INDUCTIVE CAUSATION ON STRATEGIC BEHAVIOR: THE CASE OF RETAILER AND MANUFACTURER PRICING

A Dissertation

by

FRANCISCO FRAIRE DOMINGUEZ

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

December 2009

Major Subject: Agricultural Economics

INDUCTIVE CAUSATION ON STRATEGIC BEHAVIOR: THE CASE OF

RETAILER AND MANUFACTURER PRICING

A Dissertation

by

FRANCISCO FRAIRE DOMINGUEZ

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	David J. Leatham
Committee Members,	David A. Bessler
	Oral Capps, Jr.
	Harold A. Love
	Marietta J. Tretter
Head of Department,	John P. Nichols

December 2009

Major Subject: Agricultural Economics

ABSTRACT

Inductive Causation on Strategic Behavior: The Case of Retailer

and Manufacturer Pricing. (December 2009)

Francisco Fraire Domínguez, B.S.; M.S. Instituto Tecnológico y de Estudios Superiores de Monterrey

Chair of Advisory Committee: Dr. David J. Leatham

Models of strategic behavior are usually too complex to conduct large scale analyses, and frequently rely on accurate descriptions of the strategic environment, or unrealistic assumptions which render empirical studies very sensitive to misspecification. This dissertation relates game-theoretic frameworks to models of causality inference and thus provides a reliable method to identify price leadership. Therefore, causal models can be used to study large sets of data without imposing strategic behavior a priori.

A case study is provided by analyzing the supply chain relationship among Dominick's Finer Foods and its suppliers. Although our data required aggregation, this empirical analysis successfully determined causal patterns for 60% of our sample. Of these price leaderships, 70% elicit Manufacturer Stackelberg relationships which tend to be associated with manufacturers that hold big market shares, 25% elicit Retailer Stackelbergs which seem to be associated with the biggest retailer margin profits, and

only 5% elicit a monopolistic retailer with vertical coordination. These results agree with observations made by other authors and the market structure of the 1990's.

Moreover, the strategic relationship among the suppliers is also studied. Interestingly, the dominant firms tend to isolate themselves from the price leadership, whereas the second largest firms seem to become price leaders. Our studies agree with the market literature as well. In particular, we find price leadership in a firm which was identified as a low cost leader. Finally, we discovered that the private label does not lead any firm's price unless this firm is the provider of a generic brand.

DEDICATION

This dissertation is dedicated to my wife, mother, and sister. It is my admiration for them that gives me the strength and energy to work harder. I learned from them to remain joyful and positive. But foremost, it is my love for them that keeps me from giving up.

ACKNOWLEDGEMENTS

I am deeply thankful to several people for this dissertation. First and foremost, I want to express my gratitude and admiration for every member of my research committee. I am awed not only by their knowledge and patience, but also by their outstanding quality as human beings. Their guidance was always provided with interest and motivational words, and was never restricted to academic matters. Their exemplary lives, filled with hard work and humility amidst of great achievements have reshaped my life. I look forward to the opportunity of making them proud.

Many thanks to Dr. Leatham for watching over my wife and me, for allowing me to pursue my own research interests, and for believing in my capacity even more than I. We treasure his caring actions, comments and sense of humor. I am deeply grateful to Dr. Bessler for such stimulating conversations and encouragement. I am profoundly appreciative of his guidance, support, and his interest in my research and wellbeing. My gratitude goes to Dr. Love for his patience, ideas, and advice. Many thanks as well to Dr. Capps for his patience, teachings, enthusiasm, and for the pats on the back with words of encouragement. Thanks to Dr. Tretter for her caring guidance even though she did not get to know me very well.

My appreciation is also for all the members of the faculty. In particular, I want to thank Dr. Robinson, Dr. Fuller, Dr. Woodward, and Dr. Park, who have sheltered me and provided me with such strong and valued friendship. Many thanks as well to Dr.

Victoria Salin, Dr. Eluned Jones, and Dr. Bruce McCarl for their help and advice with my research interests, studies and tests, and to Vicki Heard for all her help and smiles.

I want to express my appreciation to other professors who helped me get here. Special thanks to Dr. Antonio Arreola-Risa, who is always watching over my wife and I, and for his caring friendship from the very beginning. Thanks to the professors that helped me from Mexico: Dr. Rebeca Romero, Dr. Christian Garrigoux, Dr. Viatcheslav Kalachnikov, Dr. Lourdes Dieck, Dr. Alejandro Dieck, Dr. Irma Gómez, and Dr. José Manuel Sánchez. Thanks to CONACyT for their support. Thank you for believing in me.

My deepest gratitude goes to my wife, sister, and mother, who have sacrificed so much for me. Thanks to Lauren Dangelmayr, Jonathan Sammon, and Rollie Wicks, who are members of my family as well. Thanks to Angela Sánchez from the Office of Graduate Studies for her smiles, hugs, and jokes.

My work benefited greatly from the help provided by Randi Hughes-Fraire, my wife. She not only helped me with grammar editions, spell checks, and explanations of word meaning, but she also provided me with lessons of hard work and determination. Her kind and loving gestures helped me deal with tiredness, frustration, and lack of creativity. Her genuine interest in my work led us to many interesting discussions and often into new directions. Thanks for cheering me up and keeping me optimistic, for keeping me company while I worked, for the infinite patience. I value deeply all her sacrifices. Thank you from the bottom of my heart.

To all of you, thank you. You will remain forever in my heart.

NOMENCLATURE

DG	Directed Graph
DAG	Directed Acyclic Graph
DCG	Directed Cyclic Graph
PC	PC Algorithm
X	Capital Bold Letters are Matrices
x	Lower Case Bold and Italics Letters are Vectors
X	Capital Letters and Italics are Scalars or Random Variables
x, *a*	Lower Case Letters are Scalars or Realizations of a Random Variable
UPC	Universal Product Code

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

The search for above normal performance by firms in competitive environments leads to the use of strategic behavior. Since different strategies lead to different distribution of profits, the analysis of strategic relationships has caught the attention of both academics and practitioners. This research studies the use of causal inference models to uncover strategic behavior among competing firms. This is important because proper determination of a strategic relationship has been found to be essential for accurate modeling. Moreover, models of strategic behavior are usually too complex to conduct large scale analyses. This dissertation thus explains how the usual game theoretic frameworks imply different causal patterns and therefore can be elicited through models of causal inference. Thereafter, we apply this methodology to analyze both the horizontal relationship among several manufacturers, and their vertical relationship with a retailer.

This study addresses a common question. Simply put, who sets the price of a product? The right answer is *it depends.* On one hand, demand is determined by consumers' preferences, the variety and relationships among their feasible alternatives, and their current and expected income. On the other hand, supply is determined not only by the scarcity of the resources being used, but also by the possibility of new entrants, the existence of competitors, the relationships among them, and their relationships with

---

This dissertation follows the style of *Management Scienc*e.

their supply chains. Consequently, different environments lead to different solutions. For these reasons, researchers specify complex models that try to include all these factors. But the level of their complexity increases with their degree of comprehensiveness, which makes estimation of the underlying structure very difficult.

Indeed, complex models bring along another problem. As we will discuss later, our use of econometric models reminds us of the *observer's effect* often discussed in experimental science; it appears that the specification of structural models tends to influence the results that we obtain. Therefore, this research studies a different approach to observe the strategic relationships between firms. Such behavior has been long analyzed under the microscope of Game Theory. And, although several avenues of investigation have been used, proper identification remains a problem. Who sets the price? Is there a first mover whose actions ripple across other firms? Or is it that, when the equilibrium is disturbed, all prices just bounce up and down until a new equilibrium is reached? Are these relationships present within a supply chain? Further, does the pricing relationship between the retailer and the supplier change with different products?

Several different methods have been designed to gain insights in market power and strategic behavior. For example, New Empirical Industrial Organization (Appelbaum, 1979), Solow Residual estimation of market power (Raper, Love, Shumway, 2007), Models of Discrete Games (Bresnahan and Reiss 1990, 1991), non parametric methods (Ashenfelter and Sullivan, 1987), and many other clever configurations are either used to determine the degree of departure from perfect competition, or the nature of the underlying strategic game such as the presence of a price leader, or simultaneous movers. But each of these methodologies comes at a price. Different assumptions regarding the technology, costs, structure, or even the game itself, must be imposed which greatly complicate the estimation. Here, we identify the game, per se, with the use of Directed Graphs (DGs) by associating their causal patterns with those implied by Nash equilibriums.

In the following section, the problem to be analyzed and the dissertation objectives are stated. We review in Section 3 the most common techniques employed to address these objectives, and explain how our methodology offers useful advantages over them. Thereafter, the theoretical framework often used and how it implies the causal relationships that we elicited, is discussed in Section 4. We present in Section 5 an application of our study to the analysis of strategic behavior within a supply chain, and also learn the relationship among the suppliers of a single retailer. Finally, a summary, conclusions, and suggestions for future research are provided in Section 6.

2. STATEMENT OF THE PROBLEM

The difficulty of acquiring profits has led to the study of strategic behavior in competitive environments. Although the theoretical framework has long been established, the empirical models still need to catch up. This is because the models of strategic behavior are usually too complex to conduct large scale analyses, or they rely on accurate descriptions of the strategic environment, which may lead to misspecification. Indeed, the complexities of the analytical solutions make it infeasible to jointly study more than a few products, firms, or industries.

For instance, the Lee and Staelin (1997) study of strategic behavior in supply chains mentions the need for models that include all the dimensions of competition; horizontal and vertical at every level (manufacturer-manufacturer, retailer-retailer, and manufacturer-retailer). The authors wondered if some of their results are theoretical possibilities only.

Another example is provided by Stockton, Capps, and Bessler (2008). They used causal models to examine Samuelson's mixed demand systems, which imply that prices exogenously determine the quantities for some goods, while the opposite is true for other commodities. Thus, it would be interesting to reexamine Samuelson's duality from a strategic perspective.

The objective of this study is to relate models of causality inference to the framework used to study strategic behavior. Connecting the two frameworks enables us to characterize strategic competition and understand how pricing policies are being implemented. This is important for several reasons. First, accurate predictions are imperative for the advancement of research. Specific knowledge of the underlying strategic relationship is therefore essential for an accurate specification of structural models. Second, large scale studies are necessary for the achievement of robust conclusions. But the usual structural models are impractical in this setting. Nevertheless, causal models are algorithms that can be automated and thus facilitate the analysis of large datasets. We accomplished these objectives by revisiting the game theory analysis of strategic behavior (Nash equilibriums of the Cournot, Stackelberg, and Bertrand games), and its implications on causal models. We programmed the PC Algorithm (Sprites *et al*., 1993), Moneta's Algorithm (Moneta, 2008), and Richardson's Algorithm (Richardson, 1996) into a spreadsheet and conducted an empirical study that provides support to our theoretical findings.

The results of this research have important applications for practitioners as well. Strategic decision making requires an accurate description of the business environment before any strategies are pursued. Therefore, observing the actual pricing policies is important because competitors' optimal responses should not be assumed nor expected a priori (Lee and Staelin, 1997; Gibbons, 1992).

3. LITERATURE REVIEW

The analysis of price movements is interesting per se. The movements are not only reflections of supply and demand shocks and changes, but also consequences from strategic interactions between firms trying to secure a higher level of profits. This research studies the later. Firms' engagements into tacit or explicit agreements have produced a vast array of significant research, whether those strategies are of a horizontal (among competitors), or vertical (within a supply chain) nature.

We revisit in this section the literature that has addressed strategic behavior with the use of empirical models. It is introduced in Subsection 3.1 our review by discussing the different complexities found when eliciting strategic behavior in horizontal competition. We move on to Subsection 3.2 to explain how these problems are also present when studying vertical competition. Throughout, we describe how some of these problems can be addressed by identifying the kind of strategic interaction beforehand. Our discussion ends by proposing the use of Directed Graphs for that purpose. Although Directed Graphs have not been used in a similar setting, we review their use by economists in Subsection 3.3.

3.1 EMPIRICAL MODELS OF HORIZONTAL COMPETITION

Many different model configurations have been created to elicit strategic relationships. The most used framework was provided by Bresnahan and Reiss (1990, 1991) and Reiss (1996), which allows estimating structural models that describe simultaneous-move, sequential-move, and cooperative equilibriums. Their equations are thus used to describe the equilibrium which strongly depends on the specification of the game structure. This approach has since been extended to systems that study several different scenarios. For example, Slade (1992) analyzed a dynamic environment with the purpose of eliciting the slopes of inter temporal response functions. In this case, the Kalman filter (Kalman, 1960; Kalman and Bucy 1961) is the estimation technique used to identify the strategic behavior observed in Vancouver, British Columbia during gasoline-price wars.

Slade (1992) emphasized the need for data that covers periods with demand shocks, so the actual strategies being used can be observed. She also stressed the need of highly disaggregated data. She assumed symmetrical firms with differentiated products and identical marginal costs. But most importantly, she assumed that the observed price summarizes all its previous history. In other words, she assumes a Markov process where the response functions take the form of the expected drift. That expected drift is the one that describes the underlying game ranging from a perfect collusive behavior to a Bertrand-Nash equilibrium.

In an earlier paper, Slade (1990) considered several different models used to analyze games. These models assume infinitely repeated games with discounted future benefits embedded in the current payoff. She also focused her attention in periods of time where disequilibrium can be observed. Although her objectives were different from her 1992 paper, she emphasized that an accurate model specification is crucial. Such specification should include the choice of punishment for the cheater, the nature of the shocks that disturb the system, and what is assumed to be common knowledge to all parties.

Imposing the wrong environment on a model can lead to misspecification. As an example, Carter *et al*. (1997) argued that although only quantity-setting strategies are usually studied (since they escape the Bertrand paradox), such behavior should not be assumed a priori. Especially when differentiated products are being studied (Wolak and Kolstad, 1991). Indeed, removing restrictive assumptions is important, but difficult to successfully implement. Carter *et al*. analyzed the Japanese market for imported beef. They applied the methodology introduced by Gasmi *et al.* (1992) which consists in specifying different game structures and then evaluating which one best explains the observed data. In particular, Bertrand and Stackelberg price-based models, and quantity-based Cournot and Stackelberg models are used. First order conditions are taken on these specifications and then estimated as systems of equations. The model which provides the best statistical fit, is the one assumed to uncover the actual firms' behavior.

Aradillas-Lopez (2005) discussed a downside of this methodology. He argued that the assumption of complete knowledge of the other player's behavior, as well as economic fundamentals, will result in multiple equilibriums, which may be consistent with different model parameterizations. Ardillas-Lopez (2005) revisited Bresnahan and Reiss (1990, 1991) but avoids the complete-information assumption by allowing an environment in which the players observe a noisy signal of the game's realized payoff structure. Thereafter, the agents maximized their expected utility conditioned on imperfect information and achieved a Bayesian Nash equilibrium. An a priori nature of the game was not assumed. Although this approach does not make inferences on the nature of the game, it does emphasize that different game specifications can achieve similar results when multiple equilibriums are allowed, making it harder to accurately identify the type of strategies employed by the firms. Other authors have also paid attention to multiple equilibriums and incomplete information. Bajari *et al.* (2004) provided a simulation-based approach that determines the probability of achieving particular equilibrium points. Tamer (2003) used restrictions on the probability of the non unique outcomes to demonstrate that multiplicity of equilibriums can be exploited to achieve gains in efficiency.

There is another downside of the Gasmi *et al*. (1992) methodology. Dhar, Chavas, Cotterill, and Gould (2005) considered the leading brands sold by PepsiCo. and Coca-Cola looking for market structure and strategic pricing to study strategic games with multiple firms that promote multiple brands. The authors discussed that the accurate specification of the model is crucial, and misspecification can lead to spurious results. They argued that the main weakness of specifying profit-maximizing first-order conditions under the assumptions of Bertrand, Cournot, or Stackelberg games, relies on the simplified demand specifications used. They addressed this problem by using a fully flexible nonlinear almost ideal demand system (AIDS) and later derived the corresponding first order conditions to be estimated with full information maximum likelihood (FIML). They advocated the use of this methodology with the use of multiple products and multiple firms. However, they do not find evidence that supported either the Nash-Bertrand or the Stackelberg equilibrium in their sample.

Some interesting thoughts come also from the study of Dhar *et al.* (2005). The authors indicated that, although the same manufacturer can provide several different brands, each of them can be marketed differently. This is because the manufacturer may be the provider of a brand that dominates one segment of the market, but not be the provider of the dominating brands in different segments[1]. This is of particular interest for this dissertation. Since the structural model has to be carefully specified, it is not by any means intuitive to determine the structure of the game a priori. A firm may have a Stackelberg leading brand in one segment of the market, but a Stackelberg follower brand in a different segment of the market. They may also engage in collusive behavior for some brands, but in non cooperative competition in others.

3.2 EMPIRICAL MODELS OF VERTICAL COMPETITION

The above comments also apply to the analysis of vertical competition. Indeed, when supply chain members are independent decision makers, there are incentives to improve individual performance (Leng and Parlar, 2005). This results in an environment where a great array of strategies is pursued by firms, not only to influence the distribution of profits, but to change their power within the supply chain as well (Kadiyali, Chintagunta, Vilcassim, 2000). Therefore, game theory has also been used to analyze the relationship between retailers and manufacturers. Leng and Parlar (2005) published a very comprehensive review of more than 130 papers and classified the papers according their application areas, and although not all of them refer to pricing

[1] Thanks to Whitney Bessler for her valuable insights.

[2] By "causality" we mean that if x_i can be manipulated by changing x_j, then x_j causes x_i. In a graphical

schemes but contracts designs, most of them imply Nash (simultaneous moves) and Stackelberg equilibriums (Figure 1). As we explained above, the estimation of structural models becomes increasingly difficult. We will discuss below that a detailed characterization of the (assumed) competing environment is also required to customize the models, i.e., the number of retailers and manufacturers, the type of interaction among manufacturers and retailers, the effect of one manufacturer actions on other manufacturers, etc.

There are many papers that directly addressed the manufacturer-retailer strategic behavior by using tailored made models. Jørgensen (1986) analyzed optimal production and pricing policies of a manufacturer, whereas the retailer also wanted to determine optimal purchasing quantities and pricing policies. The problem was considered in a

Figure 1 Topical Classification of Game Theoretic Studies of Supply Chain Management. (Leng and Parlar, 2005)

1. Inventory games with fixed unit purchase cost.
2. Inventory games with quantity discounts.
3. Production and pricing competition. (Vertical competition between the retailer and manufacturers).
4. Games with other attributes.
    a. Capacity decisions.
    b. Service quality.
    c. Product quality.
    d. Advertising and new product introduction.
5. Games with joint decisions on inventory, production/pricing and other attributes.
    a. Joint inventory and production/pricing decisions.
    b. Joint inventory and capacity decisions.
    c. Joint production/pricing and capacity decisions.
    d. Joint production/pricing and service/product quality decisions.
    e. Joint production/pricing and advertising/new product introduction decisions.

continuous time setting and the complexity of its solutions derived from the use of differential game theory. McGuire and Staelin (1983) assumed static linear demands and cost functions to generate a system of equations that were later used to examine the effect of product substitutability on the way manufacturers distribute them. Once again, the business environment was highly specified. The retailer marketed only the good provided by the manufacturer such as car dealerships, gas stations, etc. Or a large wholesaler distributed only through a single retailer like Sears. These assumptions facilitated the generation of a model where manufacturers maximized profits in different scenarios of competitive behavior with different reaction functions both from competing manufacturers and retailers.

Choi (1991) studied Nash and Stackelberg equilibriums to analyze the competition problem between two competing manufacturers and a common retailer. Once again a system of equations was used to estimate the optimal behavior under very specific assumptions about how manufacturers and retailers behaved. In particular, Choi assumed that manufacturers had the same relationship with the retailer (Nash or Stackelberg in the same direction), while they hold a Bertrand-type competition among them. Choi noted that his results were greatly influenced by the specification of the demand function (linear Vs non-linear). Kadiyali, Chintagunta, and Vilcassim (2000) extended Choi's results to measure how profits are distributed within the channel and how this depended on demand factors, costs factures, and the strategic behavior. In this research, a continuum of equilibriums is allowed instead of restraining to the three types estimated by Choi.

Chiang and Hess (2003) also specified a profit maximizing system to analyze strategic behavior within a supply chain. They a priori set a Stackelberg strategy where the manufacturer was the leader. Their aim was to analyze the effect of a direct channel of distribution on the manufacturer's profits. They found that when the manufacturer and the retailer act independently, the equilibriums yield a higher price with lower sales and lower profits. When the manufacturer can sell its product directly to the consumer, a credible threat is generated that results in lower prices offered by the retailer and thereafter increasing demand for the product.

Identification of the equilibrium type is important in all these studies because different degrees of complexity are uncovered when specifically looking at the relationship between a retailer and a supplier. For instance, Martın-Herran, Taboubi, Zaccour (2006) tried to identify the optimal shelf-space allocation of products. They explicitly assumed a Stackelberg game between two manufacturers and a retailer. The manufacturers were the first movers whereas the retailer was the follower. Both manufacturers set their wholesale prices using the retailer's shelf space as second mover's reaction function. But they did not engage explicitly in a game with their competitor.

The problem just described provides another lesson. The relationship among manufacturers is as important as the relationship with the retailer, which also leads to strategic behavior. Although it is more frequently assumed that the manufacturer is the leader in the Stackelberg game (Wang, 2002; Corbett *et al*., 2004), there are several studies that challenge this assumption and analyze the possibility of a simultaneous moves. Chu and Messinger (1997) and others assumed that the retailer is the first mover (Choi, 1991; Tsay, 2002; Ertek and Griffin, 2002). But once again, no attempts are made to identify the true nature of the game.

Our problem relates very strongly to the one examined by Lau *et al*. (2007). Given a demand for a product and the cost of producing it, both the retailer and the manufacturer set their prices. Since the cost of producing the good is not known by the retailer, a bargaining process should start so as to get the lowest price possible from the manufacturer. Lau *et al*. considered how a dominant retailer should operate when there is no knowledge of the manufacturer's costs. This retailer, optimizes with only two available strategies; a Stackelberg game where the retailer is the first mover, or a Stackelberg game where the manufacturer is the first mover. Lau *et al*. find it optimal to design a "reverse quantity discount" scheme offered by the retailer to the manufacturer. But once again, a Stackelberg game is assumed a priori. Indeed, such assumption makes sense for powerful retailers as Wal-Mart or outsourcing manufacturers. But what about retailers that are not as powerful or manufacturers that have even more power?

Our problem also relates to Choi (1991) who analyzed the effect of an intermediary on the competing behavior between manufacturers. He found counterintuitive results that imply that a manufacturer should do business with an exclusive retailer, whereas the retailer should work with several manufacturers. Also, he found that channel members' profits and prices are increased as products are less differentiated. A large empirical study would provide with a more robust analysis. But a large empirical study has not been conducted because of the hardship required to specify an accurate system. Kadiyali, Chintagunta, and Vilcassim (2000) worked on Choi's findings and provided a more general system, but acknowledged that a system of equations is very sensitive to misspecification.

Lee and Staelin (1997) also studied the interactions between manufacturers and retailers. Their study directly addressed the determination of the optimal pricing margin as a response to a change in margin by another member of the supply chain. They demonstrate that the optimal pricing policy, to increase or decrease their margin, is closely related to the level of demand at a given price, and the convexity of the demand function. Therefore, they suggest that the linear (LAIDS) or nonlinear (LAIDS) specification of the demand in the structural models is not as important as determining the type of interaction present between the channel members. This is because the members of a supply chain do not always behave optimally. Therefore, the choice of a firm's optimal strategy is not constrained by the best response of the channel member, but by the actual response being observed. Indeed, Dhar *et al.* (2005) assumed a fixed markup as the retailer strategy, but they also acknowledge that studies of this nature (that assume that retailers do not engage in strategic behavior) will be biased. For this reason they recommend studying this problem with store level data and information on manufacturers-retailers contracts (a natural extension for our study as well).

But it seems that the study of strategic interactions is incomplete if only one dimension is addressed. In other words, vertical and horizontal coordination depend on characteristics of both vertical and horizontal industries, and failure to explicitly consider one dimension when studying the other could lead to inaccurate conclusions

(Lee and Staelin, 1997). Moreover, the Folk Theorems imply the existence of more equilibriums when firms engage in a repeated game (Friedman, 1971). This supports Lee and Staelin (1997) call for improvements in modeling, by including credible threats and punishment strategies to study stability of stage strategies.

Furthermore, Lee and Staelin (1997) argued that the stability of Stackelberg strategy is debatable when the retailer is the leader. This is because the game occurs in stages, where in the first one, the leader chooses its strategy based on the second mover's best response. So, the retailer would announce its margin strategy and let the manufacturer set his strategy as a response. But once the retailer owns the product, the retailer might choose a different margin from the one that was previously announced to the manufacturer. Therefore, Lee and Staelin (1997) suggest that a pre-commitment is required to restrict deviations from the equilibrium, which is enforced through punishment strategies and credible threats. We agree on the need for punishment strategies and credible threats to ensure the stability of equilibriums, but we conclude that the retailer would not find it on its best interest to deviate, since a proper pricing policy should be time consistent. Actually, it is the consistency of the strategies which allow us to use the causal models. Once the manufacturer chooses his margin, the retailer's best response should stay the same. A review of the use of causal models in economics follows. We conclude by explaining how these models are used in this research.

3.3 CAUSAL MODELS

Causal models are a research technique which identifies variables that can be manipulated by changing other variables. Although they are used extensively by statisticians, their use has been limited in the fields of economics, industrial organization, marketing, and management. Directed Graphs (DGs) allow us to stand as mere observers of price and quantities movements without assuming a priori any underlying game. This methodology can be used to observe the statistical properties of price movements and provide a simplified description of a joint probability function, useful to determine causality among variables (Frydenberg, M., 1990). Therefore, once two economic agents engage in a repeated game, such causality in price movements elicits the strategic behavior.

Causality between variables is inferred by taking a detailed look at correlations and partial correlations[2]. We can deduce that a change in one variable can be used to manipulate others if the correlation between two variables disappears (or appears) when conditioned on a third variable. This is because given a set of variables $X = \{x_1, x_2,..., x_n\}$ with a joint probability density function $\Pr[x_1, x_2,..., x_n]$, the probability function of x_i conditional on subsets of X can be obtained. Thereafter, the traditional definition of probabilistic independence can be used to infer causality among them and create graphs as an efficient representation of the underlying probability distributions. The variables are then charted with directed connectors to indicate a causal path, which in our context elicits a first mover in price setting. There are several computer algorithms that achieve

[2] By "causality" we mean that if x_i can be manipulated by changing x_j, then x_j causes x_i. In a graphical representation we have $x_j \rightarrow x_i$.

this purpose. Those algorithms that proved to be the most useful for this research will be described in detail.

As we mentioned before, the use of DGs has not been applied in strategic behavior analysis, but has been used in the field of economics. Bessler and Kergna (2002) used graphs that do not allow for recursive patterns, Directed Acyclic Graphs (DAGs), along with error correction models, to analyze how price information is discovered in one market and then transferred to others. They address questions very similar to ours but in a different environment. Specifically, they studied which markets set prices and how that information is then transferred to other markets.

Bessler and Yang (2003) applied the same methodology to analyze the structure of interdependence in international stock markets. They wondered if there are stock markets whose prices move before the rest, revealing a causal information flow among them. Bessler and Yang discussed that the recognition of non-stationarity of stock prices had induced a search for long term relationships through cointegration analysis. But previous studies usually failed to address the contemporaneous relationships among the stock markets. Specifically, Bessler and Yang questioned if innovations in a market had a causal relationship with other markets. They removed the estimated structure from the observed price data so innovations can be uncovered. These innovations were later examined with the use of DAGs which graphically showed the causality patterns that needed to be imposed on the impulse response functions.

Moneta (2008) used the same methodology to revisit the business cycle hypothesis. He introduced several small modifications to the *PC algorithm* which resulted in a more conservative analysis, but more precise DAGs too. In particular, his algorithm tests each relationship under more conditioning tests and uses a very data specific test (a Wald statistic), whereas *PC* relies on an asymptotic normal distributed statistic. In our context, we want to study how the information is transferred horizontally among manufacturers and vertically among manufacturers and retailers.

We will have one important difference from the methodology introduced by Bessler and later used by Moneta. Bessler was concerned with finding a contemporaneous causality pattern among the researched variables, whereas the dynamics were analyzed through other methodologies such as the impulse-response functions (Swanson and Granger, 1997). Therefore, the residuals of the estimated Error Correction Model (ECM) where used as inputs for the DAG algorithm. Here, we care about the dynamics per se and therefore we do not need to estimate an ECM as Bessler did, nor a VAR as in Moneta's study. Stationary series (by taking differences or differences of the logarithms of the original series) are all we need to avoid finding spurious relationships (Bessler and Kling, 1984)[3]. Thus, we started by analyzing the stationarity of the prices with Dickey Fuller tests and then transformed the data when required. These were later fed into the algorithms that create the DGs. Specifically, we used a truncated version of Richardson's Directed Cyclic Graphs algorithm (Richardson 1996) and Moneta's Algorithm, as well. However, we used the original Fisher statistic

[3] Bessler and Kling showed that wrong relationships can be inferred if the order of integration is ignored. They demonstrated their point by demonstrating (what is obviously false) that the US GDP causes sunspots, if proper treatment of integration properties are ignored.

given that the Wald test proposed by Moneta is not required here because we are not using the variance-covariance matrix of the VAR innovations as inputs.

We have discussed the literature in this section and we concluded that the use of causal models will allow us to study strategic interactions between firms from an observer's point of view. Causal models not only allow us to study pricing behavior without a priori imposing any structure in the relationship between the analyzed firms, but also permit the study of large datasets and several firms at the same time. This is imperative for robust results. The pricing behavior is explained by the usual models of strategic behavior which are reviewed in the next section. We provide an empirical case study in Section 5.

4. METHODOLOGY

We have discussed that the analysis of strategic behavior has been constrained by the complexity of the systems that have to be estimated. Research conclusions tend to be sensitive to misspecification, and broader and more robust studies are usually encouraged by the literature. Therefore, this research extends the use of causal models to the study of strategic behavior. To achieve this purpose, we now revisit the analysis of pricing strategies through the game theory framework, and explain how its findings relate to the one used for causality inference.

We start by discussing in Subsection 4.1 the analysis of Nash equilibriums of firms engaged in horizontal competition *(Manufacturer-Manufacturer)* in a static environment. Then, we discuss how this framework is extended to a dynamic setting by assuming that each of the static processes are repeated continuously in a recurring relationship. We proceed to extend this framework to that of vertical competition *(Manufacturer-Retailer)* in Subsection 4.2. Thereafter, in Subsection 4.3, we present the algorithms used in this dissertation after discussing the implications of the Nash equilibriums on causal patterns. An empirical application of these concepts is offered in Section 5.

4.1 THEORETICAL MODEL: HORIZONTAL COMPETITION

We reviewed the literature in the previous section and discussed how strategic behavior in price settings has been analyzed with the use of game theory. We concluded that industrial organization research has found it difficult to empirically observe and quantify the effects predicted by the theoretic models. This is because of the complexity of the systems that have to be estimated and the assumptions that have to be imposed on those models. Particularly, specifying the underlying game is fundamental for accurate predictions. This dissertation studies the use of causality models to identify the nature of the strategic relationships. Therefore, the theoretical models are now discussed, and how their results imply a causal relationship that can be detected by our methodology.

In this section, we link the classic horizontal strategic games to the causal graphs. Horizontal games describe scenarios where firms choose their prices or levels of production based on information collected from their competitors. We start by specifying the monopolistic decision process in a static setting, then we obtain the Nash equilibriums of the Cournot, Stackelberg, and Bertrand games (Gibbons, 1992), and we discuss their relationship with the causal models, which are fully addressed in Subsection 4.3.

As we will see, repetition of the strategic behavior is fundamental for the successful use of causality models. This is because repetition of the experiment is essential to the generation of a reliable sample. The monopolistic decision process is instructive as a benchmark and provides the incentives for collusive behavior in infinitely repeated games of horizontal competition (Gibbons, 1992). Therefore, dynamic settings of complete information are addressed, where the players engage in games that repeat either a finite or an infinite number of times. At each repetition (stage), players can observe the history of the previous plays and outcomes before they make new decisions. The problems solved here do not assume any specific functional forms and therefore yield the usual specifications found in the research literature as special cases, which pay particular attention to constant marginal costs and linear inverse demand functions.

4.1.1 Monopoly

Consider a continuous and homogeneous of degree zero demand function for a single product, $q = f^{-1}(\boldsymbol{p}, m)$, where $\boldsymbol{p}$ is the vector of competing product's prices, and m is the level of income. Ceteris Paribus, then an inverse demand function for a single product can be written as $p = f(q)$, (Deaton and Muellbauer, 1980; Gibbons, 1992). Consider also a cost function, $c(\boldsymbol{w}, q)$, concave and homogeneous of degree one in the prices of inputs, $\boldsymbol{w}$. Assuming that the monopoly's level of output has no effect on the prices of the inputs and the monopolist profits are defined by

$$\pi_m = f(q) \cdot q - c(\boldsymbol{w}, q), \tag{1}$$

then, the monopoly solves:

$$\max_q f(q) \cdot q - c(\boldsymbol{w}, q) \tag{2}$$

with optimal output q_m*, which is determined by the solution of the following first order condition:

$$\frac{\partial \pi_m}{\partial q} = f(q) + q \cdot f'(q) - \frac{\partial c(\boldsymbol{w}, q)}{\partial q} = 0 \tag{3}$$

The optimized level of profits becomes:

$$\pi_m^* = f(q_m^*) \cdot q_m^* - c(\boldsymbol{w}, q_m^*), \tag{4}$$

and the clearing monopolistic price is:

$$p_m^* = f(q_m^*). \tag{5}$$

Equation (5) is a result that needs further discussion. This implies that the monopolist chooses a quantity which has an associated price determined by the demand function. But the consumer does not directly set the price. From a strategic perspective, the monopolist sets the price and then let's the consumer purchase the quantities. Therefore, the monopolist does have the ability to manipulate the consumer's behavior. Summarizing, the monopolist produces a quantity, then sets the price, then lets the consumer buy those quantities. This implies the following causal relationship: ${q_m}^* \rightarrow {p_m}^* \rightarrow q_c$, where q_c is the consumer's actual consumption which should equal ${q_m}^*$.

4.1.2 Nash Equilibrium of the Cournot Game

But monopolies are rarely found. Cournot analyzed the possibility of two existing firms competing for the same market (Gibbons, 1992). Considering an oligopolistic scenario of *I* firms, each of them chooses how much to supply taking into account the effect of the quantities supplied by all others on the market's price through the consumer's demand function, and therefore the level of profits. Given an inverse demand function for a single product, firm *i* in a Cournot game will maximize profits π_i^c:

$$\max_q f(q^c) \cdot q_i - c_i(\boldsymbol{w}, q_i), \tag{6}$$

where q^c is the total quantity supplied to the market by all the firms:

$$q^c = \sum_{i=1}^{I} q_i \tag{7}$$

The first order condition is:

$$\frac{\partial \pi_i^c}{\partial q_i} = f(q^c) + q_i \cdot \frac{\partial f(q^c)}{\partial q^c} \cdot \left[1 + \sum_{j \neq i} \frac{\partial q_j}{\partial q_i}\right] - \frac{\partial c_i(\boldsymbol{w}, q_i)}{\partial q_i} = 0. \tag{8}$$

Assuming that no firm has an effect on the decisions of the other companies, then each firm will maximize its profits by solving for q_i from:

$$\frac{\partial \pi_i^c}{\partial q_i} = f(q^c) + q_i \cdot \frac{\partial f(q^c)}{\partial q^c} - \frac{\partial c_i(\boldsymbol{w}, q_i)}{\partial q_i} = 0. \qquad (9)$$

Given that we have I firms, then equation (9) represents a system of I equations and I unknowns that must be solved simultaneously.

The game theory literature usually resorts to a linear inverse demand function along with constant marginal costs to demonstrate that the solution to this system will yield a total level of quantities superior to the ones supplied by the monopoly, $q^c* > q_m*$ (Gibbons, 1992). This translates to a market price lower than that of the monopoly, along with lower profits to the firms, and therefore giving incentives for collusion. Such collusion would seek to restrict the output and therefore equally distribute the monopolistic profits given by equation (4).

Directly related to the objective of this dissertation, the system of equations (9) implies that the solution to each firm's choice of quantity q_i* will have the form $q_i* = g_i^c(\boldsymbol{w}, \boldsymbol{q_{-i}}*)$ for all i, were $\boldsymbol{q_{-i}}*$ is the vector that represents all the quantities supplied by all other firms not including firm i. In consequence, we could manipulate the value of q_i* by changing the value of any of the quantities supplied by any other firm, which in turn can also be manipulated by changing the value of q_i*. In other words, quantities are jointly and simultaneously determined in a Cournot game, and such joint determination should be detected by algorithms of causality that allow for cyclical structures. Such simultaneous causality will be represented by a graph that connects any two firms' production by a two-head arrow (i.e. $q_i \leftrightarrow q_j$). Moreover, if these firms decide to collude, then they would acquire the ability to manipulate the consumer's choices as it was the case with the monopoly. Assuming identical firms, then each of them would have a one directional causal relationship with the consumer $q_i \rightarrow p_m^* \rightarrow q_c$. On the other hand, failure to coordinate would imply that the consumer would not be manipulated. Therefore the causal relationship would be reversed (i.e. $q_i \leftarrow q_c$).

4.1.3 Nash Equilibrium of the Stackelberg Game

Stackelberg analyzed a different game in which there is an industry leader whose choice of quantities will cause all others to best respond (Gibbons, 1992). Then, consider firm *s* as the leader in a market of *I* firms. Each firm other than firm *s* will solve the problem in equation (6) with solution q_i* given by (9). The Stackelberg leader anticipates this and can make these responses as a part of its maximizing problem described by the equation:

$$\max_{q_s} f(q^s) \cdot q_s - c_s(\boldsymbol{w}, q_s) \tag{10}$$

where q_s is the quantity supplied by the Stackelberg leader, and q^s is the total quantity supplied to the market by all the firms:

$$q^s = q_s + \sum_{i \neq s}^{I} q_i^* \tag{11}$$

The first order condition now becomes:

$$\frac{\partial \pi_s}{\partial q_s} = f(q^s) + q_s \cdot \frac{\partial f(q^s)}{\partial q^s} \cdot \left[1 + \sum_{i \neq s} \frac{\partial q_i^*}{\partial q_s}\right] - \frac{\partial c_s(\boldsymbol{w}, q_s)}{\partial q_s} = 0 \tag{12}$$

where the derivative of q_i* can be computed with the help of the Envelope Theorem. Notice that the solution of equation (12), q_s*, is a function of the vector of input prices $\boldsymbol{w}$, but not of the quantities supplied by all the other firms, $q_s* = g_s^s(\boldsymbol{w})$. In contrast, the *I*

-1 Stackelberg follower firms do depend on the level of output provided by the Stackelberg leader, $q_i^* = g_i^s(\mathbf{w}, q_j^{s*}, q_s^*)$ for all $i \neq j, s$. This implies that although there is a bidirectional relationship between any firm other than the Stackelberg leader, $q_i \leftrightarrow q_j$, it is the Stackelberg leader who decides the output for everybody. Therefore, there should be a one way causal relationship between the Stackelberg leader and every Stackelberg follower (i.e. $q_s \rightarrow q_i$ for all $i \neq s$). Moreover, the bidirectional relationship among the followers should only be elicited when the Stackelberg leader is not present in the analysis.

In the Stackelberg game, the aggregated level of quantities supplied to the market will be even higher than that of the Cournot game, and therefore higher than those supplied by the monopoly, $q^{s*} > q^{c*} > q_m^*$. This results in an even more reduced level of profits. But this time, the profits received by the Stackelberg leader are superior to those received by the agents in a Cournot game. Although, the level of profits received by the Stackelberg followers will be lower than the π_c, the level of profits received by the leader will not make it desirable to relinquish its position. On the other hand, the Stackelberg followers have incentives to pretend that they do not see the leader’s strategy, forcing the Stackelberg leader to abandon his policy and enforce a Cournot equilibrium (Gibbons, 1992).

4.1.4 Nash Equilibrium of the Bertrand Game

The last of the games to be discussed is the one described by Bertrand (Gibbons, 1992). In this circumstance, firms use the output price as the choice variable. Whenever we consider non differentiated products, this scenario leads the Bertrand Paradox which describes a battle of prices that ends on a zero profits level for a competitive environment with non differentiated products. This is because the set of available strategies to the firms do not reach a Nash equilibrium (Equation 13).

$$q_i = \begin{cases} q(p_i) \; if \; p_i < \; p_j \; for \; all \; j \; such \; that \; i \neq j \\ \frac{q(p_i)}{I} \; if \; p_i = \; p_j \; for \; all \; j \; such \; that \; i \neq j \\ 0 \quad if \; p_i > \; p_j \; for \; any \; j \; such \; that \; i \neq j \end{cases} \tag{13}$$

The Bertrand Parodox would ensure that prices and quantities are not determined by the interactions with the competition, but by each firm's own ability to reduce their marginal costs. Therefore, the quantities/prices supplied by firms in a Bertrand game are independent of each other (i.e. $q_i \perp q_j$, $p_i \perp p_j$) and no causality pattern should be found. There are a few scenarios that provide an escape to the Bertrand Paradox (Mas-Colell *et al.*, 2004); Restrictions in the capacity of production (decreasing returns to scale), product differentiation, and adding a temporal dimension by allowing for a repeated game.

Restrictions in the capacity allow firms to price above their marginal cost, but this equilibrium is not easy to determine. This is because the lowest pricing firm will not be able to supply the whole market and rationing rules must be imposed to determine which consumers buy from each of the competing firms in this scenario. Unfortunately, none of the rationing rules is more likely than the others, and the equilibrium strongly depends on which rule is being imposed. However, the solution for the Cournot game results as a subgame perfect Nash equilibrium of the Bertrand game when firms are allowed to choose capacity first, and then compete in prices (Mas-Colell *et al.*, 2004).

Product differentiation is of special interest to our study because of the data that we have available for the Case Study offered in Section 5. Although our empirical research studied products in a highly competitive environment, their producers invest heavily in branding to set themselves apart. This is because differentiated products bestow some market power to their products given by the "elimination" of close substitutes. However, the producer's ability to manipulate its price will be bounded by the consumers' willingness to pay a premium over its closest substitute. This will result on a price slightly above their marginal cost (Mas-Colell *et al.* 2004). Therefore, the expected causal structure among different competitors remains as that of the undifferentiated products.

Finally, when repetitions of the game are allowed, then the Folk Theorems apply and therefore the monopolistic level of profits is feasible (Friedman, 1971) by collectively producing the monopoly quantity q_m* and charging the monopoly price p_m*, thus eliciting simultaneous determination of quantities. This strategy is not confined to the Bertrand game, because firms engaged in a Cournot relationship would also find it desirable to collude.

4.1.5 Game Repetition

The Folk Theorems make it imperative to pay attention to dynamic games of complete information. These are games with repetition where all previous moves by all players can be observed before the next action is chosen, and players move either in sequence or simultaneously (complete but imperfect information). As shown by Slade

(1992), players can observe past actions played by competitors. Therefore, threats and promises signaled through supplied quantities and/or prices can influence the behavior of the competitors, and more desirable equilibriums become feasible provided that players are patient enough, and the end of the economic relationship is not foreseeable. Bertrand and Cournot players can tacitly collude to achieve the monopolistic profits at each stage (repetition of the game) which still requires joint and simultaneous determination, Stackelberg leaders can learn their competitors' best response to make it endogenous in their decision process, and Stackelberg followers can observe the leader's optional choice as an exogenous determinant in their own profits maximizing process. Thus, the modification of the competitors' behavior through signals in repeated tacit relationships justifies the use of causal models to identify the nature of those relationships. Assuming that threats and promises are actually implemented, then these actions should *cause* changes in the competitors' quantities and/or prices.

Punishment strategies deserve a short discussion. Strategies that consist on punishing forever by alternating the choice of quantities (or prices) would be elicited by a causal model because one player directly causes the actions of the other. Thus, Tit-for-Tat strategies could potentially lead to misleading interpretations of the results. On the other hand, Trigger strategies just end cooperation. However, Green and Porter (1984) discussed that firms cannot always distinguish between deviations from collusion agreements and market shocks. Therefore, infinite punishment strategies would not be optimal since they could by triggered by accident (Gibbons, 1992). In any case, both Trigger and Tit-for-Tat strategies would generate an empirical structural break on the relationship among the agents. In other words, the covariance structure among the variables would elicit the change from cooperation to no-cooperation, which can be examined with log-likelihood ratio tests.

As discussed previously, regardless of being or not a repeated game, quantities are jointly and simultaneously determined when a Cournot game is played, i.e. $q_i^* = g_i^c(\boldsymbol{w}, \boldsymbol{q}_{-i}^*)$. Quantities or prices should appear independent in a Bertrand setup unless there is tacit collusion in a repeated game. In such cases, the use of threats and punishments would induce a simultaneous determination of prices. Stackelberg leaders would directly affect followers' quantities, $q_i^* = g_i^s(\boldsymbol{w}, q_j^{s*}, q_s^*)$, but quantities supplied by Stackelberg followers should not directly *cause* the Stackelberg leader's optimal choice, $q_s^* = g_s^s(\boldsymbol{w})$. These causal relationships can be elicited through causality inference models. Subsection 4.3 discusses such models in detail. Prior to that section, we now proceed to study the theoretical relationship between a retailer and its providers within this framework.

4.2 THEORETICAL MODEL: VERTICAL COMPETITION

In this section, we extend the framework previously described in Section 4.1 to analyze equilibriums associated with a perfect coordination of the supply chain, simultaneous determination (vertical Nash), a Stackelberg game where the manufacturer is the leader (manufacturer-Stackelberg), and a Stackelberg game where the retailer is the leader (retailer-Stackelberg). Finally we discuss how our results change in the presence of monopolistic behavior. The objective of this section is thus to relate each of these equilibriums to the causal inference models. Our theoretic framework is similar to the one provided by Lee and Staelin (1997), but there is a subtle yet important difference. In their study, the members of the supply chain obtain their profits by charging a margin over the cost of the product. Therefore, the margin is their decision variable which is chosen to maximize profits subject to the actions of the other members of the supply chain. For this reason, the choice of margin becomes the best response. McGuire and Staelin (1983) argued that there is no loss in generality when the framework studies margins instead of prices. This is because it only requires a rescaling of these variables if the manufacturer and the retailer have constant marginal costs. Our framework is thus different because we use the wholesale and retail prices as the choice variables in each of the maximization problems (Mcguire and Staelin, 1983; Choi, 1991; Kadiyali, Chintagunta, and Vilcassim, 2000).

Given one manufacturer, one retailer, and one product, the following (modified) assumptions observed by Lee and Staelin (1997) also apply:

Assumption 1: The demand for the product is downward slopping.

Assumption 2: Manufacturers face a constant marginal cost *c*. Retailers face a constant marginal cost *d* which can equal zero.

Assumption 3: All the firms have perfect knowledge of the demand and cost structures within the industry.

Assumption 4: Players can observe each other’s actions.

Assumption 5: There exists a unique set of equilibriums.

All of these assumptions are common in the literature, except for *Assumption 4*. But as shown by Slade (1992), not only firms can observe past actions played by their competitors, but they might also engage in periods of learning after demand or supply shocks disturb and change an old equilibrium. *Assumption 2* might deserve some attention as well. Constant marginal costs are usually assumed to focus on the interactions of the agents in the supply chain, and not on disturbances exogenous to the strategic decisions.

Before we proceed, it is imperative to have a discussion about the role of the consumer in a causality setting. Since the consumer maximizes utility by choosing how much to consume, but never by choosing prices, then the consumer will always be a source of causality in an environment with enough availability of substitutes. In other words, the retailer will choose its price (or quantities) as a best response to the demand function. This conjecture may deserve further study since it conflicts with Samuelson's mixed demand systems where prices are exogenous for some products, and quantities are exogenous for some others (Stockton, Capps, and Bessler, 2008). Nevertheless, the empirical research presented here supports our postulation that the consumer acts as a causal source. In particular, we found 93% of the times that the quantities consumed by the patrons of Dominick's Finer Foods directly cause either the retailer or the manufacturer's price. This is because the availability of substitutes restricts the retailer's ability to manipulate its prices. Therefore, although the following analytical description of the Nash equilibriums studies a demand function, $q=q(p)$, where p is a scalar instead of a vector, we must have in mind that the demand function, *q(p)*, has to be somewhat elastic.

4.2.1 Perfect Coordination

The analysis of perfect coordination of the supply chain is of special interest because the firms do not engage in double marginalization which may be a more efficient distribution of profits[4]. However, McGuire and Staelin (1983, 1986), and Coughlan and Wernerfelt (1989) showed that perfect coordination is not the best alternative for manufacturers of highly substitutable products. Therefore, a centralized decision maker, or perfect vertical integration, does not provide an adequate benchmark to compare different strategies pursued by supply chain members in a highly competitive environment. Nevertheless, the equilibrium is analyzed below for the sake of completeness.

Consider a retailer with constant marginal cost *d*, who purchases *q* quantities of a product at price $p_m{}^i$, which will be sold at price $p_r{}^i$ (*i* stands for integration). Then, the profits of this retailer are given by

$$\pi_r^i = p_r^i \cdot q - d \cdot q - p_m^i \cdot q. \tag{14}$$

Consider also a manufacturer with constant marginal cost *c*, who provides *q* quantities at price $p_m{}^i$. Then this manufacturer's profits are given by

$$\pi_m^i = p_m^i \cdot q - c \cdot q. \tag{15}$$

[4] A single profit margin is charged over the overall cost of producing and selling a product. This margin is thereafter distributed among the manufacturer and the retailer according to a joint agreement.

Then, a perfectly integrated supply chain can be modeled as a centralized decision maker with profit function:

$$\pi^i = (p_r^i \cdot q - d \cdot q - p_m^i \cdot q) + (p_m^i \cdot q - c \cdot q) \tag{16}$$

$$= (p_r^i - d - c) \cdot q, \tag{17}$$

where the manufacturer's price is irrelevant to the determination of profits since the perfectly vertically integrated firm pays the money to itself. This integrated firm maximizes profits by choosing prices (Mcguire and Staelin, 1983; Choi, 1991; Kadiyali, Chintagunta, and Vilcassim, 2000). The first order condition thus becomes:

$$\frac{\partial \pi^i}{\partial p_r^i} = q + \left(p_r^i - d - c\right)\frac{\partial q}{\partial p_r^i} = 0, \tag{18}$$

$$p_r^i = c + d - q\left(\frac{\partial q}{\partial p_r^i}\right)^{-1}. \tag{19}$$

Equation (19) is the optimal pricing policy for a perfectly integrated firm, which can be written in a more parsimonious way when multiplying the whole equation by p_r^i/q and isolating for p_r^i:

$$p_r^i = (c + d)\left(\frac{\varepsilon}{1+\varepsilon}\right), \tag{20}$$

where ε is the own price elasticity of the product. The retailer price in an integrated firm is therefore a function of the elasticity of demand, and the total marginal cost of producing and marketing the product. Substituting Equation (20) into in Equation (17) yields the profits of the perfectly coordinated supply chain:

$$\pi^i = -q(c + d)\left(\frac{1}{1+\varepsilon}\right). \tag{21}$$

This leads to the conclusion that the perfectly integrated firm will only have positive profits when the denominator in Equation (21) is negative, which is only true for an elastic demand. Interestingly, the pricing policy (Equation 20) results in an infinitely large price as the elasticity approaches -1, and negative prices thereafter. Therefore, this approach seems inadequate for this kind of products.

Considering elastic products and assuming that the manufacturer and the retailer are two different agents, we can obtain the *implied* manufacturer's profits, and the *implied* manufacturer price by dividing the integrated firm's profits according to a predetermined proportion h. Substituting Equation (19) into Equation(17), multiplying by h, and equating to Equation (15) yields:

$$\pi_m^i = -hq^2 \left(\frac{\partial q}{\partial p_r^i}\right)^{-1} = \left(p_m^i - c\right)q, \tag{22}$$

which implies:

$$p_m^i = c - hq \left(\frac{\partial q}{\partial p_r^i}\right)^{-1}. \tag{23}$$

As expected, the manufacturer's price not only depends on the cost of manufacturing, but on the consumer's reaction to the retailer's price as well.

Equations (19) and (23) are the important relationships that we have to consider for our causality analysis. Bessler and Kling (1984) discussed that only stationary series should be considered to avoid spurious results. Indeed, the changes in marginal costs could affect our conclusions. Therefore, taking differences on Equations (19) and (23), and assuming that the consumer demand function has a constant slope at a certain level, we can isolate the strategic behavior from the cost of marketing and production[5]:

$$\Delta p_r^i = -\left(\frac{\partial q}{\partial p_r^i}\right)^{-1} \cdot \Delta q, \tag{24}$$

$$\Delta p_m^i = -h\left(\frac{\partial q}{\partial p_r^i}\right)^{-1} \cdot \Delta q. \tag{25}$$

Moreover, Equations (24) and (25) were obtained by a single decision maker which divides the profits among the supply chain according to the ratio h. This implies that both prices should be jointly determined, and their correlation should be statistically significant. But as we can see through Equations (24) and (25), it is the consumer who ultimately drives both prices. Therefore, this correlation should disappear when conditioning on q. Thus, q should make those prices d-separated, and the following causal structure should be observed: ${p_r}^i \leftarrow q \rightarrow {p_m}^i$. A different conclusion is reached when the supply chain has monopolistic power. The pertinent analysis is deferred to Subsection 4.2.5.

4.2.2 Vertical Nash

Hereafter, consider the existence of independent decision makers. Thus, both the manufacturer and the retailer maximize their profits by choosing their own prices without coordination. Therefore, we can rewrite Equation (14) to obtain the retailer profit function:

$$\pi_r = p_r \cdot q - d \cdot q - p_m \cdot q, \tag{26}$$

[5] It is not unrealistic to assume a constant slope in the demand functions given the small size of the changes in price. This assumption, or constant elasticities, lies at the heart of many demand models (Deaton and Muellbauer, 1980).

where the super index *i* has been removed to indicate the absence of a centralized decision maker. For this reason, the manufacturer price is considered exogenous, leading to the following first order condition:

$$\frac{\partial \pi_r}{\partial p_r} = q + (p_r - d - p_m) \cdot \frac{\partial q}{\partial p_r} = 0, \tag{27}$$

which yields to the following pricing policy:

$$p_r = p_m + d - q \cdot \left(\frac{\partial q}{\partial p_r}\right)^{-1}. \tag{28}$$

An equivalent, and perhaps more intuitive version of this pricing policy, is obtained by multiplying Equation (28) by p_r/q and isolating for p_r , which leads to:

$$p_r = (p_m + d)\left(\frac{\varepsilon}{1+\varepsilon}\right). \tag{29}$$

Equation (29) suggests that the retailer will charge a positive margin, *rm,* over the marginal cost for elastic demands. That is, $\varepsilon/(1+\varepsilon) > 1$ when $\varepsilon < -1$, and thus $rm = p_r - p_m - d > 0$. But Equation (29) also implies that the margin is reduced as the demand becomes more elastic. In particular, a perfect elastic demand yields the no profits condition for a retailer in a perfect competitive market:

$$p_r = \lim_{\varepsilon \to -\infty} \left\{(p_m + d)\left(\frac{\varepsilon}{1+\varepsilon}\right)\right\} = p_m + d. \tag{30}$$

But this pricing policy does not work for inelastic goods. In particular, Equation (30) tends to infinity for unit elastic products, and to zero for perfectly inelastic goods:

$$p_r = \lim_{\varepsilon \to -1^-} \left\{(p_m + d)\left(\frac{\varepsilon}{1+\varepsilon}\right)\right\} = \infty, \tag{31}$$

$$p_r = \lim_{\varepsilon \to 0^-} \left\{(p_m + d)\left(\frac{\varepsilon}{1+\varepsilon}\right)\right\} = 0. \tag{32}$$

Therefore, Equations (31) and (32) strongly suggest that there cannot be an equilibrium of this type when inelastic goods are considered. This leads to the conclusion that different strategies or pricing behavior would be pursued in that scenario.

We can study the manufacturer's pricing policy in the same fashion. Rewriting Equation (15), we can state the manufacturer's profits as Equation (33), the associated first order condition in Equation (34), and the pricing policy in Equation (35):

$$\pi_m = (p_m - c) \cdot q \tag{33}$$

$$\frac{\partial \pi_m}{\partial p_m} = q + (p_m - c) \cdot \frac{\partial q}{\partial p_r}\frac{\partial p_r}{\partial p_m} = 0 \tag{34}$$

$$p_m = c - q \cdot \left(\frac{\partial q}{\partial p_r}\frac{\partial p_r}{\partial p_m}\right)^{-1}. \tag{35}$$

This is perhaps a controversial result. If the retailer and the manufacturer are independent decision makers, then $(\partial p_r)/(\partial p_m)$ equals zero, and no information can be learned from Equation (34). This suggests that the manufacturer cannot ignore the retailer's pricing policy. This is because the consumer makes his decisions based on the retailer's price, and not on the manufacturer's price. However, the manufacturer cannot directly learn the retailer's best response because the choice of prices is assumed to occur simultaneously. Given that the manufacturer knows that the retailer's benefits are described by Equation (26), then a suboptimal response by the retailer could be assumed[6]:

$$p_r = p_m + d + \frac{\pi_r}{q}, \tag{36}$$

which yields:

---

[6] Notice that this is equivalent to assuming a fixed price markup. Moreover, the manufacturer can identify the consumer's preferences and therefore infer the retailer's best response.

$$\frac{\partial p_r}{\partial p_m} = 1. \tag{37}$$

In other words, the manufacturer could expect a one-to-one relationship between his price and the one chosen by the retailer. Substituting Equation (37) into Equation (35) yields the following two optimal pricing policies for the manufacturer, which are equivalent:

$$p_m = c - q \cdot \left(\frac{dq}{dp_r}\right)^{-1}, \tag{38}$$

$$p_m = c - \frac{p_r}{\varepsilon}. \tag{39}$$

Equation (38) demonstrates that the manufacturer's choice of margin, *mm*, has fewer restrictions than the retailer when a retailer's proportional response is assumed. Given a normal product, the manufacturer's margin, $mm = p_m - c$, will be positive, but the elasticity of the product plays a different role. Specifically, Equations (40) respectively demonstrate that the manufacturer's pricing policy for perfectly elastic, unit elastic, and perfectly inelastic products, is not discontinuous as it is in the retailer's problem. Therefore, the manufacturer's pricing policy is to charge just enough to match his marginal costs for perfectly elastic products, and increase his margin as they become more inelastic. This implies that the retailer has greater incentives than the manufacturer to change the supply chain relationship as products become more inelastic.

$$p_m = \lim_{\varepsilon \to -\infty} \left\{c - \frac{p_r}{\varepsilon}\right\} = c,$$

$$p_m = \lim_{\varepsilon \to -1^-} \left\{c - \frac{p_r}{\varepsilon}\right\} = c + p_r, \tag{40}$$

$$p_m = \lim_{\varepsilon \to 0^-} \left\{c - \frac{p_r}{\varepsilon}\right\} = \infty.$$

Moreover, a closer look at Equation (38) reveals that the pricing decision of the manufacturer depends from the quantities purchased by the consumer, which in turn are affected by the retailer's price. This implies that the manufacturer's knowledge of the consumer's consumption should screen off the information transferred through the retailer. We can observe this behavior by reducing the amount of factors that affect the pricing decisions. Thus, assuming that the consumer behavior remains constant, differencing Equations (28) and (38) yields:

$$\Delta p_r = \Delta p_m - \left(\frac{\partial q}{\partial p_r^i}\right)^{-1} \cdot \Delta q, \tag{41}$$

$$\Delta p_m = -\left(\frac{\partial q}{\partial p_r^i}\right)^{-1} \cdot \Delta q. \tag{42}$$

On the other hand, if the retailer's best response is not assumed to be constant, then Equation (35) is the relevant pricing policy for the manufacturer which becomes Equation (43) when written in differences:

$$\Delta p_m = -\left(\frac{\partial q}{\partial p_r}\frac{\partial p_r}{\partial p_m}\right)^{-1} \cdot \Delta q. \tag{43}$$

These results imply the two different causal patterns shown in Figure 2. If the manufacturer expects a one-to-one response from the retailer, then Equations (41) and (42) imply the causal pattern shown in Figure 2.a. This graph represents a scenario where the consumer is the causal source for both the manufacturer and the retailer, and there exists a one-directional causal relationship from the manufacturer to the retailer. However, the structure shown in Figure 2.a cannot be identified from observational data. This is because its associated probability distribution, $\Pr(q,p_r,p_m) = \Pr(q)\ \Pr(p_m|q)\ \Pr(p_r|q,p_m)$, is the same as $\Pr(q,p_r,p_m) = \Pr(p_m)\ \Pr(q|p_m)\ \Pr(p_r|q,p_m)$ which is coupled to an

Figure 2 Vertical Nash Causal Relationships.

a)

Scenario when there is an exogenous constant response from the retailer.

- The consumer causes both prices.
- The manufacturer causes the price of the retailer.

b)

Scenario when there is an exogenous, non constant, and non independent response from the retailer.

- The consumer causes both prices.
- The manufacturer and retailer's prices are simultaneously determined.

identical graph but with a reversed causal relationship between q and p_m. Haigh and Bessler (2004) proved that these two graphs cannot be distinguished from each other because they are *Observationally Equivalent*. That is, they have the same edge structure and converging causal patterns, which render the same statistical loss function metric when the DAGs are estimated. Therefore, the algorithms will not determine the correct causal structure.

If the manufacturer assumes an exogenous but non constant response from the retailer, then equations (41) and (43) describe the equilibrium. Figure 2.b represents this scenario, where the same relationship exists with the consumer, but a bidirectional causal relationship exists between the manufacturer and the retailer. These conclusions are in line with Lee and Staelin's (1997) observation that is important to distinguish the actual pricing policy being implemented, which in this case reveals the manufacturer's decision process.

4.2.3 Manufacturer as a Leader in a Stackelberg Game

In the previous subsection, we discussed the causal patterns that are expected when the manufacturer and the retailer act with no coordination. We also concluded that there are incentives for both agents to abandon this scenario. Stackelberg games are those in which one agent acts as a leader and the other reacts as a follower, and thus the pricing process occurs in two stages. In the first stage, the leader chooses his price. In the second stage, the second agent sees the price chosen by the leader and best responds to that price, and the consumer chooses how much to purchase. Since the leader knows that the second agent will react in such way, then the leader can use the follower's best response as a part of his optimization problem.

In particular, when the manufacturer acts as the Stackelberg leader, the manufacturer sets his price in stage one and knows that the retailer will maximize his profits in stage two by implementing the pricing policy described by Equation (28). We can rewrite Equation (28) to emphasize this process:

$$p_{r,2} = p_{m,1} + d - q_2 \cdot \left(\frac{\partial q_2}{\partial p_{r,2}}\right)^{-1}, \tag{44}$$

where subscripts have been added to highlight that some variables are realized at different stages. Using the same notation, the manufacturer's profits become:

$$\pi_{m,2} = \left(p_{m,1} - c\right) \cdot q_2, \tag{45}$$

and the first order condition is:

$$\frac{\partial \pi_{m,2}}{\partial p_{m,1}} = q_2 + \left(p_{m,1} - c\right) \cdot \frac{\partial q_2}{\partial p_{r,2}} \frac{\partial p_{r,2}}{\partial p_{m,1}} = 0. \tag{46}$$

But now the manufacturer can account for the effect on the retailer's best response described by Equation (44). Therefore, the effect of the manufacturer's price on the retailer's best response is given by:

$$\frac{\partial p_{r,2}}{\partial p_{m,1}} = 1 - \left(\left(\frac{\partial q_2}{\partial p_{r,2}}\right)^{-1} \frac{\partial q_2}{\partial p_{r,2}} \frac{\partial p_{r,2}}{\partial p_{m,1}} - q_2 \left(\frac{\partial q_2}{\partial p_{r,2}}\right)^{-2} \frac{\partial^2 q_2}{\partial p_{r,2} \partial p_{m,1}} \frac{\partial p_{r,2}}{\partial p_{m,1}}\right) = 0, \tag{47}$$

which comes from the derivation of Equation (44) with respect to the manufacturer's price. The second term in the parenthesis disappears because $q_2(.)$ cannot explicitly depend on $p_{m,1}$. For this reason, Equation (47) can be reduced to:

$$\frac{\partial p_{r,2}}{\partial p_{m,1}} = \frac{1}{2},$$

which substituted in the manufacturer's first order condition yields

$$\frac{\partial \pi_{m,2}}{\partial p_{m,1}} = q_2 + \left(p_{m,1} - c\right) \cdot \frac{\partial q_2}{\partial p_{r,2}} \frac{1}{2} = 0, \tag{48}$$

which leads to:

$$p_{m,1} = c - 2q_2 \cdot \left(\frac{\partial q_2}{\partial p_{r,2}}\right)^{-1}, \tag{49}$$

$$= c - \frac{2p_{r,2}}{\varepsilon}. \tag{50}$$

Thereafter,

$$p_{m,1} = \lim_{\varepsilon \to -\infty} \left\{c - \frac{2p_{r,2}}{\varepsilon}\right\} = c,$$

$$p_{m,1} = \lim_{\varepsilon \to -1^-} \left\{c - \frac{2p_{r,2}}{\varepsilon}\right\} = c + 2p_r, \tag{51}$$

$$p_{m,1} = \lim_{\varepsilon \to 0^-} \left\{c - \frac{2p_{r,2}}{\varepsilon}\right\} = \infty.$$

The Nash equilibrium of the manufacturer-Stackelberg game is thus described by Equations (44) and (49). As it was the case for the vertical Nash equilibrium, Equations (51) show that the manufacturer's pricing policy is to charge just enough to match his marginal costs for perfectly elastic products, and increase his margin as they become more inelastic. However, Equation (49) has a larger slope than that in Equation (39), which implies a bigger manufacturer price. Therefore the retailer is obligated to reduce its margin in order to maintain a constant level of sales of a normal good. In other words, given a certain level of demand, the manufacturer is able to extract a larger share of the total supply chain profits.

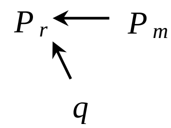
As we did before, we assume that the consumer behavior is constant and we isolate the strategic relationship among the variables by taking differences on the pricing policy functions described by Equations (44) and (49):

$$\Delta p_{r,2} = \Delta p_{m,1} - \left(\frac{\partial q_2}{\partial p_{r,2}}\right)^{-1} \cdot \Delta q_2 \, , \tag{52}$$

$$\Delta p_{m,1} = -2\left(\frac{\partial q_2}{\partial p_{r,2}}\right)^{-1} \cdot \Delta q_2, \tag{53}$$

These results lead to our conclusions. Since the retailer has to choose the price after the manufacturer has chosen his, then the retailer's price can be manipulated by changing the manufacturer's price, but not otherwise. Therefore the retailer's price is caused by the manufacturer's price. Moreover, Equation (52) indicates that the retailer's choice of price is also subject to the consumer's choice of quantities, which are determined in stage two as well. Since the decision of the manufacturer occurs in stage 1, therefore the manufacturer's price and the consumer's quantities are only related

Figure 3 **Manufacturer Stackelberg Causal Relationship.**



The retailer is subject to both the manufacturer's and the consumer's choices when the manufacturer is a Stackelberg leader. This makes the retailer's price a collider.

through Equation (44). This implies that $p_{m,1}$ and q_2 should seem independent because the quantities sought by the manufacturer may not be the actual quantities purchased by the consumer. But conditioning on $p_{r,2}$ should reveal the existing relationship described by equations (49) and (53). These relationships are best described by the *collider* shown in Figure 3.

4.2.4 Retailer as a Leader in a Stackelberg Game

We study the other direction of a Stackelberg game in this subsection. The previous subsection described the equilibrium and consequential causal pattern elicited when the manufacturer acts as a leader. Here, we study the case in which the retailer chooses which price to charge for a product, and the manufacturer best responds to whatever price was chosen by the retailer. For this reason, it is the retailer who can use the manufacturer's best response as part of its maximization process. Rewriting the maximization problems with subscripts in order to emphasize the stage process, the manufacturer's profit functions is:

$$\pi_{m,2} = (p_{m,2} - c) \cdot q_2, \tag{54}$$

and the manufacturer's first order condition and respective pricing policy are:

$$\frac{\partial \pi_{m,2}}{\partial p_{m,2}} = q_2 + (p_{m,2} - c) \cdot \left(\frac{\partial q_2}{\partial p_{r,1}}\right) \cdot \left(\frac{\partial {p_{r,1}}^*}{\partial p_{m,2}}\right) = 0. \tag{55}$$

$$p_{m,2} = c - q_2 \cdot \left(\frac{\partial q_2}{\partial p_{r,1}} \cdot \frac{\partial {p_{r,1}}^*}{\partial p_{m,2}}\right)^{-1}. \tag{56}$$

Notice in the last two equations that an asterisk (*) appears as a superscript on the retailer' price in Equations (55) and (56). This is because this price has already been chosen in stage one. Although this is an exogenous function for the manufacturer, it is not for the retailer. The retailer must anticipate the manufacturer's choice. Thus, substituting the manufacturer's best response into the retailer's profit function, the retailer's profit function becomes:

$$\pi_{r,2} = \left(p_{r,1} - d - \left(c - q_2 \cdot \left(\frac{\partial q_2}{\partial p_{r,1}}\right)^{-1} \cdot \left(\frac{\partial p_{r,1}}{\partial p_{m,2}}\right)^{-1}\right)\right) \cdot q_2, \tag{57}$$

where the asterisk has been removed from the retailer's price in the manufacturer's best response because this is an endogenous decision for the retailer. The associated first order condition is:

$$\frac{\partial \pi_{r,2}}{\partial p_{r,1}} = \left(p_{r,1} - d - c + q_2 \cdot \left(\frac{\partial q_2}{\partial p_{r,1}} \cdot \frac{\partial p_{r,1}}{\partial p_{m,2}}\right)^{-1}\right) \frac{\partial q_2}{\partial p_{r,1}} \tag{58}$$

$$+ \left(1 + \left(\frac{\partial p_{r,1}}{\partial p_{m,2}}\right)^{-1} \cdot \left[1 - q_2 \left(\frac{\partial q_2}{\partial p_{r,1}}\right)^{-2} \frac{\partial^2 q_2}{\partial {p_{r,1}}^2}\right]\right) q_2 = 0 ,$$

where the second term has been reduced since we would expect the cross partial derivative of $p_{r,2}$ with respect to $p_{m,2}$ and $p_{r,2}$ equal to zero. Isolating for the retailer's price we obtain:

$$p_{r,1} = d + c - q_2 \cdot \left(\frac{\partial q_2}{\partial p_{r,1}} \cdot \frac{\partial p_{r,1}}{\partial p_{m,2}}\right)^{-1} \tag{59}$$

$$-\left(\frac{\partial q_2}{\partial p_{r,1}}\right)^{-1}\left(1 + \left(\frac{\partial p_{r,1}}{\partial p_{m,2}}\right)^{-1} \cdot \left[1 - q_2\left(\frac{\partial q_2}{\partial p_{r,1}}\right)^{-2} \frac{\partial^2 q_2}{\partial {p_{r,1}}^2}\right]\right) q_2 \ .$$

Equations (56) and (59) reveal a theoretical contradiction. The pricing policy for the retailer involves solving a partial differential equation, and its solution implies a simultaneous determination of the retailer's and the manufacturer's price. This conflicts with our assumption that the retailer chooses price first and lets the manufacturer pick its price as a best response. Clearly, the manufacturer's actual choice of price would not affect the retailer's decision because it has already taken place. Moreover, the retailer cannot ignore the manufacturer's actions since Equations (56) and (59) would break down. Nevertheless, the manufacturer can infer an "already happened best response" from the retailer by studying the consumer behavior, and thus the manufacturer's decision has to be affected by his perception of the demand curve. On the other hand, the retailer can circumvent his problem by signaling a constant response to the manufacturer's choice. Then:

$$\frac{\partial p_{r,1}}{\partial p_{m,2}} = k, \tag{60}$$

and Equations (56) and (59) are respectively reduced to:

$$p_{m,2} = c - q_2 \cdot \left(\frac{\partial q_2}{\partial p_{r,1}} \cdot k\right)^{-1}. \tag{61}$$

$$p_{r,1} = d + c - q_2 \cdot \left(k \frac{\partial q_2}{\partial p_{r,1}}\right)^{-1} \tag{62}$$

$$-q_2\left(\frac{\partial q_2}{\partial p_{r,1}}\right)^{-1}\left(1 + k^{-1} - q_2\left(k \frac{\partial q_2}{\partial p_{r,1}}\right)^{-2} \frac{\partial^2 q_2}{\partial {p_{r,1}}^2}\right).$$

But now the retailer has a second choice variable. The optimal choice of k is given by maximizing the new profit function with respect to k:

$$\pi_{r,2} = \left(p_{r,1} - d - \left(c - q_2 \cdot \left(k \frac{\partial q_2}{\partial p_{r,1}}\right)^{-1}\right)\right) \cdot q_2, \tag{63}$$

which has the following first order condition which also is a partial differential equation:

$$\frac{\partial \pi_{r,2}}{\partial k} = q_2 \frac{\partial p_{r,1}}{\partial k} - {q_2}^2 \frac{\partial q_2}{\partial p_{r,1}} \frac{1}{k^2} = 0 \tag{64}$$

Equations (56) and (59) show that although the manufacturer is the second mover, more information other than the price is needed from the retailer. In turn, the retailer must provide a signal of his pricing policy in order to be time consistent. Nevertheless, repetition of the game ensures that k reveals the actual pricing policy pursued by the retailer. Otherwise, no equilibrium can be achieved since the manufacturer would have incentives to deviate. Taking differences on equations (61) and (62) yields:

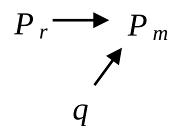
$$\Delta p_{m,2} = -\Delta q_2 \cdot \left(\frac{\partial q_2}{\partial p_{r,1}} \cdot k\right)^{-1}, \tag{65}$$

$$\Delta p_{r,1} = -\Delta q_2 \cdot \left(k \frac{\partial q_2}{\partial p_{r,1}}\right)^{-1} - \Delta q_2 \left(\frac{\partial q_2}{\partial p_{r,1}}\right)^{-1} (1 + k^{-1}) \tag{66}$$

$$+\Delta {q_2}^2 \left(\frac{\partial q_2}{\partial p_{r,1}}\right)^{-3} \left(k^{-2} \frac{\partial^2 q_2}{\partial {p_{r,1}}^2}\right) = 0 \ .$$

Equation (65) shows that the manufacturer's decision can be manipulated by both the consumer's choice and the retailer's signal, and equation (66) shows that the retailer's price can only be influenced by the consumer's choice. Nevertheless, the correlation between the retailer's choice of price and the consumed quantities at any

Figure 4 **Retailer Stackelberg Causal Relationship.**



The manufacturer is subject to both the retailer's and the consumer's choices when the retailer is a Stackelberg leader. This makes the manufacturer's price a collider.

given time should seem uncorrelated. This is because the retailer's choice of price occurs in a previous stage. However, this is a theoretical impossibility. Thereafter, the retailer's relationship with the consumer should be observed when conditioning on the manufacturer's price. This implies that the manufacturer's price must be a collider in a retailer-Stackelberg game and the causal structure in Figure 4 should be elicited.

4.2.5 Monopoly

A special case of imperfect competition deserves attention. The retailer has the ability to exert monopolistic power when the consumer is unable to find substitutes. For this reason, the retailer is able to choose the level of quantities to introduced, and lets the consumer clear the market by "choosing" the price. Moreover, the manufacturer will provide the quantities set by the retailer. For this reason, the retailer can be thought as a Stackelberg leader, where both the consumer and the manufacturer are the Stackelberg followers. Given a monotonic consumer demand function, $q = D(p)$, then the consumer's best response is the inverse demand function $p = D^{-1}(q)$. Therefore, monopolistic retailer maximizes his profits given by:

$$\pi_r^m = D^{-1}(q_1) \cdot q_1 - d \cdot q_1 - p_{m,2} \cdot q_1, \tag{67}$$

where subscripts have been used to emphasize the timing of the decision process. The first order condition becomes:

$$\frac{\partial \pi_r^m}{\partial q} = q_1 \cdot \frac{\partial D^{-1}(q_1)}{\partial q_1} + D^{-1}(q_1) - d - p_{m,2} = 0, \tag{68}$$

Isolating:

$$q_1 = \left(d + p_{m,2} - D^{-1}(q_1)\right)\left(\frac{\partial D^{-1}(q_1)}{\partial q_1}\right)^{-1}. \tag{69}$$

The price set by the retailer has to equal the consumer's best response which is given by the consumer's demand function:

$$p_{r,1}^m = D^{-1}(q_1), \tag{70}$$

where q_1 is given by Equation (69). The consumer, in turn, purchases those quantities:

$$q_2 = D\left(p_{r,1}^m\right) = q_1. \tag{71}$$

On the other hand, the manufacturer chooses his price by maximizing:

$$\pi_m = (p_{m,2} - c)q_1, \tag{72}$$

which has the following associated first order condition:

$$\frac{\partial \pi_m}{\partial p_{m,2}} = q_1 + (p_{m,2} - c)\frac{\partial q_2}{\partial p_{m,2}}. \tag{73}$$

Notice that Equation (73) is fundamentally different to all the problems that we solved before. Here, the quantities introduced to the market are chosen by the retailer and therefore the manufacturer has a direct effect on this decision through Equation (69). Solving for $p_{m,2}$ yields:

$$p_{m,2} = c - q_1\left(\frac{\partial q_2}{\partial p_{m,2}}\right)^{-1} = \frac{c - d + D^{-1}(q_1)}{2}. \tag{74}$$

These results imply that it is the retailer who acts as a causal source of information for both the manufacturer and the retailer, $q \leftarrow p_r \rightarrow p_m$. This is because both the consumer and the manufacturer best respond to the retailer's decision.

But a different conclusion is reached when there is monopolistic power along with coordination of the supply chain. If that is the case, then the problem can be modeled as a centralized decision maker with profit function:

$$\pi^{m,i} = \left(D^{-1}(q_1) \cdot q_1 - d \cdot q_1 - p_m^i \cdot q_1\right) + \left(p_m^i \cdot q_1 - c \cdot q_1\right) \tag{75}$$

$$= (D^{-1}(q_1) - d - c) \cdot q_1,$$

which implies the following supply policy:

$$q_1 = \left(d + c - D^{-1}(q_1)\right) \cdot \left(\frac{\partial D^{-1}(q_1)}{\partial q_1}\right)^{-1}. \tag{76}$$

As we did before, the price set by the retailer has to equal the consumer's best response which is given by the consumer's demand function:

$$p_r^{m,i} = D^{-1}(q_1), \tag{77}$$

where q_1 is given by Equation (76). The consumer reacts by purchasing those quantities:

$$q_2 = D\left(p_r^{m,i}\right) = q_1. \tag{78}$$

However, this time the manufacturer's price has to be obtained by equally dividing the supply chain's total profits. This implies that the manufacturer's price and the retailer's price should look unrelated, unless the quantities purchased are observed. Since the consumer is clearing the market, then the consumer acts as a collider: $p_r \rightarrow q \leftarrow p_m$. The relationships discussed in this subsection are summarized in Figure 5.

Figure 5 Retailer's Monopolistic Causal Relationship.

a) b)

The manufacturer and the consumer are both subject to the retailer's choice of quantities introduced in the market.

Scenario when the manufacturer and the retailer are integrated in the supply chain, but there is monopolistic power

We described in this subsection the causal patterns that can be expected from the different strategic relationships that occur among horizontal competition and within the supply chain. We now proceed to review how Directed Graphs can be used to elicit causal relationships among variables, and thus to imply the equilibriums described above.

4.3 DIRECTED GRAPHS

The previous subsection discussed the relationship between the determination of Nash equilibriums and causality. We concluded that Bertrand games lead to independent movements unless there is collusion in a repeated game. In such cases, the use of threats and punishments would induce a simultaneous determination of prices; Cournot games result in jointly determination of quantities and prices (i.e. $q_i \leftrightarrow q_j$; $p_i \leftrightarrow p_j$); and, Stackelberg leaders would directly affect followers' quantities, but quantities supplied by Stackelberg followers should not directly *cause* the Stackelberg leader's optimal choice (i.e. $q_s \rightarrow q_i$ for all $i \neq s$, where s represents the Stackelberg leader). These causal relationships can be elicited through causality inference models. We discuss in this subsection the machinery of causal models and the different algorithms used are outlined (Swanson and Granger, 1997; Pearl, 1995; Sprites *et al*., 1993; Richardson, 1996). Appendix A presents each of the algorithms in detail.

Let X_1 cause X_2 which causes X_3. Or, in *Causality* notation, $X_1 \rightarrow X_2 \rightarrow X_3$. Then, the joint probability distribution of these three variables can be factored as:

$$\Pr[x_1, x_2, x_3] = \Pr[x_1] \cdot \Pr[x_2 \mid x_1] \cdot \Pr[x_3 \mid x_2] \qquad (79)$$

which is known as a causal Markov condition. This implies that conditioning on a causal variable yields the full probability distribution that generates a random variable, a behavior usually observed in prices. Generalizing,

$$\Pr[x_1, x_1, \dots, x_n] = \prod_{i=1}^{n} \Pr[x_i | pa_i] \qquad (80)$$

where $\Pr[x_1, x_2, ..., x_n]$ represents the joint probability of variables $\{X_1, X_2, ..., X_n\}$ and pa_i represents the realization of some subset of variables that precede or cause X_i. Assuming the information set is both causally sufficient, i.e. there are no omitted variables that cause two included variables in the study (Sprites, 1993), and *faithful*, meaning that the relationship between two variables is not cancelled by the interactions of other variables, then Equation (79) can be used to represent causal relationships among the variables.

Three causal structures must be defined to illustrate. These are presented in Figure 6. The first one is the *Collider*: $X \rightarrow Y \leftarrow Z$. In this structure, information flows from X and Z towards Y, but not from X to Z nor Z to X. Y is the *"collider"*.

Figure 6 Basic Causal Structures in Search Algorithms.

$X \rightarrow Y \leftarrow Z$	$X \leftarrow Y \rightarrow Z$	$X \rightarrow Y \rightarrow Z$
a) Collider	b) Causal Fork	c) Causal Chain

Interestingly, if *Y* is not observed, then *X* and *Z* would be unrelated, but given that *Y* is observed, the relationship between *X* and *Z* becomes visible. A good analogy is to think of a family. A couple seems unrelated until it is known they have a child. In a more technical language, conditioning on *Y* opens a communication path between *X* and *Z*. *X* and *Z* are said to be d-separated (Pearl, 1995), unconditionally.

The second causal structure to be defined is the *causal fork*: $X \leftarrow Y \rightarrow Z$. Here *Y* causes both *X* and *Z*. Therefore, the correlation between *X* and *Z* will be nonzero since they share a common cause. But conditioning on *Y* makes the relationship between *X* and *Z* disappear. Technically, *X* and *Z* are d-connected but conditioning on *Y* makes them d-separated.

Finally, we have the *causal chain*: $X \rightarrow Y \rightarrow Z$. Now *X* causes *Y*, which in turn causes *Z*. The unconditional relationship between *X* and *Z* is non zero. But when conditioning on *Y*, the relationship disappears. It becomes d-separated.

More formally, in a set of variables {*X*, *Y*, *Z*}, the correlation between *X* and *Z* conditional on *Y* is zero if and only if *X* and *Z* are d-separated given *Y*. It is the d-separation concept that allows the creation of searching algorithms that ultimately generate the Directed Graphs. Two kinds of graphs must be distinguished; Directed Acyclic Graphs (DAGs) do not allow for a causal flow that eventually returns to the variable where the information first originated. On the other hand, Directed Cyclic Graphs (DCG) do allow for such paths to exist which sometimes imply simultaneous determination. Both the PC algorithm (Sprites *et al*., 1993) and Moneta's modification to the PC algorithm (Moneta, 2008) fall in the former category, whereas the Richardson's algorithm (Richardson, 1996) falls in the later. We created a truncated version of all these algorithms to analyze the causality of only three variables for the case of vertical competition. On the other hand, we used Tetrad 4.3.9-9 to study the horizontal relationships with the complete version of the Richardson's algorithm. In the following subsections, we explain the different algorithms and describe precisely how they were implemented. The complete algorithms are included in the Appendix A with changes in notation from their original versions to conform to the notation used in this dissertation.

4.3.1 PC Algorithm

This algorithm serves as the heart of all the procedures used in this research. Sprites *et al*. (1993), built the d-separation concept discussed above into the *PC algorithm,* which is programmed in the software TETRAD II (Scheines *et al*., 1994) and is used to generate the graphs by iteratively analyzing the correlation between two variables conditioned on another variable(s).

The algorithm initially connects all the variables with undirected edges (arrows with no heads). Then, with the aid of a Fisher significance test, edges are sequentially removed based on conditional or unconditional correlation determined to be statistically non significant. When a variable is removed, new relationships may be uncovered between the remaining variables. Therefore, the remaining edges are checked once again and direction on causality can be inferred. For instance, suppose that the edge between *X* and *Z* has been removed unconditionally on *Y*. Then the relationship *X* – *Y* – *Z* can be directed as $X \rightarrow Y \leftarrow Z$. If the edge between *X* and *Z* was removed because the correlation between them was zero when conditioned on Y, then the causality is undetermined between $X \leftarrow Y \rightarrow Z$ and $X \rightarrow Y \rightarrow Z$ since the correlation structure of causal forks is the same that that of causal chains. In such a case, PC leaves the edges undirected and another variable is required to resolve the ambiguity. The PC algorithm follows:

1. Create a complete undirected graph. That is, every variable to be studied is connected to every other variable with undirected edges.
2. Sequentially remove edges based on vanishing correlation or conditional correlation. If *i*, *j*, *k* are normally distributed and $r(i, j \mid k)$ is the sample conditional correlation of *i* and *j* given *k*, then $z(\rho(i, j \mid k), n) - z(r(i, j \mid k), n)$ is standard normal, where $\rho(i, j \mid k)$ is the correlation between variable *i* and *j* conditional on variable *k*, and |*k*| is the number of variables that we are conditioning on. $z(\cdot)$ is the Fisher statistic given by:

$$z(\rho(i,j|k),n) = \left[\frac{1}{2}\sqrt{n-|k|-3}\right] ln\left\{\frac{1+\rho(i,j|k)}{1-\rho(i,j|k)}\right\}. \quad (81)$$

3. Direct the surviving edges. Some variable(s) were used to condition the relationship in order to remove the edges. Such variable(s) are termed the *sepset* of the variables whose edge has been removed. The PC algorithm

directs an edge between X and Y into variable Z if Z is not in the *sepset* of X and Y.

4. The last part of the algorithm is intended to detect more edges and avoid the existence of cycles. But Scheines *et al*. (1994) found that this part reduces the reliability of the results and therefore are not implemented here.

The algorithm may make mistakes of two types: including or omitting an edge, and direction of an edge. In order to achieve accurate results, the significance level should be decreased as the sample size is increased (Spirtes, Glymour, and Scheines (2000)[7]. The complete algorithm has been included in Appendix A.

4.3.2 Moneta's Algorithm

Moneta (2008) argued that the *PC Algorithm* made mistakes too often, and therefore suggested that it was necessary to build a more conservative algorithm. He introduced a slightly modified version customized to analyze the innovations of a VAR. He modified the algorithm in two important ways:

1. The notion of *Sepset* is modified. Now a *Sepset* of size *n* includes all possible combinations of the *n* conditioning variables. For instance, if we are analyzing a set with 6 variables, the correlation between two of them is tested conditioned on groups of the remaining variables. That is, *sepset 1* includes 5 one-member sets (the empty set plus a 5 sets, each with one of the remaining

[7] A significance level of 0.20 is recommended at sample sizes less than 100. A significance level of 0.10 is advised if the sample size is between 100 and 300.

4 variables). *Sepset 2* includes 6 sets representing all possible combinations of 2 variables out of the 4 left for conditioning. And so forth.

2. The Wald statistic is used instead of the Fisher statistic.

For this research, we used a truncated version of the PC algorithm and we also adopted Moneta's redefinition of a *Sepset* without using a Wald statistic for this analysis because we are not analyzing VAR innovations. Appendix A presents the Moneta's modified version of the PC Algorithm provided by Sprites *et al*. (2000: 84-85), where Moneta's modifications appear in bold.

4.3.3 Richardson's Algorithm for Cyclic Graphs

We have only considered Directed Acyclic Graphs as opposed to Cyclic Graphs, which allow for a causal flow that eventually returns to the variable where the information first originated. Although Cyclic graphs are fundamentally different from Acyclic graphs, they both rely on d-separation, which Richardson (1996) used in an algorithm that allows to identify cyclic causal structures.

Richardson's algorithm relies also on the same two assumptions as PC. If the probability distribution satisfies the Faithfulness and the Markov conditions, then there exists a set of equivalent graphs, Markov Equivalent, that can represent that probability distribution and have the same d-separation relations. The algorithm searches those features that are common to all the graphs in that set, and the final graph is generated. This graph is called a Partial Ancestral Graph (PAG) and consists of edges that connect all the vertices (or variables). Each edge is allowed to have two edge-endpoints from the

Figure 7 Partial Ancestral Graphs. (Richardson, 1996)

Ψ is a PAG for Directed Cyclic Graph *G* with vertex set **V**, if and only if
(i) There is an edge between A and B in Ψ if and only if A and B are d-connected in *G* given all subsets **W** ⊆ **V**\{A,B}.
(ii) If there is an edge in Ψ out of A (not necessarily into B), A–* B, then A is an ancestor of B in every graph in Equiv(G).
(iii) If there is an edge in Ψ into B, A→B, then in every graph in Equiv(G), B is not an ancestor of A.
(iv) If there is an underlining A*—*B*—*C in Ψ then B is an ancestor of (at least one of) A or C in every graph in Equiv(G).
(v) If there is an edge from A to B, and from C to B, (A → B ← C), then the arrow heads at B in Ψ are joined by dotted underlining, thus A → B ← C, only if in every graph in Equiv(*G*), B is not a descendant of a common child of A and C.
(vi) Any edge endpoint not marked in one of the above ways is left with a small circle thus: o—*.

---

set {°,–, >} indicating respectively whether nothing can be inferred about the source and direction of the information; the information flows out of a variable, or into it. Figure 7 presents the exact definition of a Partial Ancestral Graph as defined by Richardson (1996).

Although this algorithm would allow to identify simultaneous determination of prices (or perfect coordination between the supplier and the retailer), there is not much gain when considering only three variables. Unfortunately, only one of those steps adds more information than the *PC* algorithm. That step was included in our research and allowed us to also direct single edges instead of just colliders. Appendix A presents the complete algorithm which consists of seven steps. That version was used to analyze the relationship among the manufacturers in the case study provided in Section 5. The truncated version required for the analysis of only three variables is described below:

1. Create a complete undirected graph. Every variable to be studied is connected to every other variable with undirected edges.

2. Sequentially remove edges based on vanishing correlation or conditional correlation. If *i, j, k* are normally distributed and $r(i, j \mid k)$ is the sample conditional correlation of *i* and *j* given *k*, then $z(\rho(i, j \mid k), n) - z(r(i, j \mid k), n)$ is standard normal, where $\rho(i, j \mid k)$ is the correlation between variable *i* and *j* conditional on variable *k*, and $|k|$ is the number of variables that we are conditioning on. $z(\cdot)$ is the Fisher statistic given by Equation (81).

3. Direct the surviving edges. Some variable(s) were used to condition the relationship in order to remove the edges. Such variable(s) are termed the *sepset* of the variables whose edge has been removed. The *PC Algorithm* directs an edge between X and Y into variable Z if Z is not in the *sepset* of X and Y.

4. When considering three variables where only two of them are connected with an edge, then that edge can be directed. Consider X, Y, and Z where only Y and Z are connected by an edge. Then Z → Y if Y does not belong to *Sepset(X, Z)* and X is independent from Y when conditioning on *Sepset(X, Z).*

4.3.4 Customized Algorithm

Since we are using Fisher's statistic (Equation 81) for all our tests, then the differences between the three previous algorithms is very subtle. Specifically, the *Sepsets* in the *PC Algorithm* are a subset of the *Sepsets* defined by Moneta (2008). Moreover, the *PC Algorithm* is a truncated version of R*ichardson's Algorithm* when only three variables are being considered. Therefore, the analysis of horizontal competition among manufacturers was conducted with the full version of Richardson's Cyclical Graphs. However, both *Richardson's* and *Moneta's* algorithms were used whenever only three variables were considered in the analysis of vertical competition. For this reason, our customized algorithm starts by computing all the possible correlations and partial correlations and testing them for significance to generate a large matrix that contains all Moneta's *Sepsets*. Then, the algorithm continues as Richardson's but changes the *Sepset* definition and it is truncated even further when Moneta's results are sought. The implemented algorithm follows whereas the complete algorithm is presented in the Appendix A:

1. For all pairs of vertices (y_{it}, y_{jt}), test unconditional and conditional correlations on all possible combinations of the remaining variables. Record results in a *Matrix of Sepsets*. If *i, j, k* are normally distributed and $r(i, j \mid k)$ is the sample conditional correlation of *i* and *j* given *k*, then $z(\rho(i, j \mid k), n) - z(r(i, j \mid k), n)$ is standard normal, where $\rho(i, j \mid k)$ is the correlation between variable *i* and *j* conditional on variable *k*, and |*k*| is the

number of variables that we are conditioning on. $z(\cdot)$ is the Fisher statistic given by Equation (81).

2. Create a complete undirected graph. That is, every variable to be studied is connected to every other variable with undirected edges.
3. Sequentially remove edges based on vanishing correlation or conditional correlation.
4. Direct the surviving edges. Some variable(s) were used to condition the relationship in order to remove the edges. Such variable(s) are termed the *sepset* of the variables whose edge has been removed. The algorithm directs an edge between X and Y into variable Z if Z is not in the *sepset* of X and Y. The definition of *sepset* is allowed to switch from its classical definition to Moneta's depending on which analysis is being performed.
5. When considering three variables where only two of them are connected with an edge, then that edge can be directed. Consider X, Y, and Z where only Y and Z are connected by an edge. Then Z → Y if Y does not belong to *Sepset(X, Z)* and X is independent from Y when conditioning on *Sepset(X, Z).* This step is not performed if the Moneta's definition of a *Sepset* is used.

The theoretical framework used to analyze strategic behavior between competing firms is described in Section 4. Then, the models of causality and the differences among the various algorithms used, both cyclic and acyclic, were described. We concluded by presenting the algorithm actually used in this dissertation. This algorithm was written in Visual Basic and into a spreadsheet. This enabled us to automate the analysis of 96 different manufacturers. However, the analysis of competition among manufacturers was conducted with the use of Tetrad IV. We used the presented theoretical framework and models of causality to empirically elicit the presence of strategic behavior within a supply chain, and among its manufacturers. This analysis is described in the following section.

5. CASE STUDY

It was described in detail in Section 4 how strategic interactions between competitors can be modeled through the use of game theory and elicited with the use of causal models. In particular, we described in Subsection 4.1 how Cournot, Stackelberg, and Bertrand games of horizontal competition are usually analyzed for Nash equilibriums, and how the mathematical solutions imply dependencies that can be elicited through causal inference. This framework was extended to study the (vertical) relationship between a retailer and its suppliers. Thus, we described in Subsection 4.2 the causal patterns associated to a perfect coordination of the supply chain, vertical Nash, Stackelberg leadership, and how the causal structures change in the presence of monopolistic competition by the retailer.

We provide a case study in this section. The data used for the empirical study, and how it was manipulated, is explained in the following subsection. We discuss in Subsection 5.2 the possible effects of aggregation. The analysis of the empirical results is offered in Subsection 5.3, and conclusions are offered in Subsection 5.4.

5.1 DATA

As Slade (1992) explained, highly disaggregated data is required for the type of study defined here. For this reason, we used the publicly available Dominick's database, property of the James M. Kilts Center, University of Chicago Booth School of Business (Table 1). This is the same dataset that Kadiyali, Chintagunta, and Vilcassim (2000)

Table 1 **Dominick's Finer Foods Database Categories. (James M. Kilts Center, University of Chicago Booth School of Business)**

Analgesics	Cereals	Frozen Dinners	Shampoos
Bath Soap	Cheeses	Frozen Entrees	Snack Crackers
Bathroom Tissues	Cigarettes	Frozen Juices	Soaps
Beer	Cookies	Grooming Products	Soft Drinks
Bottled Juices	Crackers	Laundry Detergents	Toothbrushes
Canned Soup	Dish Detergent	Oatmeal	Toothpastes
Canned Tuna	Fabric Softeners	Paper Towels	
Category	Front-end-candies	Refrigerated Juices	

Note. Each category contains several products.

used to analyze the vertical relationship between manufacturers and retailers by estimating a system of equations. The Dominick's database covers store-level scanner data collected in Chicago, Illinois, at Dominick's Finer Foods during the period from 1989 to 1994. It includes the weekly sales of more than 3500 UPCs which are classified in 30 categories. The dataset contains the number of units sold, number of items bundled together, price, and gross margin for each of these items.

Following our analytical discussion in Section 4, the triplet of variables, $\{P_{r,i}, P_{m,i}, q_i\}$, was created, where $P_{r,i}$ is the retailer's price of product i, $P_{m,i}$ refers to the manufacturer's price of product i, and q_i are the quantities purchased by the retailer. Only the sales of store number 123 were analyzed to minimize our exposure to price discrimination implemented in different geographic zones. This store was chosen arbitrarily but within the metropolitan area.

Unfortunately, no information was recorded unless a product was sold. This led to the presence of several missing observations in the dataset which limited the usefulness of some of the products. This inconvenience was addressed by two different avenues. First, we only analyzed those categories which were believed to have the most consumption. These categories are Beer, Cheese, Soft Drinks, Toilet Paper, and Tuna. Second, we aggregated those purchased quantities which were provided by the same manufacturer and calculated the average retailer and manufacturer prices per unit (e.g. dollars per ounces). Given that aggregation may induce further problems to the empirical analysis, a thorough discussion is provided in the following subsection, while the results of our case study are presented in Subsection 5.3.

Moreover, Dominick's database attaches a different manufacturer identification number to some products which were supplied by the same companies. As discussed in Dhar *et al.* (2005), the same manufacturer can provide several different brands, while each of the brands may be marketed differently. For this reason, we did not merge those categories. Indeed, we found evidence of a Manufacturer-Stackelberg for some products and a Retailer-Stackelberg for others, even though they were provided by the same company.

5.2 NOTES ON AGGREGATION

We have discussed that aggregation of the data was required to deal with missing observations. Certainly, the Dominick's database has no entries if a product was not sold during a week, resulting in a very low number of usable series. This is not a major concern when computing the correlations which lie at the heart of all of the algorithms used in causality analysis. However, correlations based on a small number of observations are not desired. This is because asymptotical statistical tests require large amounts of data to consider their results valid. Hence, given that a manufacturer may pursue similar strategies for some products, an aggregated product may be used to generate the triplets discussed in Section 4.2. Therefore, the triplet $\{q_i, P_{r,i}, P_{m,i},\}$ consists of the total quantities provided (and sold) by the same manufacturer regardless of the brand, the average price per unit charged by the retailer, and the average price per unit received by the manufacturer. This type of aggregation should not mask the true underlying behavior of the firms since we are aggregating over different products provided and marketed by the same firm to a single retailer. This should be true as long as each of the brands is subject to the same economic environment, and the expected returns provided by each brand are proportional to each other (Day, 1963).

Yet, it has been widely discussed that behavior at the individual level cannot always be expected to hold at the aggregated level (Deaton and Muellbauer, 1980). Although we are not analyzing a consumer demand problem, it is not obvious whether the real causal pattern for a specific brand will be elicited when several products have been aggregated together. In this context, the demand literature addresses two types of aggregation. The first one studies aggregation across different commodities and intertemporal choices through the *composite commodity theorem* and its variations (Leontief, 1936; Theil 1956, 1971; Lewbel, 1996), as well as *separability* and *two stage budgeting* (Leontief, 1947; Deaton and Muellbauer, 1980). However, our aggregation takes place over weekly observations of different versions of the same commodity (e.g. coke in cans versus coke in bottle). Thus this is not a major concern. Nevertheless, the second type of aggregation deals with grouping over consumers. This directly affects our study and a more detailed discussion is provided in subsection 5.2.2.

For these reasons, we studied the effects of aggregation on the algorithms used to elicit causal relationships in order to analyze whether aggregation may mask the true causal relationship between a manufacturer and a retailer. Each of these algorithms search for causal patterns by testing d-separation between variables with the use of the Fisher statistic shown in Equation (81). Therefore, we can reach our conclusions by determining whether the test statistic is increased (reduced) by aggregation and, as a result, gets rejected (not rejected) more than usual. Any changes on the test statistic would lead to the elicitation of a causal pattern different than that of the disaggregated data. We now proceed to describe the mathematical approach to this problem which results in unclear conclusions. Thereafter, we move forward by addressing the same problem with stochastic simulation where we conclude that aggregation has no effects on our analysis if the triplets to be aggregated have the same correlation matrix and are independent from each other. However, the correlations between the retailer's price and the wholesale prices of the different brands provided by the same manufacturer are not constant. We discuss the theoretical implications of the findings.

5.2.1 Mathematical Approach

There are only two aggregation rules used in this dissertation, although other rules may be studied[8]. The first one consists on computing simple averages for both the

[8] Nayga and Capps (1994) suggest to weight prices by quantity shares within an aggregated group.

retailer's and the manufacturer's price; $\bar{p}_r = \sum_{i=1}^{I} p_{r,i}$ and $\bar{p}_m = \sum_{i=1}^{I} \bar{p}_{m,i}$ respectively, where the index *i* denotes the different brands to be aggregated because they are provided by the same manufacturer. The other aggregation rule consists on adding all the units sold by the retailer, regardless of the brand, but provided by the same manufacturer; $\tilde{q} = \sum_{i=1}^{I} q_i$. Generalizing, consider K different types of variables. Define $\boldsymbol{x}_k$ a vector of realizations of the k^{th} type variable with dimension Tx1. Then, the squared partial correlation between $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ conditioned on $\boldsymbol{x}_3$ is given by:

$$\rho^2(\boldsymbol{x}_1, \boldsymbol{x}_2|\boldsymbol{x}_3) = \frac{(\boldsymbol{x}_1^{*\prime}\boldsymbol{x}_2^*)^2}{(\boldsymbol{x}_1^{*\prime}\boldsymbol{x}_1^*)(\boldsymbol{x}_2^{*\prime}\boldsymbol{x}_2^*)}, \tag{82}$$

where $\boldsymbol{x}_1^*$ and $\boldsymbol{x}_2^*$ are the residuals obtained from regressing $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ on $\boldsymbol{x}_3$ and a constant. Defining $\boldsymbol{W} = [\boldsymbol{\iota}\ \boldsymbol{x}_3]$ as the matrix composed by a vector of ones, $\boldsymbol{\iota}$, of dimension *T*x1, and $\boldsymbol{x}_3$, then the residual maker $\boldsymbol{M} = \boldsymbol{I} - \boldsymbol{W}(\boldsymbol{W}'\boldsymbol{W})^{-1}\boldsymbol{W}'$ can be used to rewrite Equation (82) as:

$$\rho^2(\boldsymbol{x}_1, \boldsymbol{x}_2|\boldsymbol{x}_3) = \frac{(\boldsymbol{x}_1'\boldsymbol{M}\boldsymbol{x}_2)^2}{(\boldsymbol{x}_1'\boldsymbol{M}\boldsymbol{x}_1)(\boldsymbol{x}_2'\boldsymbol{M}\boldsymbol{x}_2)}. \tag{83}$$

which has been reduced since $\boldsymbol{M}$ is symmetric and idempotent. Equation (83) presents the partial correlation coefficient of disaggregated data. Now consider the matrix $\boldsymbol{X}_k$ of dimension *T*x*I* in Equation (84) which represents *I* independent and identically distributed vectors, $\boldsymbol{x}_{i,k}$, of dimension 1x*T* of the k^{th} type:

$$\boldsymbol{X}_k = \begin{bmatrix} x_{1,1,k} & \cdots & x_{1,I,k} \\ \vdots & \ddots & \vdots \\ x_{T,1\ k} & \cdots & x_{T,I,k} \end{bmatrix}. \tag{84}$$

Then, the *I* variables can be aggregated with averages by post-multiplying $\boldsymbol{X}_k$ times the unitary vector, **1**, of dimension *I*x1, and dividing by *I*, which is the level of aggregation:

$$\overline{\boldsymbol{X}}_k = \frac{1}{I}\boldsymbol{X}_k\boldsymbol{1} = \frac{1}{I}\begin{bmatrix}\sum_i^I x_{1,i,k}\\ \vdots \\ \sum_i^I x_{T,i,k}\end{bmatrix}. \quad (85)$$

On the other hand, aggregation through summation is achieved by multiplying Equation (85) times the level of aggregation *I*:

$$I\overline{\boldsymbol{X}}_k = \boldsymbol{X}_k\boldsymbol{1} = \begin{bmatrix}\sum_i^I x_{1,i,k}\\ \vdots \\ \sum_i^I x_{T,i,k}\end{bmatrix}. \quad (86)$$

Moreover, define

$$\boldsymbol{W_A} = [\boldsymbol{\iota}\ \overline{\boldsymbol{X}}_k], \quad (87)$$

and

$$\boldsymbol{W_B} = [\boldsymbol{\iota}\ I\overline{\boldsymbol{X}}_k] \quad (88)$$

as the matrices composed by a vector of ones, $\boldsymbol{\iota}$, of dimension *T*x1, and the *I* variables of the k^{th} type aggregated either by averaging as in Equation (85), or by summation as in Equation (86). Then, the residual maker matrices associated to Equations (87) and (88) respectively are:

$$\boldsymbol{M_A} = \boldsymbol{I} - \boldsymbol{W_A}(\boldsymbol{W_A}'\boldsymbol{W_A})^{-1}\boldsymbol{W_A}' \quad (89)$$

and

$$\boldsymbol{M_B} = \boldsymbol{I} - \boldsymbol{W_B}(\boldsymbol{W_B}'\boldsymbol{W_B})^{-1}\boldsymbol{W_B}'\ . \quad (90)$$

Thus, the squared partial correlation between $\overline{\boldsymbol{X}}_l$ and $\overline{\boldsymbol{X}}_m$ conditioned on $\overline{\boldsymbol{X}}_k$ is:

$$\rho^2(\overline{\boldsymbol{X}}_l, \overline{\boldsymbol{X}}_m | \overline{\boldsymbol{X}}_k) = \frac{(\boldsymbol{1}'\boldsymbol{X}_l{}'\boldsymbol{M_A}\boldsymbol{X}_m\boldsymbol{1})^{\boldsymbol{2}}}{\boldsymbol{1}'\boldsymbol{X}_l{}'\boldsymbol{M_A}\boldsymbol{X}_l\boldsymbol{1}\boldsymbol{1}'\boldsymbol{X}_m{}'\boldsymbol{M_A}\boldsymbol{X}_m\boldsymbol{1}}\,, \quad (91)$$

which has been reduced since $\boldsymbol{M_A}$ is symmetric and idempotent. By the same method, the squared partial correlation between $\overline{\boldsymbol{X}}_l$ and $\overline{\boldsymbol{X}}_m$ conditioned on $I\overline{\boldsymbol{X}}_k$ is:

$$\rho^2(\overline{\boldsymbol{X}}_l, \overline{\boldsymbol{X}}_m | I\overline{\boldsymbol{X}}_k) = \frac{\left(\mathbf{1}' \boldsymbol{X}_l{}' \boldsymbol{M_B} \boldsymbol{X}_m \mathbf{1}\right)^2}{\mathbf{1}' \boldsymbol{X}_l{}' \boldsymbol{M_B} \boldsymbol{X}_l \mathbf{1} \mathbf{1}' \boldsymbol{X}_m{}' \boldsymbol{M_B} \boldsymbol{X}_m \mathbf{1}}\,, \tag{92}$$

and finally, the squared partial correlation between $\overline{\boldsymbol{X}}_l$ and $I\overline{\boldsymbol{X}}_m$ conditioned on $\overline{\boldsymbol{X}}_k$ is:

$$\rho^2(\overline{\boldsymbol{X}}_l, I\overline{\boldsymbol{X}}_m | \overline{\boldsymbol{X}}_k) = \frac{\left(\mathbf{1}' \boldsymbol{X}_l{}' \boldsymbol{M_A} \boldsymbol{X}_m \mathbf{1}\right)^2}{\mathbf{1}' \boldsymbol{X}_l{}' \boldsymbol{M_A} \boldsymbol{X}_l \mathbf{1} \mathbf{1}' \boldsymbol{X}_m{}' \boldsymbol{M_A} \boldsymbol{X}_m \mathbf{1}}\,. \tag{93}$$

Unfortunately, comparing Equations (91) to (93) with Equation (83) does not yield unambiguous evidence of the effect of aggregation on the partial correlation. This is because the term in parenthesis in Equations (89) and (90) cannot be reduced any further due to $\boldsymbol{W_A}$ and $\boldsymbol{W_B}$ being non square matrices.

5.2.2 Stochastic Simulations

The previous results imply that the effect of aggregation on the Fisher statistic cannot be easily observed through an analytical derivation of the test statistic. For this reason, we addressed this problem through stochastic simulation. However, we imposed a very strong assumption. We presumed that the relationship between the consumer, retailer, and manufacturer is independent for each brand to be aggregated. This implies that the manufacturer and the retailer will independently pursue the same strategies for each of the brands that fit in a certain group. This supposition is not very restrictive of the firms' behavior since those brands marketed differently can be aggregated in different groups.

However, the same assumption implies that the consumers' choice of quantity is influenced only by his relationship with the retailer and the manufacturer, and it is independent from the other brands to be grouped together. This notion can be very limiting. Deaton and Muellbauer (1980, pp149-166) address *exact aggregation* by discussing the usually strict necessary conditions in which the aggregated behavior can represent the actions of its constituents. In particular, they show that if the average of the consumers' individual demand functions for a particular good is to be represented by a single aggregate average market demand function, such function would imply that the marginal propensities to spend have to be identical for every consumer, which, of course, is not true. The average market demand thus has to be a function of the average total expenditure. This in turn implies that the individual consumer demand functions must be linear in each household's total expenditure, associated to quasi-homothetic preferences along with straight Engel curves with a constant slope across individuals, which are very rigorous, although not impossible for undifferentiated products. Generalized Linearity achieves exact *non linear aggregation* by aggregating over the expenditure patterns instead of quantities. Although this methodology results in Engel curves that need not to be straight, they need to be linearly related. Moreover, the use of a *representative* consumer is required, which brings other theoretical considerations studied by welfare economists.

Nevertheless, retailers should price in response to aggregate demand functions. Deaton and Muellbauer (1980) added that the methods discussed above render functional forms that are too specific to model actual behavior. However, less strict conditions can be achieved if the aggregate demand function is not expected to be consistent with utility maximization. In particular, the aggregated demand function should be determined by the shape of the Engel curves when all the consumers face the same prices. Then, holding constant the distribution of expenditures across individuals, the summation or average of individual demands can be represented by a single function of prices and average expenditure. This aggregate function will only need to satisfy the adding up and homogeneity restrictions of demand theory. Thus, in general, the micro and macro functional forms need not to be similar. Deaton and Muellbauer emphasize this as the reason why welfare economists should be very cautious when making inferences about microeconomic behavior based on macroeconomic observations.

Let the vector $\boldsymbol{x}_{t,i}$ consist of three observations, $(x_{t,i,k}, x_{t,i,l}, x_{t,i,m})'$, from a Multivariate Normal Distribution with mean $\boldsymbol{\mu}$ and Variance-Covariance matrix $\boldsymbol{\Sigma}$. Then, redefine $\boldsymbol{X}_t$ as an independent and identically distributed sample of size I of the vector $\boldsymbol{x}_{t,i}$:

$$\boldsymbol{X}_t = \begin{bmatrix} x_{t,1,k} & \cdots & x_{t,I,k} \\ x_{t,1,l} & \ddots & x_{t,I,l} \\ x_{t,1\,m} & \cdots & x_{t,I,m} \end{bmatrix}.$$

Let $\mathbf{1}$be a vector of ones with dimension $i \times 1$. Then,

$$\boldsymbol{X}_t\mathbf{1} = \begin{bmatrix} \sum_i^I x_{t,i,k} \\ \sum_i^I x_{t,i,l} \\ \sum_i^I x_{t,i,m} \end{bmatrix},$$

$$\mathrm{E}[\boldsymbol{X}_t\mathbf{1}] = \begin{bmatrix} I\mu_k \\ I\mu_l \\ I\mu_m \end{bmatrix},$$

and

$$\mathrm{Var}[\boldsymbol{X}_t\mathbf{1}] = \mathrm{E}[(\boldsymbol{X}_t\mathbf{1} - \mathrm{E}[\boldsymbol{X}_t\mathbf{1}])(\boldsymbol{X}_t\mathbf{1} - \mathrm{E}[\boldsymbol{X}_t\mathbf{1}])']$$

$$= \Sigma_\mathrm{i}\Sigma_\mathrm{j}\mathrm{E}\big[(\boldsymbol{x}_{t,i} - \boldsymbol{\mu}_i)(\boldsymbol{x}_{t,j} - \boldsymbol{\mu}_j)'\big].$$

But since $i \perp j$, then the covariance between vectors $i \neq j$ equals zero. Therefore,

$$= \Sigma_i \Sigma_j \mathrm{E}\left[(\boldsymbol{x}_{t,i} - \boldsymbol{\mu}_i)(\boldsymbol{x}_{t,j} - \boldsymbol{\mu}_j)'\right]$$

$$= \Sigma_i \mathrm{E}\left[(\boldsymbol{x}_{t,i} - \boldsymbol{\mu}_i)(\boldsymbol{x}_{t,i} - \boldsymbol{\mu}_i)'\right] = I\boldsymbol{\Sigma}.$$

These results facilitate the imposition of the two different aggregation rules that we have discussed. Define the linear combinations matrix $\boldsymbol{C}$ as

$$\boldsymbol{C} = \begin{bmatrix} 1/a & 0 & 0 \\ 0 & 1/b & 0 \\ 0 & 0 & 1/c \end{bmatrix},$$

where *a*, *b,* and *c* are constants to be defined later. Then,

$$\mathrm{E}[\boldsymbol{C}\boldsymbol{X}_t\boldsymbol{1}] = \boldsymbol{C}\mathrm{E}[\boldsymbol{X}_t\boldsymbol{1}] = \begin{bmatrix} I/a\ \mu_k \\ I/b\ \mu_l \\ I/c\ \mu_m \end{bmatrix} \tag{94}$$

and

$$\mathrm{Var}[\boldsymbol{C}\boldsymbol{X}_t\boldsymbol{1}] = \boldsymbol{C}'\mathrm{Var}[\boldsymbol{X}_t\boldsymbol{1}]\boldsymbol{C} = I\boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}\ . \tag{95}$$

The stochastic simulations are therefore generated as a Multivariate Normal Distribution with expected value and Variance-Covariance matrix given by Equations (94) and (95) respectively. Notice that the linear combinations matrix $\boldsymbol{C}$ allows us to impose the desired aggregation rules. In particular, making $a = 1$, and $b = c = I$ leads to the scenario where the variable of type *k* has been aggregated by summation, whereas the variables of types *l* and *m* have been aggregated with averages. Each triplet was generated one thousand times (T = 1,000) to study the effect of increasing I on both the unconditional correlation and the Fisher statistic's distribution at two different levels of correlation imposed through the Σ matrix. In every case, the expected value is set to be

Figure 8 Simulated Variance-Covariance Structures.

$$\boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

a) Uncorrelated Structure

$$\boldsymbol{\Sigma} = \begin{bmatrix} 1 & -0.8 & -0.8 \\ -0.8 & 1 & 0.8 \\ -0.8 & 0.8 & 1 \end{bmatrix}$$

b) Highly Correlated Structure.

zero and unit standard deviations are imposed to equate the Variance-Covariance matrix to the Correlation matrix. Two different scenarios were analyzed for each experiment. The first scenario studies the effects of aggregation when there is zero correlation/covariance among the variables, whereas the second scenario studies the effect of aggregation when the correlation/covariance between the triplets is big (Figure 8). Note that the correlation/covariance between the first variable and the rest was imposed to be negative to account for normal goods. These simulations were generated with *Simetar Lite 2008.* The results follow.

Table 2 displays the effects on the Fisher statistic and the unconditional correlation after increasing the correlation structure and the aggregation level. The table is divided into 5 sections. Table 2.a shows the scenario when aggregation with averages is used in all three variables (i.e. $\overline{\boldsymbol{X}}_l, \overline{\boldsymbol{X}}_m$, and $\overline{\boldsymbol{X}}_k$**)**. It shows that there is no change in the Fisher distributional statistics as *I* is increased from 1 (no aggregation) to 50 (big aggregation) regardless of the level of correlation in the data. For instance, in the case of a zero correlation structure, the standard deviation of the Fisher statistic remains at 0.99 units as the level of aggregation, *I,* is increased. Similarly, when highly correlated series are generated, the standard deviation of the Fisher statistic remains at 1 as the

Table 2 **Effects of Aggregation on the Fisher Statistic and the Unconditional Correlation.**

	Integration Level (I)					
	Uncorrelated Series			Highly Correlated Series		
	I=1	*I=10*	*I=50*	*I=1*	*I=10*	*I=50*
a) Fisher statistic based on $\rho(\bar{\boldsymbol{X}}_l, \bar{\boldsymbol{X}}_m \| \bar{\boldsymbol{X}}_k)$						
Mean	-0.02	-0.02	-0.02	15.10	15.10	15.10
StDev	0.99	0.99	0.99	1.00	1.00	1.00
Min	-2.82	-2.82	-2.82	12.09	12.09	12.09
Max	2.88	2.88	2.88	17.92	17.92	17.92
b) Fisher statistic based on $\rho(\bar{\boldsymbol{X}}_l, \bar{\boldsymbol{X}}_m \| I\bar{\boldsymbol{X}}_k)$						
Mean	0.0	0.0	0.0	15.10	15.10	15.10
StDev	1.0	1.0	1.0	1.00	1.00	1.00
Min	-2.8	-2.8	-2.8	12.09	12.09	12.09
Max	2.9	2.9	2.9	17.92	17.92	17.92
c) Fisher statistic based on $\rho(\bar{\boldsymbol{X}}_l, I\bar{\boldsymbol{X}}_m \| \bar{\boldsymbol{X}}_k)$						
Mean	-0.02	-0.02	-0.02	15.10	15.10	15.10
StDev	0.99	0.99	0.99	1.00	1.00	1.00
Min	-2.82	-2.82	-2.82	12.09	12.09	12.09
Max	2.88	2.88	2.88	17.92	17.92	17.92
d) Corr$(\bar{\boldsymbol{X}}_l, \bar{\boldsymbol{X}}_m)$						
Mean	0.00	0.00	0.00	-0.80	-0.80	-0.80
StDev	0.03	0.03	0.03	0.01	0.01	0.01
Min	-0.08	-0.08	-0.08	-0.83	-0.83	-0.83
Max	0.09	0.09	0.09	-0.76	-0.76	-0.76
e) Corr$(\bar{\boldsymbol{X}}_l, I\bar{\boldsymbol{X}}_m)$						
Mean	0.00	0.00	0.00	-0.80	-0.80	-0.80
StDev	0.03	0.03	0.03	0.01	0.01	0.01
Min	-0.08	-0.08	-0.08	-0.83	-0.83	-0.83
Max	0.09	0.09	0.09	-0.76	-0.76	-0.76

Note. $\bar{\boldsymbol{X}} = (\boldsymbol{\Sigma}_{\mathbf{i}}\boldsymbol{X}_{\boldsymbol{i}})/\boldsymbol{I}$, where $\boldsymbol{X}_i \in \mathrm{R}^3$ ~ MVN(0,Σ); Σ as in Figure 8.

aggregation level is increased. Table 2.b displays the descriptive statistics associated to the distribution of the Fisher statistic calculated for a conditional correlation whose conditional variable has been treated with a different aggregation rule. In particular, the conditional variable has been aggregated by summation only, whereas the remaining variables have been aggregated by averaging (i.e. $\rho(\bar{\boldsymbol{X}}_l, \bar{\boldsymbol{X}}_m | I\bar{\boldsymbol{X}}_k)$). The simulations show that the Fisher statistic distribution does not change with different levels of aggregation regardless of the correlated structures. The same conclusions are reached in Tables 2.c, 2.d, and 2.e which present the results of the same analysis over the Fisher statistic associated with $\rho(\bar{\boldsymbol{X}}_l, I\bar{\boldsymbol{X}}_m | \bar{\boldsymbol{X}}_k)$), and the unconditional correlations $Corr(\bar{\boldsymbol{X}}_l, \bar{\boldsymbol{X}}_m)$ and $Corr(\bar{\boldsymbol{X}}_l, I\bar{\boldsymbol{X}}_m)$, respectively.

These results indicate that the Fisher statistic remains unchanged when independent observations of a 3-variable Multivariate Normal distributed vector are aggregated by adding or averaging across different samples. This in turn implies that the algorithms would render the same causal pattern regardless of aggregation when the conditions mentioned above are met. However, those assumptions are very restrictive in particular with respect to consumer behavior.

5.2.3 Conclusions on Aggregation

We discussed in Subsection 5.2 that aggregation of the data was required to deal with missing observations. Although missing observations are not a major concern when computing the correlations which lie at the heart of all of the algorithms used in causality analysis, they do affect the size of the sample and therefore the reliability of our results.

For this reason, given that a manufacturer may pursue similar marketing strategies for some products, an aggregated product may be used to generate the triplet discussed in Subsection 4.2. Yet, it is not obvious whether the real causal pattern for a specific brand will be elicited when several products have been aggregated together.

This problem was analyzed through a mathematical approach. However, no clear insights were discovered. For this reason, the aggregation problem was explored through the stochastic simulation of a Multivariate Normal distributed vector of three variables which was subject to the two different aggregation rules used in this Dissertation. Assuming that the strategic relationship between the retailer and a manufacturer is independent for each of the manufacturer's brand, and therefore the same variance-covariance structure for each of the brands that will be aggregated, we concluded that aggregation does not influence the size of our test statistics and therefore the causal patterns remain unchanged. However, this assumption also implies that all the patrons of Dominick's Finer Foods act independently towards any of the brands being grouped together. This notion has been challenged by the demand theory literature because of its implications on the shape of the consumers' expansion paths, although empirical characterization of the aggregated demand function can be obtained if this market function is not expected to be consistent with utility maximization (Deaton and Muellbauer, 1980).

5.3 CASE STUDY RESULTS AND DISCUSSION

The results of the empirical analysis of the Dominick's database are described in this section. We analyzed only those categories believed to have more traffic in order to reduce the number of missing observations. Namely, we analyzed Beer, Cheese, Soft Drinks, Toilet Paper, and Tuna. Each of these groups contains several manufacturers, which in turn are providers of several brands in numerous presentations. Each product is classified by a *UPC* number which identifies any specific item by its last five digits, while the remaining digits identify the manufacturer. Nevertheless, there may be items produced by the same company and yet are labeled with a different manufacturer number. We interpreted this as groups of products that are marketed differently. A list of these manufacturers and a description of their products is shown in the Appendix B.

Only one store location was studied, Store 123, in order to avoid errors due to zone price discrimination. The data was first aggregated by manufacturer in each category. Both the manufacturer price and the retailer price were aggregated by averaging the per unit price of each product, while the quantities were aggregated in three different ways. Therefore, three triplets were generated for each manufacturer; $\{q_i^A, P_{r,i}, P_{m,i},\}_A$, $\{q_i^B, P_{r,i}, P_{m,i},\}_B$, $\{q_i^C, P_{r,i}, P_{m,i},\}_C$, where q_i^A represents the total units sold for manufacturer i, q_i^B represents the total number of discount coupons used to purchase the products of manufacturer i, and q_i^C represents the percent of the total sales of products supplied by the manufacturer i that were purchased with coupons. $P_{r,i}$ and $P_{m,i}$ represent the per unit average price for the retailer and the manufacturer in all instances.

In Section 5.2, we discussed that aggregation by sum or averages will not affect our causal patterns if a) the manufacturer-retailer relationship is independent, and b) similar strategies are pursued for each of the brands provided by the same manufacturer (i.e. same Variance-Covariance matrix). Hence, our aggregation occurs by manufacturer as they are identified by the UPC of each product. However, we found that the correlation between the manufacturer's and the retailer's stationary price is not constant across those brands provided by the same manufacturer. Moreover, these correlations tend to move widely across the [-1, 1] spectrum (Table 3). This may partially explain why we failed to find unambiguous causal patterns in many instances. Nevertheless, the correlations tend to cluster for some of the brands implying a closer relationship among them, but an automated cluster analysis of products for each manufacturer seems unpractical since the number of clusters is unknown and tends to vary. We now discuss this in more detail before we continue with the description of our empirical analysis.

Table 3 Distributions of the Disaggregated Brands' Correlation Among their Retailer and Manufacturer Stationary Prices.

Category/		*Correlations*			
Manufacturer	*Representative Products*	*Average*	*Std. Dev.*	*Min*	*Max*
BEER					
m1	Budweiser/Michelob	0.44	0.53	-1.00	1.00
m4	Miller	0.33	0.41	-1.00	1.00
m6	Augsburguer/Old Milwakee	-0.10	0.56	-0.94	0.27
m7	Labatta	0.81	.	0.81	0.81
m8	Molson	0.47	0.35	-0.01	0.85
m10	Coors/Keystone/Blue Moon	0.52	0.14	0.33	0.72
m11	Strohs	0.43	0.27	0.21	0.72

Table 3 Continued.

Category/ Manufacturer	*Representative Products*	*Correlations* *Average*	*Std. Dev.*	*Min*	*Max*
m12	Pilsner/Moosehead/Dos Equis	0.44	0.53	-0.06	1.00
m13	Heineken	0.35	0.41	-0.21	0.67
m14	Old Style	0.48	0.47	-0.66	1.00
m15	Corona	0.62	0.19	0.44	0.81
m18	Beck's	0.63	0.20	0.47	0.98
m20	Berghoff/Augsburguer	0.36	0.33	0.17	0.73
m22	Samuel Adams	0.07	0.02	0.06	0.09
m25	Goose Island Honkers	0.01	.	0.01	0.01
m27	Oregon Brewery India	-0.96	.	-0.96	-0.96
CHEESE					
m1	Cty Ln Colby Mild	0.24	0.19	0.04	0.59
m3	Lifeway's Farmer's C	0.01	.	0.01	0.01
m4	Kraft Colby 1/3 Less	0.40	0.29	-0.68	1.00
m9	Lol Cheddarella Chee	0.25	0.17	0.01	0.43
m12	Dom Amer Chs Food Tw	0.16	0.13	-0.21	0.38
m14	Frigo Shred Ched	0.49	0.73	-0.03	1.00
m15	Laughing Cow/Bonbel	0.30	0.21	0.15	0.45
m16	Lean N' Free /Alpine	0.31	0.14	0.16	0.48
m17	Treasure Cave Square	0.16	0.23	-0.01	0.32
m18	Sargento Wafer Thin	0.40	0.20	-0.06	0.74
m19	Borden Lt Ln Chdr	0.12	0.23	-0.07	0.51
SOFT DRINKS					
m2	Old Town Nat Sltzer	0.24	0.53	-0.25	0.80
m3	Tetley Iced Tea W/Le	-0.15	.	-0.15	-0.15
m4	Pepsi-Cola Cans	0.39	0.30	-0.25	0.98
m5	Schwepps Tonic N/R	0.35	0.19	0.08	0.69
m6	Canada Dry Ginger Al	0.70	0.25	-0.22	0.98
m7	Hawaiian Punch Gldn	0.90	0.04	0.87	0.99
m8	Fruitopia Citrus Con	0.33	0.30	0.06	0.81
m9	Royal Crown Cola	0.32	0.32	-0.32	0.96
m11	Ocean Spry	0.59	0.23	0.04	0.76
m12	World Classics Cola	0.27	0.21	-0.01	0.56
m13	Dominick's Cola 3 Lt	0.25	0.12	0.02	0.61

Table 3 Continued.

Category/ Manufacturer	*Representative Products*	*Correlations*			
		Average	*Std. Dev.*	*Min*	*Max*
m14	Liptn Brew Wild Strw	0.33	0.31	-0.46	0.64
m15	PS	1.00	.	1.00	1.00
m16	Nu Grape Soda Ppd	0.31	0.16	0.16	0.48
m17	Crush Different Flavors	0.53	0.20	0.30	0.86
m18	Country Time	0.51	0.23	0.00	0.78
m19	Sunkist	0.48	0.23	0.00	0.75
m21	Coca-Cola	0.46	0.37	-0.98	0.98
m22	New York Seltzer	1.00	.	1.00	1.00
m23	Hawiian Punch Red 8p	0.10	0.24	-0.35	0.32
m24	Hawaiian Punch /Sunny Delight	0.57	0.26	0.14	0.98
m26	Sundance Black Curra	0.40	.	0.40	0.40
m27	Dr Pepper Sugar Free	0.59	0.24	0.27	0.98
m28	Clearly Canadian	0.64	0.19	0.40	1.00
m29	A & W Root Beer (Can	0.39	0.30	-0.11	0.94
m31	Boku	0.51	0.05	0.47	0.54
m32	Vernors / Artic Twist	0.58	0.24	0.33	0.88
m33	Barq's Root Beer	0.48	0.49	-0.46	0.91
m34	Dad's Root Beer Tria	0.50	0.27	0.05	0.99
m35	Ibc Root Beer Trial	0.23	0.00	0.23	0.23
m36	Holy Cow	-1.00	.	-1.00	-1.00
m37	Seagram's Ginger Ale	0.18	0.15	-0.03	0.40
m38	Lacroix Orange Miner	0.34	0.48	-1.00	0.78
m40	Perrier Berry Nr	0.25	.	0.25	0.25
m41	C/G Reg Sprk Mineral	0.58	0.13	0.30	0.71
m43	Snapple Peach Tea	0.55	0.21	-0.11	0.76
m44	Welch Grape	0.02	0.57	-0.64	0.40
m46	A W Cream Soda Reg	0.24	0.32	-0.17	0.54
m47	7-Up	0.61	0.27	0.05	0.99
m51	Canfield Fruit Punch	0.29	0.27	-0.07	0.76
m53	Nestea 6pk Cans	0.25	0.38	-0.07	0.74
m56	Mistic	0.35	0.45	-0.03	1.00
TOILET PAPER					
m1	Angel Sft Bth Tissue	0.40	0.52	-0.25	1.00
m2	Kleenex Pp1.09	0.34	0.19	0.08	0.65

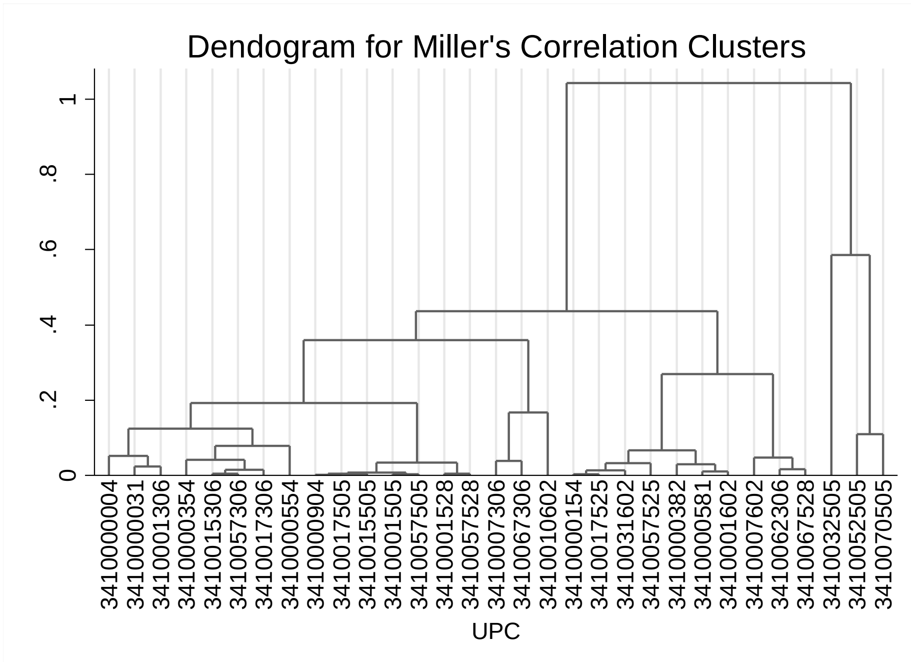
Table 3 Continued.

Category/ Manufacturer	*Representative Products*	*Correlations* *Average*	*Std. Dev.*	*Min*	*Max*
m3	Charmin Bath Tissue	0.26	0.41	-0.60	0.66
m4	Dominick's 2ply Pr B	0.16	0.40	-0.22	0.58
m5	Northern Quilted Tis	0.53	0.20	0.27	1.00
m6	Kleenex Cottonell Ba	0.49	0.23	0.00	0.82
m7	Generic Bath Tissue	-0.03	.	-0.03	-0.03

As we mentioned above, the correlation between the manufacturer’s and the retailer’s price is not constant across the different brands provided by the same manufacturer. But a cluster analysis of these correlations would help to classify those products that seem to be marketed similarly. Nevertheless, the erroneous identification of these clusters would lead to misleading conclusions. As an example, we analyzed the aggregation of those products supplied by Miller, which is the manufacturer identified as m4 in the Beer category (Table 3). More specifically, we generated the dendrogram shown in Figure 9 by using the Euclidean distance based average dissimilarity of the UPCs’ correlations to identify those products that tend to group together (Johnson and Wichern, 1998). Examination of Figure 9 led us to aggregate the UPCs in 7 clusters using the k-means method (Table 4).

It is evident that those UPCs that tend to behave similarly are precisely those of comparable brands and size or presentation (Table 4). Thereafter, a causality analysis of those clusters was individually conducted using only the triplet generated by addition of the consumer quantities. Only the cluster number two yielded an unambiguous causal

Figure 9 Identification of the Correlation Clusters for Miller Beers.



Note. Graph calculated with STATA 10.1. (StataCorp, 2007)

pattern which was the same causal pattern found when all the brands were aggregated together as a single triplet. Unfortunately, incorrect identification of the appropriate number of clusters may lead to misleading results. For instance, reducing the number of clusters from 7 to 6 in the example above leads to no identifiable causal patterns. We now continue to describe the analysis of the aggregated triplets.

Once the triplets per manufacturer were computed, each of the series were analyzed for stationarity using a Dickey Fuller test for random walks with a constant mean but without drift at the 0.05% significance level, to avoid spurious results (Bessler

Table 4 **Classification of Miller's UPCs by Nonhierarchical Clusters.**

Cluster	*UPC*	*Brand*	*Ounces*
1	3410052505	Leinenkugel Limited	72
1	3410070505	Miller Resever Amber	72
2	3410000382	Miller Lite Beer N.R	32
2	3410000581	Miller Genuine Draft	32
2	3410000154	Milwaukee's Best Bee	72
2	3410017525	Miller Genuine Draft	72
2	3410057525	Miller Lite Beer N.R	72
2	3410001602	Miller High Life Bee	144
2	3410031602	Milwaukee's Best Bee	144
3	3410000554	Miller Genuine Draft	72
3	3410000904	Miller Genuine Draft	72
3	3410001505	Miller High Life Lnn	72
3	3410015505	Miller Gen Drft Lt L	72
3	3410017505	Miller Gen Drft Lnnr	72
3	3410057505	Miller Lite Longneck	72
3	3410001528	Miller High Life N.R	144
3	3410057528	Miller Lite Beer N.R	144
4	3410007602	Milwaukee's Best Lig	144
4	3410067528	Miller Red Dog Nr Bt	144
4	3410062306	Miller Icehouse Cans	288
5	3410010602	Sharps Non-Alcoholic	144
5	3410007306	Milwaukee's Best Lig	288
5	3410067306	Miller Red Dog Cans	288
6	3410000031	Miller High Life N.R	32
6	3410000004	Miller High Life Bee	72
6	3410000354	Miller Lite Beer	72
6	3410001306	Miller High Life Bee	288
6	3410015306	Miller Genuine Drft	288
6	3410017306	Miller Genuine Draft	288
6	3410057306	Miller Lite Beer	288
7	3410032505	Miller Reserve Lnnr	72

and Kling, 1984). If the series were found to be non-stationary, then differences, logarithms, or differences of the logarithms were taken to achieve stationarity. Only those stationary triplets with the same integration order were studied. Alas, not all of the series survive this process and therefore many manufacturers were left out of the study.

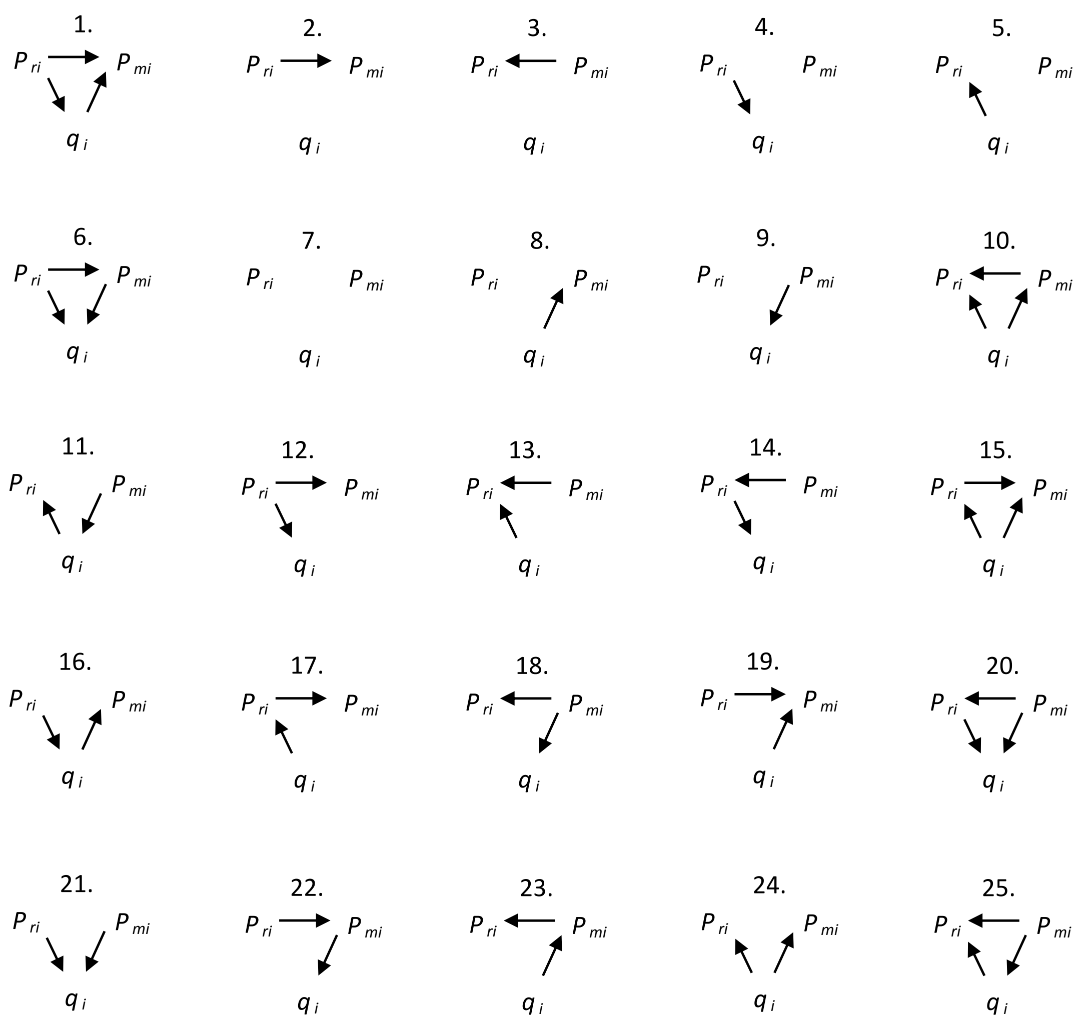
Moreover, many of these series had blocks of missing observations. This affected our methodology in two ways. First, some series cannot be made stationary because taking differences, or log differences, will never solve the problem. Second, a significant level of 0.05 is not adequate for small sample sizes. Because the Directed Graphs algorithms rely only on correlations, discontinuities in the data are not a problem as long as the series survive the stationary tests. But the length of the series is important for the accuracy of our statistical tests. Therefore, those series that had less than 30 observations were discarded. Further adjustments were required to improve the reliability of our results. These are described below and Appendix C provides a list of those manufacturers that have been excluded from the study.

Finally, the three different triplets per manufacturer were analyzed through both the Moneta's algorithm (Moneta, 2008) and a truncated version of the Richardson's algorithm (Richardson, 1996), written in Visual Basic (VBA) as a macro in MS Excel 2007; both reviewed in Subsection 4.3 and described in detail in Appendix A. Both of these algorithms rely on normally distributed data through the use of the Fisher Statistic (Equation 81). Although our series frequently rejected normality with Shapiro-Wilk tests, our aggregation methods imply asymptotic normality on all of our series through the use of the Central Limit Theorem; further requiring large datasets. Closer inspection of the p-values associated to the Fisher Statistic revealed that these p-values tend to be either absolutely small, or absolutely big, leading to no ambiguous significance tests. Indeed, changing the significance level to 1%, 5%, 10%, and 20% lead to no significant changes in the results. Therefore, the algorithms ran with a 20% significance level for those series that had less than 100 hundred observations, and 5% for the rest.

Once the different causal patterns were obtained, they were classified by type. Although there are 25 potential causal relationships among the three variables (Figure 10), both algorithms start by identifying colliders (structures 13, 19, and 21 in Figure 10), which are later used to direct the remaining causal relationships. But the presence of a collider implies that there must be only two edges. Therefore only undirected relationships and colliders could be identified. The theoretical model discussed in Subsection 4.2 implies that structures 13, 19, and 21 are respectively associated to a Manufacturer Stackelberg, Retailer Stackelberg, and retailer's monopolistic behavior with vertical coordination. Nevertheless, Richardson's algorithm contains one extra step that allowed us to identify one edge structures without requiring the previous identification of colliders. This step enabled us to identify structures 2, 3, 4, 5, 8, and 9. None of these have ambiguity in spotting a market leader.

The problem of identifying more structures through different means has already been studied. Haigh and Bessler (2004) used the Schwarz loss metrics to provide support for the Observational Equivalence Theorem which states that DAG's that have the same skeletons and converging edges are equivalent. This implies that either background

Figure 10 Possible Causal Relationships among the Manufacturer's Price (P_{mi}), Retailer's Price (P_{ri}), and Quantity Sold (q_i). (Haigh and Bessler, 2004)



information, more variables, or both, must be added to the model. Here, we could rely on economic theory which suggests that the value of a product comes from its abundance (making it cheaper and therefore increasing demand) or scarcity (making it more expensive and therefore decreasing demand). Consequently, although the algorithms cannot distinguish a *causal fork* $(X \leftarrow Y \rightarrow Z)$ from a *causal chain* $(X \rightarrow Y \rightarrow Z)$ when only three variables are studied, we can infer that the consumer quantity is always a *causal source*, making it possible to direct as causal chains those relationships where q_i is at one extreme of a two-edge structure $(q_i \rightarrow Y \rightarrow Z)$. Unfortunately, this relationship cannot be assumed with absolute certainty. In particular, we concluded that the presence of monopolistic power can lead to the interpretation of consumption quantities as causal sinks (Figure 5). Moreover, our algorithms are unable to detect simultaneous determination or perfect coordination of the supply chain $(X \leftrightarrow Y)$. Otherwise, any undirected structures of two edges with quantity at an extreme could be directed as structures 17 or 23 in Figure 10.

The results of our empirical analysis are presented in Appendix D and E, while a discussion follows. Interestingly, although the algorithms were more likely to identify a causal pattern when the quantities have been aggregated by sum, q_i^A, some of the causal patterns were elicited through one type of aggregation but not through the others. This is because two of the three aggregation rules over the consumer quantities focused their attention only on those products that were offered with a discount through the use of coupons, thus reducing the volatility of that sample. Nevertheless, the two different algorithms find supporting causal patterns for the three differently aggregated triplets except for two instances. While the Moneta Algorithm proved to be the most conservative, identifying fewer causal patterns, it was this algorithm which twice determined opposite causal directions for the same manufacturer when analyzing its three different triplets. However, the Richardson's Algorithm and the Moneta's Algorithm produced conflicting results only once out of the 93 analyzed manufacturers.

Moreover, causality was successfully determined for 60% of our sample. The Cheese category yielded unambiguous causal patterns for 73% of its manufacturers, whereas the lowest success ratio came from the Toilet Paper group with 43%. These are very high numbers considering the shortcomings in our dataset and the limits of our algorithms. Further, our results indicate that 70% of those unambiguous causal patterns elicit Manufacturer Stackelberg relationships, 25% elicit Retailer Stackelbergs, and only 5% elicit a monopolistic retailer with vertical coordination. A summary of these results are presented in Table 4, whereas the complete results are provided in the Appendix D.

These results are consistent with the market structure of the early 90s, when this sample was taken. At that time, the retail industry was highly fragmented so manufacturers tended to lead the prices. In addition, this period of time was characterized by the manufacturers' effort to increase profits due to an abundance of meaningless products, lack of innovation, and fewer private labels. A different scenario may be uncovered in more recent data because the market is dominated by fewer than 10 retailers, which enables them to resist the manufacturers' pricing policies[9].

On the other hand, Kadiyali *et al.* (2000) discussed that an increasing amount of retailers gained larger proportions of the profits earned by a supply chain during these times, suggesting more retailer "power" within the channel. They attributed this change

[9] Thanks to Mr. Stew Bishop, president of Consumer Marketing Group, Inc., and M.S. Jeff Lovering, Director of the same company, for their valuable insights.

Table 5 **Summary of Empirical Results.**

CAUSAL PATTERNS	*Beer*	*Cheese*	*Soft Drinks*	*Toilet Paper*	*Tuna*	*Total*
Total number of directed structures	13	8	22	3	10	56
Total number of manufacturers	19	11	40	7	16	93
Proportion of directed structures	68%	73%	58%	43%	63%	60%
Structure: Monopolistic Retailer with vertical coordination ($Pm \rightarrow q \leftarrow Pr$)						
Total	2	1	0	0	0	3
Proportion of the total number of manufacturers	11%	9%	0%	0%	0%	3%
Proportion of the total directed structures found	15%	13%	0%	0%	0%	5%
Structure: Manufacturer Stackelberg ($q \rightarrow Pr \leftarrow Pm$)						
Total	10	4	15	2	8	39
Proportion of the total number of manufacturers	53%	36%	38%	29%	50%	42%
Proportion of the total directed structures found	77%	50%	68%	67%	80%	70%
Structure: Retailer Stackelberg ($Pr \rightarrow Pm \leftarrow q$ or $Pr \rightarrow Pm$)						
Total	1	3	7	1	2	14
Proportion of the total number of manufacturers	5%	27%	18%	14%	13%	15%
Proportion of the total directed structures found	8%	38%	32%	33%	20%	25%
Undirected Structures						
Total	6	3	17	4	6	36
Proportion of the total number of manufacturers	32%	27%	43%	57%	38%	39%

to intense competition among manufacturers, the introduction of high quality private labels, and an increasing number of products along with a decreasing number of retailers and shelf space. Nevertheless, their study acknowledges the existence of empirical studies that provide support against these arguments.

Furthermore, our results are also consistent in regards with the low presence of monopolistic activities with supply chain coordination (Figure 5; Table 5). McGuire and Staelin (1983, 1986), and Coughlan and Wernerfelt (1989) showed that perfect coordination is not the best alternative for manufacturers of highly substitutable products. In particular, all of our categories pertain to products from highly competitive markets. Yet, monopolies with supply chain coordination were only elicited for one cheese and two beer providers. Both beer brands, Labatt and Foster's, are perceived as expensive imports with successful product differentiation through branding. However, none of these brands holds more than 1% of the market share. Nevertheless, Foster's does provide Dominicks' with the second largest profit margin among all beers, while Labatt provides an average profit margin of 16%, which is bigger than that of either Miller, Michelob, Budweiser, and Old Style which together posses 85% of the total market share. Furthermore, the cheese brand which is associated to a monopolistic game with supply chain coordination is marketed as a Dominick's private label, which may be a clear signal of channel partnerships. Moreover, this private label provides Dominicks with the largest profit margin and the second largest market share, 64% and 28% respectively. This profit margin is almost twice as big as the 36% provided by Kraft, whose market share is the biggest and doubles that of the private label.

Another interesting observation can be found when comparing our results with those of Kadiyali *et al.* (2000). Their study of tuna, with the same dataset as ours, suggested that the retailer appropriates a larger share of the channel profits, which they refer to as “pricing power”. Although they did not determine price leadership, they concluded that the retailer retains more than 50% of the channel profits obtained from Chicken of the Sea, Starkist, and BumbleBee. The authors of that study find it surprising because of the absence of a private level in this market and explained this finding by describing high competition experienced both at the retailer and manufacturer level. In particular, Chicken of the Sea (COS) seems to be the most sensitive brand to competition. On the other hand, Kadiyali *et al.* determined that the margins of these brands elicit deviations from a Bertrand - Nash equilibrium. Our study suggests that the manufacturers hold price leadership on all three brands. It is worth mentioning that although Starkist holds price leadership with its Solid White brand, it follows Dominick’s lead on the Tuna Chunk light.

Table 6 demonstrates another interesting point. On average, the brands that hold

Table 6 Average Statistics Classified by Equilibrium and Product Types.

	Average Market Share (%)		Average Retailer Profit Margin (%)	
	Manufacturer Stackelberg	Retailer Stackelberg	Manufacturer Stackelberg	Retailer Stackelberg
Beer	9.42	1.73	12.29	21.22
Cheese	16.55	1.20	48.87	38.14
Soft Drinks	3.56	3.91	22.99	26.85
Toilet Paper	7.26	2.32	25.88	36.38
Tuna	8.87	0.32	41.83	38.54

the bigger market shares tend to elicit Manufacturer Stackelberg relationships. On the other hand, on average, the brands that provide the retailer with the bigger profit margin seem to elicit Retailer Stackelberg equilibriums. However, this observation needs to be considered with caution because the Retailer Stackelberg sample is very small. In particular, this category contains only one brand of beer and only one brand of toilet paper, while the tuna group is composed by only two brands (Appendix D).

Interestingly, Dominick’s private labels are identified most of the times as Manufacturer Stackelbergs. These results are at odds with those of Cotterill and Putsis (2001) who studied a sample that covers the same time window. Their study concluded that Manufacturer Stackelbergs are more frequent for national brands while vertical Nash interactions are more common for private labels. However, our study would support Jørgensen *et al.* (2001) conclusion that Manufacturer Stackelbergs should be more frequent. His research discussed that Manufacturer Stackelbergs improve channel efficiency and consumer welfare, whereas a retailer leadership is not desirable in any of these terms. Moreover, the same study concluded that, although the retailer would be happy to assume the leadership within the channel, it would accept to be a follower in order to avoid participating in a vertical Nash. This conclusion was also reached by Lee and Staelin (1997) when increasing the profit margin comes as the best response to a supply chain member’s increase in margin.

This research is complemented with a study of the horizontal relationship among the manufacturers from the Beer, Cheese, Soft Drinks, Toilet Paper, and Tuna categories, as described in Section 4. In other words, we analyzed the causal relationship among the suppliers of Dominicks and identified the price leaders. However, some of the manufacturers were dropped from this study to avoid spurious relationships. In particular, many of the suppliers were not studied because they provided their products during different time windows or very short periods of times. The complete results are presented in Appendix E whereas a discussion is provided below.

The causal patterns were estimated two times. Once with the Richardson's algorithm, and the second time with the GES algorithm, both described in Appendix A and available through Tetrad IV (2005). Only those causal patterns common to both algorithms were retained and considered trustworthy. This is because the Richardson's algorithm presented reversals of some causal relationships when the confidence level was changed. This problem is not uncommon. Scheines *et al.* (1994) discussed that such inconsistencies are usually explained by the omission of latent variables in the models, while Demiralp, Hoover, and Perez (2008) added that adherence to the *causal sufficiency* assumption implies that arrow reversals may be sign of simultaneous determination. Spirtes *et al.* (2000, pp 115) used simulations to study the reliability of several algorithms, including the PC. They concluded that the rates of arrow and undirected edges omissions are decreased dramatically as sample sizes are increased, up to 1000 observations. Whereas the rates of arrow and undirected edges commissions are less sensitive to sample size, the frequency of the omissions and commissions will be affected by the statistical properties of the data. They suggest the use of high significance levels for small samples, and reduce the significance level as the sample increases.

Demiralp and Hoover (2003) also used Monte Carlo simulations to study the reliability of the PC Algorithm, but this time to infer the causal pattern among contemporaneous disturbances, which is required to properly specify a SVAR. They concluded that the PC Algorithm has well behaved statistical properties as long as the causal relationships are sufficiently strong (high signal-to-noise ratios). A downside of this study is that such Monte Carlo simulations are useful only if the data generation process is known, which is not common. In a subsequent study, Demiralp *et al.* (2008) complemented the previous study by providing a methodology that used bootstrap simulations which do not rely on a known data generating process.

Bryant, Bessler, and Haigh (2009) reflected on these problems and proposed a different method to test whether a variable causes another one. Their methodology suggests that inclusion of all possible related variables is not required. Equally important, their methodology provides a way to assess whether a causal relationship was rejected with weak or strong support, further improving the reliability of the results. Moreover, such strong basis rejections provide a high degree of confidence without the requirement of big samples, as long as the causal relationship is strong.

We used a different approach in this research. Appendix E shows two graphs per group. Graph *a* is the Directed Cyclic Graph obtained with the use of the Richardson's algorithm, while graph *b* shows only those edges and arrows that coincide with those obtained from the GES Algorithm (Meek, 1997; Chickering, 2002). Thus graph *b* is the one determined to be reliable and it is the one used in the following discussion.

As before, the emphasis of this study is to identify the price leaders among each category. The marketing literature identifies two different firm types that would naturally seek for price leadership. The first one is when there is the existence of a dominant or more efficient firm (Markham, 1951). This scenario exists when a single firm's choice of output can affect the market price because it supplies more than 40% of the market, leaving the residual demand to be satisfied by the rest of the industry which is characterized by one or more smaller firms that act as price takers. For this reason, the dominant firm choice of price is the one that equates its marginal benefits to its marginal costs. However, this market scenario has been deemed impractical with only a few industries that fit this model (Bain, 1960; Deneckere and Kovenock, 1992).

The second type of price leaders is the Barometric (Stigler, 1947). Barometric leaders are those firms that accurately study and interpret the market environment which results on a more precise determination of the equilibrium price. Such price is thereafter signaled and adopted by the rest of the industry because it is considered acceptable. This ability to study the economic environment is not exclusive to the dominant firms, and for this reason this leadership tends to shift if the price leader is no longer considered reliable.

Bain (1960) discussed that these two models are too simple. He explained that firms recognize the interdependence among competitors, which was considered by Deneckere and Kovenock's (1992) game-theoretic model of dominant-firm price leadership, where the leadership is completely determined by the effects of capacity constraints in a Nash equilibrium. However, relaxation of their assumptions leads to price leadership supported by brand loyalty or quality instead of size. Nevertheless, their model implies that price leadership yields more stable prices. The importance of price stability is also mentioned by Barney (2002, pp 357). Barney explains that price leaders are given by their ability to set a price that yields “acceptable” profits to the industry, and to provide environments conducive to long lasting tacit collusions. Barney further explains that price leaders are generally regarded as Stackelberg leaders.

Our empirical study thus seeks for this type of leader in our results. Although we do not have cost information from any of the firms, Kadiyali *et al.* (2000) identified Chicken of the Sea (COS) as the most efficient firm among Bumble Bee and Starkist, being Starkist the firm with the highest costs among these three. Our results show that COS is indeed a price leader and Starkist a follower. Bumble Bee appears to be independent. Interestingly, COS leads over all the firms that hold the larger market shares, including Starkist which accounts for the most sales.

The Beer category presents similar results. The firm with the greatest market share is not the price leader. Although Miller holds 42% of the market, it appears independent from the industry. However, Old Style holds the second largest market share with 27% and leads Coors/Keystone and Foster’s, which have significantly lower market shares. Budweiser also is one of the big players with 17%, and our results indicate that this firm leads Beck’s which is significantly smaller. However, Budweiser is lead by Old Milwakee/Augsburguer which is even smaller than Beck’s.

The Cheese category paints a similar picture. Kraft is the dominant manufacturer with 60% of the market and yet it fails to be the market leader. In fact, it appears to follow the prices set by Borden and Cty Ln. It is interesting to notice that the store label holds the second largest market share and appears independent from the industry's pricing decisions.

More intuitive results can be found when looking at the Soft Drinks category. Coca-Cola and Pepsi, both with the largest shares in the market, seem to be pricing leaders, although they certainly do not lead the whole market. Interestingly, Dominick's private label follows Pepsi's lead. Intuitive results are also found in the Toilet Paper category. This is because the generic brand is clearly a market follower. Once again, the market leader, Kleenex, appears to be independent whereas Charmin, the second largest firm, leads the generic brand but follows the lead of Northern Quilted, which is the third largest firm.

These results lend support to the observation that price leadership is determined by strategic behavior and not by firm size. Indeed, the largest firms seem independent from the market pricing decisions, whereas the second largest firms tend to be engaged in price leaderships. Although this is not always the case as we saw in the Soft Drinks Category. These results also support the notion that more efficient firms tend to lead. We found the leadership pattern implied by the observations offered in Kadiyali *et al.* (2000). We also found that the private label never leads over other brands, yet it did lead the generic brand toilet paper.

5. 4 CASE STUDY CONCLUSIONS

We presented in Section 5 an application of the methodology discussed in Section 4. We used Directed Graphs to study the relationship between a retailer and several manufacturers who frequently provided more than one brand each. Our study may be affected by a large number of missing observations in the data. This problem was circumvented by aggregating those brands provided by the same manufacturer to create a representative product. Although we proved that aggregation should not affect our results when a manufacturer pursues the same marketing strategy for each of its brands, we found no strong evidence that manufacturers behave this way in our study. Moreover, the same aggregation implies strong restrictions on consumer behavior.

Nevertheless, our empirical analysis unambiguously determined causal patterns for 60% of our sample. Of these, 70% elicit Manufacturer Stackelberg relationships ($q_{consumer} \rightarrow p_{retailer} \leftarrow p_{manufacturer}$), 25% elicit Retailer Stackelbergs ($q_{consumer} \rightarrow p_{manufacturer} \leftarrow p_{retailer}$), and only 5% elicit a monopolistic retailer with vertical coordination ($p_{retailer} \rightarrow q_{consumer} \leftarrow p_{manufacturer}$). These results strongly agree with the market structure of the early 1990s and predictions from other researchers. In particular, Manufacturer Stackelberg relationships are more frequently assumed (Lee and Staelin 1997; Jørgensen *et al.* 2001; Wang, 2002; Corbett *et al.*, 2004). However, there are studies were the retailer is expected or found to lead the Stackelberg relationship (Choi, 1991; Tsay, 2002; Ertek and Griffin, 2002). Interestingly, we found that the manufacturers that hold the bigger market shares tend to elicit Manufacturer Stackelberg relationships, whereas those manufacturers that provide the retailer with the bigger profit margin seem to elicit Retailer Stackelberg equilibriums. Although our algorithms are incapable of simultaneous determination, McGuire and Staelin (1983, 1986), and Coughlan and Wernerfelt (1989) showed that perfect coordination is not the best alternative for manufacturers of highly substitutable products.

We also studied the horizontal relationship among the manufacturers. The marketing literature explains that the firms that interpret the markets the best, the dominant firms, or the most efficient, are the firms that establish themselves as price leaders (Markham, 1951; Stigler; 1947). However, some authors explain that price leadership comes as a consequence of strategic behavior, which may be influenced by firm size, and are generally regarded as Stackelberg leaderships (Bain 1960; Deneckere and Kovenock, 1992; Barney, 2002, pp 357). Although we do not have any information about the cost structures of any of the manufacturers, or their ability to predict the market prices, we determined that the price leader in the tuna category is the same that Kiadiyali *et al.* (2000) determined to be the firm with the lowest costs. But our results indicate that the market leader in fact tends to be isolated from the pricing fluctuations. It is the second largest firm which tends to engage in price leadership. Finally, we discovered that the private label does not lead any firm's decisions, unless this firm is the provider of a generic brand.

6. SUMMARY AND CONCLUSIONS

This research studies the use of causal inference models to elicit strategic behavior among competing firms. This is important because proper determination of a strategic relationship has been found to be essential for accurate modeling. For instance, Lee and Staelin (1997) concluded that even the specification of the demand in the structural models is not as important as determining the type of interaction present between the channel members. Moreover, the same study demonstrated how the choice of a firm's optimal strategy is not constrained by the best response of the channel member, but by the actual response being observed. This is because of some firm's tendency to behave sub optimally. Indeed, Dhar *et al.* (2005) assumed a fixed markup as the retailer strategy, but they also acknowledge that studies of this nature will be biased.

However, identification of such underlying strategic relationships is not trivial and usually relies on structural models which are very sensitive to misspecification. Moreover, models of strategic behavior are usually too complex to conduct large scale analyses. This dissertation thus explains how the usual Nash equilibriums imply different causal patterns and therefore can be elicited through models of causal inference. The result is a different way to elicit strategic relationships with the use of algorithms that are not sensitive to misspecification and can be programmed to conduct large scale analyses.

To meet our objectives, we analytically derived the Nash equilibriums that result from the Cournot, Bertrand, and Stackelberg games in both horizontal (*manufacturer –*

manufacturer) and vertical relationships (*retailer – manufacturer*), and related them to causal patterns. Given that firms' profits depend on the decisions made by their competitors or supply chain members, they should study their competitors' and partners' prices and use this information to set their own prices or quantities (Slade, 1992; Lee and Staelin, 1997) . Therefore, these price signals can sometimes be considered as exogenously determined and for this reason as causal sources of information for the profit maximization process.

Thus, for the horizontal competition case, we concluded that the monopolist's ability to determine supply, and therefore prices, allows him to manipulate consumption. Therefore, the following pattern should be elicited: $q_{monopoly} \rightarrow p_{monopoly} \rightarrow q_{consumer}$, where $q_{monopoly}$ represents the monopolist output, and $q_{consumer}$ represents the actual consumer's choice of quantity. However, the consumer should always be a causal source of information whenever there is no monopolistic power; $q_{consmer} \rightarrow q_j \rightarrow p_j$, where q_j and p_j are firm j^{th}'s output and price in a competitive market. Firms engaged in a Cournot relationship should have a simultaneous determination of their supplies generating a bidirectional causal pattern, $q_i \leftrightarrow q_j$, whereas companies that participate in a Stackelberg game would generate an unidirectional causal pattern, where the Stackelberg leader is the source of the information: $q_{Stackelberg\ leader} \rightarrow q_i$. Finally, companies that take on Bertarnd competition should seem to be independent from each other. This is because their prices are determined by their own ability to reduce costs and lower the prices to those of a perfect competitive market. Nevertheless, superior monopolistic profits will provide incentives for collusion. Successful cooperation should result in consumer manipulation by the participant firms in the oligopoly: $q_{collusion} \rightarrow p_{monopoly} \rightarrow q_{consumer}$.

We also studied the causal patterns associated with perfect coordination of the supply chain, vertical Nash, Manufacturer Stackelberg, Retailer Stackelberg, and monopolistic power for vertical competition (competition within a supply chain). We concluded that the consumer should serve as a causal source for both the retailer's and manufacturer's price when there is vertical coordination, $p_{retailer} \leftarrow q_{consumer} \rightarrow p_{manufacturer}$, unless there exists monopolistic power, $p_{retailer} \rightarrow q_{consumer} \leftarrow p_{manufacturer}$. Different patterns should be observed whenever there is no vertical coordination. In particular, absence of coordination but presence of retailer's monopolistic power should determine the retailer's price as a causal source for both the consumers and manufacturers, $q_{consumer} \leftarrow p_{retailer} \rightarrow p_{manufacturer}$. Absence of both monopolistic power and perfect coordination would lead to either vertical Nash or Stackelberg games. The former implies simultaneous determination between the retailer and the manufacturer (when the retailer does not use a fixed mark up rule) while both are subject to the consumer choices, $p_{retailer} \leftrightarrow p_{manufacturer}$ with $p_{retailer} \leftarrow q_{consumer} \rightarrow p_{manufacturer}$. The later implies that the Stackelberg leader will act as a causal source subject to the consumer's choices. That is, the Manufacturer Stackelberg and the Retailer Stackelberg should be respectively elicited by $q_{consumer} \rightarrow p_{retailer} \leftarrow p_{manufacturer}$ and $q_{consumer} \rightarrow p_{manufacturer} \leftarrow p_{retailer}$.

These conclusions indicate the need of a causal algorithm that allows for bidirectional causality. Therefore, Directed Cyclic Graphs (DCG) as opposed to Directed

Acyclic Graphs (DAG) was considered. For this purpose, we studied one DCG algorithm (Richardson, 1996) and two DAG algorithms, PC (Sprites *et al.*, 1993) and Moneta's (Moneta, 2008). Since these algorithms are almost equivalent when studying only three variables, truncated versions of each were programmed in Visual Basic as a spreadsheet macro and used to analyze the relationship between a retailer and its suppliers. Indeed, the different algorithms provided supporting results and contradicting each other only once out of the 93 relationships that were studied.

However, the dataset had multiple missing observations which were dealt with aggregation of all those brands which were provided by the same manufacturer. Although we proved that our results should not change under the assumption that the manufacturer-retailer relationship is independent and similar for each of the manufacturer's brands, we have no evidence that this assumption is met. Cluster statistical methods can help to refine our analysis by grouping those brands that tend to act similarly, although it does not seem practical given the dimension of our study. Nevertheless, Day (1963) concluded that such type of aggregation should not mask the nature of the firm. Moreover, the same assumption places strong restrictions on the underlying consumer behavior. This is a persistent problem in consumer demand theory. Deaton and Muellbauer (1980) explained that in general, the micro and macro functional forms need not to be similar. These restrictions can be relaxed if the aggregate demand function is expected to be determined only by the shape each consumer's Engel curve in the face of the same prices and a constant distribution of expenditures.

Our results tend to agree with those of other researchers. In particular, we determined causality for 60% of our sample. This is a very high number considering the shortcomings in our dataset and the fact that our algorithms can only detect three out of the twenty five possible causal patterns that can be observed when only three variables are considered. Nevertheless, there is no ambiguity on those patterns that can be elicited. Namely we could only search for Stackelberg strategies and monopolistic activities with supply chain coordination. Our results indicate that 70% of those unambiguous causal patterns elicit Manufacturer Stackelberg relationships, 25% elicit Retailer Stackelbergs, and only 5% elicit a monopolistic retailer with vertical coordination. These results were expected because the market structure of the early 1990s was characterized by a highly fragmented retailed industry, enabling the manufacturers to lead the prices. Moreover, the convergence of the lack of private labels and the abundance of poorly developed products provided the manufacturers with incentives to pursue strategic avenues to increase profits. Furthermore, price leaderships should be frequent since they provide stability of prices (Barney, 2002 pp 357). Additionally, our results are also consistent in regards with the low presence of monopolistic activities and supply chain coordination. McGuire and Staelin (1983, 1986), and Coughlan and Wernerfelt (1989) showed that perfect coordination is not the optimal strategy when considering highly elastic products. In particular, all of our categories pertain products from highly competitive markets.

Interestingly, our study classified Dominick’s private labels as Manufacturer Stackelbergs most of the times. However, Cotterill and Putsis (2001) argued that private labels tend to achieve Vertical Nash equilibriums. Nevertheless, the same study agrees that the Manufacturer Stackelbergs should be more frequent for national brands. Moreover, Jørgensen *et al.* (2001) concluded that retailer leadership is not desirable whereas Manufacturer Stackelbergs provide several advantages. They improve supply chain efficiency and consumer welfare. Furthermore, the Vertical Nash leads to a lower level of profits. Thus, Jørgensen *et al.* concluded that the retailer would benefit from becoming a price follower in order to avoid a Vertical game of simultaneous determination. This conclusion was also reached by Lee and Staelin (1997).

Our study of the horizontal relationships among the manufacturers leads to some interesting results as well. The marketing literature explains that the dominant firms, the most efficient, or those which have a superior ability to interpret the market conditions are the ones that assume price leadership (Markham, 1951; Stigler; 1947). However, some authors explain that price leadership comes as a consequence of strategic behavior which is generally regarded as Stackelberg leaderships (Bain 1960; Deneckere and Kovenock, 1992; Barney, 2002, pp 357). Our results indicate that the market leader in fact tends to be isolated from the pricing fluctuations while the second largest firm tends to engage in price leadership. Nevertheless, we successfully identified as a price leader the only company whose costs are known to be low through Kadiyali *et al.* (2000). Finally, we discovered that the private label does not lead any firm's decisions, unless this company is the provider of a generic brand.

An immediate extension of this research would be to study a larger dataset that contains more than one retailer. A more complete dataset would enrich this research by avoiding the aggregation of brands that could lead to erroneous interpretations and likely to a higher success ratio. It would be interesting as well to study brands marketed in different presentations. Moreover, knowledge of the private characteristics of each firm would enable us to explicitly test the ability of low cost firms, barometric leaders, and dominant companies to set the price.

REFERENCES

Aradillas-Lopez, A. 2005. Semiparametric estimation of a simultaneous game with incomplete information. Working paper, Department of Economics, Princeton, NJ.

Appelbaum, E. 1979. Testing price taking behavior. *Journal of Econometrics* **9** 283-294.

Ashenfelter, O., D. Sullivan. 1987. Non-parametric tests of market structure: An application to the cigarette industry. *Journal of Industrial Economics* **35** 483-498.

Bain, J. 1960. Price leaders, barometers, and kinks. *Journal of Business* **33** 193-203.

Barney B. J. 2002. *Gaining and Sustaining Competitive Advantage*. 2nd ed. Pearson Education Inc., Patparganj, Delhi, India.

Bajari, P., H. Hong, S. Ryan, 2004. Identification and estimation of discrete games of complete information. Technical Working Papers: 0301, National Bureau of Economic Research, Inc, Cambridge, MA.

Bessler, D. A., J. L. Kling. 1984. A note on tests of Granger causality. *Applied Economics, Taylor and Francis Journals* **16**(3) 335-342.

Bessler, D. A., A. Kergna. 2002. Price discovery: The case of Millet in Bamako, Mali. *Journal of African Economies* **11** (4) 472-502.

Bessler, D. A., J. Yang. 2003. The structure of interdependence in international stock markets. *Journal of International Money and Finance* **22** 261- 287.

Bresnahan, T. F., P. C. Reiss. 1990. Entry in monopoly markets. *Review of Economic Studies* **57**(4 October) 531-553.

Bresnahan, T. F., P. Reiss. 1991. Empirical models of discrete games. *Journal of Econometrics* **48**(1-2 April-May) 57-81.

Bryant, Henry L., D. A. Bessler, M. S. Haigh. 2009. Disproving causal relationships using observational data. *Oxford Bulletin of Economics and Statistics* **71**(3) 357-374.

Carter, C. A., D. MacLaren. 1997. Price or quantity competition? Oligopolistic structures in international commodity markets. *Review of International Economics* **5**(3) 373-385.

Chiang, C., Hess. 2003. Direct-marketing, indirect profits: A strategic analysis of dual-channel supply-chain design. *Management Science* **49**(1) 1-20.

Chickering, D. M. 2002. Optimal structure identification with greedy search. *Journal of Machine Learning Research* **3** 507-554.

Choi, S.C., 1991. Price competition in a channel structure with a common retailer. *Marketing Science* **10** (4) 271–296.

Chu, W., P.R. Messinger. 1997. Information and channel profits. *Journal of Retailing* **73**(4) 487–499.

Corbett, C.J., D. Zhou, C.S. Tang. 2004. Designing supply contracts: Contract type and information asymmetry. *Management Science* **50**(4) 550–559.

Cotterill, R. W., W. P. Putsis, Jr. 2001. Do models of vertical strategic interaction for national and store brands meet the market test? *Journal of Retailing* **77** 83–109.

Coughlan, A. T., Wernerfelt, B. 1989. On credible delegation by Oligopolists: a discussion of distribution channel management. *Management Science* **35** 226-239.

Day, R. H. 1963. On aggregating linear programming models of production. *Journal of Farm Economics* **45**(4) 797-813.

Deaton, A., J. Muellbauer. 1980. *Economics and Consumer Behavior*. Cambridge University Press, Cambridge, U.K.

Deneckere, R., D. Kovenock. 1992. Price leadership. *Review of Economic Studies* **59** 143-162.

Dermilap, S., Kevin D. Hoover. 2003. Searching for the causal structure of a vector autoregression. *Oxford Bulletin of Economics and Statistics* **65** 745-767.

Dermilap, S., K. D. Hoover, S. J. Perez. 2008. A bootstrap method for identifying and evaluating a structural vector autoregression. *Oxford Bulletin of Economics and Statistics* **70**(4) 509-533.

Dhar, T., Jean-Paul Chavas, R. W. Cotterill, B. W. Gould. 2005. An econometric analysis of brand-level strategic pricing between Coca-Cola Company and PepsiCo. *Journal of Economics and Management Strategy* **14**(4) 905-931.

Ertek, G., P.M. Griffin. 2002. Supplier- and buyer-driven channels in a two-stage supply chain. *IIE Transactions* **34**(8) 691–700.

Friedman, J. 1971. A non-cooperative equilibrium for supergames. *Review of Economic Studies* **38** 1-12.

Frydenberg, M. 1990. The chain graph Markov property. *Scandinavian Journal of Statistics* **17**(4) 333-353.

Gasmi, F., J.-J. Laffomt, Quang H. Vuoing. 1992. Econometric analysis of collusive behavior in a soft-drink market. *Journal of Economics and Management* **1** 277-311.

Gibbons, Robert. 1992. *Game Theory for Applied Economists*. Princeton University Press, Princeton, NJ.

Green, E., R. Porter. 1984. Noncooperative collusion under imperfect price information. *Econometrica* **52** 87-100.

Haigh, M. S., D. A. Bessler, 2004. Causality and price discovery: An application of directed acyclic graphs. *The Journal of Business* **77**(4) 1099-1121.

Johnson, R.A., D.W. Wichern. 1999. *Applied Multivariate Statistical Analysis.* 4th ed. Prentice Hall, Upper Saddle River, NJ.

Jørgensen, S. 1986. Optimal production, purchasing and pricing: A differential game approach. *European Journal of Operational Research* **24** 64-76.

Jørgensen, S., S. P. Sigué, G. Zaccour. 2001. Stackelberg leadership in a marketing channel. *International Game Theory Review* **3**(1) 13-26.

Kadiyali, V., P. Chintagunta, N. Vilcassim. 2000. Manufacturer-retailer channel interactions and implications for channel power: An empirical investigation of pricing in a local market. *Marketing Science* **19**(2) 127-148.

Kalman, R. E. 1960. A new approach to linear filtering and prediction problems. *Transactions ASME Journal of Basic Engineering* **82** 35-45.

Kalman, R. E., R. S. Bucy. 1961. New results in linear filtering and prediction theory. *Transactions ASME Journal of Basic Engineering* **83** 95-103.

Lau, A. H. L. H. S. Lau, J.C. Wang. 2007. Pricing and volume discounting for a dominant retailer with uncertain manufacturing cost information. *European Journal of Operational Research* **183**(2) 848-870.

Lee, E., R. Staelin. 1997. Vertical strategic interaction: Implications for channel pricing strategy. *Marketing Science* **16**(3) 185-207.

Leng, M., M. Parlar, 2005. Game theoretic applications in supply chain management: A review. *INFOR*. **43**(3) 187-220.

Leontieff, W. 1936. Composite commodities and the problem of index numbers. *Econometrica* **4** 39-59.

Leontieff, W. 1947. Introduction to a theory of the internal structure of functional relationships. *Econometrica* **15**(4) 361-373.

Lewbel, A. 1996. Aggregation without separability: A generalized composite commodity theorem. *American Economic Review* **86** 524-543.

Markham, J. W. 1951. The nature and significance of price leadership. *American Economic Review* **41** 891-905.

Martın-Herran, G., S. Taboubi, G. Zaccour. 2006. The impact of manufacturers' wholesale prices on a retailer's shelf-space and pricing decisions. *Decision Sciences* **37**(1) 71-90.

Mas-Collel, A., M. D. Whinston, J.R. Green. 2004. *Microeconomic Theory*. Oxford University Press, New York Oxford.

McGuire, T.W., R. Staelin, 1983. An industry equilibrium analysis of downstream vertical integration. *Marketing Science* **2**(2) 161-191.

McGuire, T.W., Staelin, R. 1986. Channel efficiency, incentive compatibility transfer pricing, and market structure: An equilibrium analysis of channel relationships.

Research in Marketing: Distribution Channels and Institutions **8** JAI Press, L. P. Bucklin and J. M. Carman editors., Greenwich, CT.

Meek, C. 1997. Graphical Models: Selecting Causal and Statistical Models. PhD Dissertation, Carnegie Mellon University, Pittsburgh, PA.

Moneta, A. 2008. Graphical causal models and VARs: An empirical assessment of the real business cycles hypothesis. *Empirical Economics* **35**(2) 275-300.

Nayga M. R., O. Capps. 1994. Tests of weak separability in disaggregated meat products. *American Journal of Agricultural Economics* **76**(November) 800-808.

Pearl, J. 1995. Causal diagrams for empirical research. *Biometrika* **82** 669–710.

Raper, K.C., H. A. Love, C. R. Shumway. 2007. Distinguishing the source of market power. *American Journal of Agricultural Economics* **89** (1) 78–90.

Reiss, P. C. 1996. Empirical models of discrete strategic choices. *American Economic Review* **86**(2) 421-26.

Richardson, T. 1996. A Discovery Algorithm for Directed Cyclic Graphs. Proceedings of the 12th Annual Conference on Uncertainty in Artificial Intelligence *(UAI-96)*. Morgan Kaufmann, San Francisco, CA. 454-61.

Scheines, R., Spirtes, P., Glymour, C., Meek, C. 1994. *TETRAD II: User's Manual and Software.* Lawrence Erlbaum Associates Inc, Hillsdale, NJ.

Simetar Lite 2008. Simulation & Econometrics To Analyze Risk. http://simetar.com/.

Slade, M. E. 1990. Strategic pricing models and interpretation of price-war data. *European Economic Review* **34**(2-3) 524-37.

Slade, M. E. 1992.Vancouver's gasoline-price wars: An empirical exercise in uncovering supergame strategies. *Review of Economic Studies* **59**(2) 257-76.

Spirtes, P., Glymour, C., Scheines, R. 1993. *Causation, Prediction, and Search*. Springer-Verlag, New York.

Spirtes, P., Glymour, C., Scheines, R. 2000. *Causation, Prediction, and Search.* Cambridge: MIT Press, Cambridge, MA.

StataCorp. 2007. STATA 10.1. http://www.stata.com.

Stigler, G. 1947. The kinky oligopoly demand curve and rigid prices. *Journal of Political Economy* **55** 432-449.

Stockton, M. C., O. Capps, D. A. Bessler. 2008. Samuelson's full duality and the use of directed acyclic graphs. *Journal of Applied Econometrics* **11**(1) 167-191.

Swanson, N.R., C.W.J. Granger. 1997. Impulse response functions based on a causal approach to residual orthogonalization in vector autoregressions. *Journal of the American Statistical Association* **92** 357-367.

Tamer, E. 2003. Incomplete simultaneous discrete response model with multiple equilibria. *Review of Economic Studies* **70**(1) 147-65.

Tetrad. 2005. Laboratory for Symbolic and Educational Computing. Department of Philosophy, Carnegie Mellon University. http://www.phil.cmu.edu/projects/tetrad.

Theil, H. 1956. Linear aggregation of economic relations. *Economica* **23**(91) 277-279.

Tsay, A.A. 2002. Risk sensitivity in distribution channel partnerships: Implications for manufacturer return policies. *Journal of Retailing* **78**(2) 147–160.

Wang, Q. 2002. Determination of suppliers' optimal quantity discount schedules with heterogeneous buyers. *Naval Research Logistics* **49**(1) 46–59.

Wolak, F. A., C. D. Kolstad. 1991. A model of homogeneous input demand under price uncertainty. *American Economic Review* **81** 514-38.

APENDIX A

ALGORITHMS

This Appendix presents the Algorithms used for this research. The customized algorithm used for this research is also presented. The algorithms provided below are adaptations from Scheines *et al*. (1994), Moneta (2007), and Richardson (1996). Letters in bold intend to emphasize the principal differences among them. These versions have been modified from their original notation to allow ease of comparison and agreement with the notation used throughout this dissertation. A brief description of the Greedy Equivalence Search algorithm (GES) algorithm is also offered. This algorithm lies outside of the scope of this dissertation and for this reason further details are referred to the original publications (Meek, 1997; Chickering, 2002).

A.1 PC ALGORITHM

(A)
Form the complete undirected graph C on the vertex set y_{1t} , . . . , y_{kt}. Let *Adjacencies(C,* y_{it} *)* be the set of vertices adjacent to y_{it} in C; let G_{Yt} be the (unobservable) causal structure among the k elements of Y_t;

(B)
n = 0
Repeat :
Repeat :
select an ordered pairs of variables y_{ht} and y_{it} that are adjacent in C such that *Adjacencies(C,* y_{ht} *)\{*y_{it} *}* has cardinality greater than or equal to n, and a subset S of *Adjacencies(C,* y_{ht} *)\{*y_{it} *}* of cardinality n, and if y_{ht} and y_{it} are independent given S in G_{Yt}

delete edge y_{ht} — y_{it} from C and **record S in *Sepset* (y_{ht} , y_{it}) and *Sepset* (y_{it} , y_{ht});**

until all ordered pairs of adjacent variables y_{ht} and y_{it} such that *Adjacencies(C,* y_{ht} *)\{*y_{it} *}* has cardinality greater than or equal to n and all subsets S of *Adjacencies(C,* y_{ht} *) \{*y_{it} *}* of cardinality n have been tested for independence;

$n = n + 1$;

until for each ordered pair of adjacent variables y_{ht} , y_{it} , *Adjacencies(C,* y_{ht} *)\{*y_{it} *}* is of cardinality less than n;

(C)

For each triple of vertices y_{ht} , y_{it} , y_{jt} such that the pair y_{ht} , y_{it} and the pair y_{it} , y_{jt} are each adjacent in C but the pair y_{ht} , y_{jt} is not adjacent in C, orient y_{ht} — y_{it} — y_{jt} as $y_{ht} \rightarrow y_{it} \leftarrow y_{jt}$ if and only if y_{it} **is not in *Sepset* (y_{ht}, y_{jt});**

(D)

Repeat :

if $y_{at} \rightarrow y_{bt}$, y_{bt} and y_{ct} are adjacent, y_{at} and y_{ct} are not adjacent and y_{bt} belongs to **the *Sepset* (y_{at} , y_{ct})**, then orient y_{bt} — y_{ct} as $y_{bt} \rightarrow y_{ct}$;

if there is a directed path from y_{at} to y_{bt} , and an edge between y_{at} and y_{bt} , then orient y_{at} — y_{bt} as $y_{at} \rightarrow y_{bt}$;

until no more edges can be oriented.

A.2 MONETA

(A)

Form the complete undirected graph C on the vertex set y_{1t} , . . . , y_{kt}. Let *Adjacencies(C,* y_{it} *)* be the set of vertices adjacent to y_{it} in C; let G_{Yt} be the (unobservable) causal structure among the k elements of Y_t ; and let *Sepset* (y_{ht} , y_{it}) **be the set of sets of vertices** S so that y_{ht} and y_{it} are d-separated given S in G_{Yt} .

(B)

n = 0

Repeat :

Repeat :

select an ordered pairs of variables y_{ht} and y_{it} that are adjacent in C such that *Adjacencies(C,* y_{ht} *)\{*y_{it} *}* has cardinality greater than or equal to n, and a subset S of $Adjacencies(C, y_{ht})\backslash\{y_{it}\}$ of cardinality n, and if y_{ht} and y_{it} are d-separated given S in G_{Yt} delete edge $y_{ht} — y_{it}$ from C;

until all ordered pairs of adjacent variables y_{ht} and y_{it} such that $Adjacencies(C, y_{ht})\backslash\{y_{it}\}$ has cardinality greater than or equal to n and all subsets S of $Adjacencies(C, y_{ht})\backslash\{y_{it}\}$ of cardinality n have been tested for d-separation;

$n = n + 1$;

until for each ordered pair of adjacent variables y_{ht}, y_{it}, $Adjacencies(C, y_{ht})\backslash\{y_{it}\}$ is of cardinality less than n;

(C)

For each triple of vertices y_{ht}, y_{it}, y_{jt} such that the pair y_{ht}, y_{it} and the pair y_{it}, y_{jt} are each adjacent in C but the pair y_{ht}, y_{jt} is not adjacent in C, orient $y_{ht} — y_{it} — y_{jt}$ as $y_{ht} \rightarrow y_{it} \leftarrow y_{jt}$ if and only if y_{it} **does not belong to any set** of *Sepset* (y_{ht}, y_{jt});

(D)

Repeat :

if $y_{at} \rightarrow y_{bt}$, y_{bt} and y_{ct} are adjacent, y_{at} and y_{ct} are not adjacent and y_{bt} belongs to **every set of** *Sepset* (y_{at}, y_{ct}), then orient $y_{bt} — y_{ct}$ as $y_{bt} \rightarrow y_{ct}$;

if there is a directed path from y_{at} to y_{bt}, and an edge between y_{at} and y_{bt}, then orient $y_{at} — y_{bt}$ as $y_{at} \rightarrow y_{bt}$;

until no more edges can be oriented.

A.3 RICHARDSON

(A)

Form the complete undirected graph *Partial Ancestral Graphs* C on the vertex set $y_{1t}, \ldots, y_{kt}$. Let $Adjacencies(C, y_{it})$ be the set of vertices adjacent to y_{it} in C; let G_{Yt} be the (unobservable) causal structure among the k elements of Y_t;

(B)

n = 0

Repeat :

Repeat :

Select an ordered pairs of variables y_{ht} and y_{it} that are adjacent in C such that $Adjacencies(C, y_{ht})\backslash\{y_{it}\}$ has cardinality greater than or equal to n, and a subset S of $Adjacencies(C, y_{ht})\backslash\{y_{it}\}$ of cardinality n, and if y_{ht} and y_{it} are independent given S in G_{Yt} delete edge $y_{ht} \circ\!\!-\!\!\circ\ y_{it}$ from C and **record S in *Sepset* (y_{ht}, y_{it}) and *Sepset* (y_{it}, y_{ht});**

until all ordered pairs of adjacent variables y_{ht} and y_{it} such that *Adjacencies(C,* y_{ht} *)\\{*y_{it} *}* has cardinality greater than or equal to n and all subsets S of *Adjacencies(C,* y_{ht} *) \\{*y_{it} *}* of cardinality n have been tested for independence;

$n = n + 1$;

until for each ordered pair of adjacent variables y_{ht} , y_{it} , *Adjacencies(C,* y_{ht} *)\\{*y_{it} *}* is of cardinality less than n;

(C)

For each triple of vertices y_{ht} , y_{it} , y_{jt} such that the pair y_{ht} , y_{it} and the pair y_{it} , y_{jt} are each adjacent in C but the pair y_{ht} , y_{jt} is not adjacent in C, orient y_{ht} *—* y_{it} *—* y_{jt} as $y_{ht} \rightarrow y_{it} \leftarrow y_{jt}$ if and only if y_{it} **is not in *Sepset* (y_{ht}, y_{jt}); Orient y_{ht} *—* y_{it} *—* y_{jt} as y_{ht} *—*$\underline{y_{it}}$*—* y_{jt} if and only if y_{it} does belongs to the *Sepset* (y_{ht}, y_{jt});**

(D)

For each triple of vertices y_{ht} , y_{it} , y_{jt} such that (a) y_{ht} is not adjacent to y_{it} or y_{jt},(b) y_{it} and y_{jt} are adjacent, (c) y_{it} does not exist in the *Sepset* (y_{ht}, y_{jt}), then orient y_{it} *—* y_{jt} as $y_{it} \leftarrow y_{jt}$ if y_{ht} is independent of y_{it} conditioned on *Sepset* (y_{ht}, y_{jt}).

(E)

For each vertex V in *C* form the following set: y_{ht} exists in Local(*C* ,*V*) if and only if y_{ht} is adjacent to V in *C*, or there is a vertex y_{it} s.t. $y_{ht} \rightarrow y_{it} \leftarrow y_{jt}$ in *C*.

m=0

Repeat:

Repeat:

Select an ordered triple < y_{ht} , y_{it} , y_{jt} > such that $y_{ht} \rightarrow y_{it} \leftarrow y_{jt}$, y_{ht} and y_{jt} are not adjacent, and $|\text{Local}(C, y_{ht}) \setminus \{ y_{it}, y_{jt}\}| \geq$ m, and a set T which is a subset from Local(C, y_{ht})\\{ y_{it}, y_{jt}}, |T|=m, and if y_{ht} is independent of y_{jt} conditioned on T or {y_{it}}, then orient $y_{ht} \rightarrow y_{it} \leftarrow y_{jt}$ as $y_{ht} \rightarrow \underline{y_{it}} \leftarrow y_{jt}$, and record T or { y_{it} } in Supset< y_{ht} , y_{it} , y_{jt} >

Until for all triplets such that $y_{ht} \rightarrow y_{it} \leftarrow y_{jt}$, (not $y_{ht} \rightarrow \underline{y_{it}} \leftarrow y_{jt}$), y_{ht} and y_{jt} are not adjacent, |Local(C, y_{ht})\{ y_{it} }| ≥ m, every subset T which is a subset from Local(C, y_{ht}), |T| = m has been considered.

m=m+1

Until for all ordered triplets < y_{ht}, y_{it}, y_{jt} > such that $y_{ht} \rightarrow y_{it} \leftarrow y_{jt}$, y_{ht} and y_{jt} not adjacent, y_{ht} and y_{jt} are not adjacent, are such that |Local(C, y_{ht})\{ y_{it} }| < m.

(F)

If there is a quadruple < y_{ht}, y_{it}, y_{jt}, y_{kt} > of distinct vertices in C such that (i) $y_{ht} \rightarrow \underline{y_{it}} \leftarrow y_{jt}$, (ii) $y_{ht} \rightarrow y_{kt} \leftarrow y_j$ or $y_{ht} \rightarrow \underline{y_{kt}} \leftarrow y_j$, (iii) y_{it} and y_{kt} are adjacent, then orient y_{it} *—* y_{kt} as $y_{it} \rightarrow y_{kt}$ in C if y_{kt} does not exist in the Supset< y_{ht}, y_{it}, y_{jt} >

Else orient as y_{it} *—* y_{kt} as y_{it} *— y_{kt}.

(G)

For each quadruple < y_{ht}, y_{it}, y_{jt}, y_{kt} > in C of distinct vertices s.t. y_{kt} is not adjacent to both y_{ht} and y_{jt}, and $y_{ht} \rightarrow \underline{y_{it}} \leftarrow y_{jt}$, if y_{ht} is independent of y_{kt} conditioned on Supset< y_{ht}, y_{it}, y_{jt} > union with < y_{kt} >, then orient y_{it} *—* y_{kt} as $y_{it} \rightarrow y_{kt}$ in C.

A.4 GREEDY EQUIVALENCE SEARCH (GES)

The GES is a two step Algorithm that starts with a DAG representation with no edges. It then moves forward by adding edges which are scored through a Bayesian criterion. The algorithm stops when the Bayesian score can no longer be increased. Thus, in the first step, all the possible single edge additions are created. Then the resulting graphs are categorized by equivalence classes as defined by Chickering (2002), and scored with a Bayesian criterion. The equivalence class that scores the highest is retained whereas the others are discarded. There may be several graphs in the resulting equivalence group. Those edges that are common to each of these graphs, but elicit a different causal direction, will remain undirected. This is repeated until no edge additions can increase the score. The second stage starts by deleting single edges from the graph that resulted from the previous stage. If these deletions increase the score, then the change is retained and the solution is found.

A.5 ALGORITHM USED IN THIS DISSERTATION

The differences between the three previous algorithms are very subtle because we only used three variables. The *PC* and Moneta's algorithms differ only on the *sepset* definition. More precisely, the *Sepsets* in *PC* are a subset of the *Sepsets* in Moneta's. Richardson's algorithm is truncated to step (D) if only three variables are considered, which is an extension of the PC Algorithm that allows the direction of single edges. Therefore, both Richardson's algorithm and Moneta's algorithm were used in this dissertation.

To ease programming, the algorithm implemented in this research starts by computing all the possible correlations and partial correlations and testing them for significance to generate a large matrix that contains all Moneta's *Sepsets*. Then, the algorithm continues as Richardson's but changes the *Sepset* definition and avoids step (D) if Moneta's results are sought. The Algorithm is described below.

(A)
Form the complete undirected graph Partial Ancestral Graphs C on the vertex set $y_{1t}, \ldots, y_{kt}$. Let Adjacencies(C, y_{it}) be the set of vertices adjacent to y_{it} in C; let G_{Y_t} be the (unobservable) causal structure among the k elements of Y_t;

(B)
For all pairs of vertices (y_{it}, y_{jt}), test unconditional and conditional correlations on all possible combinations of the remaining variables. Record results in a *Matrix of Sepsets.*

n = 0
Repeat :
Repeat :
Select an ordered pairs of variables y_{ht} and y_{it} that are adjacent in C such that *Adjacencies(C,* y_{ht} *)\{*y_{it} *}* has cardinality greater than or equal to n, and a subset S of *Adjacencies(C,* y_{ht} *)\{*y_{it} *}* of cardinality n, and if y_{ht} and y_{it} are independent given S in G_{Yt} delete edge $y_{ht} \circ\!\!-\!\!\circ\ y_{it}$;
until all ordered pairs of adjacent variables y_{ht} and y_{it} such that *Adjacencies(C,* y_{ht} *)\{*y_{it} *}* has cardinality greater than or equal to n and all subsets S of *Adjacencies(C,* y_{ht} *) \{*y_{it} *}* of cardinality n have been tested for independence;
$n = n + 1$;
until for each ordered pair of adjacent variables y_{ht} , y_{it} , *Adjacencies(C,* y_{ht} *)\{*y_{it} *}* is of cardinality less than n;

(C)
For each triple of vertices y_{ht} , y_{it} , y_{jt} such that the pair y_{ht} , y_{it} and the pair y_{it} , y_{jt} are each adjacent in C but the pair y_{ht} , y_{jt} is not adjacent in C, orient $y_{ht} *\!\!-\!\!* y_{it} *\!\!-\!\!* y_{jt}$ as $y_{ht} \rightarrow y_{it} \leftarrow y_{jt}$ if and only if y_{it} **is not in *Sepset* (y_{ht}, y_{jt}) as defined by the *PC Algorithm* and stored in the *Matrix of Sepsets*, or as defined by Moneta and stored in the *Matrix of Sepsets*.**

(D)
If the Moneta version of *Sepsets* is not being used, then for each triple of vertices y_{ht} , y_{it} , y_{jt} such that (a) y_{ht} is not adjacent to y_{it} or y_{jt}, (b) y_{it} and y_{jt} are adjacent, (c) y_{it} does not exist in the *Sepset* (y_{ht}, y_{jt}), then orient $y_{it} *\!\!-\!\!* y_{jt}$ as $y_{it} \leftarrow y_{jt}$ if y_{ht} is independent of y_{it} conditioned on *Sepset* (y_{ht}, y_{jt}).

APPENDIX B

LIST OF MANUFACTURERS AND CHARACTERISTIC PRODUCTS

This Appendix describes the manufacturers analyzed in this dissertation. The manufacturers are categorized by product type, and information is provided to describe their representative products, number of associated UPCs, and how many of these UPC's had enough observations to be transformed into stationary series and analyzed.

Table B.1 List of Manufacturers and Characteristic Products.

Category/ Manufacturer	*Representative Products*	*Number of UPCs*	*Usable Series*
BEER			
m1	Budweiser/Michelob	88	15
m4	Miller	100	31
m6	Augsburguer/Old Milwakee	36	4
m7	Labatta	6	1
m8	Molson	9	4
m10	Coors/Keystone/Blue Moon	49	7
m11	Strohs	17	3
m12	Pilsner/Moosehead/Dos Equis	15	3
m13	Heineken	8	4
m14	Old Style	101	15
m15	Corona	14	3
m18	Beck's	11	5
m20	Berghoff/Augsburguer	8	3
m22	Samuel Adams	32	2
m25	Goose Island Honkers	6	1
m27	Oregon Brewery India	7	1
CHEESE			
m1	Cty Ln Colby Mild	10	6
m3	Lifeway's Farmer's C	2	1

Table B.1 Continued.

Category/ Manufacturer	*Representative Products*	*Number of UPCs*	*Usable Series*
m4	Kraft Colby 1/3 Less	237	115
m9	Lol Cheddarella Chee	6	5
m12	Dom Amer Chs Food Tw	59	41
m14	Frigo Shred Ched	12	2
m15	Laughing Cow/Bonbel	9	2
m16	Lean N' Free /Alpine	8	5
m17	Treasure Cave Square	4	2
m18	Sargento Wafer Thin	55	31
m19	Borden Lt Ln Chdr	19	5
SOFT DRINKS			
m2	Old Town Nat Sltzer	42	3
m3	Tetley Iced Tea W/Le	7	1
m4	Pepsi-Cola Cans	141	81
m5	Schwepps Tonic N/R	37	21
m6	Canada Dry Ginger Al	40	26
m7	Hawaiian Punch Gldn	9	6
m8	Fruitopia Citrus Con	14	6
m9	Royal Crown Cola	101	45
m11	Ocean Spry	16	9
m12	World Classics Cola	9	8
m13	Dominick's Cola 3 Lt	51	40
m14	Liptn Brew Wild Strw	27	17
m15	PS	6	1
m16	Nu Grape Soda Ppd	9	3
m17	Crush Different Flavors	24	11
m18	Country Time	16	12
m19	Sunkist	21	14
m21	Coca-Cola	143	87
m22	New York Seltzer	24	1
m23	Hawiian Punch Red 8p	12	7
m24	Hawaiian Punch /Sunny Delight	17	11
m26	Sundance Black Curra	24	1
m27	Dr Pepper Sugar Free	25	11
m28	Clearly Canadian	38	15
m29	A & W Root Beer (Can	15	12
m31	Boku	6	2
m32	Vernors / Artic Twist	22	6
m33	Barq's Root Beer	18	9

Table B.1 **Continued.**

Category/ Manufacturer	*Representative Products*	*Number of UPCs*	*Usable Series*
m34	Dad's Root Beer Tria	19	11
m35	Ibc Root Beer Trial	9	2
m36	Holy Cow	12	1
m37	Seagram's Ginger Ale	13	8
m38	Lacroix Orange Miner	50	11
m40	Perrier Berry Nr	15	1
m41	C/G Reg Sprk Mineral	32	7
m43	Snapple Peach Tea	61	38
m44	Welch Grape	11	3
m46	A W Cream Soda Reg	10	6
m47	7-Up	36	24
m51	Canfield Fruit Punch	115	19
m53	Nestea 6pk Cans	12	5
m56	Mistic	44	4
TOILET PAPER			
m1	Angel Sft Bth Tissue	8	4
m2	Kleenex Pp1.09	20	9
m3	Charmin Bath Tissue	31	14
m4	Dominick's 2ply Pr B	8	3
m5	Northern Quilted Tis	19	13
m6	Kleenex Cottonell Ba	33	19
m7	Generic Bath Tissue	3	1
TUNA			
m1	C O S	11	9
m3	King Oscar Kipper Sn	9	7
m4	Dom/Hh	9	5
m5	Orleans Dom Med Deve	14	7
m6	Booth Ezo Sardines I	5	3
m8	C O S	29	17
m9	Romanoff Iceland Lum	8	1
m10	Reese Flat Anchovies	20	5
m11	King Oscar Sardines	14	5
m12	Romanoff Whitefish C	9	2
m13	Pillar Rock Pink Sal	5	4
m14	Polar Smoked Mussels	16	8
m15	Star Kist Chunk Ligh	13	2

Table B.1 Continued.

Category/ Manufacturer	*Representative Products*	*Number of UPCs*	*Usable Series*
m16	Starkist Solid White	25	13
m17	Bum Bee Oil Chk Lite	22	15
m18	Snow's Chopped Clams	3	2

APPENDIX C

LIST OF MANUFACTURERS EXCLUDED FROM THE ANALYSIS

This Appendix provides a list of those manufacturers whose series do not meet the minimum statistical requirements for this study such as number of observations and stationarity either in levels or achieved through transformations. Therefore, these manufacturers were excluded from the analysis.

Table C.1 List of Manufacturers Excluded from the Analysis.

Category/ Manufacturer	*Representative Products*	*Manufacturer*	*Representative Products*
BEER			
m2	Hook/Ballard	m17	Warsteiner
m3	Pabst/Hamms	m20	Berghoff/Augsburguer
m5	New Amsterdam	m23	Tecate
m8	Molson	m24	Rogue
m12	Pilsner/Moosehead/Dos Equis	m25	Goose Island Honkers
m15	Corona	m26	Shipyard Export Ale
m16	Cold Springs/Naked Aspen	m27	Oregon Brewery India
CHEESE			
m2	Kaukauna Ranch Balls	m23	Maybud Gerard Brie T
m3	Lifeway's Farmer's C	m24	Friendship Farmers C
m5	Hoffman Hot Pepper S	m25	Manischewitz Red Hor
m6	Migdalamer White Che	m26	Swiss Knight Cheese
m7	Mothers White Horser	m27	Win Schuler Cheese B
m8	Schneiders String C	m28	Sorrento 16 Oz Mozza
m10	Koshure Lite Muenste	m29	W.W Low Sodium Chees
m11	Trad Cho Cho Chip Ck	m30	Stella Mozzarella Sq
m13	Dom /Generic	m31	Cf
m16	Lean N' Free /Alpine	m32	Formagg Yellow Amer
m17	Treasure Cave Square	m33	Miller's Sl American

Table C.1 Continued.

Category/ Manufacturer	*Representative Products*	*Manufacturer*	*Representative Products*
m20	Rondele-Sliced Frenc	m34	Lc Shipper/Wedge-Com
m21	County Line Indiv W	m35	Kr Velveeta Loaf Lig
m22	Dorman No Salt Swiss	m36	Trad Cho Cho Chip Ck
SOFT DRINKS			
m1	Orenzada, Lemoniada	m31	Boku
m5	Schwepps Tonic N/R	m32	Vernors / Artic Twist
m6	Canada Dry Ginger Al	m36	Holy Cow
m10	Old Tyme	m39	Faygo
m11	Ocean Spry	m42	Poland Springs
m15	PS	m43	Snapple Peach Tea
m17	Crush Different Flavors	m45	Chapelle
m18	Country Time	m46	A W Cream Soda Reg
m19	Sunkist	m47	7-Up
m20	Tropicana	m48	Naturale
m22	New York Seltzer	m49	Everlast
m23	Hawiian Punch Red 8p	m50	Corr
m24	Hawiian Punch/Sunny D	m52	Penafield
m25	Ruby Red / Squirt	m54	Quest
m26	Sundance Black Curra	m55	Arizona Teas
m28	Clearly Canadian	m56	Mistic
m30	Equator	m57	Arizona Assorted Flavors
TOILET PAPER			
m2	Kleenex Pp1.09	m5	Northern Quilted Tis
m3	Charmin Bath Tissue	m6	Kleenex Cottonell Ba
TUNA			
m2	Romanof Gldn Lumpfis	m8	C O S
m5	Orleans Dom Med Deve	m9	Romanoff Iceland Lum
m6	Booth Ezo Sardines I	m10	Reese Flat Anchovies
m7	Orleans Whole Oyster	m13	Pillar Rock Pink Sal

APPENDIX D

EMPIRICAL ANALYSIS RESULTS (VERTICAL COMPETITION)

This appendix presents the results from the empirical case study. The results are organized by the type of causal structure found in each of the five studied categories; Beer, Cheese, Soft Drinks, Toilet Paper, and Tuna. Therefore, the following tables present all the possible causal relationships that can be obtained by the algorithms used in this research. Namely, all the possible colliders Pm→q←Pr, q→Pr←Pm, Pr→Pm←q, which are respectively associated with either monopolistic retailers with vertical coordination/integration, manufacturers as a Stackelberg leader, or Retailer as a Stackelberg leader. Finally, those causal relationships which were not identified are also presented. In all instances, a sample of each of the manufacturers' representative products is indicated along with the profit margin that they provide to the retailer, which expressed as a percent of the manufacturers' average revenues. The market share, and the average value of this share, is also presented for each of the manufacturers.

Table D.1 Pm→q←Pr : Monopolistic Retailer with Vertical Coordination/Integration.

Category	*Manufacturer*	*Representative Products*	*(%) Margin*	*(%) Market Share*	*(%) Market Share Value*
BEER					
	m7	Labatt	16.20	0.03	0.05
	m9	Foster's	31.61	0.11	0.21
CHEESE					
	m12	Dom Amer Chs Food Tw	63.63	28.43	25.34
SOFT DRINKS					
	–	–	–	–	–
TOILET PAPER					
	–	–	–	–	–
TUNA					
	–	–	–	–	–

Table D.2 q→Pr←Pm : Manufacturer as a Stackelberg Leader.

Category	*Manufacturer*	*Representative Products*	*(%) Margin*	*(%) Market Share*	*(%) Market Share Value*
BEER					
	m1	Budweiser/Michelob	11.83	16.98	17.10
	m4	Miller	10.57	41.68	40.56
	m10	Coors/Keystone/Blue Moon	10.94	2.22	2.36
	m11	Strohs	14.61	1.12	0.89
	m13	Heineken	14.06	2.35	4.50
	m14	Old Style	-0.61	26.46	20.96
	m18	Beck's	10.30	2.19	3.88
	m19	Bass/Guinness/Harp/Red Stripe	15.66	0.10	0.22
	m21	Pete's	22.51	0.05	0.08
	m22	Samuel Adams	13.04	1.10	1.85
CHEESE					
	m1	Cty Ln Colby Mild	55.75	0.95	1.29
	m4	Kraft Colby 1/3 Less	35.62	59.62	59.67
	m14	~Frigo Shred Ched	53.31	0.20	0.25
	m18	Sargento Wafer Thin	50.82	5.42	8.26
SOFT DRINKS					
	m8	Fruitopia Citrus Con	43.56	0.02	0.04
	m9	Royal Crown Cola	15.23	9.09	8.42
	m12	World Classics Cola	28.42	0.17	0.13
	m13	Dominick's Cola 3 Lt	35.42	10.17	6.60

Table D.2 Continued.

Category	*Manufacturer*	*Representative Products*	*(%) Margin*	*(%) Market Share*	*(%) Market Share Value*
	m21	Coca-Cola N/R	13.10	28.63	29.10
	m27	Dr Pepper Sugar Free	9.05	1.78	1.94
	m29*	A & W Root Beer (Can	17.99	0.85	0.85
	m33	Barq's Root Beer	14.31	0.40	0.41
	m34	Dad's Root Beer Tria	9.52	0.71	0.66
	m35	Ibc Root Beer Trial	30.94	0.05	0.10
	m37	Seagram's Ginger Ale	19.98	0.14	0.15
	m38	Lacroix Orange Miner	39.61	0.19	0.36
	m41	C/G Reg Sprk Mineral	29.21	0.06	0.14
	m51	Canfield Fruit Punch	18.71	1.15	0.97
	m53	Nestea 6pk Cans	19.76	0.05	0.09
TOILET PAPER					
	m1	Angel Sft Bth Tissue	21.38	6.22	4.96
	m4	Dominick's 2ply Pr B	30.39	8.30	7.68
TUNA					
	m1	C O S	32.28	5.43	4.79
	m3	King Oscar Kipper Sn	47.04	1.72	4.79
	m4	Dom/Hh	32.03	9.80	6.78
	m11	King Oscar Sardines	36.56	4.24	6.50
	m14	Polar Smoked Mussels	44.06	1.68	2.90
	m16	Starkist Solid White	32.67	28.00	25.48

Table D.2 Continued.

Category	*Manufacturer*	*Representative Products*	*(%) Margin*	*(%) Market Share*	*(%) Market Share Value*
	m17	Bum Bee Oil Chk Lite	32.52	19.90	17.70
	m18	Snow's Chopped Clams	77.50	0.22	0.27

Note.

* Conflicting result: Each of the two algorithms elicit a different causal pattern.

Table D.3 Pr→Pm←q: Retailer as a Stackelberg Leader

Category	*Manufacturer*	*Representative Products*	*(%) Margin*	*(%) Market Share*	*(%) Market Share Value*
BEER					
	m6	Augsburger/Old Milwaukee	21.22	1.73	1.68
CHEESE					
	m9**	LOL Cheddarella Chee	42.28	0.98	0.89
	m15*	Laughing Cow/Bonbel	45.41	0.23	0.48
	m19**	Borden Lt Ln Chdr	26.71	2.40	2.05

Table D.3 Continued.

Category	*Manufacturer*	*Representative Products*	*(%) Margin*	*(%) Market Share*	*(%) Market Share Value*
SOFT DRINKS					
	m2	Old Town Nat Sltzer	28.55	0.19	0.13
	m4	Pepsi-Cola Cans	14.29	26.22	27.66
	m7	Hawaiian Punch Gldn	36.33	0.13	0.15
	m14	Liptn Brew Wild Strw	19.00	0.67	0.98
	m16	Nu Grape Soda Ppd	25.92	0.06	0.06
	m40	Perrier Berry Nr	38.66	0.03	0.09
	m44	Welch Grape	25.23	0.10	0.09
TOILET PAPER					
	m7	Generic Bath Tissue	36.38	2.32	1.30
TUNA					
	m12	Romanoff Whitefish C	43.44	0.04	0.42
	m15	Star Kist Chunk Ligh	33.64	0.60	0.38

Note.

* This manufacturer had a one-edge structure: Pr→Pm.

** Moneta's algorithm proved to be unreliable for these manufacturers.

Table D.4 Undirected Structures.

Category	*Structure*	*Manufacturer*	*Representative Products*	*(%) Margin*	*(%) Market Share*	*(%) Market Share Value*
BEER						
	q – Pr – Pm	m8	Molson	9.81	0.46	0.67
	q – Pr – Pm – q	m12	Pilsner/Moosehead/Dos Equis	13.69	0.43	0.67
	q – Pr – Pm	m15	Corona	15.19	1.45	2.57
	q – Pr – Pm – q	m20	Berghoff/Augsburger	11.50	1.20	1.35
	q – Pr	m25	Goose Island Honkers	25.99	0.10	0.17
	q – Pr – Pm – q	m27	Oregon Brewery India	11.36	0.07	0.11
CHEESE						
	Pr – Pm – q	m13	Dom /Generic	44.78	0.34	0.21
	Pr – Pm	m16	Lean N' Free /Alpine	60.70	0.18	0.26
	Pr – Pm	m17	Treasure Cave Square	62.04	0.20	0.39
SOFT DRINKS						
	q – Pr – Pm	m5	Schwepps Tonic N/R	18.28	1.29	1.44
	Pr – Pm – q	m6	Canada Dry Ginger Al	30.74	2.20	2.12
	Pr – Pm	m11	Ocean Spray Cranberr	50.06	0.03	0.08
	q – Pr – Pm – q	m17	Orange Crush	14.87	1.13	1.09
	Pr – Pm	m18	Countrytime Lemonade	34.31	0.12	0.23
	q – Pr – Pm – q	m19	Sunkist Lemonade	22.17	1.88	1.64
	Pr – Pm	m22	New York Sltzr Diet	34.45	0.01	0.02
	Pr – Pm	m23	Hawiian Punch Red 8p	3.39	0.28	0.31
	q – Pr – Pm – q	m24	Ruby Red/Squirt	17.03	0.75	0.83

Table D.4 Continued.

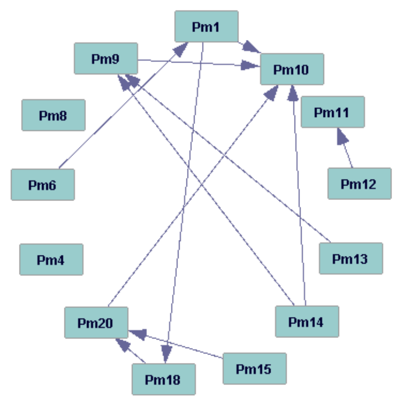
Category	*Structure*	*Manufacturer*	*Representative Products*	*(%) Margin*	*(%) Market Share*	*(%) Market Share Value*
	q – Pr – Pm	m26	Sundance Black Curra	36.05	0.01	0.03
	q – Pr – Pm – q	m28	Clearly Canadian Bla	34.76	0.09	0.33
	q – Pr – Pm – q	m32	Vernors Diet	17.50	0.15	0.17
	q – Pr – Pm – q	m43	Snapple Peach Tea	31.78	0.57	1.29
	Pr – Pm	m46	A W Cream Soda Reg	26.57	0.19	0.18
	q – Pr – Pm – q	m47	Seven-Up Diet	20.37	10.35	10.98
	q – Pr	m49	Everlast Mixed Berry	23.16	0.01	0.02
	Pr – Pm	m56	Mistic Breeze Grape/	44.99	0.02	0.04
TOILET PAPER						
	q – Pr – Pm	m2	Kleenex Pp1.09	16.88	5.73	6.77
	q – Pr – Pm – q	m3	Charmin Bath Tissue	14.98	28.10	26.21
	Pm – q – Pr	m5	Northern Quilted Tis	12.64	17.52	15.84
	q – Pr – Pm – q	m6	Kleenex Cottonell Ba	12.68	31.81	37.25
TUNA						
	Pr – Pm	m5	Orleans Dom Med Deve	34.95	0.41	1.09
	q Pr – Pm – q	m6	Booth Ezo Sardines I	51.03	0.87	1.14
	q – Pr – Pm – q	m8	C O S	36.92	19.15	16.90
	q Pr Pm	m9	Romanoff Iceland Lum	43.80	0.01	0.14
	q – Pr – Pm q	m13	Pillar Rock Pink Sal	36.38	7.22	8.39
	q – Pr – Pm – q	m10	Reese Flat Anchovies	41.35	0.71	2.31

APPENDIX E

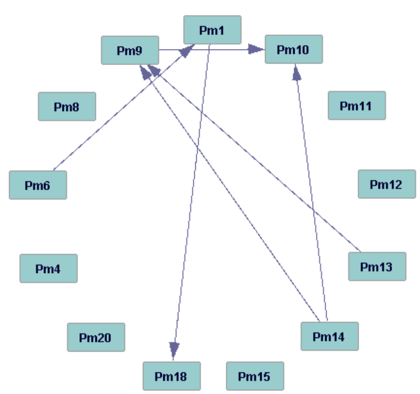
EMPIRICAL ANALYSIS RESULTS (HORIZONTAL COMPETITION)

The following pages show the results from the empirical analysis of the horizontal competition among the manufacturers. Each category contains two graphs. In all instances graph a) is obtained with the use of the CCD algorithm. This causal structure is compared with one obtained from the GES algorithm. Those edges that are common in both causal structures are shown in graph b), which are the ones considered as reliable. The representative products for each manufacturer are presented in all instances, along with their market share and their market share value, both expressed as percents.

Figure E.1 Causal Relationships among Manufacturers in the Beer Category.



a) Causality at $\alpha = 5\%$

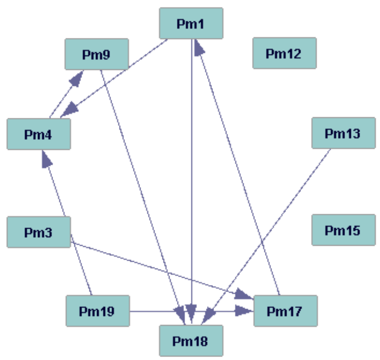


b) Reliable causal structure

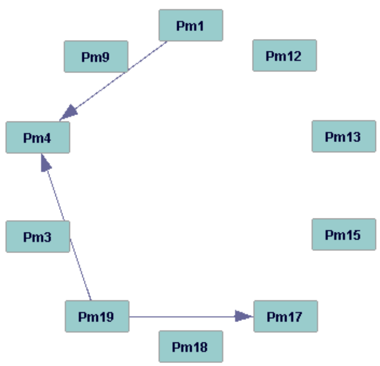
Manufacturer	*Representative Products*	*(%) Market Share*	*(%) Market Share Value*
m1	Budweiser/Michelob	16.98	17.1
m4	Miller	41.68	40.56
m6	Augsburguer/Old Milwakee	1.73	1.68
m8	Molson	0.46	0.67
m9	Foster's	0.11	0.21
m10	Coors/Keystone/Blue Moon	2.22	2.36
m11	Strohs	1.12	0.89
m12	Pilsner/Moosehead/Dos Equis	0.43	0.67
m13	Heineken	2.35	4.5
m14	Old Style	26.46	20.96
m15	Corona	1.45	2.57
m18	Beck's	2.19	3.88
m20	Berghoff/Augsburguer	1.2	1.35

Note. Sample size: 146 observations.

Figure E.2 Causal Relationships among Manufacturers in the Cheese Category.



a) Causality at α = 10%

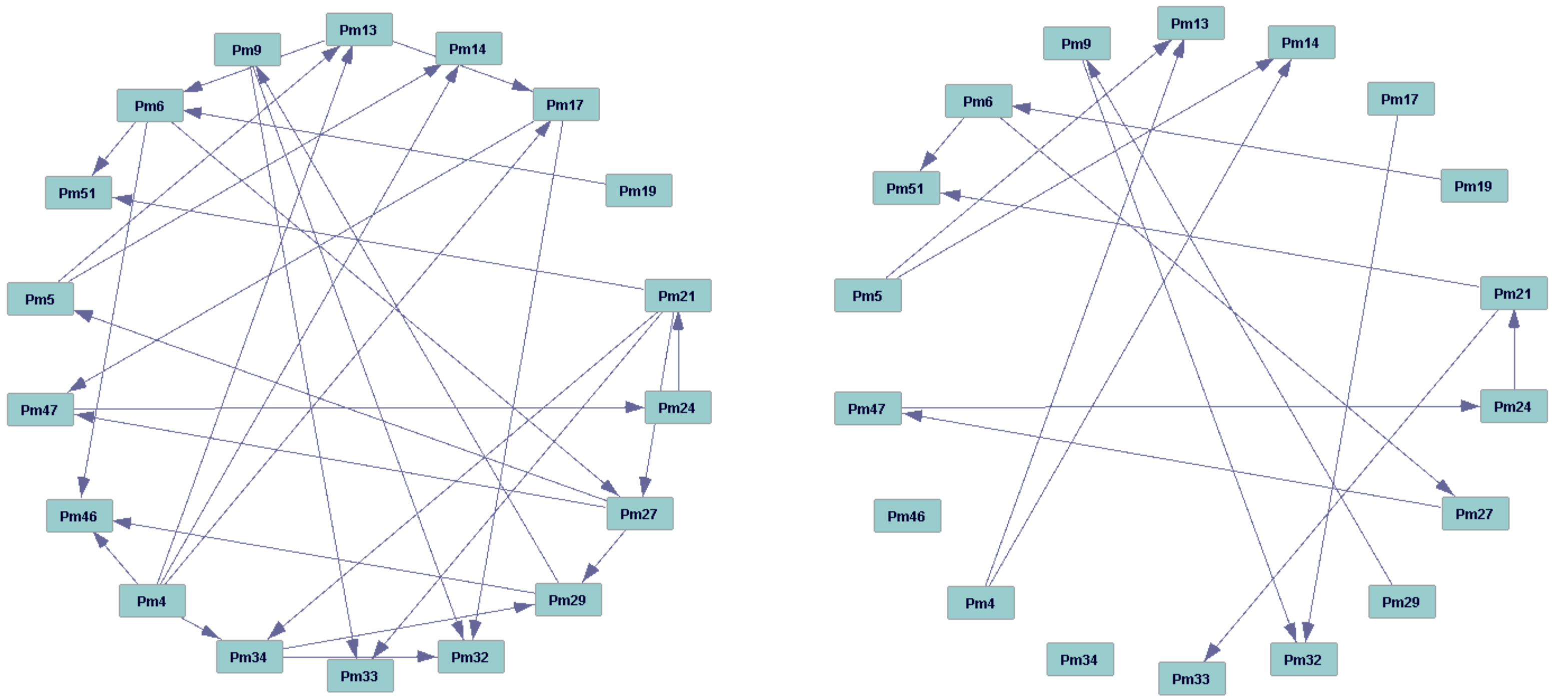


b) Reliable causal structure

Manufacturer	*Representative Products*	*(%) Market Share*	*(%) Market Share Value*
m1	Cty Ln Colby Mild	0.95	1.29
m3	Lifeway's Farmer's C	1.03	0.87
m4	Kraft Colby 1/3 Less	59.62	59.67
m9	Lol Cheddarella Chee	0.98	0.89
m12	Dom Amer Chs Food Tw	28.43	25.34
m13	Dom /Generic	0.34	0.21
m15	Laughing Cow/Bonbel	0.23	0.48
m17	Treasure Cave Square	0.2	0.39
m18	Sargento Wafer Thin	5.42	8.26
m19	Borden Lt Ln Chdr	2.4	2.05

Note. Sample size: 90 observations.

Figure E.3 Causal Relationships among Manufacturers in the Soft Drinks Category.



a) Causality at α = 5%

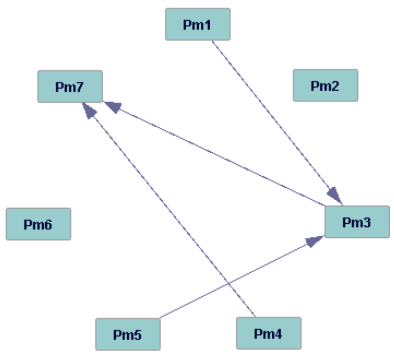
b) Reliable causal structure

Figure E.3 Continued.

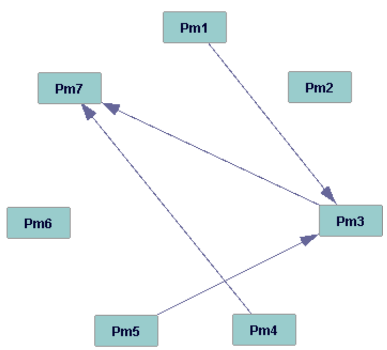
Manufacturer	*Representative Products*	*(%) Market Share*	*(%) Market Share Value*
m4	Pepsi-Cola Cans	26.22	27.66
m5	Schwepps Tonic N/R	1.29	1.44
m6	Canada Dry Ginger Al	2.2	2.12
m9	Royal Crown Cola	9.09	8.42
m13	Dominick's Cola 3 Lt	10.17	6.6
m14	Liptn Brew Wild Strw	0.67	0.98
m17	Crush Different Flavors	1.13	1.09
m19	Sunkist	1.88	1.64
m21	Coca-Cola N/R	28.63	29.1
m24	Hawaiian Punch /Sunny Delight	0.75	0.83
m27	Dr Pepper Sugar Free	1.78	1.94
m29	A & W Root Beer (Can	0.85	0.85
m32	Vernors / Artic Twist	0.15	0.17
m33	Barq's Root Beer	0.4	0.41
m34	Dad's Root Beer Tria	0.71	0.66
m46	A W Cream Soda Reg	0.19	0.18
m47	7-Up	10.35	10.98
m51	Canfield Fruit Punch	1.15	0.97

Note. Sample size: 336 observations.

Figure E.4 Causal Relationships among Manufacturers in the Toilet Paper Category.



a) Causality at α = 1%

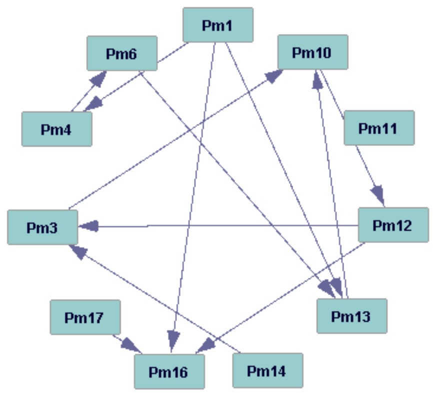


b) Reliable causal structure

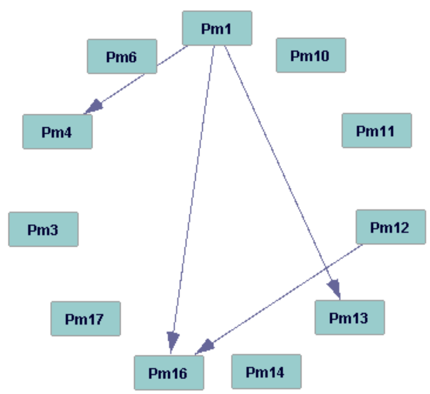
Manufacturer	*Representative Products*	*(%) Market Share*	*(%) Market Share Value*
m1	Angel Sft Bth Tissue	6.22	4.96
m2	Kleenex Pp1.09	5.73	6.77
m3	Charmin Bath Tissue	28.1	26.21
m4	Dominick's 2ply Pr B	8.3	7.68
m5	Northern Quilted Tis	17.52	15.84
m6	Kleenex Cottonell Ba	31.81	37.25
m7	Generic Bath Tissue	2.32	1.3

Note. Sample size: 158 observations.

Figure E.5 Causal Relationships among Manufacturers in the Tuna Category.



a) Causality at α = 5%



b) Reliable causal structure

Manufacturer	*Representative Products*	*(%) Market Share*	*(%) Market Share Value*
m1	C O S	5.43	4.79
m3	King Oscar Kipper Sn	1.72	4.79
m4	Dom/Hh	9.8	6.78
m6	Booth Ezo Sardines I	0.87	1.14
m10	Reese Flat Anchovies	0.71	2.31
m11	King Oscar Sardines	4.24	6.5
m12	Romanoff Whitefish C	0.04	0.42
m13	Pillar Rock Pink Sal	7.22	8.39
m14	Polar Smoked Mussels	1.68	2.9
m16	Starkist Solid White	28	25.48
m17	Bum Bee Oil Chk Lite	19.9	17.7

Note. Sample size: 86 observations.

VITA

Name: Francisco Fraire Domínguez

Address: Department of Agricultural Economics, Texas A&M University, 301D Blocker, College Station, TX 77843-2124

Email Address: pacofraire@gmail.com

Education: B.S., Physics, Instituto Tecnológico y de Estudios Superiores de Monterrey, 1999

M.S., Applied Statistics, Instituto Tecnológico y de Estudios Superiores de Monterrey, 2004

Ph.D., Agricultural Economics, Texas A&M University, 2009

Experience: Lecturer, Department of Agricultural Sciences, Sam Houston State University, Huntsville, TX, 2009

Research Assistant, Texas Extension Agricultural Economics, Texas A&M, College Station, TX, 2008, 2009

Courses Coordinator, ITESM-Thunderbird, School of Global Management, 2004

Financial Analyst and Foreign Exchange Market Trader, Prime International Analysts. C.V., 2001

Physics Laboratory Instructor, Instituto Tecnológico y de Estudios Superiores de Monterrey. 1999

Research: Eliciting strategic behavior within a supply chain

Countries economic dependencies elicitation

Least cost shipping and storage of west Texas cotton

Foreign exchange forecasting with Markov Switching Vector Error Correction models and currency baskets

Analysis of factors influencing grain traffic on St. Lawrence Seaway