = 0$ for all $k \neq 0$, and $\rho_{uv}(0) \neq 0$
10. X and Y are independent	$\rho_{uv}(k) = 0$ for all k

The advantages of univariate residual cross–correlation to distinguish causal relationships between two series are as follows. This method is easier to understand and apply than cross–spectral techniques (Miller, 1979). First, there is no need to impose a specific lag structure (k) (Griliches, 1967, re–quoted from Miller, 1979). Second, this approach can prevent spurious regression problems (Granger and Newbold, 1974) when the lead–lag relationships are determined. Third, this approach avoids the testing problems encountered when autocorrelated time series are cross–correlated.

On the other hand, some empirical disadvantages have been found. First, if time–ordered data series under study have reasonable parameters which might be represented by a linear ARIMA representation, and if, the residuals are white noise processes, then this approach is plausible to assess lead–lag relationships. But if not, a severe problem from the first step, model identification, will occur. Second, since this procedure measures the relationship between two time series using a separate ARIMA model, and residuals from each series are pre–verified to be white noise series, individual estimated cross–correlations can be misleading (Pierce, 1977). For such a problem, Pierce suggested a test statistic based on Haugh(1976), $U = n \sum_{k=1}^{m} [r_{\hat{u}\hat{v}}(k)]^2$, which is similarly distributed like the chi–square distribution with degrees of freedom equal to m. However, Sims (1977a) points out that the use of the U statistic to conclude causal relationships is unwise, because such correlations tend to be biased toward zero due to specification error.[1)]

---

1) The following is an example of specification error. Suppose Y causes X under a bivariate system. Then an ARIMA model allows the data to indicate the relative importance of past X and past Y series in forecasting X. Prewhitened X, in this time, uses a misspecified model because past Y should be included.

Out–of–Sample Forecasting Performance (1)

Ashley, Granger and Schmalensee (1980), argued that the use of univariate residual cross–correlation techniques to test Granger–type causality of economic time series is appropriate. As an alternative, they extended the Pierce procedure by using bivariate transfer functions and suggested a test method with out–of–sample forecasting performance of models related to the non–prewhitened series of interest.

Define MSE(X) as the population mean–square of the one–step forecast error of X_{n+1} forecasted from a linear structure of X_{n-i}, $i \geq 0$. Similarly, MSE(X,Y) is defined as the population mean–square of the one–step forecast error of X_{n+1} using the optimum linear forecast based on the information set, $\{X_{n-i}, Y_{n-i}, \ i \geq 0\}$. If MSE(X,Y) < MSE(X), then Y causes X. The reason for the use of mean–squared error to detect causality is because causality is basically used and defined as a statement about forecasting ability. Thus, finding causality from tests on forecasting are appropriate (Ashley et al.).

Ashley et al. presented a five–step approach for out–of sample forecasting performance to analyze causality between a pair of time series X_t and Y_t. The first two steps of the approach are analogous to the Box–Jenkins techniques. Actually, this procedure requires dividing each series, with two finite data sets of X and Y, into two categories. One is so–called within–sample or sample observation, which includes around the first 70–90 percent of the series. The remaining observations are retained as the out–of–sample or post–sample observations. The within–sample observations are used for step (a) to (d), whereas the last remaining sample series, out–of–sample, are used to evaluate out–of–sample forecasting performance in steps (e). Step (f) is added to clearly distinguish between the above procedures. Below are the five steps (Ashley at al.)

and an additional step.

(a) ARIMA models on each single series are estimated and their residuals, say μ_{X_t} and μ_{Y_t} are pre–whitened.

(b) Cross–correlations between two residual series solved in step (a) are analyzed such as in equation (3.2.3) at some lag k.

According to the univariate cross–correlation technique, every causal direction can be indicated after completing these two steps. The next three steps are different testing procedures which Ashley et al. presented.

(c) The third step is to build a bivariate model relating the residuals, and to identify, estimate, and diagnostically check the model. If in step (b), one–direction causality between two series is present, then the model, a bivariate model on the residuals, is unidirectional and identified directly from the cross–correlogram. If the relationship between the two series represents a feedback (bidirectional causality), the bivariate model should use a pair of transfer functions on the cross–correlogram.

(d) In the fourth step, the corresponding model relating the original series is constructed by combining the bivariate model for residuals with the original univariate model. On the fitted models, checking for common factors, estimation, and diagnostic checking are also carried out.

(e) The fifth step of the procedure is to generate a set of one–step forecasts for an out–of–sample period by using the bivariate model for the original series, within–sample observations. That is, this step is used to evaluate the post–sample forecasting performance of models fitted to the original series. The rationale for this step, and this entire approach, is that a bivariate regression model for Y_t and X_t improves the forecastability of Y_t if there is at least one non–zero cross–correlation of the innovations from lags for X_t leading Y_t and no–cross correlation from lags for Y_t leading X_t (Brandt and Bessler, 1982).

(f) The last step is to estimate mean square errors generated from univariate forecasts and bivariate forecasts, and to compare MSE(X) and MSE(X,Y). If MSE(X) < MSE(X,Y), then it is said that X causes Y.

Out–of–sample forecasting performance to detect causal relationships between two finite economic time series is an alternative method to a univariate residual cross–correlation model. It provides tests which have asymptotically valid significance levels (Bessler and Brandt, 1982). However, it will not give clear causal directions because it applies parameters derived from pre–whitened within–sample data to out–of–sample series, and it does not utilize the autoregressive properties in within–sample data processing (Bessler and Kling, 1984). To rule out these defects with this method, an alternative modification will be introduced in the sixth section.

Vector Autoregressive Analysis

Price discovery refers to a process of reaching a satisfactory price for both buyers and sellers. In cases where relevant information must be recognized in the system, a dynamic simultaneous equation model could be used. However, a complete dynamic simultaneous equation model requires adequately generating economic data series and it cannot be modeled with a single equation. 'Vector' in vector autoregressive (VAR) processes implies that all the variables and parameters are represented in vector notation in the simultaneous–type equation model including lags of variables. A representative of a VAR process is the general process of multiple time–series.

The Granger–causal relationship between two or more variables can be analyzed in a VAR process framework. Consider the bivariate system with variables, x and y in the system to investigate Granger causality.

$$x_t = f\{x_{t-1}, x_{t-2}, ..., x_{t-k}, y_{t-1}, y_{t-2}, ..., y_{t-p}\} \tag{3.4.1}$$

$$y_t = g\{x_{t-1}, x_{t-2}, ..., x_{t-k}, y_{t-1}, y_{t-2}, ..., y_{t-p}\} \tag{3.4.2}$$

The above two equations are a kind of simultaneous equation model including lags[1)] on each variable.

From equations (3.4.1) and (3.4.2), the general functional forms could be represented as:

$$x_t = \alpha_1 + \beta_1 x_{t-1} + ... + \beta_p x_{t-p} + \gamma_1 y_{t-1} + ... + \gamma_q y_{t-q} + \nu_{1t} \tag{3.4.3}$$

$$y_t = \alpha_2 + \eta_1 x_{t-1} + ... + \eta_p x_{t-p} + \rho_1 y_{t-1} + ... + \rho_q y_{t-q} + \nu_{2t} \tag{3.4.4}$$

where ν_{1t} and ν_{2t} are two serially uncorrelated white–noise processes.

The above system can be expressed as the general ordinary least squares functional form giving the sense of Granger's own causality test. That is,

$$x_t = \alpha_1 + \sum_{j=1}^{p} \beta_j x_{t-j} + \sum_{i=1}^{q} \gamma_i y_{t-i} + \nu_{1t} \tag{3.4.5}$$

$$y_t = \alpha_2 + \sum_{j=1}^{q} \rho_j y_{t-j} + \sum_{i=1}^{p} \eta_i x_{t-i} + \nu_{2t} \tag{3.4.6}$$

---

1) An appropriate number of lags can be selected by Akaike's concept of information criterion (AIC) or Schwarz's Bayesian Information Criterion (SIC), etc. (for more discussion, see Judge et al., 1985 and the fifth section).

If the δ's are all significantly different from zero, then it is said that the y series causes the x series.

Writing the bivariate systems, (3.4.3) and (3.4.4), in vector and matrix notation while ignoring deterministic components, e.g., constants, trend, etc., gives

$$Y_t = \Psi_i Y_{t-1} + \dots + \Psi_h Y_{t-h} + V_t, \tag{3.4.7}$$

where $Y_t = \begin{bmatrix} x_t \\ y_t \end{bmatrix}$, $\Psi_i = \begin{bmatrix} \beta_i & \gamma_j \\ \eta_l & \rho_m \end{bmatrix}$, $V_t = \begin{bmatrix} \nu_{1t} \\ \nu_{2t} \end{bmatrix}$, and $t = 1, 2, \dots, n$.

The above system (3.4.7) can be expressed by the difference equation $C(L)Y_t = V_t$, where $C(L) = \sum_{j=0}^{\infty} c_j L^j = c_0 + c_1 L + c_2 L^2 + \dots + c_j L^j$ is an n x n matrix with the back–shift operator L, Y_t is an n x 1 vector wide–sense stationary stochastic process,[2)] and V_t is now an n x 1 vector of white noise with means of zero and contemporaneous covariance matrix $E(V_t V_t') = \Sigma_V$, an n x n matrix.

A system represented by vectors is practical because all the series in the system are presumed to be stationary. Using the vector notation system, Sims (1972) set forth an alternative test of Granger causality. Sims' more general

---

2) The wide–sense stationarity property: A vector stochastic process is called wide–sense stationary if (a) V_t in expression (3.4.7) is a zero–mean white noise vector, so that $E(V_t) = 0$ and $E(V_t V_s') = \begin{cases} 0, & t \neq s, \\ \Sigma_V, & t = s \end{cases}$ (b) all the random vectors, Y_t, have the same mean, $E(Y_t) = \mu_Y$ for all t, (c) variances of Y_t must be finite, $Var(Y_{ht}) < \infty$ for all h = x, y and all t, and (d) the covariance matrices of vectors Y_t and Y_{t-k} do not depend on t but only on k, $Cov(Y_t, Y_{t-k}) = E[(Y_t - \mu_Y)(Y_{t-k} \quad \mu_Y) = \Omega_k$ for all t. These conditions imply for practical purposes that the time series under consideration must not have trends, fixed seasonal patterns, or time–varying variances.

expression of the system in practice can be formulated to identify the Granger causality testing scheme as follows:

$$Y_t = C(L)\epsilon_t, \text{ and} \tag{3.4.8}$$

$$X_t = D(L)\epsilon_t + E(L)\eta_t, \tag{3.4.9}$$

where C(L), D(L), and E(L) are ploynomials in the lag operator L, and ϵ_t and η_t are white noise processes. From equation (3.4.8), $\epsilon_t = C^{-1}(L)Y_t$. Substituting ϵ_t to equation (3.4.9) yields

$$X_t = D(L)C^{-1}(L)Y_t + E(L)\eta_t. \tag{3.4.10}$$

Equation (3.4.10) can be expressed by a scalar notation such as:

$$X_t = \sum_{j=1}^{q} \delta_j Y_{t+j} + \sum_{i=0}^{p} \gamma_i Y_{t-i} + \nu_t, \tag{3.4.11}$$

where ν_t is white noise process. The null hypothesis that X does not cause Y is equivalent to all the coefficients on the future values of Y being equal to zero. That is, $\delta_j = 0$, j = 1, 2, ..., q. Thus, Sims' causality test is to estimate a two-sided regression model.

Sims' version of Granger's causality test is different from Granger's own testing scheme. Sims considers Y series as a distributed lag function of future, past and current X's. Sims proposed the use of a pre-filter $(1-0.75B)^2$, and claimed that such a filter would reduce the serial correlation problem in the regression residuals. But the problems suggested are that no one knows which pre-filter is appropriate and which number of lags for q and p are best. With weakness of the approach, Sims' test remains doubtful.

On the other hand, Geweke (1980) suggested a test for Granger–type causality with the following specification:

$$Y_t = \alpha_1 + \sum_{j=1}^{p} \beta_j Y_{t-j} + \eta_t \tag{3.4.12}$$

and

$$Y_t = \alpha_2 + \sum_{j=1}^{p} \gamma_j Y_{t-j} + \sum_{i=1}^{q} \varphi_i X_{t-i} + \omega_t \tag{3.4.13}$$

where η_t and ω_t are white noise residuals. The null hypothesis for testing causality from these equations is that all φ's are zero. Causality from the X to Y series is not evident unless φ's are significantly different from zero.

The test statistic for testing for Granger causality in Geweke's model is the well–known F value. The estimated F statistic, F^*, is then $\dfrac{ESS(1) - ESS(2)}{q} \Big/ \dfrac{ESS(2)}{N-p-q-1}$, where ESS(1) is the error sum of squares from ordinary least squares regressions in equation (3.4.12), ESS(2) refers to the error sum of squares in equation (3.4.13), and N is the number of observations. And to get the result of X not causing Y, the F estimate, F^*, must not be large.

The Geweke test eliminates the arbitrary selection of a filter and the possibility of bias in two–sided filtering (Hudson, 1984). Geweke also explains a test method for no instantaneous causality between two series. That is, using equation (3.4.13) and adding current values of X, then one can test for no instantaneous causality.

Several studies point out some problems in Sims' test for causality. Nelson and Schwert (1982) and Geweke Meese, and Dent (1983) present evidence that Haugh's univariate residual cross–correlation analysis and Sims' VAR causality test are not as powerful as Granger's causality test using Monte Carlo techniques. The reasons are because (1) the null hypothesis that is tested is necessary but not sufficient to imply causality, and (2) any specification error renders the causality test results uninterpretable. Therefore, VAR representation by Geweke, if needed in the analysis, will be used.

Geweke–type VAR representation in equation (3.4.13) is similar to the augmented unrestricted VAR defined by Engle and Granger (1987). Engle and Granger explained several types of VAR processes, providing the test method for cointegration from them. Those representations are briefly described as follows. In general, 'restricted denoted by R', in time series implies that the error term estimated from the regression is re–utilized to a linear regression model as an independent variable. And the 'augmented denoted by A' means that given variables take into account their lagged variables as independent variables. At this time, all the lagged series used should correspond to the characteristics of the autoregressive processes of the variables.

The restricted VAR, RVAR, which means the VAR representation consists of a vector correction error, lagged independent and lagged dependent variables. Thus the augmented RVAR is the same as RVAR except that a higher order of variables is postulated as independent variables. The UVAR is the unrestricted VAR which does not include the corrected error term in dependent variables. The AUVAR is the same as UVAR except a higher order of variables is assumed (For more discussion, see Engle et al. p. 266).

Cointegration and Error Correction Models

In the analysis of price discovery using vector autoregressive representations, data differencing could result in a loss of information, which is included in the original data series being studied. Engle and Granger (1987) remark that, "Here the cointegration is implied by the presence of the levels of the variables so a pure VAR in differences will be misspecified if the variables are cointegrated. Thus vector autoregressions estimated with cointegrated data will be misspecified if the data are differenced and will have omitted important constraints if the data are used in levels. Of course, these constraints will be satisfied asymptotically but efficiency gains ... may be achieved by imposing them (p. 259)." Therefore, a technique combining an error correction model with cointegration theory will need to be applied.

Engle and Granger provide a technique for applying an error correction model (ECM) based on data characteristics, which is said to be "cointegrated". The relationship between cointegration and ECMs was considered by Granger (1981) after Sargan (1964) explained the short–run adjustment toward long–run equilibrium. Phillips (1991), Johansen (1988) and Stock (1987) provide more technical developments for the relationship between cointegration and error correction models. Engle and Yoo (1987) give some rationale about those procedures by presenting the results of simulation experiments.

According to Granger and Newbold (1974) and Phillips (1986) as it was described, the ordinary regression is possibly spurious when variables in an equation are not stationary. In that case, the correlation coefficient, which is usually used to test Granger–type causality, will have no meaning because those variables have lost some information. Specific features in spurious regressions are that correlated stochastic trends would yield a high R^2, but the Durbin–Watson statistic calculated from nonstationary residuals tends to be low. To prevent the danger of spurious regressions, most nonstationary data series are first differenced to achieve stationarity. However, first–differencing data results in model misspecification by neglecting some missing information due to differencing. Most earlier time–series approaches ignore this fact including vector autoregressive analysis.

On the other hand, cointegration provides some evidence to include the missing information by recognizing that two variables are of the same order (p) of cointegration, and that a linear relation of two variables is cointegrated of order p–1. The following are the concepts of integration, cointegration and ECMs which will be a part of techniques applied in this study, relying basically on Engle and Granger.

Consider that there are two variables, X_t and Y_t. A series is said to be I(0) if it is stationary, where 'I' implies integration, and value, 0, in parenthesis indicates the number of order differenced to obtain a stationary series. Similarly, if a series needs to be differenced by p times to become stationary, then the order of integration will be denoted by I(p). For most economic data, the stationarity condition could be achieved after first or second differencing. But there is no firm explanation for it. So, the first step is to check the stationarity condition of each raw data series under consideration. A series can be differenced unless it is believed to be stationary.

The second step is to estimate the lag length for each series proven to be nonstationary in the first step. For the Dickey–Fuller and the augmented Dickey–Fuller tests, the length of lags for each nonstationary series will be selected. This process will be accomplished by using Schwarz's Bayesian Information Criterion (SBC) or the Akaike Information Criterion (AIC).

Schwarz (1978) derives a method which can be used to choose the optimal AR order, say p, such as

$$\mathrm{SBC}(k) = \ln\tilde{\sigma}_k^2 + \frac{k\ln T}{T}, \tag{3.5.1}$$

where k is the length of lags, T is the total number of observations, and $\tilde{\sigma}_k^2$ is the maximum likekihood estimate of the residual variance when an AR(k) is fitted to the data series. Then the length of lags, the optimul AR order $\hat{p}$ is determined such that:

$$\mathrm{SBC}(\hat{p}) = \min \{\mathrm{SBC}(k) \mid k = 1, 2, ..., m\}, \tag{3.5.2}$$

where m is some upper bound for $\hat{p}$, that is $k \leq m$. Another criterion to choose the optimal AR order $\hat{p}$ was proposed by Akaike (1974). That is,

$$\mathrm{AIC}(k) = \ln\tilde{\sigma}_k^2 + \frac{2k}{T}, \text{ and} \tag{3.5.3}$$

the optimal lags are determined by

$$\mathrm{AIC}(\hat{p}) = \min \{\mathrm{AIC}(k) \mid k = 1, 2, ..., m\}, \tag{3.5.4}$$

with the same conditions that the SBC has.

In general, the value of the SBC becomes smallest in a few lag lengths, whereas AIC tends to be lower whenever the lag length becomes longer. Thus, the length of lags are determined by using SBC and/or AIC. If the minimum values of SBC and AIC are encountered at the same order of length, there is no question about the order. But if not, one of the methods is used to determine the length of lag so that the residuals in the model for the ADF test presented in the next step can be white noise.

The third step is to test for a unit root by using the DF and/or the ADF. The models of a series, X_t, for the DF and the ADF tests are as follows:

Consider a simple autoregressive model (AR(1)) for the series, X_t. The Dickey–Fuller test for the series is accomplished using a first order autoregressive model such as

$$X_t = \alpha + \rho X_{t-1} + \epsilon_t \text{ [AR(1)]}, \tag{3.5.5}$$

where X is denoted for price series of each series of each commodity used in this study, and $t = 2, 3, ..., 1{,}272$. ρ is a real number and $\{\epsilon_t\}$ is a sequence of independent normal random variables with mean zero and variance σ^2, that is $\{\epsilon_t\} \sim NID(0,\sigma^2)$. The time series X_t converges to be a stationary state if $|\rho| < 1$. If $|\rho| = 1$, the time series is not stationary and the variance of X_t is $t\sigma^2$, and it is called a random walk. The null hypothesis that $|\rho| = 1$ is of some interest in applications because it corresponds to the hypothesis that it is appropriate to transform the time series by differencing. The test hypothesis for AR(1) representation of equation (3.5.5) is $\begin{cases} H_0: |\rho| = 1 \\ H_a: |\rho| < 1 \end{cases}$. And then the estimate of the Dickey–Fuller statistic (τ_μ) would be calculated by the ratio of $\hat{\rho}$ to the standard error of $\hat{\rho}$. As it represents a kind of pseudo–t statistic, the critical values are obtained from the original DF tabled values given in Dickey and Fuller (1979). If the estimate is less than the DF tabled value, then the null hypothesis will be rejected. Otherwise, it will not be rejected. A similar DF test is performed based on the following equation that the series responds to time variables.

$$X_t = \alpha + \beta t + \rho X_{t-1} + \eta_t \tag{3.5.6}$$

The estimated DF statistic resulting from this equation is known by τ_τ rather than τ_μ.

Similarly, on the basis of the results of Fuller (1976), assume that any time series is adequately represented by the model [AR(2)]

$$X_t = \gamma_0 + \gamma_1 t + \delta_1 X_{t-1} + \delta_2 X_{t-2} + \epsilon_t, \tag{3.5.7}$$

where ϵ_t's are NID$(0,\sigma^2)$ random variables. Equation (3.5.7) can be rewritten as:

$$X_t - X_{t-1} = \gamma_0 + \gamma_1 t + (\delta_1 + \delta_2 - 1)X_{t-1} - \delta_2(X_{t-1} - X_{t-2}) + \epsilon_t. \tag{3.5.8}$$

In equation (3.5.8), the coefficient on X_{t-1} in the right–hand side regression equation can be used to test the hypothesis that $\delta_1 + \delta_2 = 1$. Moreover, to extend the results for the first order process with p = 2 to the higher order autoregressive process, one could represent the model modified as:

$$\Delta X_t = \gamma_0 + \gamma_1 t + \varphi_1 X_{t-1} + \sum_{j=1}^{p} \varphi_{j+1} \Delta X_{t-j} + \nu_t, \tag{3.5.9}$$

where p is the number of lag. The null hypothesis is that $\varphi_1 = 0$ or $\delta_1 + \delta_2 = 1$ by definition, under the condition that $\gamma_1 = 0$, whereas the alternative hypothesis is $\varphi_1 < 0$. The null hypothesis will be rejected when the ratios of the estimate of φ_1 to its standard error (test statistic) is negative and above the critical values. The test of the null hypothesis is the so–called 'augmented Dickey–Fuller' unit root test which was suggested by Engle and Granger (1987). Because the null hypothesis is verbally integrated as "the series X_t has a unit root," or the mean of a time series X_t is a linear function of time, the rejection of the null hypothesis implies that the series X_t does not have a unit root but has a fixed mean, and so is stationary. Estimate, τ_τ, is calculated when the time variable is considered in the equation for testing. If, in equations (3.5.7) and (3.5.9), the time (t) variable is not considered, the null hypothesis is simply interpreted such that the series does not have a unit root. The estimated statistic is τ_{μ} in this case. If the length of lag is selected to be one, then 'p' in (3.5.9) will be one. For this case, the test method will be Dickey–Fuller's unit root, but not the ADF.

Steps 1 to 3 explain whether a series is nonstationary, and if it is not stationary, how to generate a stationary series with first–differencing. An I(0) series is quite different from an I(1) series in time series. An I(0) series has constant mean, and tends to diverge little from the mean. But an I(1) series may have a variable mean, and also lose some information because it is differenced. And an I(1) series tends to concentrate on an earlier value. Therefore an I(1) series potentially has long–run variability. If two series, X_t and Y_t, are both I(1), frequently called 'vector–integrated', then any linear combination of X_t + AY_t will also be I(1) in general. However, if the linear relationship represents I(0), both X_t and Y_t are said to be cointegrated.

The fourth step is to generate a corrected error process so as to set up an error correction model through the cointegrated relationship. Since both series are integrated by the order of one, a cointegrating regression model will be built:

$$X_t = \alpha + \beta Y_t + u_t. \tag{3.5.10}$$

It is easy to get the estimated residual series, $\hat{u}_t$, from the above equation. For the estimated residuals, the method similar to steps 1 through 3 would be re–introduced and applied. That is, the stationarity of $\hat{u}_t$ is checked and, if it is not stationary, the length of lags is determined using SBC or AIC. If the length of lag is determined to be one, the DF unit root test procedure is used, and if it is more than one, the ADF test is utilized.

Actually, the cointegrating relationship between two series is determined by the DF or the ADF test[1] on residuals, $\hat{u}_t$. The null hypothesis for those tests is that there exists no cointegrating relationship between two series. The DF test would have a functional form such as:

$$\hat{u}_t = a + b\hat{u}_{t-1} + \epsilon_t, \tag{3.5.11}$$

and the null hypothesis is that b = 1 against b < 1. Also, the ADF test would have a form of

$$\Delta u_t = a + b\hat{u}_{t-1} + \sum_{j=1}^{p} c_j \Delta \hat{u}_{t-j} + \epsilon_t, \tag{3.5.12}$$

with the null hypothesis of b = 0. If the estimated test statistic is negative and below the critical value, then both X_t and Y_t are said to be cointegrated.

First–differencing data series results in the loss of information with regard to the long–run relationship between pairs of series. Thus, first–order differencing potentially has long run swings. Therefore, the notion of cointegration is that long–run trends of the series adjust according to an equilibrium constraint. That is, if cointegrated, the pairs of series can be expected to move together in the long run. Co–movement is closely related to the dynamic structure of two variables and so is called "synchronization" by Tung (1990). The characteristics of the short–run dynamics of both series will be adapted to the

---

1) Other methods to verify the cointegrating relationship are to use the Durbin–Watson test statistic, the Maximum Likelihood Estimator(MLE), as suggested by Johanson (1988), the restricted VAR test, the augmented restricted VAR test, the unrestricted VAR test, and the augmented unrestricted VAR test etc. (For more discussion on various VAR tests, see Engle and Granger, 1987).

error correction models. Accordingly, the long–run relationships are integrated as an equilibrium state, and if a portion of the disequilibrium moving away from such equilibrium exists, it will be adjusted by the corrected error.

There is no similar alternative method to deal with causal relationships between two series which were rejected by the cointegration test except for vector autoregressive analysis. For pairs of series that were accepted as 'cointegrated', the causal relationship between them can be tested by using the error correction model. When cointegration relationships between pairs of series exists, the estimated error, $\hat{u}_t$, from equation (3.5.10) can be applied to the unrestricted VAR equation as a vector. If applied, then it is called 'the corrected error'.

The fifth step is to build the error correction models and test causality using the corrected error. The one period lagged correction error, $\hat{u}_{t-1}$, will be an independent vector variable in the restricted VAR representation equation such as:

$$\Delta X_t = \alpha_0 + \alpha_1 \hat{u}_{t-1} + \sum_{i=1}^{p} \beta_i \Delta X_{t-i} + \sum_{j=1}^{q} \gamma_j \Delta Y_{t-j} + \epsilon_t. \qquad (3.5.13)$$

The expression (3.5.13) is called the "error correction model (ECM)", which is a kind of restricted VAR model, specifically, the augmented restricted VAR model. Similarly, another ECM can be expressed as:

$$\Delta Y_t = \varphi_0 + \varphi_1 \hat{u}_{t-1} + \sum_{k=1}^{p} \psi_k \Delta X_{t-k} + \sum_{m=1}^{q} \kappa_m \Delta Y_{t-m} + \eta_t. \qquad (3.5.14)$$

In both ECMs, (3.5.13) and (3.5.14), Δ implies 'the change' of variables, ϵ_t and η_t are are jointly white noise residuals, respectively, and at least one of α_1 and φ_1 is non–zero. Both are statistically useful to test causal relationships between two series because cointegration and error correction models capture proportionally the omitted long–run information as well as the remaining short–run trends from the differenced data.

The null hypotheses for testing causality are as follows: In equation (3.5.13), $\begin{bmatrix} H_0\colon \alpha_1 = 0, \text{ and } \gamma_j = 0, \forall\ j, \\ H_a\colon \text{not } H_0 \end{bmatrix}$ which means that variable Y does not Granger–cause variable X. Moreover, in an alternative equation (3.5.14), $\begin{bmatrix} H_0\colon \varphi_1 = 0, \text{ and } \psi_j = 0, \forall\ k, \\ H_a\colon \text{not } H_0 \end{bmatrix}$ which means that X does not Granger–cause Y. From analyzing both equations by the ordinary least squares method, the test statistic can be obtained as a standard F statistic. The estimated F statistic must be bigger than the critical value in the F table in order for the null hypothesis to be rejected.

Cointegration and ECMs discussed in the above five steps will be one of the bases for testing causality in this study. If individual economic variables under consideration in this study vary extensively over time, meaning they include long–run swings by differenced variables, pairs of series are not accepted to be cointegrated. Then causal relationships between the series would be tested by using the unrestricted vector autoregressive model. As discussed in the fourth section, the augmented unrestricted VAR representation (AUVAR) is based on a vector autoregression which is not restricted in order to satisfy the cointegration constraint. The AUVAR is simply represented and modeled by expressions (3.5.13) and (3.5.14) except that there is not a corrected error term.

Out–of–Sample Forecasting Performance (2)

In an earlier section, one method of post–sample causality testing was discussed. The first two steps of the five–step approach presented by Ashley et al. (1980) are analogous to the Box–Jenkins procedure, which means that Ashley et al.'s method is still using univariate residual cross–correlation. Several studies have pointed out its defects.

Bessler and Kling (1984) expand the new approach to causality tests developed by Ashley et al. using Sims–type vector autoregressive representation on non–stationary economic time series. In reality, estimating a regression equation requires a prerequisite that each data series must be stationary. However, few data series satisfy the stationarity condition. When data series do not meet the stationarity condition, the procedure of differencing provides a sufficient condition regarding stationarity.

Several economic data series have characteristic properties which imply nonstationarity. If in univariate residual cross–correlation analysis, those nonstationary data are identified as good models without the rectification of nonstationarity to stationarity, there is no problem. However, due to data properties, e.g., economic time series under consideration have seasonal trends, it is somewhat hard to identify well–fitted models for those data, especially daily data with trend and seasonal patterns (For more discussion, see the seventh section). In such cases, the differencing procedure generates better model results. Differencing is necessary because it causes all the data, even non–stationary data, to be pre–whitened.

For example, let's look at out–of–sample tests presented by Ashley et al.. First, they divide the data series into two categories: within–sample and out–of–sample. Then the first part of the data series, within–sample, is modeled by using ARIMA techniques. At this time, those data could be differenced to identify a suitable ARIMA model only if the data are nonstationary. After a few steps, parameters estimated by applying ARIMA techniques to differenced data would be used on the out–of–sample data to test for Granger causality or to forecast more efficient economic values in the future. The problem arising here is that results from within–sample tests with differenced data would have different properties from those derived from undifferenced out–of–sample data. That is, Granger causality tests of out–of–sample nondifferenced data cannot be compared with the same property of causality that within–sample data have. Therefore, when data series are not stationary, differencing data by using the autoregressive properties of the series studied would generate more efficient result.

Out–of–sample performance using vector autoregressive properties of the series is as follows:

(a) Nonstationary series under consideration are differenced and modeled as autoregressive representation. Again, as in the vector autoregressive analysis, the number of lags are selected by using AIC or SBC, etc.. For instance, consider two series, say X and Y. From this step, a model is fitted as

$$X_t = \alpha + \sum_{i=1}^{p} \beta_i X_{t-i} + v_t. \tag{3.6.1}$$

(b) Now, consider using both variables to set up and estimate a bivariate model based on the number of lags for each variable.

$$X_t = \gamma + \sum_{i=1}^{p} \eta_i X_{t-i} + \sum_{j=1}^{q} \varphi_j Y_{t-j} + \epsilon_t. \tag{3.6.2}$$

The second step is to estimate models (3.6.1) and (3.6.2), and mean square errors from them as MSE(X) and MSE(X,Y), respectively. Comparisons between MSE(X) and MSE(X,Y) evaluated from within–sample data will provide the a priori belief that either X or Y causes another series. Of course, a priori beliefs can be estimated by comparing F–statistics calculated from (3.6.1) and (3.6.2). The use of an F–statistic for obtaining priori beliefs about causal relationships between pairs of series indicates the general procedure of the VAR tests with within–sample data.

(c) The third step is the procedure that all coefficient parameters estimated from equations, (3.6.1) and (3.6.2) from within–sample observations should apply to the remaining post–sample observations to produce a linear combination of the coefficients and the differenced out–of–sample data with the same order as the differenced within–sample data. From expressions (3.6.1) and (3.6.2), all coefficients estimated will be applied to out–of–sample data. The estimated coefficients are denoted by $\hat{\alpha}$, $\hat{\beta}_i$, $\hat{\gamma}$, $\hat{\eta}_i$, and $\hat{\varphi}_j$. These estimated coefficients will be combined in linear regression models as

$$\overline{X}_t = \hat{\alpha} + \sum_{i=1}^{p} \hat{\beta}_i \overline{X}_{t-i} \tag{3.6.3}$$

$$\overline{X}_t = \hat{\gamma} + \sum_{i=1}^{p} \hat{\eta}_i \overline{X}_{t-i} + \sum_{j=1}^{q} \hat{\varphi}_j \overline{Y}_{t-j} + \tag{3.6.4}$$

where $\overline{X}$ and $\overline{Y}$ indicate differenced out–of sample data whereas X and Y represent the differenced within–sample data.

(d) Then, from (3.6.3), out–of–sample mean–squared errors, MSE($\overline{X}$), would be derived. Also, mean–squared errors from equation (3.6.4) are derived, MSE($\overline{X}$,$\overline{Y}$).

(e) The fifth step in testing performance is to compare the MSE($\overline{X}$) to the MSE($\overline{X}$,$\overline{Y}$). When MSE($\overline{X}$) is greater than MSE($\overline{X}$,$\overline{Y}$), there exists a causal relationship running from the Y to X series. Mean–squared errors represent the size of individual forecast errors from the actual values.

(f) The last step is to compare the results of causality tests from within–sample and out–of–sample data. If a priori beliefs about the relationship between two series obtained from within–sample data are the same as the results of out–of–sample causality tests, it provides evidence about the causal relationships. If not, causal relationships are doubtful.

Thus, causal relationships between variables are explored from the better performance test yielding the lower MSE from out–of–sample data to the worse performance test from within–sample data.

More details developed for testing the significance of improvements in mean–squared forecasting error (MSE) are as follows. The procedure is summarized into two steps. First, all relative parameters from pre–whitened within–sample observations (e.g., 1987 through 1990 in this study) are estimated. Then the obtained MSEs from within–sample univariate and bivariate models, or the estimated F statistics are compared. As a result of the comparison, a priori beliefs about causal relationships will be demonstrated. Second, those parameters would be applied to retained differenced raw data, out–of–sample observations (e.g., 1991 data in this study). Then since parameters were estimated from the univariate as well as bivariate models on within–sample observations, two different mean square errors according to different models are calculated. If mean–squared errors for the univariate model (say, $\overline{Y}_t$) are greater than that for the bivariate model (say, $\overline{X}_t$ and $\overline{Y}_t$), in the sense of out–of–sample forecasting, it is said $\overline{X}_t$ causes $\overline{Y}_t$.

Because the basic definition of causality which is employed to demonstrate price discovery in this study is a statement about forecasting ability, the comparison of mean–squared errors usually used in forecasting to test causality is more reasonable.

Checking for Seasonality

Many economic time series contain important seasonal components. Thus various models of seasonality which may differ among series could exist.

Seasonal price trends of hog–related commodities under study are usually due to the particular way in which the commodities are produced and/or distributed. Even though seasonal price trends are produced by the interaction of production and/or consumption factors for a commodity, such patterns are not described here. Seasonal patterns are defined as repeating patterns of prices that can be discovered and measured from raw data observations. Figure 3 shows each price series. Seasonal patterns cannot be seen clearly in Figure 3. In sum, for the five years under consideration, prices in early summer seemingly show a little seasonal trend in each year. For the futures market prices, no seasonal pattern is found. Prices follow the random walk hypothesis of futures prices. If futures prices are first–order differenced, then the residuals satisfy the requirements to be a stationary process.

The Geweke–type or Granger's own testing for causality are performed based on the data pre–filtered to remove trend and seasonality. Thus, although the data series have a seasonal pattern, it will not be a restriction to applying Granger–causality tests.

Most of the current literature for price discovery using cointegration fail to consider the effects of seasonal integration. If price series used in this study represent non–seasonal or non–periodic patterns, it would not be necessary to consider it. However, most agricultural commodities have seasonal factors in either demand or supply, so their price series may also exhibit seasonal patterns.

Assume that there are two economic time series of interest, X_t and Y_t. The methods to identify and test cointegrating relationships between two series

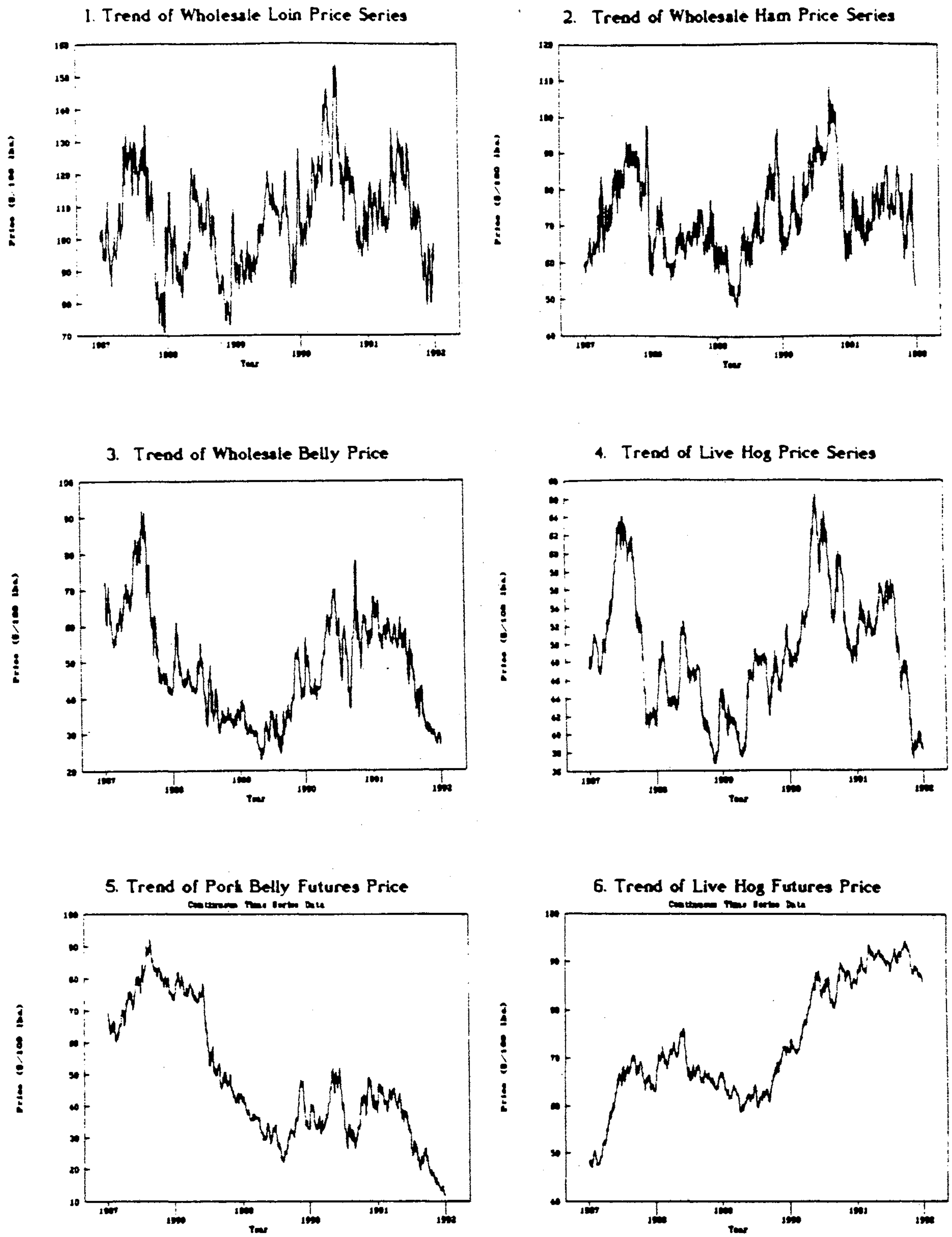

Figure 3. Trends of the Price Series

are summarized in the literature as follows.

First, if both series do not have any seasonal variation, then testing the possibility of cointegration can be accomplished as in the previous section.

Second, if two series are seasonally cointegrated at zero frequency but not seasonal frequencies, then a standard cointegration test can be performed on the revised forms of series, $\tilde{X}_t = S(L)X_t$, $\tilde{Y}_t = S(L)Y_t$, which represent removed seasonal roots by $S(L) = \dfrac{(1-L^s)}{(1-L)}$, where s is an order of seasonal differentiation. Engle et al. (1989) said that if time series under consideration are seasonally cointegrated at zero frequency but not seasonal frequencies, then the value of parameter (α) yielded from the static OLS equation $X_t = \alpha Y_t + \epsilon_t$, would not generally be consistent. But, using $\tilde{X}_t$ and $\tilde{Y}_t$ will generate a consistent parameter ($\tilde{\alpha}$).

Third, consider that both time series X_t and Y_t are integrated at some of the zero and seasonal frequencies. That is, both series have unit roots at the zero frequency (as in the first case) as well as at some seasonal frequencies. Then, seasonal unit roots must be pretested on each series in order to test cointegration on both series. If some seasonal unit roots were thought to be present in both series, then the seasonal roots must first be removed and then the standard cointegration tests can be applied, such as in the second case. The test method for this case was introduced by Hylleberg et al. (1990) and is as follows. They considered testing seasonal unit roots in quarterly data.

When the quarterly data are considered and seasonality is assumed to be present, the seasonal differencing operator introduced by Box and Jenkins (1976) is

$$(1-L^4)X_t = \epsilon_t \tag{3.7.1}$$

$$= (1-L)\,(1+L+L^2+L^3)X_t$$
$$= (1-L)\,(1+L)\,(1+L^2)X_t$$
$$= (1-L)\,S(L)X_t$$

In the above equation, the polynomial $(1-L^4)$ in the left–hand side can be solved with four unit roots equal to 1, –1, i and –i, where i represents a complex form. That is,

$$\begin{aligned}(1-L^4) &= (1-L)\,(1+L)\,(1+L^2) \\ &= (1-L)\,(1+L)\,(1-iL)\,(1+iL)\end{aligned} \tag{3.7.2}$$

They proved that a general autoregression, $\psi(L)X_t = \epsilon_t$ can be rewritten as

$$\varphi(L)Y_{4t} = \pi_1 Y_{1t-1} + \pi_2 Y_{2t-1} + \pi_3 Y_{3t-2} + \pi_4 Y_{3t-1} + \epsilon_t, \tag{3.7.3}$$

where
$$\begin{aligned}Y_{1t} &= S(L)X_t \\ Y_{2t} &= -(1-L+L^2-L^3)X_t \\ Y_{3t} &= -(1-L^2)X_t \\ Y_{4t} &= (1-L^4)X_t\end{aligned}$$

Then seasonality would be tested by the parameters of equation (3.7.3). The hypotheses on the test are i) $\pi_1 = 0$ for a unit root of 1, in equation (3.7.1), ii) $\pi_2 = 0$ for –1, iii) $\pi_3 = 0$ for i and –i. If π_2 and either π_3 or π_4 are not equal to zero (rejection of ii and iii), then there will not be seasonal unit roots. If the series has no unit roots, which means it is stationary, then all π_j (j=1, 2, 3, 4) must not be zeros.

Price series considered in this study have a seasonal pattern during the early summer period. However, since all data series are daily indices except for weekend and national holidays, and the number of yearly observations average 255, seasonality cannot be tested using current statistical methods.

Error correction models contain long–run swings against the omitted information which results from differencing the series. Therefore, even though seasonals are not considered in the process of cointegration, it will be concluded that the results from ECMs are of relatively high quality.

Bessler and Kling (1984) introduced in–sample and out–of–sample tests for Granger causality. Whether or not there is causality from Y to X is defined by using the autoregressive property of the data series. That is, the fact that an optimal forecasting model for X_t using past values of X and Y performs better than one using only past values of X was used in testing for in–sample and out–of–sample Granger causality. The data are filtered to remove trend and seasonality before those tests are applied to the data series. Therefore, even though there seems to be seasonality, it may not be problem due to the process of pre–filtering.

Summary

Price discovery processes have been analyzed by using various techniques based on Granger's causality. The use of a Box–Jenkins type model is easier to identify and to prevent spurious regression problems. This approach, however, does not consider the autoregressive properties of the data.

Meanwhile, the vector autoregressive test method cannot account for the long run state even though it responds to the property of autoregressiveness.

To avoid both defects and generate better results, cointegration and error correction models are simultaneously used in the price discovery process. But, if all data series do not satisfy suitable conditions, such as stationarity, first–order integration, and in turn, cointegration etc., the results from these methods will also be suspect. In this case, an out–of–sample causality test is conducted as an independent way to analyze the price series.

One of the goals in this study was to examine the process of price discovery in cases where the data series have exhibited seasonal patterns. Unfortunately, a seasonal cointegration study cannot be accomplished because there are operational problems associated with the daily data. If the data series have problems because of seasonality when they are analyzed in the context of price discovery, they may be analyzed by using alternative methods explained in the next chapter.

CHAPTER IV

APPLICATION OF CAUSALITY TESTS AND EMPIRICAL RESULTS

Introduction

In the previous chapter, several bivariate model approaches were discussed for studying price discovery. All are not utilized to empirically analyze discovering price in this study. This chapter moves to estimation and testing, focusing on details of the results.

The general procedure for applying causality tests is shown in Figure 4. First, all data sets will be collected and futures price series will be generated as a continuous–type data series, which is described in the following section. The second process is to test stationarity of each time series by using normality tests. That is, testing the normality of residuals will determine whether the series is necessarily and sufficiently stationary or not. If the series are stationary, then the causal relationships between a pair of series can be analyzed by the VAR process. The third step, in case the stationarity conditions of each series are not satisfied, stationarity can be re–tested by using the DF and/or the ADF tests. However, stationarity proven by these tests will be necessary but not sufficient. Since most economic time series are not stationary, the necessary stationarity state of those series can be generated.

The fourth procedure has three branches. One is to test causality using an error correction model for the series proven to have cointegrating relationships.

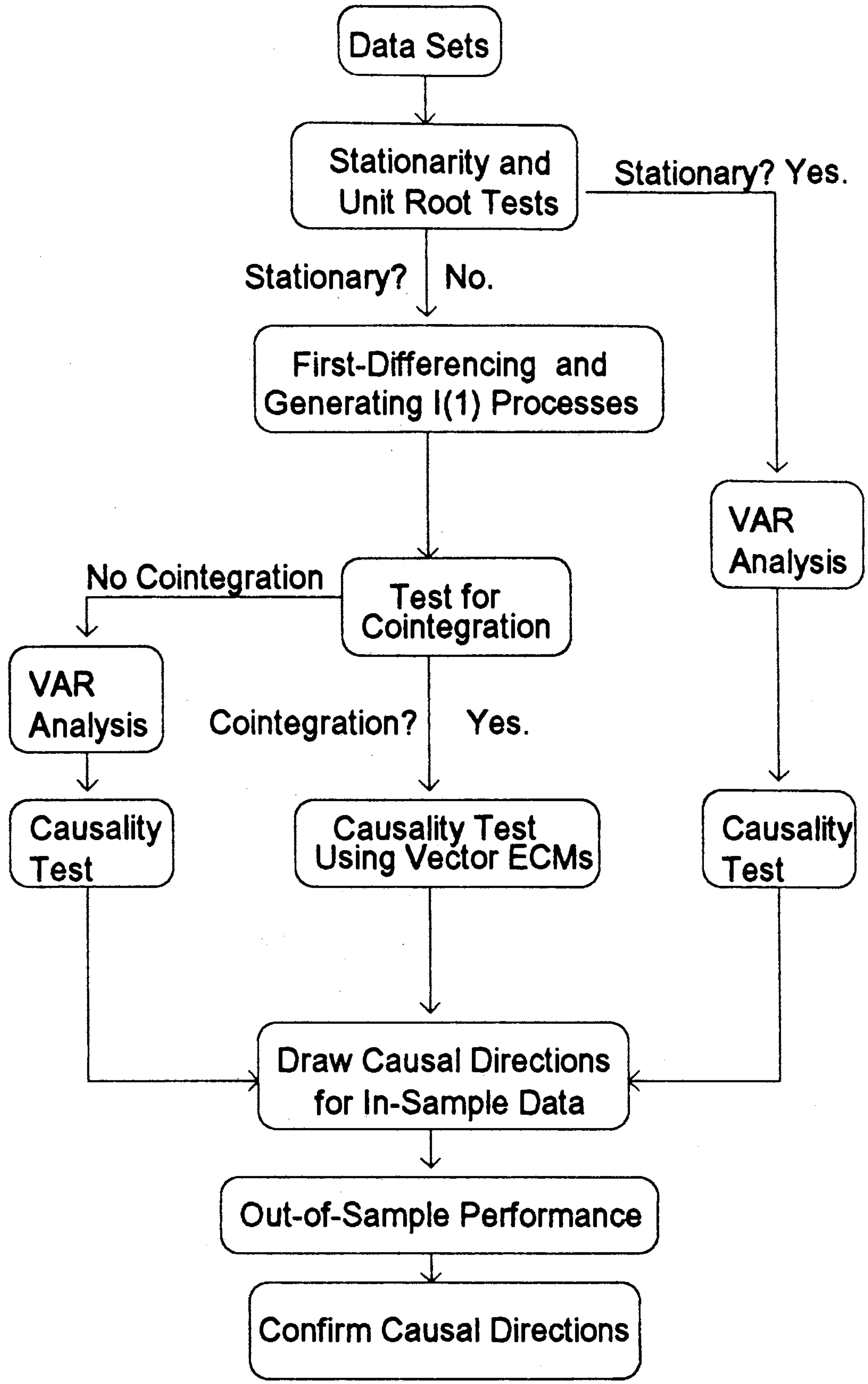

Figure 4. Procedure for Causality Tests

The second is to test the series using a VAR process for the series which are not cointegrated with each other. The third is to use a VAR model for analyzing the data series which are proven to be stationary in the second step. If the series are cointegrated, then the error term from the cointegrating regression equation will be corrected and the corrected error will be used as a vector. Because the error correction model is a kind of VAR representation using the vector error correction term, the causal relationship can be tested. The fifth step in Figure 4 is to draw conclusions and implications resulting from causality tests for the in–sample data. The sixth step is to test causality by using the out–of–sample method, which can be performed whether or not the series under consideration are cointegrated. Finally, causal directions resulting from the out–of–sample data are compared to those from the in–sample data, and the results of causal directions are confirmed.

Of the entire procedure, an important step is to check for stationarity. If the series are not stationary, then they must be made stationary. Without checking for stationarity, well–fitted results will not represent the real properties of the economic time series.

In the following section, data collection and the continuous futures market price series are discussed. Stationarity will also be examined.

Causal relationships between pairs of series for whole time–spanned sample observations are described in the third section. This section includes an explanation of test results from out–of–sample causality performance, the vector error correction model, and the unrestricted VAR analysis.

In the fourth section, whether or not causal relationships between pairs of series have changed over time will be analyzed. Only two methods are utilized, ECMs and the AUVAR process.

Finally, conclusions are found in the fifth section, including the discussion of seasonal patterns which appear in the data sets.

Data and Stationarity

Data Discussion

All the variables in this study are daily series collected from 1987 to 1991. Daily prices in general have more information on changes and trends about the original series than weekly, monthly, or quarterly. Thus, losing information due to the use of weekly or some other non–daily data will result in distorted results even though there is no problem with analyzing the data using the methodology discussed.

The daily spot price series for live hogs and for wholesale pork (loins, hams and bellies) used in this study are collected by U.S.D.A.. The daily price series for live hogs are for hogs categorized as U.S.D.A. grade 1–2 and weighing 210–240 pounds in Iowa and southern Minnesota. Wholesale pork prices are collected in Omaha. The weight of fresh wholesale pork products are 14–18 pounds for loins, 20–26 pounds for hams, and 14–16 pounds for bellies, respectively.

The daily closing futures price series for pork belly and live hog futures prices are collected from the Chicago Mercantile Exchange. The pork belly futures contract maturity months are Feburary, March, May, July, and August. The maturity months for the live hog futures contract are Feburary, April, June, July, August, October, and December.

In this study, continuous futures prices for both series are constructed in order to aviod the jumps and falls which occur when maturity months changed. The procedures to generate smoothing the series (which was introduced by Djunaidi et al., 1993) are as follows: (1) All price series for each contract are first–order differenced, $\Delta FUT_{t,h} = FUT_{t,h} - FUT_{t-1,h}$, where t is date and h is contract. (2) ΔFUT to a new contract month is switched on the last day of the month prior to contract maturity. Since the procedure (2) generates a continuous series of first–order differenced nearby futures contract, it can be used to construct a continuous futures price series. Let's assume that the first observation of each futures price series is its actual futures price. (3) Then, the modified continuous futures price series (MFUT) are created as

$$
\begin{aligned}
&\text{MFUT}_1 \text{ is assumed to be known,} \\
&\quad \text{MFUT}_2 = \text{MFUT}_1 + \Delta\text{FUT}_2, \\
&\quad \text{MFUT}_3 = \text{MFUT}_2 + \Delta\text{FUT}_3, \\
&\quad \cdot \\
&\quad \cdot \\
&\quad \text{MFUT}_t = \text{MFUT}_{t-1} + \Delta\text{FUT}_t.
\end{aligned}
\tag{4.2.1}
$$

The overall data period used is January 5, 1987 through December 27, 1991. Thus, the number of observations in each series is 1,272. A few missing values, around 15 out of 1,272 on the average in each series, were found. In those cases, the price of the previous day was used.

Variables are abbreviated as follows:

LH: cash market live slaughter hog prices.

WL: wholesale market loin prices.

WH: wholesale market ham prices.

WB: wholesale market pork belly prices.

FLH: live hog futures market prices.

FPB: pork belly futures market prices.

Furthermore, year observations for each variable are also simplified, e.g., live slaughter hog prices collected in the cash market during 1987 are abbreviated by '87LH', wholesale pork belly prices collected in the wholesale meat market for 1988 by '88WB', and pork belly futures market prices collected in CME and constructed to be continuous in 1991 by '91FPB', and so on.

Continuous futures prices for both pork bellies and live hogs for each year are yearly segments from the entire continuous data series.

Stationarity and the Dickey–Fuller Test

All live hogs, wholesale pork (loins, hams and bellies), and futures price (live hogs and pork bellies) series are tested for the normality of residuals and to determine stationarity. Checking the stationarity of residuals is the process of making them prewhitened.

For each series, the dependent variable, is regressed on time, and then residuals are tested for normality graphically and by the Blom test method. Each series was non–stationary. From the graphs of residuals, variances of each series look stationary but their means exhibit non–stationarity over time.

To remove non–stationarity, each series is first–order differenced. However, first–order differencing is often more legitimate after Dickey–Fuller's (1979) unit root test is made to check for random walk behavior.

The Dickey–Fuller test for each series was accomplished using a first–order autoregressive model such as equations (3.5.5) for $\hat{\tau}_{\mu}$, and (3.5.6) for $\hat{\tau}_{\tau}$. The DF tabled values for τ_{μ} and τ_{τ} are –3.43 and –3.96 in significant level of $\alpha =$ 0.01, –3.12 and –3.66 in significance of $\alpha = 0.025$, and –2.86 and –3.41 in $\alpha = 0.05$, respectively, when sample size is greater than 500. These tabled values represent the probability of a smaller value.

Table 3 represents the results of unit root tests on each series. The results of the Dickey–Fuller test show that live hogs, wholesale bellies, live hog futures and pork belly futures series are random walk, but wholesale loins and hams price

TABLE 3

DICKEY–FULLER'S UNIT ROOT TEST
FOR ALL OBSERVATIONS

Series	$\hat{\tau}_{\mu}$	$\hat{\tau}_{\tau}$
Live hog price	–1.46	–1.39
Wholesale loin price	–3.45	–3.48
Wholesale ham price	–3.69	–3.61
Wholesale belly price	–2.26	–2.31
Futures live hog price	–1.79	–1.77
Futures pork belly price	–0.53	–1.74

series are not random walk even at a significant level of $\alpha = 0.01$ (cf. the calculated $\hat{\tau}_{\mu,\mathrm{Ham}} = -3.69$, and $\hat{\tau}_{\mu,\mathrm{Loin}} = -3.445$) (Table 3).

Visually, both mean and variance of the first–order differenced series appeared to be stationary. However, testing for random walk for each series shows that all except two series are random walk.

Another important question described in the problem statement section was whether price discovery has changed over time or not. As a preliminary condition for the question, stationarity and Dickey–Fuller tests were applied to each year's price series under consideration. Total observations for each series number 1,272, and each series may be divided into 5 years, by year, or about 255 observations per year.

Normality tests on the residuals of yearly–segmented data series for each price series except for 1988 wholesale ham prices are rejected by the Blom test. Because, in 1988, the wholesale ham price series is stationary in mean and variance over time, it cannot necessarily be first–order differenced. To accomplish cointegration tests between pairs of series, however, it was also first–order differenced after conducting the DF test. Table 4 show the results of the unit root test on each yearly data series.

TABLE 4

THE DF TEST FOR YEARLY–SEGMENTED SAMPLE OBSERVATIONS

Series	H_0	Year				
		87	88	89	90	91
LH	$\hat{\tau}_\mu$	–0.28	–1.47	–1.11	–1.28	–0.06
	$\hat{\tau}_\tau$	–0.72	–2.19	–2.02	–1.07	–1.98
WL	$\hat{\tau}_\mu$	–1.66	–1.73	–1.77	–1.54	–1.87
	$\hat{\tau}_\tau$	–1.73	–1.85	–2.1	–1.61	–2.48
WH	$\hat{\tau}_\mu$	–2.09	–3.03	–1.69	–1.57	–2.40
	$\hat{\tau}_\tau$	–1.57	–3.17	–2.16	–0.7	–2.15
WB	$\hat{\tau}_\mu$	–0.80	–1.71	–1.15	–2.16	–0.48
	$\hat{\tau}_\tau$	–1.14	–3.49	–1.94	–2.21	–3.54
FLH	$\hat{\tau}_\mu$	–1.71	–1.85	–0.31	–1.90	–2.53
	$\hat{\tau}_\tau$	–0.37	–3.36	–2.1	–1.47	–2.66
FPB	$\hat{\tau}_\mu$	–1.35	–0.16	–1.67	–1.72	–0.10
	$\hat{\tau}_\tau$	–1.22	–2.55	–1.65	–1.69	–4.17

This test is based on the following first–order autoregressive model.

$$X_{t,h} = \beta_1 + \beta_2 X_{t-1,h} + \epsilon_{t,h} \quad [\text{AR}(1)], \tag{4.2.2}$$

$$\begin{aligned} t &= 1,\ 2,\ \ldots,\ 255 \text{ (for 1987, 88, 90)} \\ & 254 \text{ (for 1989)} \\ & 253 \text{ (for 1991), and} \\ h &= 1987,\ 88,\ ..,\ 91. \end{aligned}$$

The test hypothesis is $\begin{bmatrix} H_0:\ |\beta| = 1. \\ H_a:\ |\beta| < 1 \end{bmatrix}$

Similarly, values in parenthesis in Table 2 indicate another test statistic (τ_τ), which is calculated from equation (4.2.3).

$$X_{t,h} = \gamma_1 + \gamma_2 X_{t-1,h} + ct + \epsilon_{t,h} \ [\text{AR}(1)], \tag{4.2.3}$$

$$\begin{aligned} t &= 1,\ 2,\ \ldots,\ 255 \text{ (for 1987, 88, 90)} \\ & 254 \text{ (for 1989)} \\ & 253 \text{ (for 1991), and} \\ h &= 1987,\ 88,\ ..,\ 91. \end{aligned}$$

The test hypothesis is $\begin{bmatrix} H_0:\ |\beta| = 1 \text{ conditional on } c = 0. \\ H_a:\ \text{not } H_0 \end{bmatrix}$

The cumulative critical values for τ_μ and τ_τ at a 2.5 % confidence level for sample sizes of 250 are –3.14 and –3.69, respectively. At a 5 % confidence level with same sample size, the cumulative critical values for τ_μ and τ_τ are –2.88 and –3.43, respectively. The null hypothesis is rejected if the estimated $\hat{\tau}_\mu$ or $\hat{\tau}_\tau$ are smaller than the cumulative critical values at a given significant level. That is, it can be concluded that all series in every year exhibit a random walk.

The Augmented Dickey–Fuller (ADF) Test

The previous Dickey–Fuller unit root test results assumed that all the series under consideration in this study are of the first–order autoregressive model [AR(1)]. Similarly, consider that any time series is represented by the second– or higher–order autoregressive model. In that case, the augmented Dickey–Fuller tests was accomplished based on equations (3.5.7) or (3.5.9).

To appropriately satisfy the ADF test, a prerequisite is to determine the number of lags which must be used for the regression model (3.5.9). The number of lags chosen as the maximum possible order to consider is shown in Tables 5 and 6 for all observations and yearly observations, respectively, and was determined by using Schwarz's Bayesian Criterion (SBC) shown in equations (3.5.1) and (3.5.2) and/or Akaike Information Criterion (AIC) given in equations (3.5.3) and (3.5.4). The critical values of Dickey–Fuller's pseudo t–statistics for a sample size greater than 500 and at a 2.5 % confidence level are $\tau_{\mu} = -3.12$, and $\tau_{\tau} = -3.66$. With sample size of 250, $\tau_{\mu} = -3.14$ and $\tau_{\tau} = -3.69$.

Thus, Table 7 shows that for no series, the null hypothesis is rejected. That is, price series for wholesale loins, hams and bellies, and for cash slaughter hogs can be first–order differenced because all series are proven to have a unit root. Since, according to the SBC and AIC, pork belly futures and live hog futures prices have only one lag in the autoregressive representation (Table 5), these two series will be tested by using the Dickey–Fuller test instead of the augmented Dickey–Fuller test. This means we cannot apply it to equation (3.5.9) but to equation (3.5.5). Similarly, the results of unit root tests on the yearly segmented data series are shown in Table 8. All calculated pseudo t–statistics in the table imply that the null hypotheses cannot be rejected. Statistics in parentheses in tables 7 and 8 represent the formal Dickey–Fuller test statistics,

TABLE 5

THE LENGTHS OF LAGS – ALL OBSERVATIONS

Series	Length of Lags for Max. AR Order
Live hog	5
Wholesale loin	17
Wholesale ham	13
Wholesale belly	5
Pork belly futures	1
Live hog futures	1

TABLE 6

THE LENGTHS OF LAGS – YEARLY OBSERVATIONS

Series	Year				
	87	88	89	90	91
LH	3	6	3	3	8
WL	2	2	2	2	1
WH	2	1	3	3	5
WB	1	4	2	2	8
FPB	1	2	1	2	1
FLH	1	1	1	1	1

TABLE 7

THE UNIT ROOT TESTS – ALL OBSERVATIONS

H_0	Calculated Pseudo t–Statistics					
	WL	WH	WB	LH	FPB	FLH
$\hat{\tau}_{\mu}$	–3.06	–3.19	–2.32	–2.24	–0.53*	–1.46*
$\hat{\tau}_{\tau}$	–3.06	–3.02	–2.37	–2.19	–1.74*	–1.39*

Values with asterisk represent the DF test results, otherwise, values indicate the ADF test results. A one–sided test is conducted. At a 2.5 % confidence level, the critical values for the statistic for sample sizes of over 500 are $\tau_{\mu} = -3.12$, and $\tau_{\tau} = -3.66$.

TABLE 8

THE UNIT ROOT TESTS – YEARLY OBSERVATIONS

Series	H_0	Calculated Pseudo t–Statistic				
		87	88	89	90	91
LH	$\hat{\tau}_{\mu}$	–0.06	–2.11	–0.80	–1.18	–0.32
	$\hat{\tau}_{\tau}$	–0.49	–2.56	–1.88	–0.97	–1.63
WL	$\hat{\tau}_{\mu}$	–2.08	–2.63	–2.36	–2.47	–1.88
	$\hat{\tau}_{\tau}$	–2.15	–2.79	–2.55	–2.54	–2.48
WH	$\hat{\tau}_{\mu}$	–2.56	(–3.03)	–1.59	–2.14	–2.99
	$\hat{\tau}_{\tau}$	–2.46	(–3.17)	–2.53	–1.67	–2.80
WB	$\hat{\tau}_{\mu}$	(–1.8)	–1.66	–1.61	–2.77	–0.52
	$\hat{\tau}_{\tau}$	(–1.1)	–3.32	2.50	–2.95	–2.20
FPB	$\hat{\tau}_{\mu}$	(–1.4)	–0.38	(–1.7)	–2.06	(0.1)
	$\hat{\tau}_{\tau}$	(–1.2)	–2.75	(–1.7)	–2.03	(–4.1)
FLH	$\hat{\tau}_{\mu}$	(–1.7)	(–1.8)	(–0.3)	(–1.9)	(–2.5)
	$\hat{\tau}_{\tau}$	(–0.4)	(–3.3)	(–2.1)	(–1.4)	(–2.6)

A one–sided test is conducted. At a 2.5 % confidence level, the critical values of the statistic for sample sizes of over 250 are τ_{μ} = –3.14, and τ_{τ} = –3.69. Values in parenthesis represent the DF test results, otherwise, values indicate the ADF test results.

whereas other values indicate the standard augmented Dickey—Fuller test statistics. The rejection of the unit root hypothesis provides the necessary but not sufficient condition to conclude the series is stationary.

Tables 3, 4, 7, and 8, in conclusion, indicate that all sample series and yearly data series can be first—order differenced in order to satisfy the condition of necessary stationarity. For a few series, either the τ_μ or τ_τ statistic is smaller than the critical value, which means that first differencing is doubtful. In these cases, first differencing was achieved according to either one of the τ_μ or τ_τ, which satisfies the condition that the series are random walk.

Results for 1987—1991

Results from the DF tests showed that wholesale loins and hams price series were not a random walk. Even though they are not random walk, results do not indicate they are stationary. Although the other series except for those two series satisfy the random walk criteria, test results do not imply that first—differenced series will be sufficiently stationary. Augmented DF tests confirmed for all variables that every time series had a unit root, which means all series are nonstationary. Therefore, all the series are first—order differenced so that the first—differenced series can meet the necessary stationarity condition.

There are eleven combinations or pairs of series which will be modeled as simple linear regression models (see Table 9).

Testing for Cointegration

In Table 9, when the X_t series is considered the dependent variable and the Y_t series is the independent variable, the ordinary linear regression models are called 'the cointegrating regression model', because all the series are integrated of order one, I(1). In the cointegrating regressions, according to previous beliefs, cash prices will be the dependent variable and futures prices the independent variable (model numbers 1 and 5); cash prices are the dependent variable and wholesale pork prices the independent variable (model numbers 9—11); and wholesale meat prices are the dependent variable and futures prices the dependent variable (model numbers 2—4 and 6—8).

TABLE 9

PAIRS OF TWO SERIES MODELED

No. of Model	X_t	Y_t
1	Live hog	Pork belly futures
2	Wholesale loin	Pork belly futures
3	Wholesale ham	Pork belly futures
4	Wholesale belly	Pork belly futures
5	Live hog	Live hog futures
6	Wholesale loin	Live hog futures
7	Wholesale ham	Live hog futures
8	Wholesale belly	Live hog futures
9	Live hog	Wholesale loin
10	Live hog	Wholesale ham
11	Live hog	Wholesale belly

Two variables are said to be cointegrated when a linear combination of the two is stationary, even if each variable is not stationary. Whether or not two series shown in Table 9 are cointegrated should be tested .

The cointegration tests of the price series are conducted to determine whether a linear combination of the two series is integrated of order zero, I(0). There are several test methods for cointegration such as the Durbin—Watson (DW) statistic, the DF test, the ADF test, and MLE.

The DW test is based on the estimated cointegrating linear regression model. Estimated cointegration models are as follows:

$$\underset{\substack{(0.48)\\ [95.4]}}{LH_t = 46.09} + \underset{\substack{(0.009)\\ [7.25]}}{0.067FPB_t} + u_{1,t} \tag{4.3.1}$$

$$\underset{\substack{(1.09)\\ [98.7]}}{WL_t = 107.69} + \underset{\substack{(0.021)\\ [-1.61]}}{0.034FPB_t} + u_{2,t} \tag{4.3.2}$$

$$\underset{\substack{(0.80)\\ [89.9]}}{WH_t = 72.15} + \underset{\substack{(0.02)\\ [1.43]}}{0.022FPB_t} + u_{3,t} \tag{4.3.3}$$

$$\underset{\substack{(0.86)\\ [33.96]}}{WB_t = 29.20} + \underset{\substack{(0.017)\\ [23.98]}}{0.396FPB_t} + u_{4,t} \tag{4.3.4}$$

$$\underset{\substack{(1.09)\\ [32.58]}}{LH_t = 35.53} + \underset{\substack{(0.015)\\ [12.83]}}{0.188FLH_t} + u_{5,t} \tag{4.3.5}$$

$$\underset{\substack{(2.38)\\ [30.39]}}{WL_t = 72.31} + \underset{\substack{(0.032)\\ [14.38]}}{0.461FLH_t} + u_{6,t} \tag{4.3.6}$$

$$\underset{\substack{(1.72)\\ [26.78]}}{WH_t = 46.11} + \underset{\substack{(0.023)\\ [15.96]}}{0.370FLH_t} + u_{7,t} \tag{4.3.7}$$

$$\underset{\substack{(2.42)\\ [15.44]}}{WB_t = 37.29} + \underset{\substack{(0.033)\\ [4.59]}}{0.149FLH_t} + u_{8,t} \tag{4.3.8}$$

$$\underset{\substack{(0.71)\\ [11.98]}}{LH_t = 8.51} + \underset{\substack{(0.007)\\ [57.99]}}{0.385WL_t} + u_{9,t} \tag{4.3.9}$$

$$LH_t = \underset{\substack{(1.02) \\ [21.88]}}{22.34} + \underset{\substack{(0.014) \\ [26.73]}}{0.370}WH_t + u_{10,t} \tag{4.3.10}$$

$$LH_t = \underset{\substack{(0.42) \\ [75.04]}}{31.38} + \underset{\substack{(0.008) \\ [44.74]}}{0.372}WB_t + u_{11,t} \tag{4.3.11}$$

According to the DW test results, all combinations of pairs under consideration have cointegrating relationships (Table 10).

TABLE 10

COINTEGRATION TESTS USING DW STATISTICS[1)]

NO. of model	X_t	Y_t	DW statistic
1	LH	FPB	0.011
2	WL	FPB	0.037
3	WH	FPB	0.035
4	WB	FPB	0.020
5	LH	FLH	0.011
6	WL	FLH	0.043
7	WH	FLH	0.053
8	WB	FLH	0.015
9	LH	WL	0.107
10	LH	WH	0.032
11	LH	WB	0.038

1) The null hypothesis is that there is no cointegration. The critical value given by Engle and Yoo (1987, p.158) to reject the null hypothesis is 0.2 at 5 % confidence level with a sample size of 200. Since all calculated DW statistics are less than the critical values, cointegrating relationships between each two series exists.

However, Engle and Yoo stated that the DW statistic does not appear to be very useful for testing cointegration. Therefore, whether a linear relationship is cointegrated will be determined by other test methods besides the DW statistic.

The use of the DF or ADF tests to test cointegration must presume that the stationarity condition on the residuals from the cointegrating regression model will be checked. If not, the lag length of residuals must be pre–determined. If the residual series is stationary, both series are believed to be cointegrated. Checking stationarity on the residuals indicates that no residual series is stationary. Thus, the lengths of lags are determined by using the minimum value of the SBC. Table 11 shows the lengths of lags for residuals to be used to test for cointegration. When the lengths of lags is known to be one, cointegration will be tested by the DF test. If it is more than one, the ADF test will be used to test for cointegration.

TABLE 11

THE LENGTHS OF LAGS ON RESIDUALS
– ALL OBSERVATIONS –

Y_t Series	X_t Series				
	LH	WL	WH	WB	FPB
FPB	4	2	4	2	—
FLH	3	2	4	2	1
WL	2	—	3	2	—
WH	3	—	—	3	—
WB	1	—	—	—	—

The DF or ADF test of the residuals for the estimated equations from (4.3.1) through (4.3.11) are as follows:

$$\Delta\hat{u}_{1,t} = \underset{(-0.25)}{-0.005} - \underset{(-1.81)}{0.005}\hat{u}_{1,t-1} - \underset{(-5.26)}{0.147}\Delta\hat{u}_{1,t-1} + \underset{(7.91)}{0.223}\Delta\hat{u}_{1,t-2}$$

$$- \underset{(-3.47)}{0.097}\Delta\hat{u}_{1,t-3} - \underset{(-4.16)}{0.116}\Delta\hat{u}_{1,t-4} \qquad (4.3.12)$$

$$\Delta\hat{u}_{2,t} = \underset{(-0.02)}{-0.001} - \underset{(-4.70)}{0.025}\hat{u}_{2,t-1} - \underset{(-8.68)}{0.243}\hat{u}_{2,t-1} + \underset{(-2.14)}{0.060}\Delta\hat{u}_{2,t-2} \qquad (4.3.13)$$

$$\Delta\hat{u}_{3,t} = \underset{(0.005)}{0.0003} - \underset{(-3.82)}{0.022}\hat{u}_{3,t-1} - \underset{(-11.99)}{0.337}\Delta\hat{u}_{3,t-1} + \underset{(5.265)}{0.155}\Delta\hat{u}_{3,t-2}$$

$$+ \underset{(2.711)}{0.079}\Delta\hat{u}_{3,t-3} - \underset{(0.239)}{0.007}\Delta\hat{u}_{3,t-4} \qquad (4.3.14)$$

$$\Delta\hat{u}_{4,t} = \underset{(-0.38)}{-0.018} - \underset{(-2.83)}{0.011}\hat{u}_{4,t-1} - \underset{(-3.33)}{0.093}\hat{u}_{4,t-1} + \underset{(1.17)}{0.033}\Delta\hat{u}_{4,t-2} \qquad (4.3.15)$$

$$\Delta\hat{u}_{5,t} = \underset{(-0.75)}{-0.014} - \underset{(-1.07)}{0.003}\hat{u}_{5,t-1} - \underset{(-4.19)}{0.118}\Delta\hat{u}_{5,t-1} + \underset{(9.73)}{0.265}\Delta\hat{u}_{5,t-2}$$

$$- \underset{(-3.37)}{0.095}\Delta\hat{u}_{5,t-3} \qquad (4.3.16)$$

$$\Delta\hat{u}_{6,t} = \underset{(-0.14)}{-0.011} - \underset{(-4.97)}{0.028}\hat{u}_{6,t-1} - \underset{(-8.46)}{0.236}\hat{u}_{6,t-1} + \underset{(-2.09)}{0.059}\Delta\hat{u}_{6,t-2} \qquad (4.3.17)$$

$$\Delta \hat{u}_{7,t} = \underset{(-0.19)}{-0.012} - \underset{(-3.84)}{0.024\hat{u}_{7,t-1}} - \underset{(-11.5)}{0.324\Delta\hat{u}_{7,t-1}} + \underset{(5.01)}{0.147\Delta\hat{u}_{7,t-2}}$$

$$- \underset{(-2.74)}{0.080\Delta\hat{u}_{7,t-3}} - \underset{(-0.25)}{0.007\Delta\hat{u}_{7,t-4}} \qquad (4.3.18)$$

$$\Delta \hat{u}_{8,t} = \underset{(-0.72)}{-0.035} - \underset{(-2.68)}{0.009\hat{u}_{8,t-1}} - \underset{(-6.57)}{0.184\hat{u}_{8,t-1}} - \underset{(-1.36)}{0.038\Delta\hat{u}_{8,t-2}} \qquad (4.3.19)$$

$$\Delta \hat{u}_{9,t} = \underset{(-0.20)}{-0.007} - \underset{(-6.19)}{0.058\hat{u}_{9,t-1}} - \underset{(-3.91)}{0.110\Delta\hat{u}_{9,t-1}} - \underset{(-0.12)}{0.003\Delta\hat{u}_{9,t-2}} \qquad (4.3.20)$$

$$\Delta \hat{u}_{10,t} = \underset{(-0.22)}{-0.006} - \underset{(-3.25)}{0.016\hat{u}_{10,t-1}} - \underset{(-5.04)}{0.142\hat{u}_{10,t-1}} + \underset{(4.81)}{0.135\Delta\hat{u}_{10,t-2}}$$

$$- \underset{(-1.66)}{0.047\Delta\hat{u}_{10,t-3}} \qquad (4.3.21)$$

$$\Delta \hat{u}_{11,t} = \underset{(-0.23)}{-0.005} - \underset{(-3.83)}{0.021\hat{u}_{11,t-1}} \qquad (4.3.22)$$

where values in parentheses represent the Student's t–statistic.

Assume that the null hypothesis for cointegration is that there is no cointegrating relationship between pairs of series. If the t–statistic on the one–period lagged residual term from equations (4.3.12) through (4.3.22) is less than the critical value (–3.37) given by Engle and Yoo (1987), the null hypothesis will be rejected at the 5 % significance level. The rejection of the null hypothesis suggests that both pairs of series are cointegrated and the residuals from the cointegrating linear regression models are stationary.

Tests for cointegration were conducted. Models of the cointegration relationship must show improved long–run forecasts compared to models not cointegrated. Results of the DF or the ADF tests on residuals generated from the cointegrating models are summarized in Table 12. Only six pairs of series out of 11 were found to be cointegrated. Tests support cointegration relationship between live hogs and wholesale loin prices. Live hog and wholesale ham prices seem to be cointegrated. The cointegration relationships between wholesale loins and either of the futures market prices, and between wholesale hams and either of the futures market prices are found. Moreover, tests for cointegration indicate that live hogs and wholesale bellies prices are also cointegrated. Both wholesale bellies and either of the futures market prices, and live hogs and either of the futures price series appear not to be cointegrated.

Even if the variables are nonstationary, they may contain long run components. The cointegration test detects the long run equilibrium relationship between two variables. That is, the meaning given when one individual

TABLE 12

RESULTS OF THE DF AND ADF TESTS[1)] ON RESIDUALS

Y_t Series	X_t Series			
	LH	WL	WH	WB
FPB	–1.81	–4.7	–3.82	–2.83
FLH	–1.07	–4.97	–3.84	–2.68
WL	–6.19	–	–	–
WH	–3.25	–	–	–
WB	(–3.83)	–	–	–

1) The null hypothesis is 'there is no cointegration'. Values in parenthesis indicate the estimated DF test statistic. Otherwise, all are the ADF test statistics. If values are less than the critical value (–3.37), the null hypothesis is rejected.

nonstationary series is integrated of order one, I(1), and a linear relationship of any two series is integrated of order one, is that each the individual series have long run swings but that the two individual series do not have long run characteristics. If it is integrated of order zero, it will be demonstrated that both series are in the long run state.

Tests for cointegration is a procedure for relationships between two variables for the same asset. For the price series which are cointegrated, it is stated that these markets are operating efficiently. Cointegration between two prices is a necessary condition for market efficiency. Thus, a price series (dependent variable) of one can consistently predict the other price series (independent variable). For the price series which are not cointegrated, i.e., live hogs vs pork belly futures, live hogs vs live hog futures, and wholesale pork bellies vs either of the futures market prices, it can be concluded that they are not operating efficiently relative to each other. That is, at least one market is said to be inefficient. In other words, information about a price series of one (independent variable in the cointegrating regression model) cannot be used in predicting the other price series (dependent variable in the model).

In conclusion, pairs of series cointegrated of order zero imply that they have information about prices of a commodity in one market that would be used in predicting prices of a commodity in other markets.

Price Discovery Results using ECMs

An error correction model for two variables which are cointegrated was applied to examine the price discovery process. An error correction model is used to eliminate or correct the equilibrium error.

Granger (1986) said that if two series are cointegrated, "there must be Granger causality in at least one direction, as one variable can help forecast the other" (p.218), even though the reverse is not necessary true.

The lead/lag relationships between two series cointegrated of order zero, I(0), are estimated using the augmented restricted VAR representations of equation (3.5.13) and (3.5.14). The vector corrected error term from the cointegrating regression is restricted and included as an independent variable. And the lagged variable of each series in the cointegrating model are selected by SBC and/or AIC in order to augment to the models as independent variables. The chosen lengths of lags of each series, shown in Table 5, will be augmented to build an error correction model.

The results for price discovery using ECMs are described in Table 13. According to the results, both live hog and pork belly futures market prices lead wholesale ham prices and wholesale belly prices cause live slaughter hog cash prices. However even though they are cointegrated of order zero, wholesale loin prices have a feedback relationship with both pork belly and live hog futures market prices. Also, loin prices in the wholesale market and live hog prices in the cash market have a bidirectional causal relationship.

One can infer the following from the results. (1) Relationship between live hog prices and wholesale pork prices: Since pork belly prices cause live slaughter hog prices, only wholesale belly prices of the wholesale meat prices and the two futures contract prices provide information for predicting cash hog prices. There is no significant domination between wholesale loin prices and live hog prices because of their feedback relationship. (2) Relationship between either of the futures contract prices and wholesale pork prices: The prediction of ham prices in wholesale pork markets can be enhanced by either pork belly or live hog futures market prices. However, wholesale loin prices and either of the futures contract prices cause each other. (3) Wholesale ham prices and both futures price series: It is concluded that the changes in wholesale ham prices are well explained by the

TABLE 13

CAUSAL RELATIONSHIPS BASED ON COINTEGRATION AND ECMs

X_t	Y_t	F_1^*	F_2^*	Causal Direction
WL	FPB	5.38(0.005) [2,1234]	1.91(0.012) [18,1234]	WL <–> FPB
WL	FLH	11.58(0.0001) [2,1234]	1.68(0.037) [18,1234]	WL <–> FLH
WH	FPB	6.12(0.002) [2,1242]	1.38(0.155) [14,1242]	FPB –> WH
WH	FLH	15.5(0.0001) [2,1242]	1.34(0.179) [14,1242]	FLH –> WH
LH	WL	3.48(0.0001) [18,1230]	8.4(0.0001) [6,1230]	LH <–> WL
LH	WB	5.08(0.0001) [5,1255]	2.11(0.049) [6,1266]	LH <–> WB

F_1^* is calculated by using Equation (3.5.13) and F_2^* is calculated based on equation (3.5.14). Values in (.) represent the p–value and values in [.] indicate the degrees of freedom. The significance level is considered at 5 percent. If F_1^* > the critical value in F table, then Y_t causes X_t series. If F_2^* > F table value, X leads Y.

corrected error from the cointegrating regression model (4.3.3) /or (4.3.7), the changes in the lagged pork belly futures prices/ or live hog futures prices, and the changes in the lagged wholesale ham prices. (4) The results show that changes in the lagged live hog futures prices do not effectively and consistently influence changes in the live hog prices in one direction. Similarly, wholesale pork belly prices are not influenced by pork belly futures prices. In other words, futures market prices for live hog or wholesale pork bellies are not the center of price discovery for them, respectively.

As a whole, this result does not provide clear evidence of the general view that futures prices cause spot prices. Furthermore, wholesale ham prices reliably follow both pork belly futures as well as live hog futures contract prices.

Unrestricted VAR Causality Tests

When pairs of series are integrated, application of vector autoregression for differenced variables is incompatible because it omits the error correction term. If some pairs of series are to be cointegrated but have the properties of a vector autoregressive process, then they can be analyzed with the equation for the augmented unrestricted VAR.

Causality tests using the AUVAR are conducted, based on equations (3.5.13) and (3.5.14) except that the vector corrected error term is excluded from the independent variables.

Overall results of the VAR analysis are given in Table 14. Results indicate that the role of futures markets in providing price information to wholesale pork belly prices seemingly exists. Price changes in pork belly futures lead price changes in the live hog cash market, not the reverse. The price discovery function of the live hog futures market to cash hog markets is not clear, but feedback exists. Between wholesale meat markets and futures markets, pork belly futures market prices strongly lead pricing in wholesale pork belly prices. There is no dominant market in pricing between the wholesale belly market and the live hog futures market. Their relationship is bidirectional. Also a strong feedback relationship exists between cash hog and wholesale ham prices.

There is no unidirectional causality between live hog and live hog futures prices. A large number for the F_1^* statistic (315.45), implying live hog futures

TABLE 14

CAUSAL RELATIONSHIPS BASED ON
THE VAR MODELS

X_t	Y_t	F_1^*	F_2^*	Causal Direction
LH	FPB	90.83(0.001)	0.81(0.543)	LH <– FPB
LH	FLH	315.45(0.0001)	5.23(0.001)	LH <–> FLH
WB	FPB	19.07(0.001)	0.96(0.427)	WB <– FPB
WB	FLH	81.80(0.0001)	2.93(0.02)	WB <–> FLH
LH	WH	5.23(0.0001)	5.6(0.0002)	LH <–> WH

Values in (.) represent the p–value. The significance level is considered at 5 percent. If $F_1^* >$ the critical value in F table, then Y_t causes X_t series. If $F_2^* >$ F tabled value, X leads Y.

prices cause cash hog prices, cannot be interpreted to be more powerful than for the F_2^* statistic (5.23), indicating that live hog prices lead live hog futures market prices, because the p–values for both statistics are the same. Therefore, it cannot be concluded that a strong causal relationship is found running from live hog futures prices to cash hog prices.

In sum, causal directions analyzed from the VAR process do not agree generally with the common view that live hog futures market prices are expected to lead cash hog market prices. However, for pork belly prices, futures market prices strongly cause wholesale pork belly prices. In the case of live hog prices and pork belly futures prices, strong lead–lag evidence is from pork belly futures market prices to live hog prices. Also, wholesale ham prices do not seem to lead cash hog prices.

In–Sample Results for 1987–1990

Bradshaw and Orden (1990) stated that in–sample and out–of–sample tests for Granger causality whether or not there is causality from Y to X is defined by whether or not an optimal forecasting model for current X using past values of X and Y performs better than one using only past values of X. In–sample tests for pairs of series provide general beliefs on the causal directions. Thus, using ECMs, AUVAR, and Geweke–type models were utilized to draw causal directions for the in–sample data.

Since all data series are not stationary, they are first–difference transformed. Out–of–sample causality testing is not the best technique. A better one was suggested by Ashley et al. and Bessler and Kling. The a priori belief resulting from within–sample tests will be re–applied to out–of–sample data. Then causality tests can be performed using out–of–sample tests, and compared with the a priori belief.

To conduct causality tests on out–of–sample data, first, general knowledge about the causal relationships of the within–sample data (1987–1990) must be found. Such knowledgement is called a priori belief about the lead or lag relationship. The previous methodology to evaluate causal relationship of within–sample data was a Sims– or Geweke–type vector autoregressive model. As discussed in previous sections, however, if a linear relationship of two series is integrated of order zero, then the results from the VAR process do not express all the information, e.g., long run and short run properties of the original series, and so it will be meaningless. Therefore, a priori beliefs are first found by using error correction models when the series are cointegrated, and secondly by using VAR representations for non–cointegrated series. Table 15 represents the directions of

TABLE 15

CAUSAL DIRECTIONS FROM WITHIN–SAMPLE AS A PRIORI BELIEFS(I)

Method	X_t	Y_t	F_1^*	F_2^*	Causal Direction
VAR	LH	FPB	83.23(0.0001)	1.18(0.316)	FPB –> LH
VAR	LH	FLH	312.82(0.0001)	6.0(0.0001)	FLH <–>LH
ECM	WL	FPB	5.03(0.007)	1.74(0.043)	FPB –> WL
ECM	WL	FLH	11.0(0.0001)	1.61(0.069)	FLH –> WL
ECM	WH	FPB	5.41(0.0046)	1.03(0.389)	FPB –> WH
ECM	WH	FLH	15.27(0.0001)	0.68(0.641)	FLH –> WH
VAR	WB	FPB	15.51(0.0001)	1.18(0.316)	FPB –> WB
VAR	WB	FLH	18.56(0.0001)	0.21(0.645)	FLH –> WB
ECM	LH	WL	3.80(0.0001)	9.2(0.0001)	WL <–> LH
VAR	LH	WH	10.71(0.0001)	6.9(0.0001)	WH <–> LH
ECM	LH	WB	4.14(0.0010)	2.1(0.0661)	WB –> LH

F_1^* is calculated by using Equation (3.5.13) and F_2^* is calculated based on Equation (3.5.14) using ECMs or AUVAR. Values in (.) represents the p–value. The significance level is considered at 5 percent. If F_1^* > the critical value in F table, then Y_t causes X_t series. If F_2^* > F table value, X leads Y.

causality for within–sample data (1987–1990) which will be the a priori beliefs.

Cointegrating relationships between pairs of series for 4 years data are the same as for 5 years of data (see Tables 12 and 15). It confirms that both cointegrated markets are efficient, their prices are each nonstationary, and both efficient markets have an equilibrium relationship which is stationary. For pairs of series which are cointegrated, an error correction model was used to test Granger causality. Otherwise, the augmented UVAR was used for 4–year pairs of series. Causal directions for in–sample data are similar to those for the entire data. Only three pairs of series have different directions from the previous results. That is, results show that either of the futures market prices cause wholesale loin prices unidirectionally but not bidirectionally, and wholesale pork belly prices also lead live hog futures prices but there is not feedback.

On the other hand, the use of ECMs to obtain the a priori beliefs about causal directions is not available for out–of–sample causality tests because out–of–sample performance is achieved by using all parameters estimated from the prewhitened within–sample data. That is, a parameter of the vector error correction term cannot be directly used with the out–of–sample data. For this, a priori beliefs are obtained from the standard Geweke–type causality test method, following equations (3.6.1) and (3.6.2). A priori beliefs from Geweke's causal test are shown in Table 16.

Results of the Geweke test indicate that both pork belly futures contract and live hog futures contract prices lead cash slaughter hog and wholesale pork belly prices. All wholesale pork prices also cause live hog prices. Live hog futures prices also cause wholesale loin and ham prices, whereas pork belly futures contract prices do not lead wholesale loin and ham prices in pricing. These results are interpreted somewhat differently from the a priori beliefs indicated by ECMs and AUVAR analyses in Table 15.

From both Tables 15 and 16, within–sample causality tests result in strong evidence of causal relationships. Results support our general beliefs, that is, 'futures lead cash market prices,' 'wholesale meat market prices cause cash hog prices,' and 'futures prices have information for pricing in wholesale meat markets,' etc..

TABLE 16

CAUSAL DIRECTIONS FROM WITHIN–SAMPLE
AS A PRIORI BELIEFS(II)

X_t	Y_t	F^*	Causal Direction
LH	FPB	40.96(0.0001)	FPB –> LH
LH	FLH	156.1(0.0001)	FLH –> LH
WL	FPB	0.461(0.6308)	NO
WL	FLH	6.377(0.0018)	FLH –> WL
WH	FPB	0.266(0.7668)	NO
WH	FLH	10.25(0.0001)	FLH –> WH
WB	FPB	64.40(0.0001)	FPB –> WB
WB	FLH	35.49(0.0001)	FLH –> WB
LH	WL	11.90(0.0006)	WL –> LH
LH	WH	15.08(0.0001)	WH –> LH
LH	WB	4.14(0.0010)	WB –> LH

F^* is calculated by using Equation (3.6.2). Values in (.) represents the p–value. The significance level is considered at 5 percent. If F^* > the critical value in F table, then Y_t causes X_t series.

Out–of–Sample Causality Tests

Based on these causal results from within–sample tests, out–of–sample tests were performed. As explained, out–of–sample causality tests were conducted by comparing the MSEs that are estimated from two Geweke–type VAR equations (3.6.3) and (3.6.4). Table 17 reports the results of the test.

TABLE 17

OUT–OF–SAMPLE CAUSALITY TESTS

$\overline{X}_t$	$\overline{Y}_t$	MSE($\overline{X}$)	MSE($\overline{X}$,$\overline{Y}$)	Causal Direction
LH	FPB	0.4340	0.4197	FPB –> LH
LH	FLH	0.4340	0.4094	FLH –> LH
WL	FPB	8.2745	8.2636	FPB –> WL
WL	FLH	8.2745	8.2524	FLH –> WL
WH	FPB	4.2131	4.1963	FPB –> WH
WH	FLH	4.2131	4.2063	FLH –> WH
WB	FPB	2.4218	2.1362	FPB –> WB
WB	FLH	2.4218	2.3405	FLH –> WB
LH	WL	0.4340	0.4306	WL –> LH
LH	WH	0.4340	0.4268	WH –> LH
LH	WB	0.4340	0.4335	WB –> LH

If MSE($\overline{X}$) is greater than MSE($\overline{X}$,$\overline{Y}$), then $\overline{Y}$ causes $\overline{X}$. It is not clear that a series causes the other series using these MSEs, because there is little change in MSEs between uni– and bi– variate models.

Since, by construction, $MSE(\overline{X}) > MSE(\overline{X},\overline{Y})$ is defined by '$\overline{Y}$ causing $\overline{X}$,' every pair of series has a unidirectional causal relationship. According to the results, the a priori beliefs are true, that is, futures contract prices cause cash prices, futures contract prices cause wholesale pork prices, and wholesale pork prices cause cash market prices, respectively. The post sample MSE for the bivariate model, $MSE(\overline{X},\overline{Y})$, for live hogs and its futures contract prices, and for pork bellies and its futures contract prices were increased by 6 percent and 13 percent, which are respectively higher, compared to those resulting from the univariate model. But the post sample tests show that pork belly futures contract prices weakly cause wholesale loin and ham prices, which differs from the results in Table 16.

In conclusion, out–of–sample causality tests are consistent with a priori beliefs that both futures price series are leading cash hog prices and wholesale belly prices. But if the degree of lead or lag depends on the magnitude of MSE, we can conclude that causal relationships are strong. And also, even though wholesale loin and ham prices influence live hog prices, their role in pricing live hogs seems to be weak. It is also suggested that futures prices play a small role in forecasting and pricing wholesale loin and ham prices.

Year by Year Trend in Causality

So far, the price discovery process was studied for all observations in the data series under consideration in this study. Results suggest that futures market prices cause cash slaughter hog prices as well as wholesale meat market prices. Moreover, pork meat prices in wholesale markets also provide some information to cash hog market in pricing. A goal in this section is to determine whether such relationships have changed over time.

To see the existence of a general trend in hog and pork market pricing, two of the several methods discussed in Chapter 3 were used: ECMs and VAR. All yearly data series were proven to be integrated of order one after their original series had not been found to be stationary. Each integrated series can be combined in paris of two series to create a cointegrating regression model. At least 55 ordinary least squares linear regression models were built. For every pair of series, cointegration was tested. The DW statistics estimated from the cointegrating regression models indicate that only 3 pairs of series were not cointegrated, wholesale ham and pork belly futures series in 1987, wholesale ham and live hog futures series in 1987, and live hog and wholesale loin prices in 1987. Since the DW statistic is not preferred to test for cointegration, the DF or ADF tests were applied. The process of choosing one of the DF or ADF tests is determined by the lengths of lags on residuals from the cointegrating regressions. Table 18 presents the lengths of lags chosen by SBC.

Whether pairs of series have a cointegrating relationship is determined by the DF or ADF test statistics in Table 19. Of course, the DF test was applied to pairs of series in which the error term of their cointegrating regression model has the lag length of one. Otherwise, the ADF test was used.

For pairs of series which are cointegrated, causality tests were conducted on the error correction models, equations (3.5.13) and (3.5.14). Otherwise, the augmented unrestricted VAR test was used. The AUVAR test is based on the same equations as the ECMs except that the corrected vector error term is excluded.

Results of causal direction using both ECMs as well as AUVAR are reported in Table 20. Table 20 shows the tendency toward causal relationships over time. (1) Live hog and wholesale loins or hams: They are bidirectionally causing each other in general except for the 1991 series. During 1991, which is

TABLE 18

THE LENGTHS OF LAGS ON RESIDUALS
– YEARLY DATA

Year		87	88	89	90	91
Y_t	X_t					
FPB	LH	1	6	3	1	6
	WL	2	2	2	2	1
	WH	2	2	3	2	3
	WB	1	1	1	1	1
FLH	LH	1	3	1	3	1
	WL	2	2	2	2	1
	WH	2	3	3	2	2
	WB	1	1	1	2	1
WL	LH	1	1	1	1	1
WH	LH	1	6	3	1	3
WB	LH	1	1	1	1	1

Length of lags was determined by the smallest SBC.

TABLE 19

TEST RESULTS OF COINTEGRATION
– YEARLY DATA

Y_t	X_t	87	88	89	90	91
FPB	LH	–1.97	–2.31	–1.02	–1.22	–1.66
	WL	–2.38	–2.23	–3.61	–2.44	–2.27
	WH	–4.01	–3.87	–1.67	–1.80	–2.59
	WB	–0.90	–3.71	–1.69	–2.86	–3.28
FLH	LH	0.07	–1.20	–0.86	–1.28	–1.04
	WL	–2.22	–2.26	–2.96	–2.27	–2.1
	WH	–3.68	–4.00	–2.26	–1.53	–2.75
	WB	–0.84	–3.32	–3.14	–3.55	–0.92
WL	LH	–3.55	–2.97	–2.35	–3.2	–3.01
WH	LH	–0.71	–2.06	–1.8	–1.63	–0.92
WB	LH	–3.38	–2.15	–1.43	–1.49	–2.29

The null hypothesis is that there is no cointegration. The test ststistic for sample size of 200 at 5 % level is –3.37. If the calculated statistic is less than the critical value, then the null hypothesis will be rejected.

TABLE 20

CAUSALITY TEST RESULTS FOR YEARLY DATA

			Xt																			
			LH					WL					WH					WB				
		Yr.	87	88	89	90	91	87	88	89	90	91	87	88	89	90	91	87	88	89	90	91
Yt	WL	87	B																			
		88		B																		
		89			B																	
		90				B																
		91					YX															
	WH	87	B																			
		88		B																		
		89			B																	
		90				B																
		91					YX															
	WB	87	YX																			
		88		YX																		
		89			B																	
		90				B																
		91					YX															
	FLH	87	B					B					YX					B				
		88		B					No					B					YX			
		89			B					No					No					B		
		90				B					B					YX					YX	
		91					YX					No					YX					YX
	FPB	87	YX					B					YX					B				
		88		B					No					YX					B			
		89			YX					YX					No					B		
		90				YX					XY					No					B	
		91					YX					No					YX					B

Notion 'XY' implies X series causes Y series, and 'YX' means series Y causes series X. 'B' denote that X and Y are caused bidirectionally, and 'No' represents there is no certain causal relationship between the two series.

the same period for out–of–sample data in the previous section, wholesale prices led live hog prices. (2) Live hog and wholesale belly: Generally, wholesale belly prices cause cash hog prices. (3) Live hog and live hog futures market price series: There is no clear change of causality between live hog and its futures prices. Both have a bidirectional causal relationship in each year except 1991. So, there are mainly bidirectional causal relationships between live hog and live hog futures market prices over time. However, pork belly prices led live hog prices with a change once in 1988. (4) Live hog and pork belly futures market prices: Pork belly futures market prices are mainly causing live hog futures market prices, except for 1988. (5) Wholesale pork prices vs futures contract prices: Several changes occurred in their relationships, but they were not consistent over time. Especially, there is no certain causal relationship between wholesale loin and live hog futures contract prices. But it is concluded that live hog futures market prices are chiefly causing wholesale ham prices, even though there are sometimes exceptions. Furthermore, wholesale pork belly prices are mainly caused by live hog futures contract prices. (6) Wholesale pork and pork belly futures market prices: It cannot be concluded that there is clear causal relationship between wholesale loin and pork belly futures market prices. And pork belly futures contract prices mainly cause wholesale ham prices. Out of 55 pairs of series considered, only five pairs, pairs of wholesale belly and belly futures prices for five years, consistently represent bidirectional causality.

There is only one pair of series which violates the normal viewpoint. In 1990, wholesale loin prices appear to affect pork belly futures. For pairs of series for some years, causal relationships do not exist. From the results for yearly data, it can be concluded that, even if it is not consistent, causal directions are true from live hog futures to wholesale pork belly prices, and from wholesale pork belly prices to live hog prices sequentially. It is interesting because, in general, beliefs for live hog futures contract prices causing cash hog prices are not accepted by the results. Instead, sequential causal directions from live hog futures to wholesale pork belly, and from wholesale pork belly to live hog, are appear dominant. Therefore, it is summarized that wholesale loin, ham, and live hog futures contract prices are not the center for pricing live hogs.

Accordingly, there are changes in causal relationships over time, but changes are not consistent. Year by year causal relationships, in general, do not correspond to causal relationships yielded from out–of–sample tests, but correspond to causal directions for for all observations in this study. That is, wholesale pork belly and pork belly futures contract prices are more important to price live hogs in the future spot market than live hog futures market prices.

Summary

Most of the series used are nonstationary. Most of them are integrated of order one, and some linear relationships of two integrated series represent cointegration. The techniques outlined in chapters 3 and 4 involve using the Granger causality definition. So, for cointegrated pairs of series, causality tests were conducted by the ECMs. For other pairs of series, not cointegrated, the AUVAR test was used. Moreover, as an independent method, post–sample causality tests were used to confirm the price discovery process.

Price discovery results for the entire data period are reported in Table 21. The table describes the causal directions generated by ECMs or VAR. The notations in the table indicate that causal directions run from wholesale bellies to live hogs, from pork belly futures to live hogs, from both futures series to hams, and from pork belly futures to wholesale belly prices. Other pairs of series have bidirectional causal relationships.

TABLE 21

CAUSAL DIRECTION

Y_t	X_t			
	LH	WL	WH	WB
WL	B	–	–	–
WH	B	–	–	–
WB	YX	–	–	–
FLH	B	B	YX	B
FPB	YX	B	YX	YX

'YX' denotes that Y causes X, and 'B' implies 'bidirectional feedback.'

Results from the post–sample test presented in Table 17 do not permit interpreting the causal relationships as the above. The differences between MSEs solved from a univariate model and the bivariate model are small, and the changes in them from uni– ($\overline{X}$) and bi– ($\overline{X},\overline{Y}$) variate models are also small. It is somewhat impossible to ascertain that $\overline{X}$ causes $\overline{Y}$ or $\overline{Y}$ causes $\overline{X}$. When small changes in MSEs are assumed to be very important based on theory, it can be stated that all futures series lead cash hog prices as well as wholesale meat prices, and every wholesale pork series also precede live hog prices. Otherwise, causal directions cannot be distinguished except for two pairs of series: live hog and live hog futures, and wholesale pork bellies and its futures prices.

An interesting observation is found from Table 15 compared with Table 17. Table 15 describes the results of causality tests conducted on four years of data between 1987 and 1990. The processes and methodologies used are the same as those used to generate the reports in Table 21. For 4 years, causal relationships are clearly shown: wholesale pork bellies and pork belly futures lead cash hogs, and every futures price series causes every wholesale meat price series.

It is reported that both live hog and wholesale loin and ham prices are caused bidirectionally. Still, this study does not confirm that live hog futures prices are much stronger for pricing cash hog market, but that both represent a bidirectional causal relationship.

The Geweke test results for 4 years also represented clear evidence of causal relationships between pairs of variables (Table 16). Nevertheless, out–of–sample performance does not provide more clear causal directions with big MSE values. As a whole, it can be inferred that futures prices are leading cash hog prices, and wholesale belly prices are clearly causing cash slaughter hog prices.

There is no clear trend in causal direction over time. Moreover, causality test results for each year are not similar to those from whole–sample observations or from the four year sample period.

CHAPTER V

CONCLUSIONS

Introduction

The previous chapters described the theoretical background and the concept of causality, procedures for analyzing price discovery, and empirical results with respect to this research. This study includes defining price discovery, giving a rationale for price discovery, and examining the concept of Granger causality.

Theory ascertained a general view point about the importance of futures markets in production decisions and price formation. Price discovery is defined as the process of finding a price which is generated in a specified market based on the available price related information. That is, price discovery is the process of seeking a market–clearing price based on given information. When the related past and present information is assumed to be two relevant price variables, price discovery is accomplished by using the concept of correlation between a pair of random variables. That is, this study searched for weakly efficient market prices based on current available information, say, past market prices.

To estimate price discovery lead–lag relationships, several methods were employed. They were focused on two intermediate goals which are short run as well as long run phenomena. For pairs of series, cointegration with error correction models are popularly used in research. A specific feature for using an error correction model is that the current change of a variable is affected first by the immediate short–run effect from last period's error based on the cointegrating equilibrium regression, and second by the change of another variable in the bivariate system model. Since the short–run error term was already adjusted to past disequilibrium, an error correction model represents both short–run and long–run effects. However, despite such an advantage, an error correction model has a disadvantage in application. That is, cointegration with an ECM is not available if two integrated series with the same order are not cointegrated. As an alternative, the augmented unrestricted vector autoregressive process is applied. The price discovery process is also examined by post–sample causality tests, which allow verifying whether or not a priori beliefs from the within–sample data are true.

Using these models, three hog–related markets are analyzed: live hogs, wholesale pork – loins, hams, and pork bellies, and futures markets live hog futures and pork belly futures. Short–term as well as long–term price relationships among producers, wholesalers, and futures markets were empirically analyzed based on a specified theoretical background. The process of price discovery in the short term was investigated by the VAR analysis and out–of–sample forecasting method.

Each series was collected from different markets and locations, e.g., live hog prices from the Iowa and Southern Minnesota area, wholesale pork prices from Omaha, and futures prices from the CME. Daily data were used because better results can be obtained when data containing more relevant information is utilized. That is, daily data excludes some plausible sources of variation in weekly, monthly, or quarterly data.

Using these theories, concepts, methodologies, and data, the price discovery process in hog and pork markets was investigated. In the following section, theoretical and empirical results are summarized.

Findings

A price discovery study involves pricing efficiency. Price discovery begins with a general price level and concludes with transaction prices (Ward, 1988). The role of futures market prices in price discovery is commonly recognized by several analysts. Findings in this study are divided into two categories; theoretical and empirical ones.

Chapter 2 demonstrated some theoretically important evidence.

(1) As generally known, futures prices lead cash prices in producer's decision–making. This is applicable to other market agents, wholesalers, packers, retailers, etc..

(2) Price discovery provides some benefit to each market participant because it increases market efficiency. That is, if a person has more information resulting from price discovery, about pricing a product or a commodity, his profit would increase more than others who do not have such information.

Cross–spectral analysis allows examining price discovery lead–lag relationships. Also univariate ARIMA techniques have been applied to price discovery research. However, researchers argue that they have disadvantages related to interpreting and understanding results and statistically lower power. Furthermore, even though VAR analysis provides a method to describe lead–lag relationships, it disregards the long–run properties of cointegrated series. Therefore, price discovery results were derived by cointegrating with ECMs, the AUVAR and out–of–sample causality tests, following the advantages of each method. Chapter 4 released various results of price discovery in hog and pork markets. Followings are results for the 1987–1991 sample period.

(1) Relationships between wholesale loin prices and both pork belly and live hog futures market prices have a long–run equilibrium state. Relationships between wholesale ham prices and futures market prices, and between live hog prices and both wholesale loin and pork belly prices also have a long–run equilibrium relationship. Cointegration between two market prices indicates that those two markets are operating efficiently. The other 5 pairs of series, out of 11, are not cointegrated even though each individual series is integrated of order one.

(2) For cointegrated series, an analysis of error correction models results in unidirectional causal relationships between wholesale ham price series and both futures market prices. Otherwise, results indicate the bidirectional causal relationships between pairs of series (Table 13). Therefore, the hypothesis, which was that there exists co–movement between pairs of series, is rejected. The condition for co–movement of the series is that both series must be cointegrated and also that they do not have a causal relationship.

(3) For non–cointegrated series, the AUVAR analyses were applied. Pork belly futures prices led live hog prices and wholesale pork belly prices. The other 3 pairs of series explicitly showed feedback relationships (Table 14).

(4) In sum, live hog pricing was not affected by live hog futures contract prices, but by pork belly futures market prices. Both pork belly futures prices and live hog futures also led wholesale meat prices. Out of three different wholesale meat prices, wholesale pork belly prices followed both futures market prices much more than wholesale loin or ham prices did. Relationships between live hog prices and wholesale meat prices did not appear to have a strong unidirectional causality. Both wholesale loins, hams, and bellies have feedback relationships with live hog prices.

The following results are from the 4 years data period (1987–1990), which was stated as 'within–sample.' For 4 year samples, results were used as a priori beliefs to test post–sample data.

(1) Results from ECMs and the AUVAR were similar to previous results for 1987–1991. Rather, results for 4 years were clearer than for the 5 year sample analysis (Table 15). Pork belly futures market prices led live hog prices as well as wholesale loin, ham, and pork belly prices. Live hog futures contract prices also cause wholesale loin, ham, and pork belly prices. Live hog prices have feedback relationship with live hog futures market prices. But, live hog prices followed wholesale pork belly and pork belly futures market prices.

(2) To follow the general process of post–sample tests, Geweke–type tests were conducted on within–sample data regardless of cointegrating properties. Except for wholesale loin vs pork belly futures, and wholesale ham vs pork belly futures prices, both futures prices caused live hog prices as well as wholesale meat prices. Also, wholesale meat prices led live hog prices (Table 16).

Using the a priori beliefs, causality was tested for out–of–sample data (1991). Even though the differences of estimated MSEs are small and thus the the strengths of causality are also weak, causal directions appear unidirectional. Out–of–sample test results confirmed the results from causal directions which were generated by Geweke–type tests for within–sample data. But it did not provide clear correspondence to the results from ECMs and AUVAR.

One hypothesis was whether there is yearly trend in causality over time. However, there was no outstanding trend found over time (Table 19, 20). But the analysis for yearly data provided three important features of price discovery. First, a sequential causal chain was found from live hog futures market prices to wholesale belly prices, and from wholesale belly prices to live hog prices, even though live hog futures market prices have had feedback relationship with live hog prices. Second, wholesale loin, ham, and live hog futures contract prices did not lead cash hog prices whereas wholesale pork belly and pork belly futures market prices mainly caused live hog prices. Third, pork belly futures market prices did not lead wholesale pork belly prices but feedback existed consistently over time.

As a whole, some hypotheses described in chapter 1 were rejected by results for 5 years data and yearly data, but not rejected by results for 4 years. Live hog futures prices did not lead the pricing of cash slaughter hog prices, but pork belly futures market prices also preceded wholesale belly prices. It should be noted that wholesale pork belly and pork belly futures prices also play an important role in pricing cash hogs, which differs from wholesale loin or ham prices. In general, it can be concluded from this study that futures market prices affect wholesale loin and ham prices, even though the strength of causation is relatively weak. There seems to be no yearly trend in causality over time. Results from analyzing yearly data have little correspondence to those of the entire five–year period by out–of–sample tests.

Limitations and Concluding Remarks

This study has a number of limitations.

First, since all the data series are those of agricultural commodities, they may exhibit seasonal patterns. Such seasonal patterns were not considered in this study. But, it may not be a severe limitation because differencing the data series provided a necessary condition for stationarity. For more accuracy, however, it may be required.

Second, in within–sample tests, causal relationships were derived from the formal VAR analysis, even though the results were compared to those of ECMs and AUVAR. And, all the parameters derived from the Geweke–type causality test were used to test causality for post–sample causality in the data. As stated, the Geweke–type VAR representation does not consider the long–run equilibrium state.

Third, for the absence of cointegration relationships, the AUVAR tests were applied. As a result, some difficulties in interpretation for all pairs of series occurred. For example, comparing the causal relationships between live hog futures prices and wholesale loin prices resulting from ECMs between live hog futures market prices and wholesale ham prices generated from the AUVAR, can be obscure.

Fourth, the loss of some information from generating the continuous nearby futures contract prices can occur.

Fifth, this study focussed only on bivariate model approaches. The order of causal strength can be derived from other methods, that is, trivariate model approach. If more relevant variables are included, it would increase the efficiency of the results because information will increase.

Further research on price discovery is recommended, overcoming limitations stated above.

First, theoretically, out–of–sample causality tests can be conducted by using univariate residual cross–correlation models or a vector autoregressive process. However, if a priori beliefs as a general view are drawn from cointegration and error correction models such as Table 15, they cannot be accepted because of the vector error correction. Trying to find a new approach for out–of–sample testing using ECMs would be an interesting project.

This study considers only relationships between pairs of series (bivariate). The next research could analyze relationships among numerous variables, e.g., live hogs – wholesale loins – and live hog futures prices (trivariate), or more.

This research was accomplished, considering only three related market prices. Expanding commodities and including more information, such as cash corn prices and corn futures prices, could be developed. When corn prices are included, more clear relationships will be defined according to a priori corn—hog relationships.

BIBLIOGRAPHY

Akaike, H., A New Look at the Statistical Identification Model, *IEEE Trans. Auto. Control* 19, 1974. 716–723.

Antonovitz, F. and T. Roe, Effects of Expected Cash and Futures Prices on Hedging and Production, *The J. of Futures Markets* 6, 1986. 187–205.

Ashley, R., C.W.J. Granger and R. Schmalensee, Advertising and Aggregate Consumption: An Analysis of Causality, *Econometrica* 48, 1980. 1149–1167.

Atkin, M., *Agricultural Commodity Markets, A Guide to Futures Trading*, Routledge, New York, 1989.

Bailey, D. and B.W. Brorsen, Hedging Carcass Beef to Reduce the Short–Term Price Risk of Meat Packers, *W. J. of Agri. Econ.*, 1985. 330–337.

Barksdale, H.C., Hilliard, J.E., and M.C. Ahlund, A Cross–Spectral Analysis of Beef Prices, *Amer. J. of Agri. Econ.*, 1975. 309–315.

Barnhart, S.W. and A.C. Szakmary, Testing the Unbiased Forward Rate Hypothesis: Evidence on Unit Roots, Cointegration, and Stochastic Coefficients, *Journal of Financial and Quantitative Analysis* 26, 1991. 245–267.

Benninga, S., R.Eldor, and I. Zilcha, Optimal Hedging in the Futures Market under Price Uncertainty, *Economic Letters* 13, 1983. 141–145.

Bessler, D.A., Relative Prices and Money: A Vector Autoregression on Brazilian Data, *Amer. J. of Agri. Econ.*, 1984. 25–30.

Bessler, D.A. and J.A. Brandt, Causality Tests in Livestock Markets, *Amer. J. of Agri. Econ.* 1982. 140–144.

Bessler, D.A., and T. Covey, Cointegration: Some Results on U.S. Cattle Prices, *The Journal of Futures Markets* 11, 1991. 461–474.

Bessler, D.A. and J.L. Kling, A Note on Tests of Granger Causality, *Applied Economics* 16, 1984. 335–345.

Box, G.E.O. and G.M. Jenkins, *Time Series Analysis*, Holden–Day Inc., San Francisco, 1976.

Bradshaw, G.W. and D. Orden, Granger Causality from the Exchange Rate to Agricultural Prices and Export Sales, *W. Journal of Agri. Econ.* 15, 1990. 100–110.

Brandt, J.A. and D.A. Bessler, Composite Forecasting: An Application with U.S. Hog Prices, *Amer. J. of Agri. Econ.*, 1981. 135–140.

———, Forecasting with a Dynamic Regression Model: A Heuristic Approach, *N. Central J. of Agri. Econ.* 4, 1982. 27–33.

———, Price Forecasting and Evaluation: An Application in Agriculture, *Journal of Forecasting* 2, 1983. 237–248.

Brorsen, B.W., D. Bailey, and J.W. Richardson, Investigation of Price Discovery and Efficiency for Cash and Futures Cotton Prices, *W. Journal of Agri. Econ.* 9, 1984. 170–176.

Caines, P.E., C.W. Keng and S.P. Sethi, Causality Analysis and Multivariate Autoregressive Modeling with An Application to Supermarket Sales Analysis, *Journal of Economic Dynamics and Control* 3, 1981. 267–298.

Chowdbury, A.R., Futures Market Efficiency: Evidence from Cointegration Tests, *The Journal of Futures Markets* 11, 1981. 577–589.

Currie, D., Some Long Run Features of Dynamic Time Series Models, *The Economic Journal* 1981. 704–715.

Dantine, J–P, Information, Futures Prices, and Stabilizing Speculation, *J. of Economic Theory*, 1978. 79–98.

Davidson, J.E.H., D.F. Hendry, F. Srba, and S. Yeo, Econometric Modeling of the Aggregate Time Series Relationship between Consumer's Expenditure and Income in the United Kingdom, *Economic Journal* 88, 1978. 661–692.

DeCoster, G.P, W.C. Labys, and D.W. Mitchell, Evidence of Chaos in Commodity Futures Prices, *The Journal of Futures Markets* 12, 1982. 291–305.

Denbaly, M. and Vroomen H., Dynamic Fertilizer Nutrient Demands for Corn: A Cointegrated and Error Correcting System, *Amer. J. of Agri. Econ.*, 1983. 203–209.

Diamond, P. and M. Rothschild, *Uncertainty in Economics*, Academic Press, Inc., New York, 1978.

Dickey, D.A. and W.A. Fuller, Distribution of the Estimators for Autoregression Time Series with a Unit Root, *J. of the Amer. Stat. Assoc.* 74, 1979. 427–431.

Djunaidi, H., B.W. Brorsen, and D.S. Tilley, Cointegration and Causality Tests in the Cocoa Market, unpublished draft, Oklahoma State University, 1992.

Engle, R.F., Band Spectrum Regression, *International Economic Review* 15, 1974. 1–11.

Engle, R.F. and C.W.J. Granger, Cointegration and Error Corrections: Representation, Estimation and Testing, *Econometrica* 55, 1987. 251–276.

Engle, R.F., C.W.J. Granger and J.J. Hallman, Merging Short– and Long–run

Forecasts, *J. of Econometrics* 40, 1989. 45–62.

Engle, R.F. and B.S. Yoo, Forecasting and Testing in Cointegrated Systems, *J of the Econometrics* 35, 1987. 143–159.

Evans, E.V., *Country Elevator Use of the Market*, in A. Peck (ed.), Views from the Trade, Chicago, CBOT, 1978.

Fama, E.F., Efficient Capital Markets: A Review of Theory and Empirical Work, *J. of Finance*, 25, 1970. 383–423.

Faminow, M.D., Analysis of the Lead–Lag Structure of Two Wholesale Beef Price Quotes using Residual Cross Correlation, *North Central Journal of Agricultural Economics* 3, 1981. 89–94.

Feder, G., R.E. Just and A. Schmitz, Futures Markets and the Theory of the Firm under Price Uncertainty, *The Quarterly J. of Economics*, 1980. 317–328.

Feige, E.L. and D.K. Pearce, The Causal Relationship Between Money and Income: Some Caveats for Time Series Analysis, *The Review of Economics and Statistics*, 1979. 521–533.

Fisher, P.G., Seasonality in Large–scale Macroeconometric Models, *Journal of Forecasting* 11, 1992. 255–270.

Franzmann, J.R. and R.L. Walker, Trend Models of Feeder, Slaughter, and Wholesale Beef Cattle Prices, *Amer. J. of Agri. Econ.*, 1972. 507–512.

Fuller, W.A., *Introduction to Statistical Time Series*, John Wiley & Sons, New York, 1976.

Garbade K.D. and W.L. Silber, Price Movements and Price Discovery in Futures and Cash Markets, *Review of Economics and Statistics*, 1983. 289–297.

Gardner, B.L, Futures Prices in Supply Analysis, *Amer. J. of Agri. Econ.*, 1976, 81–84.

Geweke, J., Testing the Exogeneity Specification in the Complete Dynamic Simultaneous Equation Model, *Journal of Econometrics* 7, 1978. 163–185.

———, Inference and Causality in Economic Time–Series Models, *Handbook of Econometrics*, ed. Zvi Griliches and M. Intrilligator, chap. 17, Amsterdom, North–Holland Publishin Co. 1980.

———[a], Latent Variable Models for Time Series, *Journal of Econometrics* 17, 1981. 287–304.

———[b], A Comparison of Test of the Independence of Two Covariance–Stationary Time Series, *Journal of the Amer. Stat. Assoc.* 76, 1981. 363–372.

Geweke, J., R. Meese, and W. Dent, Comparing Alternative Tests of Causality in Temporal Systems: Analytic Results and Experimental Evidence, *Journal* *of Econometrics* 21, 1983. 161–194.

Granger, C.W.J., The Typical Spectral Shape of an Economic Variable, *Econometrica* 34, 1966. 150–161.

———, Investigating Causal Relations by Econometric Models and Cross–spectral Methods, *Econometrica* 37, 1969. 424–438.

———, Testing for Causality: A Personal Viewpoint, *Journal of Economic Dynamics & Control* 2, 1980. 329–352.

———, Some Properties of Time Series Data and Their use in Econometric Model Specification, *Journal of Econometrics* 16, 1981. 121–130.

Granger, C.W.J., and T.H. Lee, Investigation of Production, Sales and Inventory Relationships Using Multicointegration and Non–symmetric Error Correction Models, *Journal of Applied Econometrics* 4, 1989. s145–s159.

Granger, C.W.J., and P. Newbold, Spurious Regressions in Ecoonometrics, *Journal of Econometrics* 2, 1974. 111–120.

———, *Forecasting Economic Time Series*, Academic Press, New York, 1977.

Griliches, Z., Distributed Lags: A Survey, *Econometrica*, 1967. 16–49.

Hakkio, C.S., and M. Rush, Market Efficiency and Cointegration: an Application to the Sterling and Deutschemark Exchange Markets, *Journal of International Money and Finance* 8, 1989. 75–88.

Hannan, E.J. *Multiple Time Series*, New York, John Wiley & Sons, Inc., 1970.

Hardle, W., *Applied Nonparametric Regression*, Cambridge University Press, New York, 1989.

Harvey, A.C., Linear Regression in the Frequency Domain, *International Economic Review* 19, 1978. 507–512.

Haugh, L.D., Checking the Independence of Two Covariance–Stationary Time Series: A Univariate Residual Cross–Correlation Approach, *J. of the American Stat. Assoc.* 71, 1976. 378–385.

Haugh L.D. and G.E.P. Box, Identification of Dynamic Regression (Distributed Lag) Models Connecting Two Time Series, *Journal of the Amer. Stat. Assoc.* 72, 1977. 121–130.

Hey, J.D., *Uncertainty in Microeconomics*, New York University Press, New York, 1979.

Hicks, J., *Causality in Economics*, Basic Books, Inc., New York, 1979.

Hudson, M.A., *Price Discovery and Leading Relationships between Cash and Futures Prices for Agricultural Commodities: Theory and Applications*, unpublished Ph.D Dissertation, Virginia Tech, Blacksburg, Virginia; 1984.

———, Cash–Futures Causal Flows and Marketing Efficiency, In W. Purcell &

Purcell & J. Rowsell, eds., *Key Issues in Livestock Pricing: A Perspective for the 1990s*, Blacksburg, Virginia: Research Institute on Livestock Pricing, 1987. 178–210.

Hume, D., *A treatise of Human Nature*, Edited by L.A. Selby–Bigge, part 3, 2nd edition, at the Clarendon Press, Oxford, 1978. 69–179.

Hylleberg, S., R.F. Engle, C.W.J. Granger, and B.S. Yoo, Seasonal Integration and Cointegration, *Journal of Econometrics* 44, 1990. 215–238.

Jacobs, R.L., E.D. Leamer, and M.P. Ward, Difficulties with Testing for Causation, *Economic Inquiry* 17, 1979. 401–413.

Johansen, S, Statistical Analysis of Cointegration Vectors, *Journal of Economic Dynamics and Controls* 12, 1988. 231–254.

Johansen, S., and K. Julelius, Maximum Likelihood Estimation and Inference on Cointegration – with Applications to the Demand for Money, *Oxford Bulletin of Economics and Statistics* 52, 1990. 169–210.

Judge, G.G., W.E. Griffiths, R.C. Hill, H. Lutkepohl and T.C. Lee, *The Theory and Practice of Econometrics*, 2nd edition, Wiley, New York, 1985.

Kahl, K.H., M.A. Hudson, and C.E. Ward, Cash Settlements Issues for Live Cattle Futures Contracts, *Journal of Futures Markets* 9, 1989. 237–248.

Kant, I., *Critique of Pure Reason*, Translated by J.M.D. Meiklejohn, revised edition, Willey Book Co., New York, 1943.

Kaylen, M.S., Vector Autoregression Forecasting Models: Recent Developments Applied to the U.S. Hog Market, *Amer. J. of Agri. Econ.* 1988. 701–712.

Koontz, S.R., P.Garcia and M.A. Hudson, Dominant–Satellite Relationships Between Futures and Live Cattle Cash and Futures Markets, *The Journal of Futures Markets* 10, 1990. 123–136.

Laffont, J–J., *The Economics of Uncertainty and Information*, The MIT Press, Massachusetts, 1989.

Leistikow, D., Announcements and Futures Price Variability, *The Journal of Futures Markets* 9, 1989. 477–486.

Marsh, J.M. and G.W. Brester, Intertemporal Price Adjustments in the Beef Market: A Reduced Form Analysis of Weekly Data, *W. Journal of Agri. Econ.* 14, 1989. 235–245.

McCoy, J.H., *Livestock and Meat Marketing*, Westport, Connecticut, The AVI Publishing Co, Inc., 1972.

McKenna, C.J., *The Economics of Uncertainty*, Oxford Press, New York, 1986.

McKinnon, R.I., Futures Markets, Buffer Stocks, and Income Stability for Primary Producers, *J. of Political Economy*, 1967. 844–61.

Miller, S.E., Univariate Residual Cross–Correlation Analysis: An Application to

Beef Prices, *North Central Journal of Agricultural Economics* 1, 1979. 141–146.

———, Lead–lag Relationships between Pork Prices at the Retail, Wholesale, and Farm Levels, *S. Journal of Agricultural Economics*, 1980. 73–76.

Naik, G. and R.M. Leuthold, Cash and Futures Price Relationships for Nonstroable Commodities: An Empirical Analysis Using a General Theory, *W. Journal of Agri. Econ.* 13, 1988. 327–338.

Nelson, C.R., and G.W. Schwert, Tests for Predictive Relationships between Time Series Variables: A Monte Carlo Investigation, *Joutnal of Amer. Stat. Assoc.* 77, 1982. 11–18.

Oellermann, C.M., B.W. Brorsen, and P.L. Farris, Price Discovery for Feeder Cattle, *The Journal of Futures Markets* 9, 1989. 113–121.

Oellermann, C.M. and P.L. Farris, Futures or Cash: Which Market leads Live Beef Cattle Prices?, *The Journal of Futures Markets* 5, 1985. 529–538.

Phillips, P.C.B., Understanding Spurious Regressions in Econometrics, *J. of Econometrics* 33, 1986. 311–340.

———, Optimal Inference in Cointegrated Systems, *Econometrica* 59, 1991. 286–306.

Pierce, D.A., Relationships — and the Lack of Theoref — between Economic Time Series with Special Reference to Money and Interest Rates, *Journal of the Amer. Stat. Assoc.* 72, 1977. 11–22.

Rausser, G. and T.F. Cargill, The Existence of Broiler Cycles: An Application of Spectral Analysis, *Amer. J. of Agri. Econ.* 52, 1970. 109–121.

Robertson, J.C. and D. Orden, Monetary Impacts on Prices in the Short and Long Run: Some Evidence from New Zealand, *Amer. J. of Agri. Econ.*, 1990. 160–171.

Sandmo, A., On the Theory of the Competitive Firm under Price Uncertainty, *American Economic Review* 61, 1971. 65–73.

Sargan, J.D., Wages and Prices in the United Kingdom: A Study in Econometric Methodology, in Harts, Mills, and Whitaker (eds.), *Econometric Analysis for National Economic Planning*, London, Butterworth, 1964. 22–54.

Sargent, T., A Classical Econometric Model of the U.S., *Journal of Political Economy 84*, 1976. 207–237.

Sargent, T., *Macroeconomic Theory*, 2nd edition, Academic Press, New York, 1987.

Sargent, T. and C.A. Sims, Business Cycle Modeling without presenting to have Too Much a Priori Economic Theory, in C.A. Sims ed., *New Methods in Business Cycle Research*, Minneapolis, FRB of Minneapolis, 1977. 45–109.

Saunders, P., Causality of U.S. Agricultural Prices and the Money Supply: Further Empirical Evidence, *Amer. J. of Agri. Econ.* 1988. 588–596.

Schroeder, T.C. and B.K. Goodwin, Price Discovery and Cointegration for Live Hogs, *The Journal of Futures Markets* 11, 1991. 685–696.

Schwarz, G. Estimating the Dimension of a Model, *The Annals of Statistic* 6, 1978. 461–464.

———, Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root, *Econometrica* 49, 1981. 1057–1072.

Sengupta, J.K., Non–parametric Approach to Dynamic Efficiency: A Non–parametric Application of Cointegration to Production Frontiers, *Applied Economics* 24, 1992. 153–159.

Sheehan, R.G., and R. Grieves, Sunspots and Cycles: A Test of Causation, *Southern Economic Journal* 48, 1982. 775–777.

Shoesmith, G., Cointegration, Error Correction and Improved Medium–Term Regional VAR Forecasting, *Journal of Forecasting* 11, 1992. 91–109.

Sims, C. A., Money, Income and Causality, *American Economic Review* 62, 1972. 540–562.

———[a], Comment, *J. of the Amer. Stat. Assoc.* 72, 1977. 23–24.

———[b], Exogeneity and Causal Ordering in Macroeconomic Models, in C.A. Sims ed., *New Methods in Business Cycle Research*, Minneapolis, FRB of Minneapolis, 1977. 23–43.

Singleton, R.C., An Algorithm for Computing the Mixed Radix Fast Fourier Transform, *IEEE Transactions on Audio and Electroacoustics* 17, 1969. 93–103.

Stein, J., Speculative Price: Economic Welfare and the Idiot of Chance, *Review of Economics and Statistics* 63, 1981. 223–232.

Stock, J.H., Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors, *Econometrica* 55, 1987. 1035–1056.

Tiao, G.C., and G.E.P. Box, Modeling Multiple Time Series with Applications, *J. of the Amer. Stat. Assoc.* 76, 1981. 802–816.

Tomek, W.G. and R.W. Gray, Temporal Relationships among Prices on Commodity Futures Markets: Their Allocative and Stabilizing Roles, *Amer. J. of Agri. Econ.* 52, 1970. 372–380.

Tung, L., *Time Series of Synchronization and Chaos*, unpublished Ph.D dissertation, University of California, San Diego, 1990.

Ward, C.E., Relationship between Fed Cattle Market Shares and Prices paid by Beefpackers in Localized Markets, *W. Journal of Agri. Econ.* 7, 1982. 79–86.

———, An Empirical Study of Price Discovery and Competition for Slaughter Lambs, *W. Journal of Agri. Econ.* 9, 1984. 135–44.

———, *Meatpacking Competition and Pricing*, Research Institute on Livestock Pricing, Blaksburg, Virginia, 1988.

Working, H., Theory of the Inverse Carrying Charges in Futures Markets, *Journal of Farm Economics* 30, 1948. 1–28.

———, The Theory of Price Storage, *American Economic Review* 39, 1949. 1254–1262.

———, New Concepts Concerning Futures Markets and Prices, *American Economic Review* 52, 1962. 431–459.

———, Economic Functions of Futures Markets, in '*Futures Trading in Livestock – Origins and Concepts*, Bakken, H (ed.)', Chicago, CME, 1970.

———, Price Effects of Futures Trading (1960) 'represented in "*Selected Writings of Holbrook Working*, Peck, A.E. (ed.)', Chicago, CBOT, 1977. 45–75.

Zellner, A., Causality and Econometrics, in three Aspects of Policy and Policymaking: Knowledge, Data, and Institutions, *Journal of Monetary Economics* 10, supplementary series, 1979. 9–54.

2

VITA

Boo–Yong Song

Candidate for the Degree of

Doctor of Philosophy

Thesis: A PRICE DISCOVERY STUDY OF CASH SLAUGHTER HOG PRICES, WHOLESALE PORK PRICES, AND LIVE HOG AND PORK BELLY FUTURES MARKET PRICES

Major Field: Agricultural Economics

Biographical:

Personal Data: Born in Hadong, Kyeong–Nam in Korea, February 10, 1959, the second son of Mrs. O. S. Kim and the late Mr. I. S. Song; served in Korean Army, August, 1984 through August, 1986; married with Ms. S. K. Lee, August, 1987; son, E. Junhyuk was born, August, 1988; daughter, J. Jinha was born, April, 1990.

Education: Graduated from Dae–A High School, Jinju, Korea, in February, 1978: received Bachelor of Science Degree in Agricultural Economics from Dongguk University, Korea, in February, 1982; received Master of Science Degree in Agricultural Economics from Dongguk University in February, 1984; studied Economics and Agricultural Economics at Iowa State University, August, 1987 through August, 1990; completed requirements for the Doctor of Philosophy Degree at Oklahoma State University in July, 1993.

Professional Experience: Graduate Assistant, Department of Agricultural Economics at Dongguk University, March, 1983 through February, 1984; Graduate Research Assistant, Agricultural Experimental Station, Department of Agricultural Economics, Oklahoma State University, November, 1991 through present, 1993.