KENNYA BEATRIZ SIQUEIRA

THE DYNAMICS OF FARM MILK PRICE FORMATION IN BRAZIL

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Economia Aplicada, para obtenção do título de *Doctor Scientiae*.

VIÇOSA
MINAS GERAIS – BRASIL
2007

To my beloved family and everyone that, somehow, helped me to get through in the USA.

ACKNOWLEDGE

To God for giving me the energy, courage, determination, and providing me with such good friends.

To my family for the love, care, and attention.

To my advisor Professor Antônio Carvalho Campos for supporting all my decisions and giving me the freedom to conduct my work.

To Dr Richard L. Kilmer for the great opportunity to work together, for the professional and emotional support, and for the example of diplomacy, politeness and responsibility.

To Professor and eternal advisor Carlos Arthur Barbosa da Silva for the friendship, partnership, attention, and orientation. To Professor Sebastião Teixeira Gomes for sharing his experience and knowledge about the milk market. To Professors Danilo Rolim Dias de Aguiar, João Eustáquio de Lima, Fátima Marília de A. Carvalho, Brício dos Santos Reis, and all the professors of the Agricultural Economics Department of the Federal University of Viçosa for the wonderful teaching and guidance.

To Dr. Charles Moss for his math mind and methodological support. To Dr Schmidt, Dr Ward, and Dr Spreen for the friendly reception. To Dr William Bomberger for providing software.

Special thanks to Dr David Bessler, Clark Gymour, Joseph Ramsey, and Dr Bruce L. Benson for all the marvelous support provided in the execution of the dissertation.

To Vick, Jess, Bodo, Carole, and Donna for the academic and legal support at the University of Florida.

To friends Robert and family, Christie and family, Jessica and the Myslikowski family, Yuko, Vivian, Wen, Zack, Nestor, Sofiya, Yuri, Patricia, Fabio, Naoko, Jessica, Nadia, Sabine, Hyung Kyung, Hong Tao, Lan, Yvonne, Ivone, Bonnie, Bee, Marco, Caroline, Evandro, Marta, Graziela and family, Ms Ellen, Ilka, Maria, JG, Mohammad Ali, Antonella and Guido, Lucio, and Valéria for being my family in US, and Eduardo, Ritaumária, Renata, and André for our eternal friendliness.

To my beloved aunts Edna, Ione and Dalva for being like a mother for me.

To those from the Departamento de Economia Rural of Universidade Federal de Viçosa: Graça, Cida, Carminha, Luíza, Tedinha, and Rosângela.

To the Universidade Federal de Viçosa and Departamento de Economia Rural for the oportunity to achieve professional growth.

To Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), and Conselho Nacional de Desenvolvimento Científico e Tecnológico for the financial support.

BIOGRAPHY

KENNYA BEATRIZ SIQUEIRA, daughter of Venício Siqueira and Márcia Regina da Silva Siqueira, was born in Varginha, Minas Gerais, in November 4, 1978.

In May 10, 2002, she received her degree in Food Engineering at the Federal University of Viçosa, in Viçosa, Minas Gerais.

In April of 2002, she started the Master Program in Applied Economics at the Federal University of Viçosa. On December 9 of 2003, she defended her dissertation titled "Feasibility of introducing of a milk future contract in Brazil."

In August of 2004, the Brazilian Mercantile & Futures Exchange presented her the award for the best dissertation regarding futures markets in Brazil.

In February of 2004, she started a PHD Program Program in Applied Economics at the Federal University of Viçosa. In February 27, 2006, she attended the University of Florida, in Gainesville, Florida, United States of America, in order to do a Sandwich Program at the Food and Resource Economic Department. On April 18, 2007, she received a Ph.D. degree in Applied Economics.

CONTENTS

Página

LIST OF TABLES

LIST OF FIGURES

RESUMO

SIQUEIRA, Kennya Beatriz, M. Sc., Universidade Federal de Viçosa, abril de 2007. **A dinâmica da formação do preço do leite no Brasil**. Orientador: Antônio Carvalho Campos. Co-orientadores: Richard L. Kilmer, Danilo Rolim Dias Aguiar e Sebastião Teixeira Gomes.

O setor lácteo é um dos maiores setores da economia agrícola do País. Entretanto, este setor tem sofrido significativas mudanças no período pós-liberalização. Por isso, é importante saber quais mudanças ocorreram na integração espacial do mercado e na formação do preço do leite no nível de produtor. Este problema é importante para o desenvolvimento do setor lácteo e ainda não foi estudado no Brasil. O objetivo geral deste trabalho é analisar a dinâmica de formação do preço do leite no Brasil. A teoria adotada é uma versão de Faminow e Benson (1990), desenvolvida para um mercado olipsionista no Brasil. A metodologia é uma modificação de Gonzalez-Rivera & Helfand (2001), a qual é composta de extensão do mercado, padrão de interdependência e grau de integração do mercado. A extensão do mercado é determinada através da medida do índice de auto-suficiência, teste de raiz unitária e procedimento de Johansen. Este último é focado na busca de uma tendência comum entre as séries temporais. O padrão de interdependência é estudado usando a análise do vetor de correção de erros, em associação com Directed Acyclic Graphs. Por último, o grau de integração é medido pelas funções de impulso-resposta derivadas da decomposição de Bernanke. Como resultado, verificou-se que o

mercado lácteo no Brasil é composto por Goiás, Mato Grosso, Mato Grosso do Sul, Minas Gerais, Paraná, Rio de Janeiro, Rio Grande do Sul, Santa Catarina e São Paulo. Nas regiões Norte e Nordeste existem vários mercados locais. Constatou-se também que este mercado tem uma baixa velocidade de ajustamento a choques e não existe um líder na formação de preço no Brasil.

ABSTRACT

SIQUEIRA, Kennya Beatriz, M.Sc. Universidade Federal de Viçosa, April 2007. **The dynamics of farm milk price formation in Brazil**. Adviser: Antônio Carvalho Campos. Co-Advisers: Richard L. Kilmer, Danilo Rolim Dias de Aguiar and Sebastião Teixeira Gomes

Dairy is a highly relevant segment of the Brazilian agribusiness economy. However, this segment has changed significantly after deregulation Thus, it is worth to know what the changes are in the spatial integration of the market and in milk price formation at the farm level. This problem is important to the development of the dairy sector and has not been studied in Brazil. The general objective of this work is to analyze the dynamics of milk price formation in Brazil. The theory adopted is a version of Faminow & Benson (1990), applied for an oligopsony market in Brazil. The methodology is a modification of Gonzalez-Rivera & Helfand (2001), which is compounded in extension of the market, pattern of integration, and degree of integration. The extension of the market is determined through the measure of self-sufficiency index, unit root test, and Johansen test. The last one is focused on the searching for a common trend between the time series. The pattern of interdependence is studied using the VEC/VAR analysis in association with the DAG. Lastly, the degree of integration is measured by the impulse response functions derived from the Bernanke decomposition. As a result, we found that the milk market in Brazil is composed by Goiás, Mato Grosso, Mato Grosso do Sul, Minas

Gerais, Paraná, Rio de Janeiro, Rio Grande do Sul, Santa Catarina, and São Paulo. In the North and Northeast of Brazil, there are local milk markets. We also discovered that these markets have small speed of adjustment to shocks and there is no leader in milk price formation.

1. INTRODUCTION

Dairy is a highly relevant segment of the Brazilian agribusiness economy. In 2004, the Brazilian milk production was roughly 23.48 billion liters, equivalent to US$ 27.45 billion, corresponding to 1.19% of the Gross Domestic Product (GDP) for Brazil. So, dairy farm production is one of the highest generators of employment and growth per unit of investment capital (Embrapa, 2005).

However, there is a significant heterogeneity among dairy farms. Large dairy farms, with elevated productivity indexes, are located among small traditional dairy farms that work with rudimentary techniques of production. It results in great differences between the production systems used by small and large dairy farms. The small dairy farms represent 80% of the producers and are responsible for 20% of national production, while 20% of the dairy farms are classified as large and represent 80% of national production (Gomes, 2003).

Historically, domestic milk consumption has exceeded domestic production. Brazil is characterized as a net importer of dairy products with annual purchases in the order of one billion liters (Milkpoint, 2005). More recently, Brazilian dairy exports, which have been traditionally sporadic and of little significance, have experienced growth.

1.1. Governmental intervention on the Brazilian milk market

Federal governmental intervention in the dairy sector is an important part of Brazilian milk history. In 1945, the government began the intervention at both the farm and retail level. Then, the government began to control the margins of factories profit (Meirelles, 1989). The regulation of the milk price, in association with a closed economy, had a protectionist character, which delayed the modernization of the dairy industry (Gomes, 2003). The absence of competition from imported products and the low level of requirements of consumers resulted in stagnation of the dairy sector. It persisted with lower levels of productivity, elevated costs of production and lower quality products (Martins, 1999).

In 1991, the government decided for the liberalization of the milk segment. After deregulation, the dairy farm production sector showed relevant technological development, which is verified by increased rates of growth in productivity and milk production (Gomes, 2003). However, in the context of post liberalization, high levels of price instability started to be observed at the farm level, leading to repeated demands for government intervention in the transactions between farmers and processors.

Indeed, even after more than one decade of deregulation, significant advances have not yet been observed in the vertical coordination between farm milk production and the distribution chain (Gomes et al., 2002). Typically, there is no price, quantity or any other production or marketing commitment formalized in the supply chain relationships between farmers and processors, like contracts, options, or futures markets. As such, the traditional forms of price risk management cannot exist, and this hinders planning and decision making for all players in the milk chain. By and large, these chain actors have to cope permanently with uncertainty regarding quantities, quality, prices and other related variables (Gomes et al., 2002).

On the other hand, there was a change in milk production regions. In 2002, the largest growth in farm milk production occurred in Northern Brazil, with an increased of 325 million liters. With a yearly output of 29.16% of national production in 2001, Minas Gerais ranks 1st among the country's largest milk producers, followed by Goiás with 11.32%, Rio Grande do Sul with 10.83%, Paraná with 9.21% and São Paulo with 8.69%. Therefore, São Paulo, the traditional milk producer, has decreased in rank (Embrapa, 2002).

The region Center-West has been notable in milk production. This region, mainly Goiás, has shown increasing rates of growth in milk production. In Minas Gerais (region Southeast), farm milk production has relocated from traditional dairy regions, like the South, Southeast, and Zona da Mata to the Triângulo Mineiro and Alto Parnaíba. This relocation was caused by a reduction in production costs, which made Brazilian dairy products more competitive with imports (Tannús, 2001). It represents a possibility of acquiring inputs at lower prices, because dairy farms are closer to the largest grain production regions (Gomes, 1997).

Another important part of the dairy sector is long-life milk, called Ultra High Temperature milk (UHT). UHT was introduced in Brazil in 1972, but it only began to be successful in the 1990s, after the government deregulated the milk markets. Since then, sales have increased an average of 27% per year. The market share of UHT milk in the fluid milk market rose from 4.4% in 1990 to 74% in 2002 and there are presently more than 100 brands on the market (Embrapa, 2003).

So, with the government deregulation, changes have taken place in the dairy industry, in farms (i.e., production location and farm size) and consumers' profile (i.e., product quality), and in processing techniques (i.e., UHT milk). In this context, it is important to identify how the integration is between different regions after the price liberalization.

1.3. Problem statement

What are the changes in the spatial integration of the market and in milk price formation at the farm level after deregulation? This problem is important to the development of the dairy sector and has not been studied in Brazil.

1.4. Objectives

The general objective of this work is to analyze the dynamics of milk price formation in Brazil by identifying where the milk price is formed and what the relationship is between milk prices in different Brazilian states.

The specific objectives are:

- To determine the extention of the milk market in Brazil,

- To evaluate the pattern of interdependence between the Brazilian states in milk production, and
- To measure the degree of integration of the Brazilian milk market.

1.5. Hypothesis

I hypothesize that all the states in Brazil are grouped in only one big market, and Minas Gerais is the state that controls the milk price formation in Brazil.

2. BRAZILIAN MILK MARKET

2.1. World dairy economy

Milk is produced in almost all countries of the world, and has a great economic and nutritional importance. According to FAO (2006), world dairy production in 2005 is estimated in 530,718 million tons, a value that is increasing over the years, due to trade liberalization, decreasing subsidies, and world growth economic.

On the world ranking for milk production in 2005, the seven largest producers are the United States, India, Russia, Germany, France, China, and Brazil (Table 1). Together, these seven countries accounted for 47.1% of global milk production. The United States is the leading producer of milk, followed by India (USDA, 2006). However, the milk production on India has few commercial qualifications, inasmuch as about 50% thereof is derived from buffalo (Hemme et al., 2003).

Table 1 – World ranking for milk production in 2005*.

	Countries	Milk production (million ton)	Percentage of Total	Accumulated
1	United States	80,150	15.1	15.1
2	India	38,500	7.2	22.3
3	Russia	30,600	5.8	28.1
4	Germany	27,600	5.2	33.3
5	France	25,282	4.8	38.1
6	China	24,530	4.6	42.7
7	**Brazil**	23,320	4.4	47.1
8	New Zealand	14,625	2.7	49.8
9	United Kingdom	14,577	2.7	52.5
10	Ukraine	14,000	2.6	55.1
11	Poland	12,400	2.3	57.4
12	The Netherlands	10,531	2.0	59.4
13	Italy	10,500	2.0	61.4
14	Australia	10,150	1.9	63.3
15	México	9,873	1.9	65.2
16	Turkey	9,500	1.8	67.0
17	Pakistan	9,082	1.7	68.7
18	Japan	8,255	1.5	70.2
19	Argentina	8,100	1.5	71.7
20	Canada	8,100	1.5	73.2
	Other countries	141,042	26.8	100,0
	T O T A L	530,718	100,0	

Source: USDA (2006).

*These values include cow, buffalo, sheep, goat, and camel.

In this ranking, Brazil is in the seventh position with 23,320 thousand tons of milk in 2005. The relative share of Brazil's production in the world total has increased in recent decades, from 2.1% in 1970 to 3.1% in 1990 (Souza, 2000). In the 1990s and early 2000s, this growth was further accentuated, reaching 4.6% in 2001. However, in 2005, it is equivalent to 4.4% of the world production of cow milk.

Up to 2004, Brazil had been the sixth in the world ranking. Nowadays, the sixth place belongs to China, which increased production by more than 300% between 2000 and 2005. The growth rate of China is the highest. However, excluding China, Brazil has realized the highest growth rates in the last 10 years. It represents a growth rate 73% higher than the U.S. On the other hand, Russia, Germany, and France have shown negative rates (Gomes, 2006).

According to Meireles (2003), between 1995 and 2001, milk production in Brazil increased 4% per year, while the other countries of South America have a growth rate of 2%, Asia 3.4%, Africa 1.9%, North, and Central America 1.9%, and Europe has decreased 1% per year. So, Brazil was exceeded only by New Zealand and Australia (5.5% per year), which became the most important exporter block on the world[1].

In contrast, the number of cows has decreased in all countries, except for India. Nonetheless, in 2004, the number of cows in Brazil was still 69% higher than in the United States, while the American production was 235% higher than the Brazilian production (Gomes, 2006). It can be explained by the differences in the production systems. In Brazil, extensive systems predominate with supplementary food in the winter. In other words, in Brazil the farms are composed by big green field where the cattle can graze. In the USA, the production is intensive, based on ensilage and concentrate cow feed. However, the average cost of milk production in Brazil is lower than in the USA.

On the other hand, Brazil has a lower rate of productivity with 1,534 liters/cow/year, while the USA presents a productivity of 8,703 liters/cow/year. This occurs because of a large number of small farmers in Brazil, who produce less than 50 L/day, which affects the total productivity (Gomes, 2006).

Comparing milk consumption, Brazil presents a similar consumption pattern to the other largest world producers. In 2004, the Brazilian consumption was 68.4 kg/person/year, compared to 65.3 kg/person/year in France and 61.4 kg/person/year in Germany. Only the USA and Russia have a consumption just over 90 kg/person/year, while India had a lower level of consumption (32.4 kg/person/year).

In the international trade market, New Zealand, European Union, Australia, and United States are the greatest exporters of milk with 34%, 31%, 15%, and 4%, respectively. However, it is one of the commodities that has more subsidies. According to Neves & Consoli (2006), the subsidies for milk are approximately 59% more than for meat.

Brazil has always been characterized as milk importer. But, in 2004 for the first time in history, Brazilian exports surpassed the Brazilian imports (Table 2).

[1] These values include cow, buffalo, sheep, goat, and camel milk.

Table 2 – Milk imports, exports, and per capita consumption between 1990 and 2004.

Year	Imports (thousand liters)	Exports (thousand liters)	Per capita consumption (liters/person/year)
1990	906,000	737	106.34
1991	1,313,000	3,502	111.62
1992	276,000	17,411	107.50
1993	632,000	66,908	106.59
1994	1,250,000	6,312	110.76
1995	3,200,000	9,650	126.20
1996	2,450,000	36,699	133.24
1997	1,930,000	20,284	128.89
1998	2,270,000	14,140	129.49
1999	2,410,000	20,731	130.89
2000	1,800,000	42,080	126.77
2001	808,000	84,270	123.18
2002	1,468,000	142,340	131.53
2003	554,000	173,360	129.90
2004	350,000	400,000	130.48

Source: Embrapa Gado de Leite (2006).

Between 1994 and 2000, Brazil had a period of high milk imports (Table 2). However, after 2002, it has changed abruptly. During the same period, exports have increased in response to the global consumption growth and favorable exchange taxes. Among the main buyers of the Brazilian milk are Iraq, Algeria, and Angola (Neves & Consoli, 2006).

On the other hand, per capita consumption had an increase in 1994 after the Plano Real[2]. It is explained by the increase in income, caused by the Plano Real. Nevertheless, during the remainder of the period, per capita consumption has plateaued.

2.2. National dairy production

In Brazil, the milk production generated approximately R$ 23.48 billion liters, in 2004 (Neves & Consoli, 2006). It is present on approximately 40% of the farms and distributed in all the national territory (Nogueira Netto et al., 2003, cited by Fassio et al,

[2] Plano Real was a political plan implemented in July 1, 1994. The objective of this plan was to stabilize the economy, through the control of the inflation.

2005). According to Leite Brasil (2006), the milk segment is so significant that just cheese consumers spent R$ 5.54 billions, which is equivalent to more than half of the consumer's citrus value. Figure 1 shows the production of milk in Brazil.

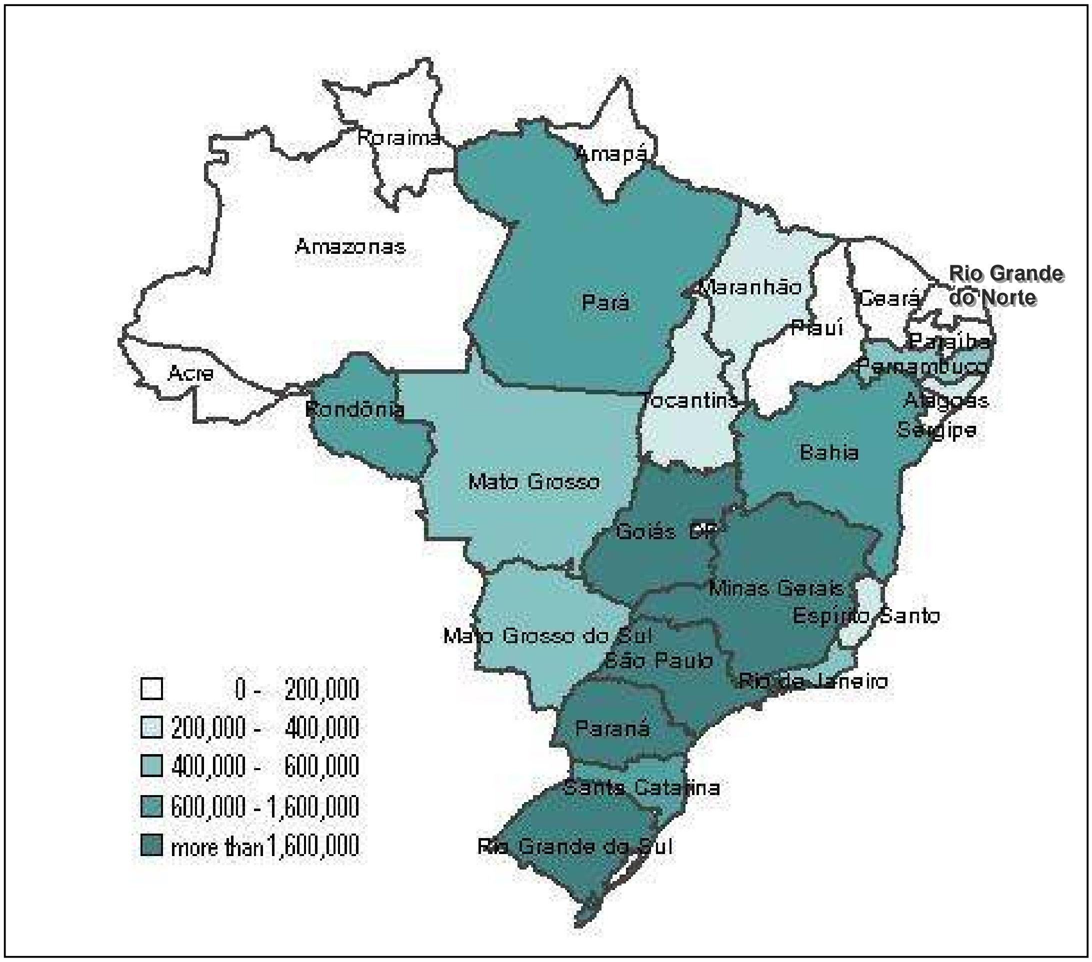


Figure 1 – Production of milk production in the Brazilian states, 2004.
Source: Developed by the author

In the last 10 years, the milk production has increased considerably in Brazil along with the number of cows and productivity in Brazil (Table 3).

Table 3 – Milk production, number of cows and productivity between 1980 and 2005.

Year	Milk production (thousand liters/year)	Cows (thousand heads)	Productivity (liters/cow/year)
1980	11,956	16,513	724
1981	11,675	16,492	708
1982	11,816	16,387	721
1983	11,818	16,276	726
1984	12,303	16,743	735
1985	12,453	17,000	732
1986	12,879	17,600	732
1987	13,399	17,774	754
1988	13,941	18,054	772
1989	14,532	18,673	778
1990	14,933	19,073	783
1991	15,547	19,964	779
1992	16,273	20,476	795
1993	16,074	20,023	803
1994	16,273	20,068	811
1995	16,985	20,579	825
1996	19,089	16,274	1,173
1997	19,245	17,048	1,129
1998	19,273	17,281	1,115
1999	19,661	17,396	1,130
2000	20,380	17,885	1,139
2001	21,146	18,194	1,162
2002	21,643	19,005	1,139
2003	22,254	19,256	1,156
2004	23,475	20,023	1,172
2005 (*)	24,762	20,820	1,189

Source: Embrapa Gado de Leite (2006).

* estimative

According to Gomes (2006), the geometric growth rate of milk production in the 1990's was 3.26%, while between 2000 and 2004 it was 4.48%. Productivity increased 3.98% between 1995 and 2004.

Milk production in Brazil is characterized by heterogeneity among the farmers. There are traditional and small farmers who utilize rudimentary techniques of production, as well as modern farmers, which use high technology and have higher indexes of productivity.

According to Sbrissia (2005), it is hard to know the exact number of dairy farmers in Brazil because of the heterogeneity among them and because of the huge dimension of the country. The rural census of 1996 in Brazil declared that there were 1.81 thousand dairy farmers. However, data from 15 leading dairy companies in Brazil asserted that there were 95,847 dairy farmers in 2002. Nowadays, the modern farmers correspond to 20% of the total farmers and are responsible for 80% of the production. On the other hand, the traditional farmers are 80% of the total and produce about 20% of the national production (Table 4).

Table 4 – Number and production of the farmers associated to Itambé, in 2004.*

Range of production (liters/day)	Farmers (%)	Production (%)
Less than 50	21.44	1.72
50 a 100	19.79	4.21
100 a 200	22.45	9.36
200 a 500	19.63	17.55
500 a 1.000	8.20	16.86
More than1.000	8.49	50.30

Source: Gomes (2006).

* Biggest Brazilian dairy cooperative, located in Minas Gerais.

In 2004, 41% of the Itambé's farmers produced up to 100 liters/day, which is equivalent to only 6% of total production. Nevertheless, 17% of the farmers produced more than 500 liters/day, which corresponds to 67% of the volume of milk received by Itambé. So, it is a good example of the inequality of farmers in Brazil.

In addition, studies point out that this heterogeneity has increased in the last years. Gomes (2006) states that from 1980 to 2004, the percentage of the farmers who produce up to 50 liters/day changed from 27% to 2%. In the same period, the relative share of farmers that produce more than 500 liters/day changed from 25% to 67%.

Sbrissia (2005) affirms that the milk segment in Brazil is in a period of change through which other countries such as New Zealand had already passed. The number of farms is dropping, the number of cows and size of the farms are increasing, while the price is decreasing and more efficiency and scale economies are required. So, a concentration process is happening in milk production in Brazil. Nevertheless, it may be a result of the new system of milk production and trade in Brazil.

First of all, the new rules to collect and cool the milk imposed by the Normative Instructive 51 (IN 51) (promoted by Agriculture Ministry) caused an impact on milk production. This rule aimed to improve quality of the milk in Brazil. One of the requirements of the IN 51 is that milk be kept in cooling tanks on farms. So, some small farmers choose to associate with others in using of the same cooling tank. It sounds like a good alternative inasmuch as small farmers have neither the income nor volume of milk to keep their own tank.

Furthermore, the processors in Brazil started to pay farmers for milk volume and quality. It means that if a farmer delivers a higher quantity of milk periodically, he/she will receive a better price for the milk. Because of this, some small farmers decided to associate and deliver milk together (using the name of only one farmer) in order to take advantage of this system of pricing.

As a result, the number of small farmers that deliver milk to processors decreased and the number of big farmers increased. However, it does not mean that the total number of farmers decreased. It is just a result of the association on the cooling and on the delivery of milk.

2.3. Dairy companies

Neves & Consoli (2006) affirms that Brazil has 1,973 dairy processors with 34.4% of them in Minas Gerais, 13% in São Paulo, and 10.4% in Goiás. In this sense, 55% of the processors work with less than 10,000 liters/day. Only 5.3% of them work with more than 100,000 liters/day of which 28.8% are in Minas Gerais and 20.2% in São Paulo. It is an indication that there is some market power since the 14 leading companies processed 40% of the legal milk in Brazil in 2004 (Neves & Consoli, 2006). It was 21% for the 3 greatest companies and 10% for the first one. It means that there is an oligopsony situation. Table 5 shows the reception of milk for the 14 leading companies in Brazil. According to Gomes (2006), the processors segment became more concentrate in the early years due to growth of powder milk plants, which is feasible only on a big scale.

Table 5 - Leading companies of dairy products in Brazil, 2004.

Company/ Brand	Quantity receipted (thousand liters)			Number of farmers	Liters/day
	Farmers	Third market	Total		
DPA (3)	1,246,000	462,000	1,708,000	6,110	557
ITAMBÉ	982,000	23,000	1,005,000	7,325	366
ELEGÊ	737,782	103,767	841,549	25,001	81
PARMALAT	388,117	203,730	591,847	4,400	241
CCL	254,057	106,067	360,124	4,388	158
EMBARÉ	250,867	55,382	306,249	2,380	288
LATICÍNIOS MORRINHOS	233,310	66,134	299,444	3,200	199
CENTROLEITE	258,195	10,073	268,268	5,049	140
SUDCOOP	234,876	31,385	266,261	5,998	107
CONFEPAR	210,543	51,690	262,233	6,152	94
BATÁVIA	224,561	0	224,561	4,019	153
LIDER ALIMENTOS	184,240	18,439	202,679	5,243	96
DANONE	134,575	61,824	196,399	605	608
GRUPO VIGOR	171,009	20,913	191,922	996	469
TOTAL	5,510,132	1,214,404	6,724,536	80,866	186

Source: Leite Brasil (2006).

Another fact that deserves attention in the Brazilian milk market is called the third market or spot market. It refers to small companies or cooperatives that buy milk and resell it to bigger companies. According to Gomes (2006), it is becoming very popular in Brazil, because of the necessity of delivering bigger volumes of milk to processors in order to get a good price. In addition, there is the modernization of the milk industry, which is very costly for small companies.

In 2004, the reception of milk for spot market was 900,490 million liters, corresponding to 19% of the total milk received by the industry. In big companies, such as DPA, this volume was equivalent to 25% of the group, while for Parmalat it was 29% (Gomes, 2006). According to Nogueira (2006), about 30% of the milk of the cooperatives is sold on the spot market. In Santa Catarina, for instance, approximately 70% of the milk of the cooperatives is sold on the spot market.

Another important fact related to milk industry is that the processors are not equally distributed over the Brazilian territory. They are more concentrated in the South and Southeast of the country. Terraviva (2006) affirms that the only five states – Minas Gerais, São Paulo, Rio Grande do Sul, Goiás, and Paraná are responsible for 80% of the Brazilian processed milk (Table 6).

Table 6 – Production and reception of milk by the Brazilian states in 1997*.

	Reception of milk		Milk production	
	Quantity (thousand liters)	Share (%)	Quantity (thousand liters)	Share (%)
Rondônia	261,918	2.46	335,913	1.80
Acre	8,437	0.08	31,831	0.17
Amazonas	159	0.00	32,487	0.17
Roraima**	-	-	9,523	0.05
Pará	44,740	0.42	290,210	1.55
Tocantins	12,759	0.12	138,083	0.74
North	**328,013**	**3.08**	**838,047**	**4.49**
Maranhão	15,080	0.14	138,961	0.74
Piauí	12,522	0.12	75,504	0.40
Ceará	88,634	0.83	387,990	2.08
Rio Grande do Norte	49,995	0.47	161,629	0.87
Paraíba	5,604	0.05	149,802	0.80
Pernambuco	56,834	0.53	357,853	1.92
Alagoas	58,158	0.55	301,614	1.62
Sergipe	12,648	0.12	127,228	0.68
Bahia	257,465	2.42	688,475	3.69
Northeast	**556,940**	**5.24**	**2,389,056**	**12.80**
Minas Gerais	2,919,135	27.44	5,602,015	30.02
Espírito Santo	197,627	1.86	339,339	1.82
Rio de Janeiro	420,299	3.95	451,223	2.42
São Paulo	1,942,548	18.26	2,003,165	10.73
Southeast	**5,479,609**	**51.52**	**8,395,742**	**44.99**
Paraná	835,171	7.85	1,579,837	8.47
Santa Catarina	292,257	2.75	852,169	4.57
Rio Grande do Sul	1,291,684	12.14	1,913,124	10.25
South	**2,419,112**	**22.74**	**4,345,130**	**23.28**
Mato Grosso do Sul	177,042	1.66	414,947	2.22
Mato Grosso	155,478	1.46	380,517	2.04
Goiás	1,446,304	13.60	1,868,976	10.01
Distrito Federal	74,102	0.70	30,749	0.16
Center-West	**1,852,926**	**17.42**	**2,695,189**	**14.44**
Total	**10,636,600**	**100**	**18,663,164**	**100**

Source: Terra Viva.

* Fabrics with federal, state, or municipal inspection

** State with less than 4 informants.

Table 6 is interesting, but it can be interpreted in various ways. The data about reception may simply present the level of industrialization of each state. In this case, Minas Gerais and São Paulo confirm their hegemony with 27% and 18%, respectively, in the total milk received by processors. However, if we consider that there is a strong interstate trade in milk, this first interpretation for the Table 6 may be wrong (Terraviva, 2006).

On the other side, the difference between the produced quantity and received quantity may indicate the level of informality in each state. In this sense, it seems more plausible to believe that the Northeast has a large quantity of informal milk processing. However, one can see on Table 6 the importer or exporter capacity in each state. Therefore, Rio Grande do Sul, Goiás, and São Paulo would be probably milk importers. This last one clearly is a milk importer because of its huge consumer market (Terraviva, 2006). But, utilizing only this table we cannot affirm anything about the milk industry. It would be necessary to have a time series on exports and imports, and quantity of milk received by the processors, which is unavailable in Brazil.

In relation of the retail, Neves & Consoli (2006) states that there were 71,951 groceries selling milk in Brazil in 2004. It means that 55% of the processed milk is traded by the groceries and 13% by the bakeries. However, the big retailers can exert certain level of local market power. On the other hand, Meireles (2003) affirms that because of the changes during the 1990's, there was a declined in relative importance of small retails.

3. LITERATURE REVIEW

Studies about market analysis have become very popular in the academic environment. In the competitive context of globalization of markets, studies about market integration, law of one price, price transmission and others are quite important for the understanding and proceedings in agricultural markets. So, in this section, we are going to show the state-of-the-art market integration with the purpose of identifying the appropriate theory and methodology for this work.

3.1. Theoretical definitions and controversies about market integration

First of all, it is necessary to define some concepts related to market integration, as market efficiency, transaction cost, Law of One Price, etc. However, the discordances in this area are in a large scale. The controversies over market integration analysis began with the definition of a market, as referred by Fackler & Goodwin (2000).

One of the first definitions of market was given by Cournot in 1838. Afterward, prominent economists as Marshall (1947), Cassell (1918), and Stigler (1969) provided different definitions. In this sense, Stigler's definition seems more acceptable for the academy. Stigler (1969, p. 85) relied on the idea that the market is "the area within which the price of a commodity tends to uniformity, allowance being made for transportation costs".

Nevertheless, the divergences do not stop at this point. Several meanings are given for market integration. The mainstream in economics defines market integration as the satisfaction of the Law of One Price (LOP) or as the existence of cointegration between price series. But, McNew (1996) argued that market integration is a concept based more on statistical criteria rather than on economic facts. Moreover, according to Li & Barrett (1999), tests of the LOP are a test of the perfect integration hypothesis, not a test of imperfect integration.

Monke & Petzel (1984), cited by Baek & Koo (2005), specified an integrated market as those markets in which prices of homogeneous products move together, which is similar to the concept of cointegration. On the other hand, Fackler & Goodwin (2000) adopted the idea that market integration is a measure of the degree to which demand and supply shocks in one region are transmitted to another.

For Barrett (2001, p. 20), "market integration represents the Walrasian transfer of excess demand from one market to another as captured in actual or potential physical flows". This sight does not require that prices are equilibrated. Thereby, two regions can be (imperfectly) integrated even if they are imperfectly competitive or inefficiently restricted, if price in one region responds or not to shocks in the other, and if physical flow of the good occurs or not between these locations (Li & Barrett, 1999). The last point is interesting and was discussed by Stigler & Sherwin (1985), cited by Massey (2000). They affirmed that the existence of trade flows between two areas is neither a necessary nor a sufficient condition to conclude that markets are integrated. In areas where there is a large flow of goods, price discrimination can lead to differences in prices between these markets, and hence, to segmentation of the markets.

In contrast to Barret (2001) and following the neo-classical theory of the firm, Goodwin & Djunaidi (2000) said that market integration does not occur if the prices in geographically separated markets are not at equilibrium price. They also pointed out that the price should be results of demand and supply forces in the output markets.

On the other hand, Fackler & Goodwin (2000) considered the following hypothetical model for the price transmission ratio:

$$R_{AB} = \frac{\dfrac{\partial p_B}{\partial \varepsilon_A}}{\dfrac{\partial p_A}{\partial \varepsilon_A}} \qquad (1)$$

It represents a shock, ε_A, that shifts the excess demand for a good in region *A* but not in the region *B*. They affirmed that market integration is a measure of expectation of the price transmission ratio. In this case, perfect integration happens when the expected price transmission ratio equals one. Besides, it is affirmed that the ratio may not be symmetric (i.e., $R_{AB} \neq R_{BA}$) so that it is possible for one region to be more integrated with another than vice-versa. In addition, if two regions *A* and *B* trade with *C*, they can be integrated as strongly as if they were direct trading partners. It is not necessary for two regions to be direct trading partners for a high degree of integration to be present. But it is necessary that the regions are part of a common trading network.

On the other hand, a great number of studies asserted that for existing market integration, the locations should present high price correlations. However, Gonzalez-Rivera & Helfand (2001) interpreted that the existence of market integration is related to two conditions. Firstly, the set of locations should be connected by a physical flow of goods. It is necessary to ensure that arbitrage occurs, but by itself, it does not guarantee market integration. Secondly, the prices between the regions should share the same long-run information. In a cointegration framework, it is equivalent to requiring the existence of one and only one integrating factor that is common to all series of prices[3].

Gonzalez-Rivera & Helfand (2001) also distinguished three different concepts within market integration, which include extension of market, pattern of integration and degree of integration. Extension of market refers to geographic boundaries of the integrated market. The pattern of integration is a measure of how the information contained in the prices is transmitted among each region of the market. The degree of integration is defined as the time of reaction so that the long-run relationship gets to absorb a shock in the whole system.

Still related with market integration is the concept of market efficiency. Fama's efficiency definition refers to a market in which prices fully reflect available information (Dahlgran, 2000). Notwithstanding, in some studies, market efficiency is taken as synonymous of the spatial arbitrage condition. But, Barrett (1996) believed that market integration is neither a necessary nor sufficient condition for market efficiency, because the identity "price differences equal transfer costs" can be consistent with perfect competition, monopoly or Pareto inferior trade. On the other hand, according to Facker & Goodwin

[3] The existence of just one integrating factor, as well as the explanations about cointegration will be better discussed ahead.

(2000), failure to find market integration can indicate both breakdown in market integration and market inefficiency.

Faminow & Benson (1990) also disagreed on the usual association with market integration and efficiency market and perfect competition. They argued that market integration can occur even in oligopoly or oligopsony. Moreover, Dahlgran (2000) pointed out that markets that present either spatial or intertemporal arbitrage opportunities are considered inefficient since reallocating resources away from low-priced locations to high-priced locations will increase social welfare. Thus, in a spatial context, allocative efficiency which requests that price differences between two points be less than or equal to arbitrage costs, reflects a segmented market and not integrated market.

It is also important to discuss the concept of arbitrage. In accordance with Fackler & Goodwin (2000), the arbitrage condition is the starting point for any model of spatial price behavior and they also believed that market integration is a different concept from arbitrage. Li & Barrett (1999) explained that arbitrage does not ensure long-run competitive equilibrium. If markets are imperfectly competitive, whether due to coordination between firms or market restrictions emplaced by government, profits may persist.

So, arbitrage is a mechanism where the prices of a homogeneous good at any two or more locations will differ by, at most, the cost of moving the good from the region with the lower price to the region with the higher price. So, it can be noted that the concepts of market integration, arbitrage and transaction or transportation costs are closely related. Consequently, transaction costs have been a reason for controversies as well.

Following Fackler & Goodwin (2000), when transaction costs are high, they can hinder market integration. They also believe that it is most real in developing countries, where poor contract enforcement, inadequate policy protection, high taxes, inadequate transport and problems with the communications infrastructure increase the transaction costs. However, Miljkovic (1999) affirmed that, following the literature of spillover and geographic concentration, the effects of distance on trade in spatial analysis of market are much greater than transportation costs.

3.2. Law of One Price

The Law of One Price (LOP) is one of the most important theoretical concepts associated with market integration. The LOP states that if there is trade between two or more spatially separated regions and all profitable arbitrage opportunities are extinguished, price differences are equal to transaction cost (Barrett, 2001). So, many of the previous studies on market analysis were about LOP; however, the support for the LOP is limited, especially in the short-run. Miljkovic (1999), for example, declared that there are a great number of reasons for the failure of the LOP. Some of the most important reasons are pricing-to-market[4] (price discrimination), exchange rate risk, transportation costs, and institutional factors that pressure price in different markets.

Sexton *et al.* (1991) affirmed that there is no support for the LOP when first, regions are not linked by arbitrage, i.e., they represent autarkic markets. This point is obvious because if there is no trade there is no arbitrage, and no support for LOP. Second, there are impediments to efficient arbitrage such as trade barriers, imperfect information, or risk aversion (Ravallion, 1983). Trade barriers are so common in international trades. Imperfect information and risk aversion are factors present in every trade. Third, there is imperfect competition in one or more of the markets (Stigler & Sherwin, 1985, and Faminow & Benson, 1990). On the other hand, Fackler & Goodwin (2000) recognized that LOP is more for traded than non-traded goods and it happens more in the long-run than in the short-run period.

Ardeni (1989) wrote a classic paper about LOP. He analyzed LOP for a set of seven commodities in four countries. He argued that denial of the LOP was due to econometric shortcomings, like spurious regressions[5], nonstatitonarity in the data, high serial correlation and inappropriate use of first differences[6]. Then, in order to correct these shortcomings, he used cointegration to test the long-run relationships. As a result, Ardeni (1989) reported a lack of empirical support for the LOP as a long-run relationship. He pointed out that, in

---

[4] Pricing-to-market is a form of imperfect competition in which exporters price discriminate across destination markets and export prices depend on bilateral exchange rates (Krugman & Obstfeld, 2003).

[5] Spurious regression occurs when a correlation between data exists because of a statistical fluke, rather than true causality.

[6] It is relevant to say that this kind of shortcoming happened because the cointegration methods became popular with Engle & Granger (1987). So, Ardeni (1989) is an important study because he employed the innovative technique for this period in market analysis.

this case, failures in LOP can be caused by institutional factors, high arbitrage costs, measurements errors, imperfect information, explicit and implicit contracts, etc.

In a chronologic sequence, there is the work of Goodwin *et al.* (1990). According to classical theory, if LOP holds contemporaneously, then price parity should hold contemporaneously. Nevertheless, Goodwin *et al.* (1990) stated that this assumption ignores the temporal elements of trade, i.e., it overlooks the fact that international commodity arbitrage and trade occur over time. Moreover, they believed that there are some shortcomings in the standard approach to testing the LOP such as (1) information about transportation costs is hardly ever available and assuming that these costs are constant is questionable, (2) the standard approach requires that commodity prices be exogenous. This may be an error because the prices in two countries, for example, can be determined simultaneously due to information sharing across markets, and that agents can work in various markets at the same time, and (3) the standard LOP uses contemporaneous domestic and foreign prices in the empirical analysis.

In order to treat these limitations, Goodwin *et al.* (1990) included expectations in the analysis of LOP. They tested the LOP in international markets for U.S. agricultural commodities (wheat, oilseed products, corn, and grain sorghum). For this, Goodwin *et al.* (1990) employed two procedures. First of all, they made use of the generalized method of moments (GMM) procedure and applied nonparametric tests of parity conditions. They also utilized actual freight rates for wheat as a proxy measure of transaction costs. The empirical results supported the hypothesis that the expectations-augmented version of LOP is better than the standard version, but this outcome depends on the correct measurement of transaction costs and delivery lags. According to Goodwin & Djunaidi (2000), transaction costs are very important especially for goods that have a high relation between weight-volume.

Therefore, Goodwin *et al.* (1990) and Mohanty *et al.* (1998) explored a new methodology to evaluate if LOP holds in international commodity markets. They experimented with fractional cointegration analysis to determine whether this method is more appropriate than standard cointegration for validating the LOP. They believed that LOP could have been falsely rejected in many studies both because factors such transportation costs, price expectations, and market power were not taken into consideration, or because the methods utilized were incapable of capturing the real relationship between prices.

Mohanty *et al.* (1998) employed the same data used by Ardeni (1989), but made use of fractional cointegrating, because they affirmed that weak empirical support for the LOP is due to fact that the standard cointegration method is quite restrictive. Standard cointegration determines if the order of integration of the equilibrium errors in two price series add up to zero. This guarantees that the LOP holds. On the other hand, if the order of integration is one, LOP is not supported. So, the advantage of fractional cointegration is that it is able to discern long-run price behavior.

The results of Mohanty *et al.* (1998) confirmed that the fractional cointegration supported the existence of LOP in the majority of analyzed cases. However, this result did not differ very much from Ardeni's. The difference between these two studies is that Mohanty *et al.* (1998) found that some commodity price series were fractionally cointegrated even if cointegration is rejected. It means that the results of Mohanty *et al.* (1998) are not contradictory of the Ardeni's. Just in the cases in that there is not full cointegration Mohanty *et al.* (1998) established that markets still remain in a long-run relationship.

In another study with the same data, Baffes (1991) considered that a malfunction of the LOP would be caused by transportation costs. Therefore, with the objective to evaluate this hypothesis, Baffes (1991) introduced a restricted cointegration test, i.e., he tested for cointegration by taking the difference of two variables. In addition, the author utilized tabulated "t-values". These values were lower (in absolute terms) than the ones usually used in cointegration regressions, because this work used known cointegration parameters[7]. The results were, in most cases, unfavorable to the LOP. For the remaining cases, he believed that the LOP did not hold because of transportation cost. Therefore, Baffes (1991) tested whether freight rates were cointegrated with price differentials. The evidence showed that the stationarity[8] of the freight rate series varied with the size of sample. So, there was weak evidence of cointegration. Nevertheless, the author asserted that the nonstationarity of the freight rates would be the main source of nonstationarity of the price differentials. It is important to remember that transaction costs have components other than freight rates. Therefore, freight rates may not be a good proxy for it. Furthermore, Baffes (1991) believed that the nonstationarity of a tax/subsidy may cause the divergence from a

[7] For more details, consult Baffes (1991).

[8] According to Gujarati (2000), a series is consider stationary if its mean and variance are constant over the time; and if the value of covariance between two periods of time just depends on distance among these periods.

long-run equilibrium and the heterogeneity of data can be another important factor to be considered in future studies. Baffes (1991)'s conclusions are problematic because the results suggest a negative relationship between prices.

Following the LOP approach, Asche *et al.* (2004) examined the relationship between causality models and cointegration models in testing for the LOP. These authors expected that the choice of modeling method depends on the time series properties of the data. To illustrate this point, they studied the whitefish market in France. For this market, the causality approach rejected the hypothesis of the LOP, while cointegration models provided evidence for a well-integrated whitefish market.

3.3. Theoretical models for market integration

For the purpose of studying the market integration subject, economists have utilized basically two models, the point-space framework of Takayama and Judge and the agents-on-links model.

3.3.1. Point-space model

The most common model used in studies about international or interregional trade analysis is called the point-space trade model, developed by Takayama and Judge.

In this model, production and consumption regions are considered in separated locations (points) and there is no intraregional transaction (i.e., all intraregional transfer costs add up to zero). In addition, it associate market integration with market efficiency and perfect competition, and regional boundaries are considered fixed. So, the satisfaction of the spatial equilibrium condition in this case implies market efficiency (Barrett, 2001). Besides, it assumes instantaneous price adjustment with no feedback and trade is neither necessary, nor sufficient conditions for spatial equilibrium.

Then, the equilibrium condition of the point-space model considers the price variables, transaction costs, and trade volume. Because of this, Barrett (2001) criticized studies that utilize only the price variable in order to analyze market integration.

According to Barrett (2001), one shortcoming of the point-space model is that it assumes continuous trade, while in real agricultural markets it is quite common that there are discontinuous trade, seasonality and contemporaneously bidirectional trade. Besides, the fact that it considers market efficiency and perfect competition is also quite far from reality.

On the other hand, McNew & Fackler (1997) contended that the point-space model is adequate from cases, which many bulk goods are collected at a small number of points and distributed, to a small number of major centers. It is true especially for gasoline, grain, coal, and international trade, which involve ocean freight or tariffs imposed at the border.

3.3.2. Agents-on-links model

The opposite of the point-space model, the agents-on-links models are more used to model spatial oligopoly or monopoly situations (McNew & Fackler, 1997). It was developed by Hotelling and Smithies and is built as a network structure with markets or firms located at network nodes and consumers or commodity producers located along network links. In this case, agents can choose to transact at the node by providing the best price (McNew & Fackler, 1997).

According to McNew & Fackler (1997) in studies that utilized prices paid to producers at the plant, the agents-on-links models would be more appropriate, since agricultural products are generally produced in isolated rural areas, and in such cases, the plants have local monopsony power. Besides, they believe that in a competitive environment, one can use agents-on-links models when various plants are located at each of the major distribution centers.

3.3.3. Other alternative models

Faminow & Benson (1990) offered an alternative theory assuming that agents of a market are spatially distributed, and intraregional transportation cost exist and are significant. In this case, spatial markets are characterized as oligopolistic/oligopsonistic competition, and the regional boundaries are a function of relative prices. In this model, arbitrage only guarantees that prices at different spatial buying sites will not differ by more than the transport cost of shipping the product from one site to another. So, price differentials do not necessarily have to be identical to transportation costs. Besides, spatial competitors can price discrimate in order to invade a rival's markets (Faminow & Benson, 1990).

On the other hand, Kawaguchi *et al.,* (1997) developed a model based on the point-space model that can be applied to any degree of market structure. As the model was the built for milk market in Japan, they added government intervention (payment quota) in the model. However, when a competitive market is considered, this model is similar to the Takayama-Judge model. Nevertheless, for the case of the Japan milk market, the authors considered oligopolistic consignment sellers and many perfectly competitive small-scale producers with pooled returns. For the purpose of comparing the results, they also tested the competitive structure for the Japanese market. Therefore, the results were compatible to the theory. It means that under perfect competition, more milk was transacted, and consequently smaller prices occur than in an oligopolistic case.

3.4. Methods: advantages and disadvantages

As it can be noted previously, there is a huge quantity of discordance in market integration theory. Because of this, in recent years, a lot of work has been done in order to identify the best methodology for capturing the interaction between markets and price behavior. Thus, here we are going to review the evolution of the research along this line and compare methods and results in order to identify the more appropriate methodology for this work.

3.4.1. Price Correlation

Most of the first works about market integration utilized the idea that the prices in separated markets must be correlated. It is one of the most popular methods for measuring market integration. Price correlations are a measure of co-movements of prices and a typical bivariate correlation model can be expressed by:

$$P_{it} = \beta_0 + \beta_1 P_{jt} + \beta_2 T_t + e_t \quad (2)$$

where P_{it} is the price in region *i* at time *t* for a homogeneous good; P_{jt} is the price in region *j* at time *t* ; T_t is the transportation cost at time *t* between two regions and e_t is a random error term. This is a version of a model introduced by Richardson (1978). In this case, markets are taken to be perfectly integrated if $\beta_1 = \beta_2 = 1$ and $\beta_0 = 0$.

Although it is simple and easy to apply, this methodology presents some limitations, because it is difficult to separate the co-movement between prices from the long-run trends and other effects. Goletti & Christina-Tsigas (1995) and Fackler & Goodwin (2000) believed that parallel movements in price can occur for several reasons other than market integration. It happens mainly for agricultural commodities because surplus and deficit seasons are common. So, in order to avoid some of these spurious correlations, some studies preferred to employ price differences instead of price levels.

Even when utilizing price differences, price correlation studies have critics. Li & Barrett (1999), for example, affirmed that co-movements between prices can occur due to exogenous factors, such as inflation, population growth, seasonality or climatic conditions. Then, the existence of price co-movement between spatial markets is not an indication of market integration.

On the other hand, Jordan & VanSickle (1995) verified that the dynamic nature of spatial price adjustment is ignored in price correlation models. In addition, they pointed out the existence of other problems like simultaneity bias (since one of the prices should be considered exogenous in the model) and bias from omitted variables.

Barrett (2001) also agreed with previous authors and added that the existence of interseasonal flow reversals can lead to unreliable indicators and tests may overestimate segmentation lags in information, delivery, or contract expiration. Besides, it is not possible to verify the existence or nonexistence of heteroskedasticity common in price series. Harris (1979), cited by Fackler & Goodwin (2000), introduced a number of other shortcomings relative to price correlation. She stated that price discrimination could result in correlation coefficients of 1, which indicates market integration. Another problem is related to transaction costs. According to Fackler & Goodwin (2000), it is easy to show that any value of a correlation coefficient can indicate integrated markets when transaction costs are large enough to prohibit profitable trade and then permit independent variation of price differences within the transactions cost band. Finally, it is important to remember that all of the limitations inherent in the simple correlation coefficient approach are equally applicable to the simple regression model.

3.4.2. Ravallion model

After price correlation models, some advances were made in order to improve the market integration analysis. The Ravallion model is an interesting metodology. The Ravallion model applies a Vector Error Correction Model (VEC) that allows autocorrelation, distinguishes short-run and long-run dynamics, and can consider common inflationary and seasonal components (Barrett, 2001).

The basic Ravallion model had the form

$$P_{it} = \sum_{j=1}^{n} a_{ij} P_{it-j} + \sum_{j=0}^{n} b_{ij} P_{1t-j} + X_{it} c_i + e_{ij} \qquad (3)$$

where P_{it} refers to the price in the *i*th local (urban) marke in time *t*; P_1 refers to the price in a central market; *X* refers to the vector of other influences (inflation, seasonal dummies) and the *e'*s are white noise error terms. This model assumes that the shock in prices originate in a central market. However, the Ravallion model does not incorporate interseasonal flow reversals and direct links between markets. Besides, it is assumed that transaction costs are constant, i.e., additive or proportional (Barrett, 2001). Structural complexities led this model to disuse.

Jordan & VanSickle (1995) made use of Ravallion's model for evaluating the U.S. winter market for fresh tomatoes. According to the authors, this model had the advantage of distinguish among market segmentation, strong short-run market integration, weak short-run market integration, and long-run market integration. Because of this, Jordan & VanSickle (1995) used a two-stage least squares model in restricted and unrestricted forms, and Granger causality. They applied Granger causality in order to verify the hypothesis of price leadership. As a result, they concluded that Florida and Mexico were integrated in the same market, but Florida appeared as a leader in the fresh tomato price formation process.

3.4.3. Switching regime model

Switching regime models compare observed price differentials with observed transportation costs, estimating the probability (i.e., function of distribution of probability) that markets are in competitive equilibrium (Li & Barrett, 1999). The most common switching regime models are the Parity Bounds Model (PBM), Barrett-Li Method (BLM), Sexton, Kling and Carman Model (SKC) and Pricing-to-market (PTM) model. They are very similar in structural form and differ only in a few details.

This kind of model eliminates the problems of discontinuous trade, time varying and potentially nonstationary transaction costs that perturb other price analysis methods. Sexton *et al.* (1991) and Baulch (1997) applied endogenous switching models, which account for the multiple regimes that may be a result from transaction costs (Goodwin & Piggott, 1999). However, these models provide only static comparisons. Moreover, they presuppose serial dependence of price and transaction costs and are sensible to underlying distributional assumptions.

So, Li & Barrett (1999) said that the results of PBM are suspect, because the markets are considered to be integrated even when there is no physical trade flow of goods and no transmission of price shocks between two locations. In addition, they affirmed that PBM does not make use of trade flow data, which means that it studies only equilibrium conditions, not integration. Then, in their analysis, Li & Barrett (1999) extended PBM through complementary trade flow data.

Li & Barrett (1999) introduced a market analysis methodology based upon maximum likelihood estimation combined with a distribution model incorporating price, transfer costs, and trade flow data. According to these authors, this model differentiated between market integration and competitive market equilibrium. Besides, the model provided intuitive measures of inter-market tradebility, competitive market equilibrium, perfect integration, segmented equilibrium, and segmented disequilibrium.

Li & Barrett (1999) believed that other market analysis price models were biased, since they assumed trades are continuous or unidirectional, and transaction costs were stationary. However, the opposite of these assumptions take place regularly. Therefore, they used data including bidirectional and discontinuous trades and nonstationary transaction costs to analyze the soybean meal market among Pacific Rim economies. Nevertheless, the problem of the Li & Barrett (1999) model is that it is static and does not answer simple questions about the market such as speed of adjustment or extension of convergence to tradebility of equilibrium.

In addition, according to Rashid (2004), the bivariate nature of the PBMs can result in wrong inferences about the multivariate environment of market integration. Besides this, the author affirmed that the PBM is incapable of analyzing market integration in the absence of trade flow data. This last point is very interesting, since this kind of data is so difficult to obtain.

On the other hand, the SKC is known for the following advantages, which include (1) it eliminates the necessity of choose arbitrarily the price in one region, (2) transportation costs are estimated inside of the model, (3) the model considers that, due to arbitrage, the regions can be integrated in some periods, but not others, which makes possible the existence of three regimes whichever efficient arbitrage, shortage and glut, and

(4) the results allow inferring about efficiency of arbitrage, geographic dimensions of the market, and product substitutability (Sexton *et al.*, 2003). Other interesting points in the SKC model is the criterion of choice of the regions involved in analysis. The authors used the criterion of the quantity of trade (volume) and the distance from producer center.

Besides the models mentioned, there is the Pricing-to-market model (PTM). This model compares FOB export prices for a homogeneous good in a single source country to multiple destination markets (Barrett, 2001). Actually, the PTM tries to analyze the response of trade good prices to exchange rate changes under constant marginal cost and considering different kinds of market structure (imperfect competition) (Miljkovic, 1999).

3.4.4. Cointegration models

In recent years, cointegration has become the most popular method in market analysis. It occurred because cointegration models are applicable to nonstationary price series. Since cointegration refers to long-run relationship among prices, the existence of cointegration is an indication of interdependence on prices (Goletti & Christina-Tsigas, 1995). It means that, if in the long-run prices exhibit a linear, constant relationship, they are cointegrated.

Earlier empirical works were focused on the use of a bivariate cointegration test. The bivariate method consists of a cointegration test with two price series by assuming other prices have no effect on the market. But, this method is criticized for omitting important variables (i.e., prices) because this omission neglects indirect linkages between markets, so that the results are not reliable (Gonzalez-Rivera & Helfand, 2001). These authors also affirmed that is so difficult to determine which locations belong to the same market using a bivariate model. Furthermore, bivariate approach requires one of the two prices to be considered exogenous (Miljkovic, 1999).

So, in order to include more prices in cointegration analysis, the economists started to use Johansen's multivariate cointegration procedure. According to Asche *et al.* (2004), the multivariate approach is more adequate to analyze price series because with n price series, it can have at most $n-1$ cointegration vectors, although there are $n^2-n/2$ possible pairs. Therefore, all pairs will be redundant, except $n-1$ pairs. As exposed before, with a bivariate analysis one might obtain different results depending on which pairs are chosen in analysis. However, with a multivariate specification, one can estimate no more than $n-1$ cointegration vectors.

But the same authors also discussed that the results of cointegration are sensitive to the dimensionality of the system, so that the reliability of the results is a decreasing function of the number of parameters estimated. They believed that bivariate models have an advantage because they use all the relevant structural information and do not present the dimensional problem.

On the other hand, Barrett (2001) affirmed that cointegration is neither necessary nor sufficient condition for existence of market integration. First of all, he stated that if transaction costs are nonstationary, the price series can be cointegrated even if the markets are not integrated, which shows that cointegration is an unnecessary condition for market integration. It means that refusal of cointegration tests cannot necessarily mean absence of market integration. It may just be a sign that transaction costs are nonstationary. Secondly, cointegration can be consistent with a negative relationship between prices while the LOP suggests a positive correlation. In other words, the existence of cointegration coefficients with magnitudes far from unity is inconsistence with rates of prices change. It is a very important point, because in agricultural markets, trade flows are sometimes discontinuous, which causes break points in cointegration tests and can bias the results. The author still asserted that market segmentation can result either from inter market margins larger or less than transaction costs.

Barrett (2001) stated that when transaction costs are explicitly account for, a variable cost becomes a part of the error term in the regression equation, causing bias and inconsistency in parameter estimates. In this case, even if the freight rates are taken into consideration, it may not solve the problem of correlation between the unmeasured components of transaction costs and prices. Besides, Barrett (2001) affirmed that cointegration, Granger causality, and VEC have strong, simplying assumptions, making it impossible to measure market efficiency. Miljkovic (1999, p. 133) also declared that "the greater the transfer costs and more frequent the discontinuities, the more pronounced the nonlinearity and the more suspect the findings of linear cointegrating regressions".

A great number of studies that affirm that transaction costs are nonstationary cite the work of Engle & Granger (1987). It is one of the first and most important works of these authors about cointegration. The authors analyzed the cointegration between short and long interest rates. In this case, and consider an efficient market, the theoretical relationship between the two interest rates is given by

$$EHY = DR_{t-1} - (D-1)R_t - r_t \quad (4)$$

where *EHY* is a representation of a risk premium, *D* is the duration of the bond, R_{t-1} is the long term rate, and r_t is the short rate.

As a result, they found that *R* and *r* were cointegrated. However, the authors said that "If the risk premium is varying over time but is *I(0)* already, then it need not be included in the test of cointegration". In sum, it is the only thing that the Engle & Granger (1987) affirmed about the risk premium. Of course, it can be applicable to transaction costs, but it does not mean that transaction costs are nonstationary. In other words, they said that if the transaction costs are nonstationary, they should be included in the cointegration analysis. Nevertheless, they do not affirm that transaction costs are nonstationary.

In the same line, Rashid affirmed that the results of various studies is in favor of cointegration, and consequently to market integration. Even in countries with different level of development, such as Brazil, Indonesia, Ethiopia, the Philippines, and Bangladesh, cointegration has been found.

Although some works criticize the cointegration method for not regarding trade flows, Rashid (2004) says that trade flow information is essentially contained in price data. He explained it using a case of bidirectional trade. In this situation, if traders are assumed to be profit maximizers, a trade flow reversal between two markets should cause a price reversal movement, at least in the long-run. For McNew & Fackler (1997), Fackler & Goodwin (2002) and Barrett & Li (2002), the presumption that transaction costs are stationary and that trade flows are continuous and not reversal can bias the results of cointegration. But, Rashid (2004) conclude that the results of various studies do not match with that. The majority of the cointegration studies have found the existence of market integration. Davutyan & Pippenger (1990), cited by McNew & Fackler (1997), also asserted that one is less likely to find market integration or cointegration when transaction costs are large. This statement is according to arbitrage theory, but it does not mean that the transaction costs can affect cointegration models.

Beyond that, Fafchamps & Gavian (1996), cited by Rashid (2004) conclude that when markets are not integration, both PBM and cointegration methods may lead to the same conclusions. Otherwise, Maddala & Kim (1998) assert that the Johansen approach is sensitive to specifications, mainly the selection of the number of lags. However, cointegration is still the most used method for studying market integration. Some works, like Jarvis *et al.* (2005), utilized cointegration to evaluate price convergence. In this study, beef export prices of different countries were examined. The authors provided methodological improvements for making use of a more disaggregated level of data (i.e., bone in and boneless beef price). Moreover, they made comparisons between groups of countries, instead of pairwise comparisons. As a result, they found that there is convergence between beef prices, but the protectionism exercised by some countries has been disturbing it.

On the other hand, Pendell & Schroeder (2004) investigated how mandatory price reporting has influenced the degree of market integration between regional fed cattle markets. For this, they compared before and after the implementation of the Livestock Mandatory Price Reporting Act in April 2001. Employing cointegration models and VEC, they observed that reporting does not increase market integration and that three periods is necessary to allow a market to adjust.

Alternatively, Baek & Koo (2003) introduced a model specification in order to find out the correct VAR model. These authors worked to adopt a reasonable specification for the cointegration model based on the VAR number, lag selection criteria and diagnostic tests. They also applied the concept of the general-to-specific procedure[9] to construct a VEC model. Standard tests are used to eliminate the statistically insignificant variables.

Another interesting characteristic of the Baek & Koo (2003) model is that they incorporated a structural break in the cointegration test. The inclusion of a structural break was made by dividing the sample in two sub samples (before and after the structural change) to do the cointegration. Besides, they paid more attention to determination of the lag length than previous works. Then the VEC was estimated with the same number of lags used in the cointegration test. Thus, using this methodology, Baek & Koo (2003) studied the dynamics of prices in U.S. and Canadian hard red spring wheat and durum wheat markets. The tests showed that Canada has been the price leader in the North American wheat market. The authors explained this result because, in Canada, the Canadian Wheat Board (CWB) controlled the prices and, the quality of wheat from Canadian was better than the U.S.

[9] General-to-specific procedure means that, firstly, it is realized general statistical model to capture essentials characteristics of data.

Following the cointegration framework, Baffes & Gohou (2005) examined the nature and degree of price linkages among polyester, cotton, and crude oil over time, in order to determine whether one of these prices causes variability in the others. The results confirmed that (1) there was strong co-movement between cotton and polyester prices, (2) crude oil prices had a stronger effect on polyester prices than cotton prices, and (3) price shocks originating in the polyester market were transmitted at a much higher speed to the cotton market than vice-versa.

Another interesting work is Liang *et al.* (1997). They analyzed two groups of dry bean producers regions. In the first place, the authors were interested in verifying whether the prices for the same variety are integrated between different regions. In the second place, they analyzed if the prices for different varieties of beans in the same region are integrated. Using cointegration, they found out that the commodities are almost perfect substitutes in the first group, while, in the second group, the behavior of prices indicated that the market does not identify different bean varieties.

3.4.5. Threshold cointegration models

According to Goodwin & Piggott (1999), threshold effects or nontradebility bands happen when higher shocks (i.e., shocks above some threshold) cause a different response than do smaller shocks. As noted by Ardeni (1989), import quotas, tariffs, variable levies, and other variables can create consistent and permanent price differentials between countries. Although at that time Ardeni (1989) did not realize that this kind of price differential can be correct by the threshold model.

Sephton (2003, p. 1041) affirmed that "threshold cointegration allows the equilibrating relationship to change if the series exhibits a different behavior beyond a threshold". According to Balke and Fomby (1997), cited by Saghaian (2006), cointegration refers to a global characteristic of the time series while the threshold regimes refer to the local characteristics.

A graphic analysis allows a better visualization about the necessity of threshold model in certain cases.

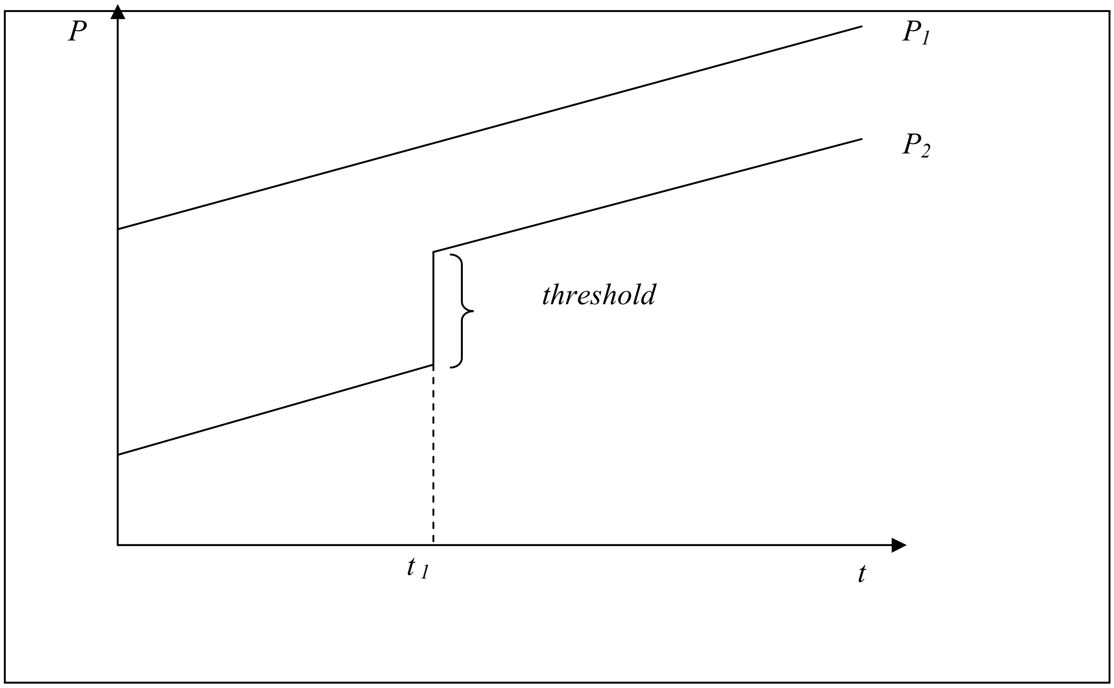


Figure 2 – Threshold effect between two price series.
Source: Developed by the author.

The Figure 1 shows the threshold effects between two price series. As can be noted, both series exhibit a common tendency after and before the period t_1. However, in the time t_1, this tendency changes. It means that there is one cointegrated relationship before t_1 and another cointegration relationship after this period. What happens is that, if one studies this case using a simple cointegration model, he/she can find that there is no cointegration. On the other hand, if he/she includes the threshold effect, the existence of cointegration will be the result. So, the threshold model appears like an alternative to the cointegration model, when one does not find a cointegration relationship between series analyzed, it is recommended to check the existence of threshold effects).

Consider a standard cointegration relationship representing an economic equilibrium

$$y_{1t} - \beta_1 y_{2t} - .... - \beta_k y_{kt} = v_t \qquad (5)$$

where y_{it} represents a price series i over time t; β's are the estimated parameters, and, v_t is an autoregressive process given by $v_t = \rho v_{t-1} + e_t$. As ρ approaches one, deviations from the equilibrium become nonstationary, and thus the y_{it} variables are not cointegrated (Goodwin & Piggott, 1999). When v_t follows a threshold autoregression, it has the following expression

$$\rho = \{ \begin{matrix} \rho^{(1)} \; if \left| v_{t-1} \right| \leq c \\ \rho^{(2)} \; if \left| v_{t-1} \right| > c \end{matrix} , \tag{6}$$

where c represents the threshold (Goodwin & Piggott, 1999).

The key to these methods is that they make use of unobservable threshold effects due to transaction costs, which reduces the likelihood that one rejects the hypothesis of market efficiency. It is a quite attractive characteristic for methods used to analyze market integration because, for many markets, prices are the only data readily available to examine spatial relationships (i.e., information about transaction costs are commonly unavailable, mainly in developing countries). So, utilizing threshold models, one can measure market integration based on price data alone, but also consider effects of transaction costs (Meyer, 2003).

One of the first and classic works about the threshold model is Goodwin & Piggott (1999). They employed the threshold cointegration model to evaluate price dynamics among regional corn and soybean markets in North Carolina. Goodwin & Piggott (1999) used two thresholds in their study and found a stable long-run equilibrium in North Carolina. They also used a nonlinear impulse response function to evaluate the dynamics of price adjustments. They discovered that, in most cases, the adjustments were completed in fifteen days.

Sephton (2003) tried to improve the study of Goodwin & Piggott (1999). He employed the same data that the precedent work used, but extended the analysis. Sephton (2003) utilized a multivariate approach to test for threshold cointegration and nonlinear cointegration. However, taking an opposite approach of Goodwin & Piggott (1999), he made use of a direct multivariate test of threshold cointegration and tested for the appropriate number of thresholds. He also considered the nonlinear relationship across a price series, which depends on transportation costs. As a result, Sephton (2003) pointed out the existence of one threshold for most of the commodities studied. Nevertheless, the outcomes were not much different from that found by Goodwin & Piggott (1999).

Meyer (2003) also made use of threshold cointegration model. For him, threshold models are interesting because they can account for the effects of transaction costs without directly using information about these costs, which are often unavailable. But, the disadvantage of this methodology is that it considers constant transaction costs. This author discussed the use of one or two thresholds for market analysis. According to him, using two thresholds may be more accurate, but appropriate econometric tests for two thresholds significance do not exist. Therefore, he used one threshold variant, which allows the existence of a "band of non-adjustment". As a selection criterion to choose the threshold, he applied a variance-covariance matrix[10]. This criterion differs from Goodwin & Piggott's. They utilized a variance-covariance matrix of the residuals, but it ignores the potential cross correlation between equations. The results of Meyer's variant threshold model indicated that the pig price in Dutch and in Germany react to deviations from the long-run disequilibrium.

In the same line of research, but with different results, there is the work of Pede & McKenzie (2005). They developed the multivariate threshold cointegration model for analyzing the integration of maize markets in Benin. They applied the empirical threshold model and compared real transaction costs with the estimated threshold coefficient. These authors believed that the two-regime threshold model allowed them to characterize a trading environment in which trade between spatially separated markets only occurs when relative price differences exceeded some level of transaction costs.

Using the Engle-Granger test[11] they found five of the twenty-one markets to be cointegrated, while Johansen test results showed nine markets to be cointegrated. Regarding threshold cointegration, Pede & McKenzie (2005) found mixed evidence. Two regime threshold models may not adequately capture dynamics between two markets when trade flow is bidirectional. Also slow moving price adjustments require more than a two weeks lag period to capture the full dynamic price movements. Further analysis in terms of VEC parameters implied little support for threshold effects. Moreover, the thresholds estimated were not consistent with observed transaction costs. Therefore, threshold models also have problems, since they still assume constant transaction costs and continuous trade flows.

[10] The variance-covariance matrix consists of the variance of the variables along the main diagonal and the covariance between each pair of variables in the other matrix positions (Gujarati, 2000).

[11] Engle-Granger test and Johansen test are tests of cointegration. Engle-Granger is the most simple of the tests, while Johansen has become more popular recently because of computer advance.

3.4.6. Complementary methodologies

In addition to all the methodologies presented, there are some other methods used to complement the market integration analysis. In this sense, the most important are Granger causality, impulse response functions, persistence profile, and directed acyclic graphs.

3.4.6.1. Granger causality

In the market integration environment, Granger causality is a tool to verify price leadership and it is usually used in association with cointegration models. Granger (1988) said that the existence of cointegration implies the existence of Granger causality in at least one direction, in a bivariate model.

Silveira (2004), for example, analyzed relationships between domestic and international sugar prices for Brazil. Furthermore, he verified what impact Brazilian sugar production has on the international price. For this, he had employed cointegration and Granger causality. The results of this study showed that there is contemporaneous correlation and Granger causality of international prices on Brazilian sugar prices; however, Brazilian sugar prices did not influence international prices.

Fardmanesh & Douglas (2003) also used cointegration and Granger causality to examine the relationship between the official and parallel exchange rates in three Caribbean countries, Guyana, Jamaica, and Trinidad, during the 1985-1993. They found that the official and parallel rates are cointegrated in all three countries, and the causation is bidirectional in the case of Jamaica and unidirectional in the cases of Guyana and Trinidad.

With the intention of improving the methodology, Schroeder (1997) also applied Granger causality, adjustment speed, deterministic models (regression equations), hedonic price models (for estimating transaction costs), and bootstrapping procedures. He used price data from 28 U.S. fed cattle slaughter plants to determine the extent of this geographic market. Therefore, with a complete and complex methodology, Schroeder (1997) found out there was a national fed cattle slaughter market in United States and the price discovery originated predominantly at plants located in Nebraska. Besides, he observed that almost 1/3 of total price adjustment to spatial integration occurs in one day. However, Asche et al. (2004), and Rashid (2004) showed that information about causality is already contained in the adjustment parameters of the cointegration models. It means that the test for weak exogeneity offers information about causality, so that it is not necessary to apply Granger causality in market integration analysis.

Moreover, in this test, causality means solely prediction. According to Bunge (1959), cited by Bessler (2006) prediction is just one of the attributes of the word cause. Granger causality focuses only on prediction. In other words, if we can predict *Y* better by using past values of *X* than by not using past values of *X*, then *X* Granger-causes *Y.* Then, the Granger causality is not the most appropriate way to treat the market integration.

3.4.6.2. Impulse response functions

Impulse response functions as well as persistence profile are methods used to define the degree of market integration. However, impulse response functions are measures of reaction that is more frequently used in literature. Impulse response functions measure the response of the endogenous variables over time to the impact of a shock to any variable represented in a VAR model (Arnade, 2006). Therefore, it is used with cointegration or directed acyclic graphs (DAG).

Haigh *et al.* (2001), for example, applied impulse response with DAG. They believed that understanding the impulse response analysis is very important for this kind of analysis. Nonetheless, few studies have paid attention to this. The majority of the studies about market integration used a cointegration model and impulse response functions and as a technique to decompose the VAR model, they generally use Choleski factorization. Nevertheless, Haigh *et al.* (2001) affirmed that Choleski factorization can bias the results because the correct structural model is not known.

Instead of the Choleski factorization, which imposes causal ordering on a VAR's variance/covariance matrix, there is the Bernanke decomposition in which causal ordering may be imposed on a VAR's endogenous variables. However, Bessler & Akleman (1998) stated that the Choleski decomposition presents the problem that a recursive ordering may be overly restrictive, while in Bernanke's approach, the true contemporaneous ordering may be unknown.

Franken & Parcell (2003) analyzed impacts of structural changes in grain/oilseed markets employing a three-tier statistical analysis of cointegration tests, flexible least squares analysis, and impulse response functions derived from a vector autoregressive model. The results seem consistent with the LOP. Goodwin e Piggott (1999) used a nonlinear impulse response function to evaluate the dynamics of price adjustments, tested shocks, and discovered that, in most of cases, the adjustments were completed in fifteen days.

Abreu *et al.* (2003) compared the domestic prices of six commodities with the international prices. They utilized cointegration, but they complemented it with impulse response functions to determine if the domestic prices of each commodity vary in the same direction as the international prices. In most of the cases, Johansen procedure indicated the existence of cointegration, i.e., a high level of linkages among domestic prices and international prices. The study also evidenced that the non-integrated prices represented markets that had governmental intervention through the period analyzed.

3.4.6.3. Persistence profile

Persistence profiles are functions that allow quantifying the degree of integration between all of the regions of the market (Pereira, 2005). Although impulse response functions is a method more used than the persistence profile, Gonzalez-Rivera & Helfand (2001) affirmed that impulse response functions are not unique functions when shocks to the system are correlated, and it probably occurs frequently with spatial prices. Therefore, the literature recommends using a Choleski decomposition to impose a recursive ordering on the variables. However, the impulse response functions can vary according to the ordering chosen. Because of this, they decided to use a persistence profile with cointegration for the Brazilian rice market.

Pereira (2005) utilized this methodology in association with cointegration in order to analyze the cattle market in Brazil. She utilize the same methodology that Gonzalez-Rivera & Helfand (2001) (i.e., cointegration and persistence profile) with one modification. The author employed the persistence profile as a dependent variable in analysis of determinants of market integration. As explicative variables, it was considered physical capital, production, consumption, etc. As a result, Pereira (2005) found that eleven states of Brazil are integrated and São Paulo is a leader of price in the short-run. The persistence profile showed that Goiás, Mato Grosso do Sul, Minas Gerais and Paraná have higher degrees of integration, i.e., they adjusted faster in a disequilibrium in a system.

3.4.6.4. Directed Acyclic Graphs (DAG's)

Directed Acyclic Graphs (DAG) is a diagram illustrating causal flow between variables no directed path from a variable that returns to itself (Haigh *et al.*, 2001). So, DAG represents conditional independence relations as implied by recursive product decomposition such as

$$Pr(v_1, v_2, v_3, ..., v_n) = \prod_{i=1}^{n} Pr(v_i / pa_i) \quad (7)$$

where Pr refers to the probability, v_i is a set of variables, pa_i refers to the realization of some subset of the variables that precede (come before in a causal sense) v_i in order; and Π refers to the product operator. This equation is called d-separation, and symbolizes a DAG in which the variables corresponding to pa_i are represented as direct causes of v_i . The Fisher's z is used to test whether conditional correlations are significantly different from zero (Haigh *et al.*, 2001).

According to Haigh et al. (2001), DAG is used to make definitive statements about contemporaneous correlations between prices. The DAG is an alternative dynamic model for analyzing price discovery over time and it is a PC algorithm implemented by the software TETRAD[12]. Haigh et al. (2001) utilized a Directed Acyclic Graphs (DAG's) and VEC in order to study the dynamics of volatile international freight prices that include the Baltic Panamax Index (PBI)[13]. Then, Haigh et al. (2001) studied the contemporaneous relationships among variables based on the variance-covariance matrix from the residuals from a VEC.

The adopted methodology consists of implementing a DAG followed by a VEC approach for estimating forecast error decompositions and impulse response. This mixture of techniques allows obtaining information on market linkages in short, intermediate, and long-run. Besides, it permits evaluating whether PBI is correctly composed (Haigh *et al.*, 2001).

The objective of Haigh *et al.* (2001) was to evaluate the degree of interconnectivity between the major shipping routes over time. The results of this study showed that, over the long-run, all freight prices are interconnected. However, it suggested that PBI might not be appropriately comprised and weighted. With DAG, Haigh *et al.* (2001) found a strong geographical pattern to information linkages and that some routes were dominant in

[12] The PC algorithm is described in detail in Spirtes, Glymour, and Scheines (1993).

[13] PBI is an index in which freight futures trading is based.

terms of price leadership. On the other hand, some routes were redundant in terms of the flow of information and, because of this, they would be removed from the index.

Yang *et al.* (2000) is another example of a paper that uses DAG. This study investigated the variant of LOP in developed and developing countries. They employed price dynamics by combining directed graphs and VEC for soybeans. As a result, they found evidence that one price predominates in developed countries while a different price exists in developing countries.

3.4.7. Other Interesting works

Beyond the presented methods, there is a long list of other innovative methodologies in the market analysis area. Some of these studies will be presented here. However, it is valid to remember that each framework was developed for a specific case and the choice of the more appropriate method will depend on the data, market, and problem under study.

Fabiosa (2000) used market analysis for evaluating if GATT reforms improved beef and wheat market efficiency. Using a cointegration model, he studied the price transmission and speed of adjustment. According to Fabiosa (2000), a variance decomposition analysis exposed larger and faster impacts in prices of non-fundamental shocks originating from other markets. So, opposite of earlier studies, Fabiosa (2000) employed time series methods with minimum structural specifications. This work also used innovation accounting to directly measure market integration. The method measured directly the variability of price and its decomposition. The test for market efficiency is based on the adjustment speed and elasticity implied in the cointegration vector. With this, Fabiosa (2000) found evidence that GATT promoted market efficiency, i.e., in post-GATT period the markets became more integrated.

On the other hand, Miljkovic (1999) studied the railroad industry by analyzing export-bound grain transportation. He used a spatial analysis for four states (Illinois, Iowa, Minnesota, and Nebraska) and two destinations (Mexican Gulf and Pacific Northwest) to determine if pricing practices by the same or different railroads in different regions are consistent. For this, Miljkovic (1999) estimated a system of structural equations and dynamics tests of regression. The results showed that the grain transportation market by rail is not perfectly integrated.

This work differs from others because all time series are stationary. So, it is not possible to use cointegration. Because of this, Miljkovic (1999) had to use three stage least squares to estimate demand and supply equations.

An interesting framework was developed by Gonzalez-Rivera & Helfand (2001) to study the Brazilian rice market. These authors divided the analysis in three steps, which include measuring the extension of the market, the pattern of integration, and the degree of integration. As extension of the market, Gonzalez-Rivera & Helfand (2001) understand the regions that present continuous flow of information and trade and it was identified by the long-run relationship on the cointegration model. The pattern of integration means interdependence on locations and was measured by the VEC. The degree of integration captures the speed of price adjustment and was calculated with persistence profile.

Gonzalez-Rivera & Helfand (2001) presented a quite complete methodology and, because of this, various works utilized a similar framework. One example is Liu & Wang (2003). They used a similar model of Gonzalez-Rivera & Helfand's for studying the egg market in six Pacific States. The results indicated that there was cointegration between the prices and the LOP holds perfectly. Besides, they found that Arizona, California, and Washington lead the egg prices in Pacific States.

The work of Rashid (2004) also incorporated the methodology of Gonzalez-Rivera & Helfand (2001). This study looks like one of the most complete in this area. Rashid (2004) studied Uganda maize market after liberalization. He analyzed extension of integration, causality among spatial locations, and the relative importance of spatial regions in price formation. Therefore, he made use of Gonzalez-Rivera & Helfand (2001), Gonzalo & Granger (1995), Masconi & Gianninni (1992), and Hall & Milne (1994). The results suggested that a small number of market locations, mainly large consumption and production districts, dominate the long-run price transmission.

Finally, there is Moser *et al.* (2005) that is one of the papers that was rich in data. They had data on rice prices, transaction costs, and infrastructure availability for Madagascar. This data permitted them to measure market integration across space, time, and form. This work examined the extension of the market in sub-regional, regional, and national scales, using Gonzalez-Rivera & Helfand (2001) model. They tried to determine if non-integration was because of high transaction costs or lack of competition. These authors observed that markets were well integrated at a sub-regional level and that factors such as high crime, remoteness, and lack of information were among the factors limiting competition. A lack of competition persisted at a regional level and high transfer costs impeded spatial market integration at a national level.

3.5. Summary

As can be noted, there is a lot of discussion about the subject of market integration. First of all, we are going to use Stigler's market definition in which the market represents the geographic area within which the price of a good is the same, except for transportation costs. Then, the Gonzalez-Rivera & Helfand's market integration definition seems like the most adequate for this study. They believed that market integration involve flows of goods and information. This is an interesting definition because it requires that both goods and information be correlated between the regions. In this case, transaction costs are very important. However, we do not believe that they are nonstationary, since that nobody could measured. It is assumed as the majority of the literature does, because this kind of data is unobservable. It seems like a misunderstanding between theory and methodology. According to arbitrage theory, high transaction costs can affect the market integration by reducing or even impeding trade. However, it does not mean that transaction cost can affect the cointegration results. The cointegration model measures if two or more variables move together in the long-run. So, even if these economic variables are influenced by various different factors (i.e., political shocks, weather, regional subsidies or levies, etc), they have still been cointegrated or not, independent of the influence of other factors. The theory of cointegration says that there is an attractor force between the variables cointegrated and nothing can change or affect it (even transaction costs). So we really believe that transaction costs can be a cause for market non-integrated or segmented, but not a critic for cointegration models.

As affirmed by Ardeni (1989) there is some kind of systematic distortions that can occur between commodity prices in different countries. However, we believe that it is more likely to happen among countries than states. In our case, we are going to analyze price between states in Brazil, and except by ICMS[14], that differs among states, the other variables like political and economy factors that affect the transaction costs and prices in

[14] ICMS is a kind of tax that should be paid by every good or service that passes through boundaries of a state in Brazil.

one state, affect all the other states too. Different from the USA, the Brazilian states have the same rules and laws, and hence, are equally influenced.

Therefore, we are going to use cointegration methods in order to analyze the extension and pattern of integration. Besides this, we are going to use the bivariate model because the multivariate model presents problems with dimensionality, i.e., the power of test decreases gradually when the number of variables increases. It means a problem for our study, because we intend to work with the 26 states of Brazil. So, our idea is work with a bivariate model first in order to identify the states that are cointegration in pairs. This procedure will allow excluding some locations, and after we can apply the multivariate model with the other states. As a complementary method, we will utilize the impulse response functions associate with the DAG. Although this method presents problems, it is the only one that measures the reaction of one market in response to shocks in another market. The persistence profile seems like an interesting model, but it can just measure the effects of a wide-shock, instead of an individual shock. In order to treat the constraints of impulse response functions, we chose to use the Bernanke's decomposition.

Besides, we assume that integrated markets are consistent with imperfect competition, since in reality, competitive markets are almost absence. Therefore, we agreed with Barrett (1996) that market integration is neither necessary nor sufficient condition for market efficiency. It was also supported by Facker & Goodwin (2000) and Faminow & Benson (1990). In our point of view, market efficiency and market integration are different things and one should use different methods to measure each one.

Finally, we can conclude that a mix of price and trade flow data in a model should be the best alternative for analyzing market integration. So, we decided to use the extension of the point-space model presented by Faminow and Benson (1990), because this model has more realistic assumptions, like imperfect competition. The agents-on-links also is an interesting model, but it is necessary to use data (price series) for cities and this kind of data is unavailable for Brazil.

4. THEORY

Some theoretical concepts were already discussed in the last chapter. Here, we are going to broach the Law of One Price, the Point-space model, and the Faminow and Benson Model.

4.1. Law of One Price

According to McNew (1996), the Law of One Price is the cornerstone of studies about market integration. It is the equilibrium condition, which guarantees no arbitrage opportunities and is necessary for spatial price efficiency.

The LOP is a concept created for international trade analysis. In this sense, the primary version of LOP states that when there is not transportation costs or barriers to trade, homogeneous commodity prices in two different markets expressed in a common currency are equated according to

$$P = P^* E \tag{8}$$

where P and P^* are domestic and foreign prices and E is the exchange rate (units domestic currency per unit of foreign currency) (Krugman & Obstfeld, 2003). This is the strict form of the LOP. More generally, the LOP states that, abstracting transportation costs, the same good is sold for the same price in spatially separated markets. According to Fackler & Goodwin (2000), the LOP has three versions:

- Weak version: in this case, it is not possible to distinguish between LOP and arbitrage;
- Strong version: the spatial arbitrage condition holds as an equality (the presumption being that trade is continuous). It is generally the version tested;
- Aggregate version: stated in terms of price indices, is called Purchasing Power Parity (PPP). It states that the exchange rate is proportional to the ratio of price levels in two countries.

The assumption needed for PPP to hold are far more restrictive than for the LOP, even if PPP is applied only to traded goods. Under PPP, the exchange rate is determined simultaneously by the ratio or the aggregate price levels P and P^* in the two countries (i.e., $E = P/P^*$), while under the LOP, P and P^* are prices for specific commodities and E is an exogenous variable (Krugman & Obstfeld, 2003).

Basically, the LOP states that price differences between separated locations that trade with each other will equal transfer costs. While price differences between locations that are not engaged in trade with each other will be less than or equal to transfer costs (Tomek & Robinson, 1990).

The equation (8) refers to the form of LOP that is often tested; therefore such tests must be interpreted not so much as tests of equilibrium conditions, but tests that are conditional on assumptions regarding trade linkages. Thus, violations of the LOP can indicate both a lack of a stable trading relationship and a disequilibrium situation (Fackler & Goodwin, 2000). On the other hand, Ardeni (1989) stated that deviations from the equilibrium might occur in the short-run; however, in the long-run exchange rate and prices should be proportional.

This kind of test distinguishes two extreme situations. At one extreme are completely separated markets (i.e., markets where the LOP does not hold) and at the other are perfectly integrated markets (i.e., those markets which exhibit the strong form of the LOP) (Fackler & Goodwin, 2000). However, it is important to remember that perfect integration of a market is different from the strong version of LOP. The LOP holds even though regions have price transmission ratios less than 1. The price transmission ratios equaling one imply a strong form of LOP. The LOP is the concept used in a spatial equilibrium of Takayama and Judge Model.

4.2. Takayama and Judge Model

The Takayama and Judge Model or point-space model defines the relationship between prices in separated geographic markets. Samuelson (1952) first studied this model. He worked with two markets, and analyzed this problem using the maximization of area under all the excess demand curves minus the area under the excess supply curves, minus transportation costs. This process of maximization results in a spatial equilibrium solution.

Takayama & Judge (1971) extended it through algorithms using a linear dependent price, and demand and supply functions. They developed a model considering a perfect competition market, without barriers or government interventions. They also considered that production and consumption occur in separated regions and the boundaries of the market are fixed. Therefore, there are inter market transactions depending on the transportation cost, but no intra market trades are allowed. It means that all transactions take place at a point, and all buyers and sellers are located at a single point.

Basically, the Takayama and Judge Model tests the LOP. In other words, it tests if price differences between two or more geographically separated markets are greater than, less than, or equal to transfer costs (Faminow & Benson, 1990). Summarizing this model, supply and demand in each region can be presented, respectively, as

$$P_{si} = s_i(Q_{si}) \tag{9}$$

$$P_{di} = f_i(Q_{di}) \tag{10}$$

where P is price, s represents supply, d refers to demand, i represents each region, and Q is quantity.

Samuelson (1952) shows that market equilibrium is achieved through the maximization of a net social payoff function, given by the sum of producer surplus and consumer surplus. In this case, the quasi-welfare function or net social payoff (as defined by Samuelson) is given by

$$W_i(Q_{si}^*, Q_{di}^*) = \int_0^{Q_{di}^*} P_{di}\, dQ_{di} - \int_0^{Q_{si}^*} P_{si}\, dQ_{si} \tag{11}$$

where W refers to welfare function. Therefore, the total welfare function is

$$NW = \sum_i W_i(Q_{di}, Q_{si}) - \sum_i \sum_j c_{ij} T_{ij} \tag{12}$$

where c represents the transport rate between regions i and j, T refers to quantity of good transported between regions i and j. Therefore, we have

$$Q_{di} \leq \sum_j T_{ji} \text{ for all } i \tag{13}$$

$$Q_{si} \geq \sum_j T_{ij} \text{ for all } i \tag{14}$$

The next step in the Takayama and Judge Model applies the maximization of net welfare (NW) as

$$\text{Max } \sum_i \left(\int_0^{Q_{di}^*} P_{di} dQ_{di} - \int_0^{Q_{si}^*} P_{si} dQ_{si} \right) - \sum_i \sum_j c_{ij} T_{ij} \tag{15}$$

$$\text{s.t. } Q_{di} - \sum_j T_{ji} \leq 0$$

$$-Q_{si} + \sum_j T_{ij} \leq 0$$

$$Q_{si}, Q_{di}, T_{ij} \geq 0$$

Using the Lagrange-multiplier and Kuhn Tucker conditions, the authors solved this problem and found the spatial equilibrium price and quantity. However, this model is very restricted, because it considers only perfect competition, inter market trades, efficient market, continuous trade, etc. Therefore, in order to apply assumptions that are more realistic and consider intra market trades, we are going to use an adaptation of the Faminow and Benson Model.

4.3. Theoretical model of the milk market

This model is based on Faminow & Benson (1990). We consider that in the geographic separated markets, transportation (transaction) costs have an important role. The fundament of this theory is that farmers differentiate between buyers based on locations. It means that, even if there are a large number of processors in a country, farmers choose to trade with those buyers located near them. Faminow & Benson (1990) explained that even if many buyers are located at the same geographic area, each farmer considers only nearest rivals as major competitors.

It happens because when there is locational interdependence and buyers and sellers are spatially distributed, transfer costs affect the net price received or paid (Faminow & Benson, 1990). So, as affirmed by Greenhut (1971), a spatial market should not be characterized as perfectly competitive when either buyers or sellers are dispersed and transportation costs are relevant. This situation is what generally happens in food and agricultural markets.

Our model assumes that farmers and processors are both dispersed in a region, and imperfect competition can exist. Besides, the regional boundaries are a function of relative prices and not a fixed limit as presupposed by the Takayama and Judge Model. In addition, this model allows intraregional transactions, which is a feature of agricultural markets.

Thus, the milk market integration in Brazil can be studied by developing a model of spatial oligopsonistic competition. This model is interesting because it allows visualization of price reaction functions, which is one of the most important characteristics of market integration.

The model consists of a few processors buying milk from several dispersed farmers located on a linear market. So, assuming n identical farmers uniformly and continuously distributed between m spatially separated buyers, where $m < n$. Figure 3 can illustrate this situation considering three processors (*X, Y*, and *Z*).

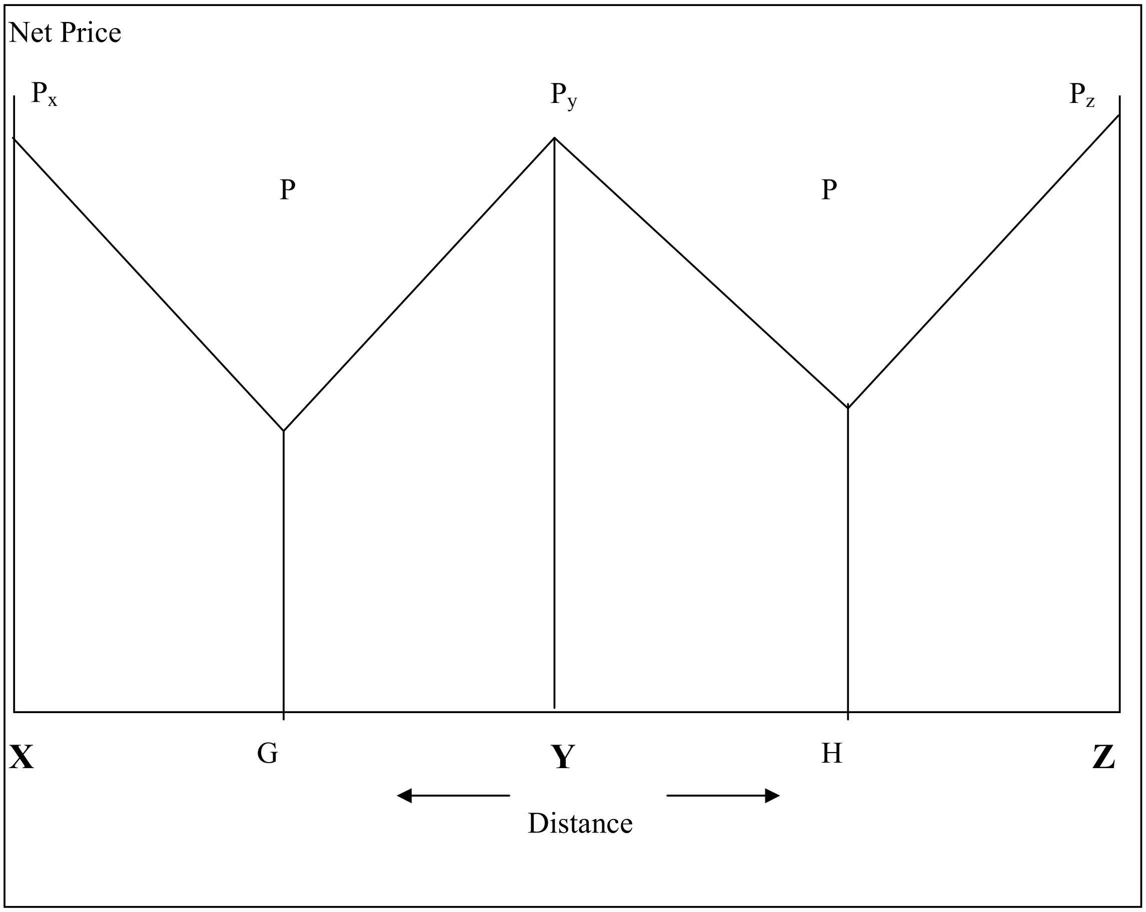


Figure 3 – Spatial price representation for an oligopsonistic situation.
Source: Developed by the author.

Figure 3 is based on the diagrams used by Ohta (1988) and Faminow & Benson (1990) to analyze the relationship between spatial markets. It shows how the prices vary in an oligopsonistic market. Basically, as the distance from buyers increase, the price net of transportation cost of delivery decreases. Therefore, for each processor, the price paid for milk at the plant is $P_{i,}$ but farmers must pay delivery costs. So the net price to the farmers $(\overline{P})$is

$$\overline{P} = P_i - u\,t \qquad (16)$$

where P_i is the price paid by the processors; u represents the units of distance between farmers and processors; and t refers to transport rate per mile (u x t means transportation cost). The boundaries between the areas from which farmers deliver to each processor occur at the points where net prices are equal, at points *G* and *H*. It means that the point *G* is a geographic limit among processors *X* and *Y*, while the point *H* is the boundary between processors *Y* and *Z*. However, it is important to reaffirm that these boundaries are variable as P_i as *t* vary (Figure 3).

In order to simplify the model, we considered only three processors or buyers with market power, while farmers are price takers (i.e., it represents oligopsonistic competition). We are working with oligopsony because the Brazilian milk market has this characteristic (Chapter 2).

Therefore, the inverse supply equation for milk is given by

$$\overline{P} = a + bQ \tag{17}$$

where $\overline{P}$ is the net price (i.e., price paid to farmers minus transportation costs); *a* and *b* are positive constants, and *Q* refers to the quantity supplied.

In order to simplify the analysis, we assumed *t* equals 1 to carry the product one unit of distance (Faminow & Benson, 1990). So, equation (16) becomes

$$\overline{P} = P_i - u \tag{18}$$

Combining (17) and (18), we find the supply equation

$$P_i - u = a + bQ$$

$$Q = \frac{P_i - u - a}{b} \tag{19}$$

The individual farm supply function (equation (19)) is aggregated according to each processor's area (Figure 3). In other words, the individual farm supply function is integrated with respect to the distance *u*, because the farmers differentiate buyers through the distance. Therefore, the milk price varies according to distance (Faminow & Benson, 1990). In this case, the aggregate supply is given by

$$Q_x = \int_0^G \left(\frac{P_x - u - a}{b} \right) du = \frac{(P_x - a)u}{b}\Big|_0^G - \frac{u^2}{2b}\Big|_0^G = \frac{u(2P_x - 2a - u)}{2b}\Big|_0^G = \tag{20}$$

$$Q_x = \frac{G(2P_x - 2a - G)}{2b}$$

$$Q_y = \int_0^{D-G} \left(\frac{P_y - u - a}{b} \right) du + \int_0^{D-H} \left(\frac{P_y u - a}{b} \right) du$$

$$Q_y = \int_0^{D-G} \left(\frac{P_y - a}{b} \right) du - \int_0^{D-G} \left(\frac{u}{b} \right) du + \int_0^{D-H} \left(\frac{P_y - a}{b} \right) du - \int_0^{D-H} \left(\frac{u}{b} \right) du$$

$$Q_y = \left(\frac{P_y - a}{b} \right) u \Big|_0^{D-G} - \left(\frac{u^2}{2b} \right)\Big|_0^{D-G} + \left(\frac{P_y - a}{b} \right) u \Big|_0^{D-H} - \left(\frac{u^2}{2b} \right)\Big|_0^{D-H}$$

$$Q_y = \left(\frac{P_y - a}{b} \right)(D-G) - \left(\frac{1}{2b} \right)(D-G)^2 + \left(\frac{P_y - a}{b} \right)(D-H) - \left(\frac{1}{2b} \right)(D-H)^2$$

$$Q_y = \frac{DP_y - Da - GP_y + Ga}{b} - \frac{D^2 - 2DG + G^2}{2b} +$$

$$\frac{DP_y - Da - HP_y + Ha}{b} - \frac{D^2 - 2DH + H^2}{2b}$$

$$Q_y = \frac{2DP_y - 2Da - 2GP_y + 2Ga - D^2 + 2DG - G^2}{2b} +$$

$$\frac{2DP_y - 2Da - 2HP_y + 2Ha - D^2 - 2DH + H^2}{b}$$

$$Q_y = \frac{D(2P_y - 2a - D)}{2b} - \frac{\backslash G(2P_y - 2a + G)}{2b} + \frac{DG}{b} +$$

$$\frac{D(2P_y - 2a - D)}{2b} - \frac{\backslash H(2P_y - 2a + H)}{2b} + \frac{DH}{b}$$

$$Q_y = \frac{D(2P_y - 2a - D + G + H)}{b} - \frac{G(2P_y - 2a - G)}{2b} - \frac{H(2P_y - 2a + H)}{2b} \tag{21}$$

$$Q_z = \int_0^H \left(\frac{P_z - u - a}{b} \right) du = \frac{H(2P_z - 2a - H)}{2b} \tag{22}$$

where Q_i represents the aggregate supply function facing each processor; *G* and *H* are distances over which firms *X* and *Z*, buy milk; and *D* is the total area or distance between processor *X* and *Y*, and also between *Y* and *Z* (Faminow & Benson, 1990). The distance *D* between processors is the same because previously we considered that the firms are uniformly distributed.

At this point, we are going to do the profit-maximization between monopsony and perfect competition, since each processor acts as monopsony for the closest farmers. As we affirmed before, although there are many processors in the market, farmers choose to deal with the nearest buyers. Therefore, even though we are considering the global processor market as an oligopsony, the local market for each processor is characterized as a monopsony. Moreover, the other side of the transaction is taken as a perfect competition, i.e., the relationship between processors and consumers and/or retailers is considered as a perfect market. Although the big supermarkets in Brazil exert certain power, they act only in the big cities so that, in the rest of the country, a large number of small groceries and bakeries position themselves on this side of the trade (i.e., they buy from the processors). So, the profit-maximization can also be given by

$$Max\ \pi_i = P_Q Q_i - P_i Q_i - c_i Q_i - F_i = \left(P_Q - P_i - c_i\right)Q_i - F_i \tag{23}$$

where π_i refers to the profit for processor *i*; p is the milk price received by processor *i*, which is fixed in this case because the processor sells in a competitive market; c_i is the marginal processing cost for processor *i* not including the cost paid to farmers for milk; and F_i is the fixed cost for processor *i*. The profits are maximized when $\partial \pi_i / \partial P_i = 0$, as

$$\frac{\partial \pi_x}{\partial P_x} = \frac{d\left[\left(P_Q - P_x - c_x\right)Q_x\right]}{dP_x} = 0$$

$$\frac{d\pi_x}{\partial P_x} = \frac{d\left\{\left(P_Q - P_x - c_x\right)\left[\frac{G(2P_x - 2a - G)}{2b}\right]\right\}}{dP_x} = 0$$

$$\frac{\partial \pi_x}{\partial P_x} = -\left[\frac{G(2P_x - 2a - G)}{2b}\right] + \left(P_Q - P_x - c_x\right)\left[\frac{\partial G}{\partial P_x}\frac{(2P_x - 2a - G)}{2b} + \frac{G\left(2 - \frac{\partial G}{\partial P_x}\right)}{2b}\right] = 0 \tag{24}$$

$$\frac{d\pi_y}{\partial P_y} = \frac{d\left[\left(P_Q - P_y - c_y\right)Q_y\right]}{dP_y} = 0$$

$$\frac{d\pi_y}{\partial P_y} = \frac{d\left\{\left(P_Q - P_y - c_y\right)\left[\frac{D(2P_y - 2a - D + G + H)}{b} - \frac{G(2P_y - 2a - G)}{2b} - \frac{H(2P_y - 2a + H)}{2b}\right]\right\}}{dP_y} = 0$$

$$\frac{\partial\pi_y}{\partial P_y} = -\left[\frac{D(2P_y - 2a - D + G + H)}{b} - \frac{G(2P_y - 2a - G)}{2b} - \frac{H(2P_y - 2a + H)}{2b}\right] + \tag{25}$$

$$\left(P_Q - P_y - c_y\right)\left\{\frac{D\left(2 + \frac{\partial G}{\partial P_y} + \frac{\partial H}{\partial P_y}\right)}{b} - \left[\frac{\partial G}{\partial P_y}\frac{(2P_y - 2a - G)}{2b} + G\frac{\left(2 - \frac{\partial G}{\partial P_y}\right)}{2b}\right] - \left[\frac{\partial H}{\partial P_y}\frac{(2P_y - 2a + H)}{2b} + G\frac{\left(2 + \frac{\partial H}{\partial P_y}\right)}{2b}\right]\right\}$$

$$\frac{\partial\pi_z}{\partial P_z} = -\left[\frac{G(2P_z - 2a - G)}{2b}\right] + \left(P_Q - P_z - c_z\right)\left[\frac{\partial G}{\partial P_z}\frac{(2P_z - 2a - G)}{2b} + \frac{G\left(2 - \frac{\partial G}{\partial P_z}\right)}{2b}\right] = 0 \tag{26}$$

The equations (24), (25), and (26) will be used further to discuss the geographic limits of the market. Analyzing the geographic boundaries between two buyers, we can say that they occur where net prices are equal (Faminow & Benson, 1990). So

$$P_x - G = P_y - (D - G)$$

$$G = \frac{1}{2}(P_x - P_y + D) \tag{27}$$

and

$$P_z - H = P_y - (D - H)$$

$$H = \frac{1}{2}(P_z - P_y + D) \tag{28}$$

It means that the oligopsonistic situation can be analyzed by the boundary conjecture (Faminow & Benson, 1990)

$$\frac{\partial G}{\partial P_x} = \frac{1}{2}\left(1 - \frac{\partial P_y}{\partial P_x}\right) \quad (29)$$

$$\frac{\partial H}{\partial P_z} = \frac{1}{2}\left(1 - \frac{\partial P_y}{\partial P_z}\right) \quad (30)$$

It suggests that boundaries conjecture on the oligopsony depends on *X*'s expectations regarding *Y*'s price response (Faminow & Benson, 1990). The exact value of P_x, P_y, P_z obtained from equations (24), (25), and (26) depends on supply and demand parameters as well as on the size of the firms' purchase area (*G* and, or *H*) and the boundary conjecture. So, based on these equations we can affirm that the prices reaction are characterized as

$$p_x = p_x\left(a, b, p, G, \frac{\partial G}{\partial P_x}\right) \quad (31)$$

$$p_y = p_y\left(a, b, p, G, H, \frac{\partial G}{\partial P_y}, \frac{\partial H}{\partial P_y}\right) \quad (32)$$

$$p_z = p_z\left(a, b, p, H, \frac{\partial H}{\partial P_z}\right) \quad (33)$$

In this case, since *G* is a function of P_x and P_y, the price of milk paid by the firm *X* depends also on the price set by the firm *Y*. At the same way, *H* is a function of P_z and P_y, so that the price of milk paid by the firm *Z* depends on the price set by the firm *Y*. Hence, indirectly, P_x and P_z are interconnected as well. They are integrated through *G*, P_y and *H*. Then, we can conclude that these three processors are in the same market, or that they are integrated, since the variation in price of one of the firms causes changes in the prices of the others. In other words, when one of these buyers decides to reduce the milk price, this information is transmitted through the price for the other buyers, causing a change in their prices as well. It is the genuine meaning of market integration: flows of goods and information.

By this analysis, we can also conclude that the variation in the processor price is a function of the distance between farmers and processors, as well as the transportation costs. So, as other studies affirm, we showed that transportation costs can be responsible for integration between separated markets. It means that these two factors can be determinant of the market integration.

Otherwise, as it is a dynamic process, feedbacks can cause additional changes on the prices; some parameters can vary over time; and different lags can occur. It is the price formation process.

5. METHODOLOGY

As emphasized before, the utilized analytical model is based on Gonzalez-Rivera and Helfand (2001), but with some modifications. Differently from other econometric models, this methodology is a stepwise procedure compounded by several equations and tests.

This methodology is divided in three parts: extension of the market, pattern of integration, and degree of integration. The extension of the market is determined through the measure of a self-sufficiency index, unit root test, and Johansen test. The last one is focused on the searching for a common trend between the time series. The pattern of interdependence is studied using the VEC/VAR analysis in association with the DAG. Lastly, the degree of integration is measured by the impulse response functions derived by the Bernanke decomposition.

5.1. Extension of the market

As specified previously, the extension of the market refers to geographical boundaries of the market. It means that we are going to identify which state is included in the Brazilian milk market. For that, according to Gonzalez-Rivera & Helfand (2001) the first step is to identify the set of locations that are interconnected either directly or indirectly by continuous unidirectional trade using the Index of Self-Sufficiency.

5.1.1. Index of Self-Sufficiency

The Index of Self-Sufficiency (ISS) will present an estimate of the quantity of milk commercialized between the Brazilian states. However, the trade data are not available in Brazil. Because of this, we will estimate the annual trade flow for each state of Brazil, in order to exclude the regions that have trade reversals (i.e., milk exporters that become importers and vice versa) as well as identifying the regions that are close to self-sufficiency, and consequently are likely to present discontinuous trade.[15].

So, the definition of the size of the market evolves the analysis of the spatial pattern of the production, consumption and trade on Brazilian states, and afterward the search for those states that share a common integrating factor[16].

The annual trade flow can be obtained by annual milk consumption in each state of Brazil. The proxy for state consumption is made by multiplication between population and per capita consumption in each state. Associating this information with the production data in each Brazilian state we can calculate the ISS that is given by ratio of the state's production share to its consumption share as

$$ISS = \frac{state\ \ production}{state\ consumption} \tag{34}$$

The closer to 1, the closer to self-sufficiency is a state. Values larger than 1 indicate that a state is a net milk exporter, while values smaller than 1 indicate that the state is a milk importer.

5.1.2. Common integrating factor

After identifying the states that are connected by trade, we begin the search for those states that share the common integrating factor, i.e., those states that have the same trend in the long-run. The existence of one and only one integrating factor for all series is obtained when the prices are cointegrated, and there are *n - 1* cointegrating vectors (Gonzalez-Rivera & Helfand, 2001).

---

[15] The exclusion of regions that have discontinuous and/or reverse trade is justified due to the fact that these situations are only analyzed by using switching regime models.

[16] The definition and importance of the common integrating factor or common trend will be present ahead in this chapter.

In order to understand the importance of the existence of a common integrating factor, it is useful to consider the following situation, proposed by Gonzalez-Rivera & Helfand (2001). A *n x 1* nonstationary *I (1)* vector of log-prices[17] $P_t = \{ p_{1t}, p_{2t}, ..., p_{nt} \}$ where p_{it} refers to the log-price of a commodity at time *t* in market *i*. As a time series, P_t can be decomposed into two components as

$$P_t = A_{nxs} f_t + \widetilde{P}_t \qquad (35)$$

where A_{nxs} is a matrix of coefficients, f_t is an *s x 1* vector of *s (s < n)* common unit root factors and $\widetilde{P}$ is a *n x 1* vector of stationary components. In other words, P_t is composed of a permanent and a transitory component (Engle & Granger, 1991). It means that each element in the vector P_t is formed by a linear combination of a smaller number of *I(1)* common factors (permanent component) plus an *I(0)* or transitory component (for example $p_{it} = \sum_{j=1}^{s} a_{ij} f_{jt} + \widetilde{p}_{it}$) (Gonzalez-Rivera & Helfand, 2001).

The existence of the same stochastic trend among the series is the reason for the coincidence between the movements of the variables p_{it} in the long-run. Equation (35) is called a common factor representation and it exists if and only if there are $n - s$ cointegrating vectors among the elements of the vector P_t (Engle & Granger, 1991). It is the Granger representation theorem.

In this sense, if there are *k* variables, there can exist *r* cointegrating vectors and *s* common trends (common factors), so that $s = n - r$. Doldado et al. (1999), cited by Seddighi et al. (2004), affirmed that, in a multivariate situation, verifying the existence of $(n - r)$ common trends is equal to testing for *r* cointegrating vectors.

So, our definition of the extent of an integrated market requires that $s = 1$ because we are searching for locations that share the common tendency in the long-run. Therefore, the search for only one common integrating factor between the time series is equivalent to searching for $n - 1$ cointegrating vectors (Gonzalez-Rivera & Helfand, 2001). It can be made by the application of the Johansen procedure.

[17] It is common in market integration studies to perform analysis on the log-prices.

5.1.2.1. Cointegration test

Cointegration is a statistical framework to test for long-run equilibrium relationships among nonstationary series and is used to search for one common integrating factor between regions. The starting point in the analysis of cointegration is the identification of the nonstationarity of the series.

Firstly, we are going to use the correlogram, i.e., the plot of the autocorrelation functions of the time series. The correlogram shows the values of the autocorrelation coefficients for ($\hat{\rho}$) against the lags (*k*). The literature states that the correlogram of a stationary series should converge to zero geometrically, as occur with a white noise's correlogram. However, Enders (2004) affirms that the correlogram is inexact, because what can appear as a unit root process for one can be seen as a stationary series for another. For this reason, we are going to consider other unit root tests.

5.1.2.1.1. Unit root test

5.1.2.1.1.1. ADF test

According to Gujarati (2000) a time series is weakly stationary if the mean $E(X_t) = \mu$ is constant for all *t*, the variance $var(X_t) = E(X_t - \mu)^2 = \sigma^2$ is constant for all *t*, and the covariance $cov(X_t, X_{t+k}) = E[(X_t - \mu)(X_{t-k} - \mu)] = \gamma_k$ is constant for all *t* and $k \neq 0$. This signifies that the mean and variance are constant over time and the (auto)covariance between two time periods , such *t* and *t* + *k*, depends only on the distance *k* between these time periods and not on the actual time period *t* at which these covariances are calculated (Gujarati, 2000).

One of the most used tests for stationary analysis is the Augmented Dickey Fuller (ADF), which is given by

$$P_t - P_{t-1} = \Delta P_t = \alpha + \beta t + \delta Y_{t-1} + \sum_{k=1}^{p-1} \sigma \Delta Y_{t-k} + \varepsilon_t \quad (36)$$

where P_t represents the milk price series in each state of Brazil; α refers to drift; β refers to a deterministic trend coefficient; *t* is time; δ refers to the existence of a unit root in the series P_t; *k* is the number of lags; and ε_t is a errors vector. The ADF tests the hypothesis H_0: $\delta = 0$ against H_a: $\delta > 0$ in equation (36). This equation is estimated by ordinary least squares (OLS). If the null hypothesis is accepted, the series has a unit root and is nonstationary. Otherwise, the series does not have a unit root and is stationary. The statistical significance of estimated coefficients can be verify by Dickey & Fuller (1976) or MacKinnon (1991)'s tables[18].

5.1.2.1.1.1.1. Lag length selection

One important point in unit root analysis is the selection of lag length. Too few lags cannot capture the actual error process, and therefore misestimate the model. On the other hand, too many lags reduce the power of the test. Therefore, the ADF test can point out the existence of a unit root for some lag lengths but not for others. In addition, small models tend to have better out-of-sample performance than large models (Enders, 2004).

Moreover, as affirmed by Enders (2004) and Manual RATS (2004) the Principle of Parsimony, associated with Box-Jenkins approach, is the primary idea to select lag lengths and models. This principle states that parsimonious models produce better results than overparameterized models.

The *Q*-statistic (Portmanteau test, Box-Pierce test, and Ljung-Box test) has been used to select the correct lag length. It tests whether a set of autocorrelations are significantly different from zero. According to Enders (2004), high sample autocorrelations lead to large values of *Q*, so that a white noise process (i.e, a process in which all correlations are zero) would have a *Q* value of zero. However, Enders (2004) and Davis & Newbold (1979) affirm that this test works poorly even in moderately large samples.

Enders (2004) recommends starting the process of selecting the lag with a relative long lag length and pare down by using a *t*-test, Akaike Information Criterion (AIC) or Schwartz Criterion (SC) [19]. In very large samples, these methods should select the same lag length.

[18] Standard values of the t-statistic cannot be used because in a nonstationary process, the variance becomes infinitely large as *t* increases. So Dickey-Fuller developed a finite sample and asymptotic distributions of the *t*-statistic. It is a one-sided test, setting the critical region on the values smaller than 1. In this case, if the series presents values higher than these critical values, we can accept the null hypothesis, and, consequently conclude that the series is nonstationary. For example, if the *t*-statistical is -3 and considering the critical values -2.89 (5%) and -3.51 (1%), you can reject the null hypothesis at 5%, but not at 1% (Enders, 2004).

[19] In order to select the lag length, AIC works by minimizing the sum $T\ln(SSR) + 2(1-p-q)$, while SC minimizes $T\ln(SSR) + (1-p-q)\ln(T)$. The benefits to include more lags are measured by SSR, and the costs are given by the second term. So, the SC incorporates a larger penalty, so that the marginal cost to add lags using SC exceeds that of the AIC. Because of this, SC selects a more parsimonious model.

However, often the SC selects a more parsimonious model than the AIC or *t*-test. Because of this, the SC is preferred when they disagree. Nevertheless, when these two tests select different lags, we are going to check the Hannan-Quinn criterion (HQC) and the statistical significance[20] of the coefficients using both selected lags. The HQC will be used because its results are very similar to SC, and according to Liew (2006), with a relatively large sample (120 or more observations), HQC is found to outdo the rest in correctly identifying the true lag length.

5.1.2.1.1.1.2. Intercept and trend

Moreover the lag length selection, there is another problem on the unit root test. The tests for identifying the presence of unit roots in a variable are conditional on the presence of drifts and deterministic trends and the tests for the presence of drifts and deterministic trends in a variable are conditional on the presence of a unit root (Seddighi et al., 2000).

In addition, we are going to follow the Enders (1995)'s procedure. It starts with the most general case (i.e., including trend and drift), and moves toward the most specific case. If it is known that the series has a drift or trend, then the null hypothesis of a unit root can be tested using the standard normal distribution. However, due the fact that the unit root tests have a low power of rejecting the null hypothesis of a unit root, these procedures stop anytime that the null hypothesis is rejected. It is then concluded that the time series under investigation is stationary (Seddighi et al., 2000).

Moreover, Coelho (2002) states that the presence of drift and trend can be determined through the level of significance of these elements. Therefore, we are going to start with the Enders' procedure (model with intercept and trend), and then check the significance of these factors to determine the most appropriate model.

5.1.2.1.2. Johansen test

If we find out that the price series in every state of Brazil are nonstationary, we can proceed to the Johansen test. In opposition, those milk prices that show stationarity will be exclude from the analysis, since the cointegration analysis requires that all the time series present the same order of integration.

[20] The *t*-statistical will be analyzed considered 5% of significance or the value of 1.96. In absolute values, when the *t* calculated is greater than *t* tabulated, the coefficient is considered significantly different from zero.

Johansen's approach is one-stage procedure. It has the advantage that, when the number of cointegrating relationships is identified, we have both the cointegrating vectors and the short-run dynamics of the system. The Johansen procedure identifies the cointegrating rank *r* and provides estimates of the cointegrating and adjustment matrices, using the Maximum Likelihood Method.

The initial point is the vector autoregressive model (VAR)

$$\Delta P_t = \alpha + \Gamma_1 P_{t-1} + \Gamma_2 P_{t-2} + ... + \Gamma_{n-1} P_{t-n+1} + \varepsilon_t \qquad (37)$$

where P_t is a vector of a milk price series at time *t*; α is a vector of constants; Γ_i is a matrix of coefficients relating series changes at lagged *i* period to current changes in the series; and ε_t is a vector of independent identically distributed (IID) errors. As occurred with the ADF test, the cointegration analysis is also vulnerable to the lag length and the presence or not of the intercept and trend. Then, we are going to use the same criteria described before to select the elements in the cointegration test.

However, if some series are nonstationary and cointegrated, a pure VAR in difference will be misspecified. Under this circumstance, the vector error correction model (VEC) is appropriate to study the behavior of the series in the short and long-run. The VEC model with *n* lags is given by

$$\Delta P_t = \alpha + \Gamma_1 \Delta P_{t-1} + \Gamma_2 \Delta P_{t-2} + ... + \Gamma_{n-1} \Delta P_{t-n+1} + \prod P_{t-1} + \varepsilon_t \qquad (38)$$

where P_t is a vector of milk price series; α is a vector of constants; Γ_i is a matrix of coefficients relating series changes at lagged *i* period to current changes in the series; $\prod$ is a matrix of rank $r < k$; and ε_t is a IID vector of errors. The matrix $\prod$ can be written as $\Pi = \alpha\beta'$; α is an matrix of coefficients or matrix of adjustment coefficients, and β is an matrix of cointegrating vectors. Using this expression for $\prod$ we can rewrite the equation as $\Pi P_{t-1} = \alpha\beta' P_{t-1} = \alpha Z_{t-1}$. The error correction term, also know as short-run disequilibrium, is Z_{t-1}.

The rank of Π, *r*, which is equal to the number of linearly independent cointegrating vectors, also adds up to the number of characteristic roots, λ, and is used to investigate the cointegration relationships. Three situations can occur (Johansen, 1995):

- If $r = n$, where n is the number of series analyzed, the variables in levels are stationary, and hence there is not cointegration. So, the VAR model could be formulated in terms of the levels of all variables.
- if $r = 0$, none of the linear combinations are stationary (i.e., there is not cointegration between the variables). In this case, the VAR model could be formulated in terms of the first differences of the variables.
- When $0 < r < n$, there exists r cointegrating vectors or r stationary linear combinations between the stationary series. It means that there exist r cointegration relationships among the time series.

We are going to use the trace test to identify the number of cointegrating vectors in the system. The trace test (λ_{trace}) is written as

$$\lambda_{trace} = -\mathrm{T}\sum_{i=r+1}^{n} \ln(1-\lambda_i) \tag{39}$$

where T is the number of observation, and λ_i refers to the i^{th} eigenvalue.

It tests the hypothesis that there is at most r cointegrating vectors. The alternative hypothesis is that there exists more than r cointegrating vectors. Critical values are found in Maddala & Kim (1998).

With the Johansen procedure it is possible to determine the locations that belong to the same market. In this sense, we started the search for a common factor using a bivariate model[21], because it is not possible to work with all states together. So, we are going to test

[21] We choose to work with pairs of price series, and later, with small groups (five series by group) because the cointegration approach does not allow a high number of variables. It is a numerical methods problem. As the cointegration works with matrices, this problem comes from the condition number of the matrix. All matrices have a condition number defined by the maximum eigenvalue over the minimum eigenvalue. The condition number gives a bound on how inaccurate the solution of a regression will be after the approximate solution. "The condition number is a measure of stability or sensitivity of a matrix (or the linear system it represents) to numerical operations. It means that we may not be able to trust the results of computations on an ill-conditioned matrix. Matrices with condition numbers near 1 are said to be *well-conditioned*. Matrices with condition numbers greater than one (such as around $10^5$for a 5 x 5 Hilbert matrix) are said to be *ill-conditioned*" (PlanetMath, 2006). Thus, if the condition number is large, even a small error in b may cause a large error in x (Wikipedia, 2006). For more information about condition numbers, see Wikipedia (2006), Holistic Numerical Methods Institute (2006), and PlanetMath (2006). However, this kind of problem depends on the data set. Some data sets can handle a higher number of variables, while others (the majority) cannot.

the cointegration between pairs of states. With this result, we can exclude those states that are not cointegrated with the same order. Afterward, we can test the multivariate model with the states that were cointegrated in a bivariate model, and verify if they present a common trend in the long-run.

At this point, we are going to exclude those states that are completely independent of the others, i.e., those states that are not cointegrated with any other states. On the other hand, those states that present cointegration will be grouped. Again, we are going to test the existence of a common trend on these groups. Each group will be tested for the presence of *k-1* cointegrating vectors. If the group denotes *k-1* cointegrating vectors, it means that the states in that group have a common trend.

If we find out that all the states within the groups are cointegrating, the next step is selecting a series in each group and do the cointegration test in order to check if the groups have a common trend. It is possible, because if the series in a group have a common trend, any time series of this group can represent the group. So, any series can be chosen to check the cointegration. Finally, if the different groups denote a common trend it means that all the states analyzed represent one market. On the other hand, if they are not cointegrated, it signifies that each group is an independent milk market.

5.2. Pattern of integration

The pattern of integration or interdependence refers to the set of relationships among the states that constitute the market as revealed through the VEC and the DAG analysis.

5.2.1. Vector error correction model

As explained previously on the equation (38), the matrix $\prod$ that has information about the long-run relationship between the variables analyzed can be decomposed by $\prod = \alpha\beta'$. The matrix α represents the speed of adjustment to equilibrium and is a matrix of long-run coefficients, while the matrix β is a cointegrating matrix, and provides information about the long-run structure. So, the matrix α contains the information to uncover the spatial structure of the market or short-run relationships (Gonzalez-Rivera & Helfand, 2001). On the other hand, the contemporaneous structure can be summarized through structural analysis of the errors.

The matrix α is the measure of the average speed of convergence towards the long-run equilibrium and plays a crucial role in analyzing how each of the price series will respond to deviations from the long-run equilibrium relationship (Haigh & Bessler, 2004). It provides an indication regarding the short-run adjustment processes of each series. In this sense, smaller values of α means that, in a sudden disequilibrium, the variable considered adjusts slowly to return to the long-run equilibrium. On the other hand, a higher value of α indicates that the adjustment is fast.

As affirmed by Gonzalez-Rivera & Helfand (2001), there are different patterns that could be observed in a VEC model. If all elements of the matrix α are statistically significant, it means that each region reacts to every single disequilibrium of each location. According to Gonzalez-Rivera & Helfand (2001), it is a case of extreme interdependence among regions, where the information contained in prices is generated in every single location. In other words, the price is formed by all regions of the market.

Another extreme situation is the case where an exogenous central location dominates the long-run behavior of the system, i.e., one central location is responsible for the price formation. In this case, all α_{ij}, $j = 1, \ldots, n - 1$ should be statistically zero. This is called the weak exogeneity test, and the null hypothesis is given by $H_0 : \alpha_{ij} = 0,\ j=1, \ldots, n-1$. A failure to reject the null hypothesis implies that there exists an exogenous location that by itself would be the integrating factor of the system (Gonzalez-Rivera & Helfand, 2001). In addition, Haigh et al. (2001) stated that it can be used to test if a market is weakly exogenous or unresponsive with regard to other markets.

However, it is possible to find many different patterns between these two extremes. In order to determine the pattern of integration in a market, Gonzalez-Rivera & Helfand (2001) affirms that it is necessary to begin with a multivariate VEC model, because the test for weak exogeneity can appropriately reduce the system. This test is applied utilizing the likelihood ratio (LR)

$$LR = 2[ln\,\ell_u - ln\,\ell_r] \sim \chi^2(v) \qquad (40)$$

where $ln\,\ell_u$ is a log of the likelihood of all coefficients (unrestricted) in the equation; $ln\,\ell_r$ is a log of the likelihood of a subset of the coefficients (restricted) in the equation; and χ^2

is the chi square distribution where v is the number of restrictions imposed (Seddighi et al., 2000). The pattern of interdependence can also be analyzed through DAG.

5.2.2. Directed acyclic graphs[22]

DAG is a relative new methodology in the economics environment. We chose to work with DAG models, because they can determine in which market innovations in price discovery happen, as well as the directions in which prices shock flows.

The main idea of the DAG is to represent the causal relationships among a set of variables using an arrow graph or picture. In our case, the causation represents the possibility of changing the effect variable *Y* by changing the cause variable *X*. It is based on identifying restrictions in the innovation correlation matrix (Σ) from the VAR (Bessler, 2006).

For definition, "a directed graph *G* is a causal graph for *C* if there is a directed edge or node from *X* to *Y* in *G* if and only if *X* is a direct cause of *Y* relative to *C* "(Spirtes et al., 2006). According to Bessler (2006), using a mathematical definition, graph is an ordered triple representend by<*V, M, E*>, where *V* is a non-empty set of vertices (variables); *M* is a non-empty set of marks (symbols attached to the end of undirected edges); and *E* is a set of ordered pairs.

In this kind of graph, the arrows represent cause and effect flows. So, an arrow placed with its base at *X* and head at *Y* indicates that *X* causes *Y*: $X \rightarrow Y$. However, the causal relation is not symmetric. In other words, if *X* causes *Y*, it does not mean that *Y* causes *X* (Bessler, 2006). In addition, the graphs methodology uses the genealogical terminology in referring to variables, as children, parents, grandparents, ancestors, etc (Bessler, 2006). So for instance, on the graph $A \rightarrow B \rightarrow C$, A is the parent of *B* and grandparent of *C*, while *C* is the child of *B*. A node (variable) is said to be a root if it has no parents, a sink if it has children (Bessler, 2006).

In a graph, each member of *E* is an edge. Vertices connected by an edge are called adjacent. In a set of three variables, for example, one is a collider if arrows converge on it: $A \rightarrow B \leftarrow C$. On the other hand, B is called a common cause or a causal fork if

[22] This section was based on Scheines et al. (1994).

$A \leftarrow B \rightarrow C$, i.e., B causes A and C. Also, the following situation is possible $A \rightarrow B \rightarrow C$ [23] (Bessler, 2006).

5.2.2.1. Tetrad

In this study, we are going to use the TETRAD 4.3.7-3 to work with DAG. The TETRAD has three distinct parts: a picture or graph to represent the specifying hypothetical causal relations among the variables; a specification of the family of probability distributions and kinds of parameters associated with the graphical model; and a specification of the numerical values of those parameters (Bessler, 2006).

This program was developed with support from the National Aeronautics and Space Administration and the Office of Naval Research (Tetrad, 2006). It is supported by two statistical sources: recursive linear structural equation models, and Bayesian networks, which calculates the maximum likelihood estimates of the parameters (Spirtes et al., 2006). The software is based on statistic causal inferences.

In our case, a recursive structural equation model (RSEM) refers to VEC model or corresponding VAR. So, the price in each state of Brazil included in the VAR will be a node in the DAG, and the covariance matrix generate by the VAR will be the input for the DAG. It is assumed that the error terms are independently and identically distributed; the first and second moments of all error terms exist and are finite; and the second moment (variance) of an error term is different from zero (Spirtes et al., 2006). So, as the error terms are assumed to be independent, they are not included in the graph. Instead, the program works with latent variables representing the correlated errors. These latent variables are detected by considering the Markov condition and Faithfulness[24] (Spirtes et al., 2006).

---

[23] In this case, *A* and *C* are d-separated. The exactly definition of d-separation will be exposed ahead, but it is relevant to give an example now. Considering two trains: one starts at *A*, the other at *C*. Both move toward *B*. Unconditionally, they will crash at *B*. However, if we condition on *B*, (if we build a switch station at *B* with side tracks), we open-up the flow from *A* to *C*. So, conditioning on *B* makes *A* and *C* d-connected (directionally connected) (Bessler, 2006). However, the unconditional association (correlation) between *A* and *C* will be non-zero, if they have a common cause *B*. So, if we condition on *B*, the association between *A* and *C* disappears (Pearl, 2000, p.17). The last situation represents a chain where the unconditional correlation between *A* and *C* will be non-zero, but the same association conditional on *B* will be zero (Bessler, 2006).

[24] Faithfulness means that all independence and conditional independence relations between observed variables are a consequence of the Markov condition applied to the true causal structure.

Then, the TETRAD uses the idea of a Markov condition to associate "the causally sufficient causal structures[25] and the set of independence and conditional independence relations in the probability distributions" (Spirtes et al., 2000). Two intuitive consequences of the Markov condition are that an effect is independent of its indirect causes conditional on its direct causes, and variables are independent conditional on their common causes (Spirtes et al., 2000).

"The joint distribution among the non error variables *V* in an RSEM is determined by the triple $<G, D(\varepsilon_t), x_t>$, where *G* is the causal graph over *V*, $D(\varepsilon_t)$ is the joint distribution among the error terms ε_t, and x_t is the linear coefficients that corresponds to each arrow in the path diagram" (Spirtes et al., 2006). The RSEMs take as either input a covariance matrix, or raw data, and assume a multivariate normal distribution (Spirtes et al., 2000). When the covariance matrix is not put in directly, the raw data is converted to a correlation matrix by the software. In our case, the covariance matrix will be the input. With this, a maximum likelihood test is applied. It is a test of independence or conditional independence when zero correlation is analyzed. In order to construct graphs representing causal structures, the program uses judgments about independence constraints in the population (Spirtes et al., 2006).

In a Bayesian network, instead of using error terms, the program expresses the probability distribution of each effect as a function its direct causes (Spirtes et al., 2006). According to Spirtes et al. (2006) the joint distribution over the variables *V* in a discrete Bayesian network can be factored according to the causal structure in the following way

$$P(V) = \prod_{x \in V} P(x \,/\, direct\ causes\ of\ x) \qquad (41)$$

where *P* is probability of a non-empty set of vertices *V* and *x* is the variable. This equation is called d-separation, and allows writing the probability of the variables in terms of the product of the condition independence relations among the variables (Haigh et al., 2001). One variable is said to be d-separated another variable *X* from *Y* in a graph G, $(X \perp Y / Z)_G$, if and only if *Z* blocks the only path between *X* and *Y*. Pearl (2000), through

[25] A variable is assumed to be causally efficient if it includes all the common causes of variables. In other words, causal sufficiency means that there are no omitted variables that cause two or more of the included variables.

d-separation, showed that there is a link between the causal graphs and the underlying probability distribution of the data generating process. For example, if G is a directed acyclic graph with vertex set V, if A and B are in V and if H is also in V, then G linearly implies the correlation between A and B conditional on H is zero if and only if A and B are d-separated given H (Bessler, 2006).

Greedy equivalence search (GES) algorithm is utilized in TETRAD IV. The GES algorithm searches in a stepwise manner using a Bayesian scoring criterion to score all possible causal flows between variables to obtain the best graph (Yu & Bessler, 2006). This algorithm is described in Chickering (2002). The software also uses the PC algorithm that begins with an undirected graph in which all the variables are originally connected. The program proceeds stepwise to remove adjacent edges when partial (conditional) correlations are not statistically significant from zero at an identified significance level and assigns causal flow directions for the remaining edges. In other words, "each edge or node is subjected to tests that the correlation between its endpoints is zero. Edges surviving these unconditional correlation tests are then subjected to conditional correlation tests" (Bessler, 2006). The Fischer's z is used to test the significance of the conditional correlation (i.e., to test the significance from zero) (Bessler, 2006).

Directed graphs allow writing the price vectors in terms of orthogonalized innovations. Nonzero, off-diagonal elements of the residual matrix allow for a shock in one variable to affect other variables in the model contemporaneously, which determines the causal structure behind the correlation in innovations (Swanson & Granger, 1997).

In sum, DAG allows us to see the causal relationship between the prices, evidencing the correlation and interdependence pattern between the regions. Moreover, it provides information about the true contemporaneous ordering of the variables, which is used in the Bernanke decomposition to create the impulse response functions. In this sense, we are going to use the VAR results obtained in JMulti as an input to the DAG analysis with Tetrad. Therefore, the results of DAG will work as information to help the placing of zeros on the VAR innovations, and consequently, to create the impulse response functions[26] at JMulti.

[26] The impulse response analysis and Bernanke decomposition will be discussed in detail ahead.

5.3. Degree of integration

The degree of integration is analyzed via the impulse response functions. However, in order to do the impulse response functions, we need to identify restrictions on relationships among contemporaneous innovations. In other words, we need to orthogonalize the innovation accountings from the VAR.

There are many equivalent representations of the VAR model. However, if we orthogonalize the innovations, the covariance matrix equals the identity matrix. If we do not orthogonalize, in general, the covariance matrix of the errors is not diagonal. So, orthogonalized innovations have two principal advantages over non-orthogonal ones. First, they are uncorrelated across both time and equations, and second, it is preferable and reliable to examine a shock to a single variable in a cointegrated system (RATS Manual, 2004). Therefore, we chose the Bernanke decomposition, or structural VAR to treat the innovation accounting from the VAR.

5.3.1. Bernanke decomposition

According to Awokuse & Duke (2006), once the DAG is obtained, the causal interrelationship can then be used to specify the ordering of the Bernanke factorization of the VAR.

Bernanke decomposition is an orthogonalization based on the assumptions of distinct, mutually orthogonal, behavioral shocks drive the model, and that lagged relationships among the variables are not restricted (Awokuse & Duke, 2006). In opposite of Choleski decomposition, Bernanke relaxes the assumption of a just-identified structure of the innovation accounting, allowing the imposition of over-identifying restrictions on the model (RATS Manual, 2004).

This decomposition is particularly interesting since it relies on prior theory as the source of their identifying restrictions. It means that the cointegration test and the correlation among the variables is used as information for the Bernanke decomposition. However, as several researchers (Awokuse & Duke (2006), Awokuse & Bessler (2003), and Haigh et al. (2001)) affirm, there is no easy way for identifying the VAR innovations.

Therefore, following Spirtes et al. (2000) we can identify the contemporaneous relationships among the variables based on the covariance matrix from residuals from the VAR by using DAG. DAG offers help in providing database evidence of ordering in contemporaneous time t, assuming the information set on variance-covariance matrix is causally sufficient (Yu & Bessler, 2006). It means that DAG will give us the information to place zeros on the VAR innovations. However, Bernanke (1986) states that for a VAR in n variables, if we leave more than n (n − 1)/2 parameters free (to be estimated) the model is not identified. It means that in a (6 x 6) matrix, we are able to identify 21 parameters, i.e., the 6 diagonal elements and 15 nonzero elements.

5.3.2. Impulse response functions

According to Gonzalez-Rivera and Helfand (2001), degree of integration refers to the reaction time for each of the long-run equilibrium relationships to absorb a system-wide shock.

As suggested by Haigh et al. (2001), the matrix Γ_i in the equation (38) also can be used to analyze the short-run relationship between data series. However, the results can be difficult to interpret. Because of this, the same authors recommended the employment of the impulse response functions based on the VAR model in order to determine the short-run interrelationship among the price series.

Impulse response functions provide the effect of a one-time shock in one of the system's series on itself and on other series in the system (Hamilton, 1994). In other words, this concept allows analyzing the impact of shocks and the way in which shocks are transmitted among markets.

The impulse response function is a concept that gives additional information about the dynamic interrelationships among prices. It is employed to investigate the mechanism of shocks, i.e., it allows tracing out the path of the various shocks on the variables contained in the VAR system. For this, one has to convert the VAR model into its moving average (MA) representation (Hamilton, 1994), which, following Enders (1995) is given by

$$P_t = \sigma + \sum_{i=0}^{\infty} A_t^i e_{t-i} \tag{42}$$

where P_t is the price series; $\sigma = E[P_t]$, which is the expected value of P_t; $A = B^{-1}\Gamma_i$, where B^{-1} is the inverse matrix of Band e_{t-1} is an error term. To better understanding, we are going to consider the bivariate case in matrix form, which is

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} a_{10} \\ a_{20} \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} . \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} \quad (43)$$

where y_t and z_t represent a time series; the a's are coefficients; and e is an error term. In terms of a Moving Average (MA), equation (9) becomes

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \bar{y} \\ \bar{z} \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^i . \begin{bmatrix} e_{1t-i} \\ e_{2t-i} \end{bmatrix} \quad (44)$$

In this case, the error terms (e_{1t} and e_{2t}) are composites of two shocks ε_{yt} and ε_{zt}. Since $e_t = B^{-1}\varepsilon_t$, we can write

$$\begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} = \left[1/(1 - b_{12}b_{21})\right] + \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix} . \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix} \quad (45)$$

Then, we can combine equation (44) and (45) to form

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \bar{y} \\ \bar{z} \end{bmatrix} + \left[1/(1 - b_{12}b_{21})\right] . \sum_{i=0}^{\infty} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^i . \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix} . \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix} \quad (46)$$

In order to simplify equation (13), we can consider the matrix ϕ_i composed by

$$\phi_i = \left[A_1^i /(1 - b_{12}b_{21})\right] + \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix} \quad (47)$$

Rewrite the equation (46) as

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \bar{y} \\ \bar{z} \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} \phi_{11}(i) & \phi_{12}(i) \\ \phi_{21}(i) & \phi_{22}(i) \end{bmatrix} . \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix} \quad (48)$$

More efficiently,

$$x_t = \mu + \sum_{i=0}^{\infty} \phi_i \varepsilon_{t-i} \tag{49}$$

The coefficients ϕ_i are used to generate the effects of ε_{zt} and ε_{yt} shocks on the entire time paths of the $\{z_t\}$ and $\{y_t\}$ sequences (Enders, 1995). It means that the elements of ϕ_i are impact multipliers or impulse response functions. The accumulated effects of shocks can be obtained by the summation of the coefficients of the impulse response functions.

Therefore, the impulse response functions consist to calculate multipliers from each simulation’s statistically relevant responses. This kind of multiplier indicates the long-run behavior average percentage change in a responding variable per percentage change in a shock variable (Babula et al., 2003). In this case, a positive result implies that each percentage change in the shock variable directionally coincides with the shock variable changes; while a negative value indicates that a variable response is in the opposite direction of the shock (Babula et al., 2003).

A better way to visualize the impulse response is plotting the results. In this case, some results can happen, as affirmed by Abreu et al. (2003):

- If the paths tend to a constant different from zero, it means that a new equilibrium was established.
- Lower/higher values of the equilibrium indicate that the prices are increasing more/less than in the long-run equilibrium.
- Oscillatory values imply a cyclic effect of shock, what means that a tendency of changing prices is varying around the equilibrium, and, in some instant, it will converge to equilibrium.
- Negative values indicate that other prices react in the opposite direction of the price tested.

In this study, we are going to consider a one-time shock in one variable (price) to see the response in 24 months. Since the data utilized in this work is monthly, we considered 2 years a relatively reasonable period to see the behavior of the prices subjected to shocks.

5.4. Data

The nominal price series will be collected in the Fundação Getúlio Vargas (FGV). We will make use of the monthly milk price for every Brazilian state between July 1994 and September 2005.

According to the theory chapter, the ideal data to analyze the market integration would be data by production zone (group of closest cities that have similar characteristics of milk production) or data by processors. However, these type of data are not available in Brazil.

We are not using the real prices because we believe that if we deflate the prices, we are already creating a tendency. So, since the cointegration test tries to find a common tendency between multiple time series, it does not make sense to input a trend before running the cointegration test. Consider two time series *P1* and *P2*. If we deflate the series, we have $P1,2 = P1,1.\ I$ and $P2,2 = P2,1.I$, where I is the deflator index. Combining them as occurs in a cointegration test, it is easy to find that $P1,2 - P2,2 = (P1,1 - P2,1).\ I$. Therefore, deflating a price series creates a trend between them.

The time period was established based on the following information. Firstly, it includes the pos-liberalization period of the milk market in Brazil (1991). Secondly, it is after the Plano Real. Plano Real was the most important economic plan in Brazil (Chapter 2). It was implemented in July, 1994 and had a big effect on the economy, including agriculture and livestock.

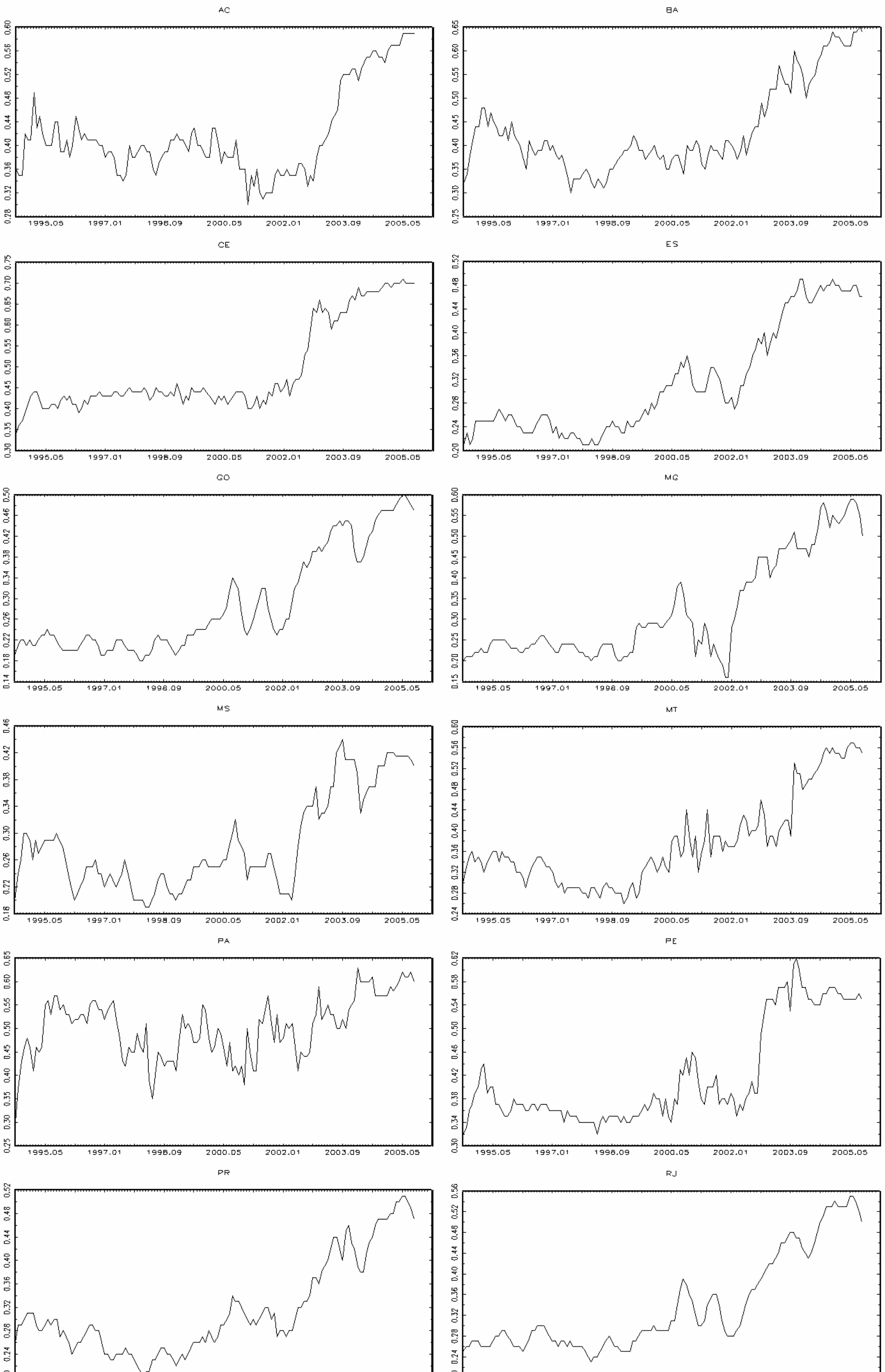
Consumption, population, and production data will be collected at Instituto Brasileiro de Geografia e Estatística (IBGE). Data referring to dairy consumption are available on Pesquisas de Orçamento Familiar (POF) for 1987, 1996, and 2002. However, for 1987 and 1996, there is only data for the main capitals of Brazil (São Paulo, Rio de Janeiro, Belo Horizonte, Belém, Curitiba, Fortaleza, Porto Alegre, Recife, and Salvador). Therefore, we decided to work solely with consumption data in 2002, because it is the only data that approaches exactly what we need. Otherwise, we would have to make many assumptions that could cause bias in the results, since it will be the first step on the analysis.

We are going to use JMULTI 4.14 and TETRAD 4.3.7-3. JMulti will be used to run the ADF, cointegration, and VAR/VEC analysis, including Bernanke decomposition and impulse response functions. The Tetrad will run the DAG.

6. RESULTS

As the literature recommended, the study of the price series started with a graphical analysis. It is a way to see the overall behavior of the series on the period studied. It is useful to identify the presence or absence of important elements, as a trend, structural breaks, seasonal behavior, outlier, etc. Even evidences of stationarity can be noted by a visual analysis of the time series.

All monthly milk prices in Brazil are plotted from 1994.07 to 2005.09, except Tocantins, Maranhão, Amazonas, Amapá, Alagoas, Piauí, and Roraima. They presented significant discontinuities in the time series, which makes them unacceptable for econometric modeling. The abbreviations for the states are Acre (AC), Bahia (BA), Ceará (CE), Espírito Santo (ES), Goiás (GO), Minas Gerais (MG), Mato Grosso (MT), Mato Grosso do Sul (MS), Pará (PA), Paraná (PR), Pernambuco (PE), Rio de Janeiro (RJ), Rio Grande do Norte (RN), Rio Grande do Sul (RS), Rondônia (RO), Paraíba (PB), Santa Catarina (SC), Sergipe (SE), and São Paulo (SP). Figure 4 shows the behavior of the milk prices in 19 states of Brazil.


AC
BA
CE
ES
GO
MG
MS
MT
PA
PE
PR
RJ
1995.05
1997.01
1998.09
2000.05
2002.01
2003.09
2005.05

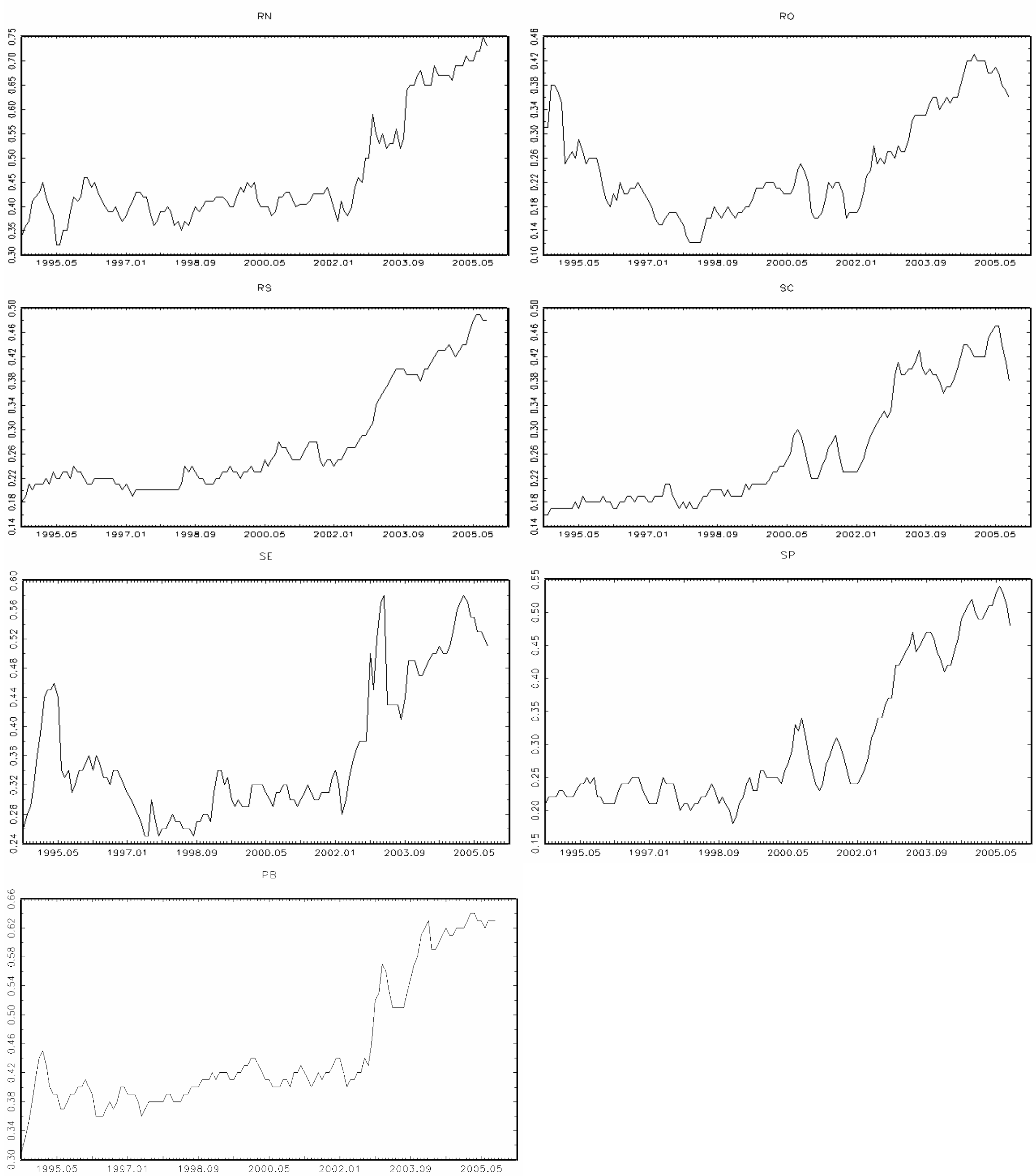


Figure 4 – Nominal price received by milk farmers in 19 Brazilian states, between 1994 and 2005.

Source: Results of the research.

The visual analysis of the series shows the existence of a similar behavior among the series over time. It may be an indication of cointegration among the prices. In addition, the series look nonstationary inasmuch as they do not deviate around an average. All the price series seem to have an increasing trend. This fact was taken into consideration in the unit root tests.

6.1. Extension of the market

As described on Chapter 5, the first step in this methodology is to calculate the ISS. With this index, we can eliminate some states from the analysis. In other words, those states that present ISS close to the unity will be excluded from this analysis because they are close to the self-sufficiency in milk. Table 7 presents the ISS for each state.

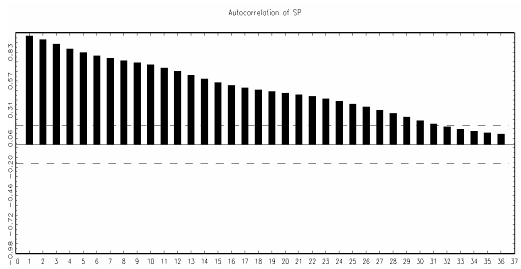
Table 7 – Index of self-sufficiency (ISS) for each Brazilian state in 2002

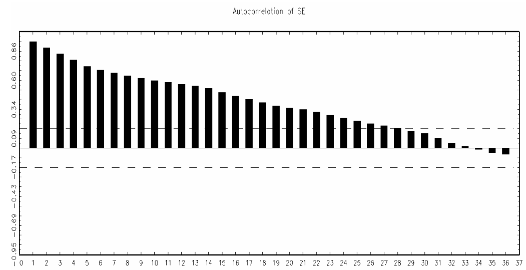
	2002	**E=exporter/I=importer**
Rondônia	7.346016	E
Acre	7.043135	E
Pará	4.502715	E
Ceará	1.243894	E
Rio Grande do Norte	1.562246	E
Paraíba	**0.977158**	I
Pernambuco	1.928356	E
Sergipe	1.695826	E
Bahia	1.975390	E
Minas Gerais	4.939201	E
Espírito Santo	2.626919	E
Rio de Janeiro	0.645445	I
São Paulo	0.739482	I
Paraná	3.398532	E
Santa Catarina	2.676032	E
Rio Grande do Sul	2.934572	E
Mato Grosso do Sul	3.400182	E
Mato Grosso	3.116471	E
Goiás	8.576831	E

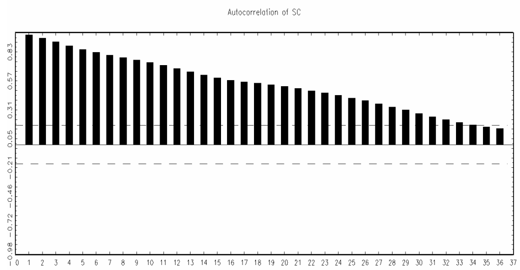
Source: Results of the research.

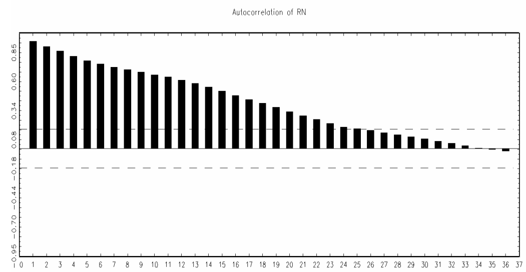
Table 7 shows that only Paraíba is close to self-sufficiency in milk (0.977158). It means that Paraíba does not need to trade (sell or buy) milk with other states. As a result, Paraíba is out of our analysis, because we are interested in finding those states that are integrated by trade and milk price formation.

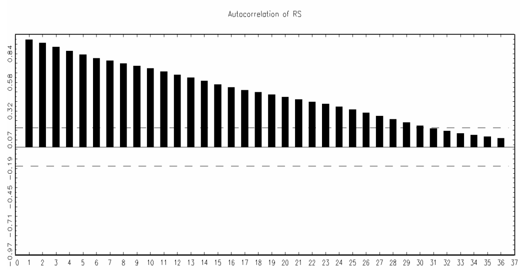
The correlogram of the 18 remaining states is presented in Figure 5.

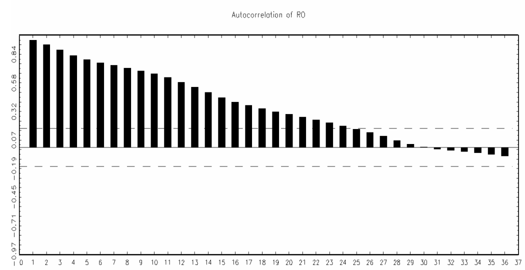

Autocorrelation of SP

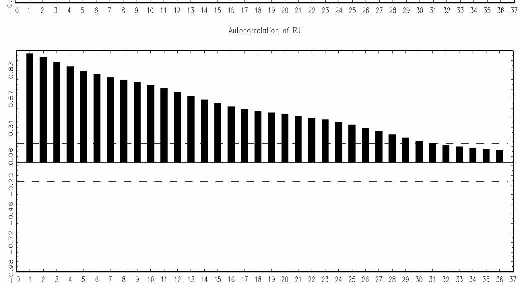

Autocorrelation of SE

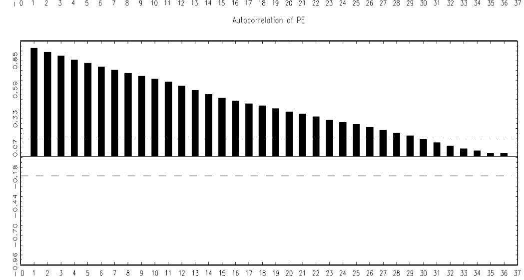

Autocorrelation of SC

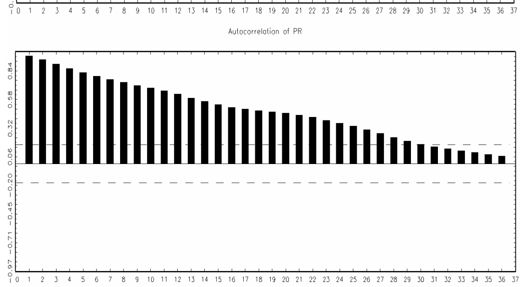

Autocorrelation of RN

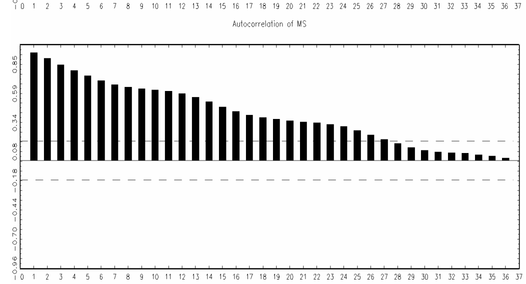

Autocorrelation of RS

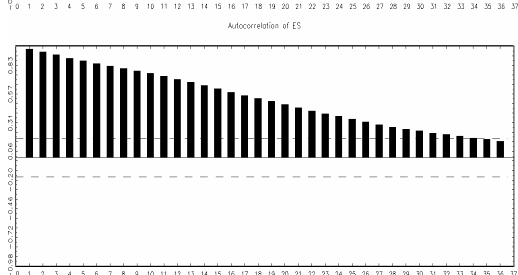

Autocorrelation of RO

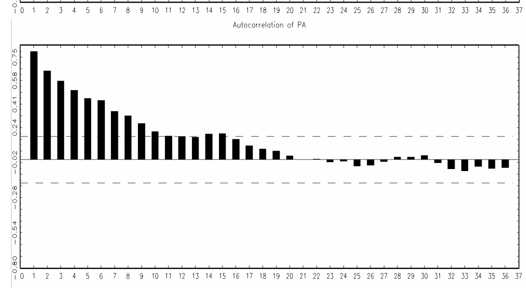

Autocorrelation of RJ


Autocorrelation of PE


Autocorrelation of PR


Autocorrelation of MS


Autocorrelation of ES


Autocorrelation of PA

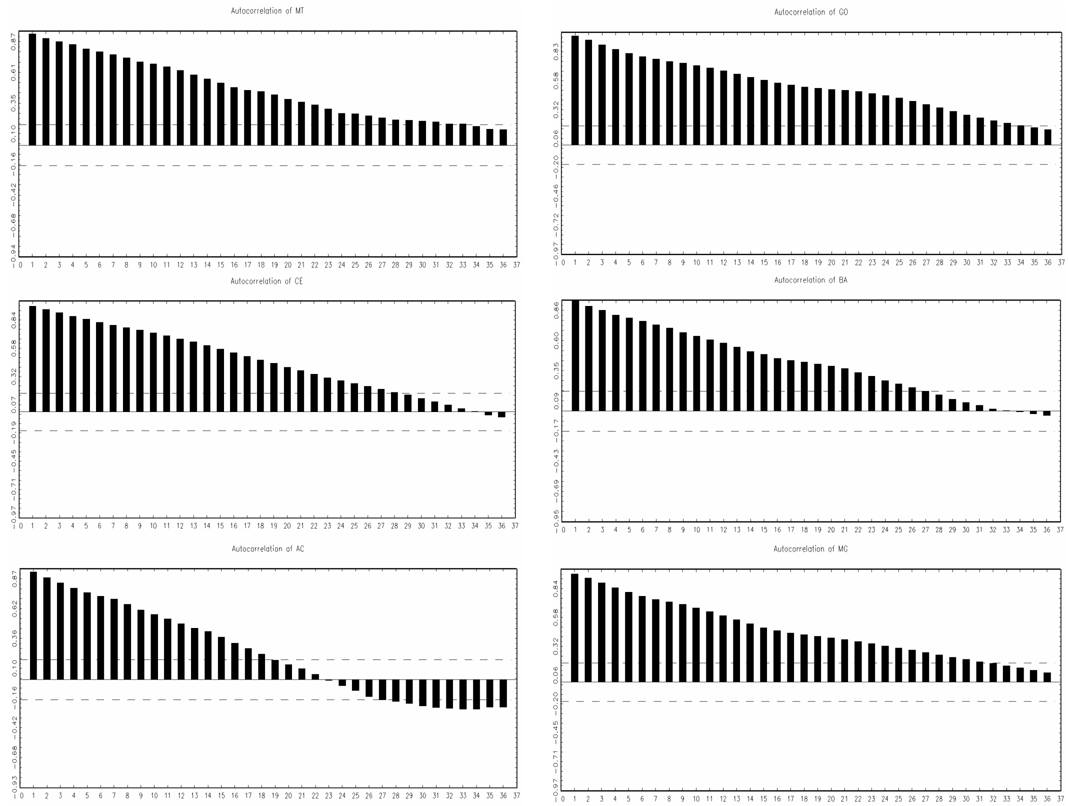


Figure 5 – Correlogram of the milk price in 18 states of Brazil.

Source: Results of the research.

The correlogram of the series indicates clearly that they are nonstationary, since they behave as a random walk series, i.e., they have a slow decay. On the other hand, stationary series behave as a white noise process, i.e., they decrease quickly. Another characteristic observed in Figure 5 is that the series do not appear to have seasonal components, since there are no peaks and dips in the seasonal lags, for example, lags 4, 8, 12, or 6, 12, 18, etc.

However, even though the correlograms of the milk prices in Brazil are very clearly about the nonstationarity of the series, the literature recommends doing the unit root test. The summary of the ADF test is exhibited in Table 8.

Table 8 – Results of an ADF test for the milk price series in selected states of Brazil

States	Lags	Test statistics
Acre	8	-1.3887
Bahia	2	-1.1626
Ceará	1	-1.1994
Espírito Santo	1	-1.9380
Goiás	1	-3.1549*
Mato Grosso	2	-1.8740
Mato Grosso do Sul	1	-2.3159
Minas Gerais	9	-2.4126
Paraná	1	-2.0235
Pará	1	-3.5989**
Pernambuco	1	-1.8906
Rio de Janeiro	4	-1.8478
Rio Grande do Norte	1	-1.4695
Rio Grande do Sul	1	-0.7437
Rondônia	1	-2.2793
Santa Catarina	1	-2.7754
Sergipe	5	-1.8902
São Paulo	1	-2.3566

* Statistically significant at 10% level, and ** at 5%.
Source: Results of the research.

According to Enders (2005), the ADF test has a low power of rejecting the null hypothesis. It means that it is necessary to have a higher level of significance (5% or 1%) to assume that a series is stationary. However, the literature varies in using 5% or 1% as a significance level. Asche et al. (2004), Rashid (2004), and Mattos & Garcia (2004) used 5%. On the other hand, Gonzalez-Rivera & Helfand (2001), Pendell & Schroeder (2004), and Saghaiana et al. (2006) used 1%, which indicates that the choice of the level of significance depends on the features of the study. We opted for using 5%, which means that Pará is considered stationary. Our decision was made based on the plot (Figure 4) and on the correlogram (Figure 5) of the price series in Pará. Since Pará is the only state that is stationary at the level of 5% of significance, we agree that the milk price in this state seems to behave slightly different from the others. Looking at the Figure 4, we can notice that the milk price in Pará crosses the mean of the sample many times, which is an indication of stationarity. In addition, the correlogram of this series is a little different from the other states, since it decreases faster than the others do.

Therefore, considering 5% level of significance, all the series are nonstationary, except Pará. It means that Pará was excluded from the rest of the analysis, because the cointegration test requires that all the series be nonstationary. Consequently, the cointegration in pairs was tested with 17 states.

6.1.1. Cointegration in pairs

As described at Chapter 5, we cannot test the cointegration of 17 states simultaneously. It is a numerical methods problem and the results would not be reliable. In this case, we chose to use the bivariate model, i.e., we tested the cointegration between pairs of states. As the number of cointegration tests between pairs of 17 states is quite large we are going to present a summary of the cointegration results in Table 9.

Table 9 – Summary of the cointegration test between the pairs of 17 states of Brazil*.

	AC	BA	CE	ES	GO	MS	MT	MG	PR	PE	RJ	RN	RO	RS	SC	SE	SP
AC		■										■		■			
BA	■				■	■	■		■		■		■	■	■	■	■
CE					■												
ES					■	■	■	■		■			■		■		
GO		■	■	■		■	■	■	■	■	■	■	■	■	■	■	■
MS		■		■	■		■	■	■	■	■	■	■		■		■
MT		■		■	■	■		■	■		■		■	■			
MG				■	■	■	■		■	■	■			■	■		■
PR		■			■	■	■	■			■		■		■		■
PE				■	■	■		■			■		■		■		■
RJ		■			■	■	■	■	■	■		■			■	■	■
RN	■				■	■					■		■		■		■
RO		■		■	■	■	■		■	■		■				■	
RS	■	■			■		■	■							■		■
SC		■		■	■	■		■	■	■	■	■		■		■	■
SE		■			■						■		■		■		■
SP		■			■	■		■	■	■	■	■		■	■	■	

Source: Results of the research.

* Shaded boxes mean cointegration, while white boxes mean that the states are not cointegrated.

According to Table 9, Goiás is the state that has the higher number of cointegrating relationships (15). It is cointegrated with all the states, except Acre. It is followed by Mato Grosso do Sul and Santa Catarina with 12 cointegrating relationship; São Paulo, Rio de Janeiro, and Bahia, with 11; Minas Gerais with 10; Mato Grosso, Paraná, and Rondônia with 9; Pernambuco with 8; Rio Grande do Sul, Espírito Santo, and Rio Grande do Norte, with 7; Sergipe, with 6; Acre with 3; and in the last position is Ceará, with 1.

So, in relation to the number of cointegrating interactions between the prices, Goiás deserves attention. Goiás is the second leading state in milk production in Brazil, although its production has been developed recently. According to experts, the development of the milk production in this state is caused by the proximity of the production centers of cow feed and other inputs for the dairy farms (Chapter 2). Another important characteristic of this state is its central location in the Brazilian territory, which facilitates the delivering of the production to all the other regions.

Table 9 also shows a certain level of division between the Brazilian states. Except for Bahia, the first states on the list of the number of cointegrating relationship between the prices are all from the regions Center-West, South, and Southeast. However, it is an expected result, since the production, consumption, and processing of milk occurs in the Southeast, Center-West, and South of Brazil. On the other side, the last positions belong to states from North and Northeast, which are not traditional regions in dairy production or processing. It is also important to remember that North and Northeast are responsible for the highest level of informal milk in Brazil.

Another interesting feature of the table above is that there is a significant number of cointegrating relationships among all Brazilian states. It indicates that there is flow of information in the Brazilian dairy market, and there is a flow of milk between the states as well. The number of cointegrating relationship show us that the states in Brazil share a similar movement in the prices, i.e, a considerable number of states follows the information (as production, consumption, price, etc) in other states in order to price the milk. It does not make the price of the milk, or the market to be national, but indicates that there is integration in the market.

6.1.2. Cointegration in groups

After the pairwise test, we are going to make groups of states in order to analyze the market integration. The main idea now is to make groups based on the results of the cointegration in pairs, i.e., those states that were cointegrated in pairs will be put together in order to make groups. Our tests indicated that those states that are cointegrated in pairs are more likely to be cointegrated in groups.

However, it is possible that some states that are not cointegrated with one of the others could be cointegrated with the group. Because of this, we are going to reconsider some non-cointegrated states when we make the groups. For example, in the region South, Rio Grande do Sul and Paraná were not cointegrated in pairs. However, since Rio Grande do Sul is a very important state in milk production, we are going to incorporate it in the group in order to test the cointegration. It is almost a trial and error game, and if we do not find cointegration in the groups, the first state to be taken off will be Rio Grande do Sul. The same will happen in each group. In other words, we are going to test the cointegration in a group, but if we do not find *(k-1)* cointegration relationships, we are going to exclude or replace the state that was not cointegrated with some of the others in the cointegration in pairs.

In contrast, there are states that are not cointegrated with 2 or 3 states in the group. When it happens, we are not going to test this state in group, because our experience indicates that the group will be non-cointegrated. It occurs, for instance, in the region Northeast, where Bahia, Ceará, Pernambuco, Rio Grande do Norte, and Sergipe is in. Nevertheless, only Bahia and Sergipe are cointegrated in pairs. In this case, it is not worth to try the cointegration in group. It indicates that, in this region, the milk market is local, instead of global, which makes sense since they produce and consume a small quantity of milk, i.e., they have few milk trading with others. The same thing is valid for region North (Acre and Rondônia).

At this point, we are going to use different criterion to group the states. The first groups were made following the official Brazilian grouping (regions): North, Northeast, Center-West, South, and Southeast. However, in the region North and Northeast, the states were not cointegrated in pairs. So, we are going to concentrate our attention to regions Southeast, South, and Center-West. Therefore, we can test the cointegration among states within regions South, Southeast, and Center-West. So that the first three groups will be:

1. Paraná, Santa Catarina, and Rio Grande do Sul, representing the region South.
2. Minas Gerais, São Paulo, Rio de Janeiro, and Espírito Santo, corresponding to region Southeast.
3. Goiás, Mato Grosso, and Mato Grosso do Sul, as region Center-West.

We can also make groups associating regions as

4. Center-West and South: Goiás, Mato Grosso or Rio Grande do Sul, Mato Grosso do Sul, Santa Catarina, and Paraná.

5. South and Southeast: Santa Catarina, Paraná, Minas Gerais, São Paulo, and Rio de Janeiro or Rio Grande do Sul or Espírito Santo.

6. Center-West and Southeast: Goiás, Mato Grosso do Sul, Minas Gerais, São Paulo, and Rio de Janeiro or Mato Grosso.

In addition, some more groups were made using data from IBGE (2004):

7. The group of five states leading in milk production and according to Terraviva (2006) the most processing capacity, which are Minas Gerais, Goiás, Paraná, Rio Grande do Sul, and São Paulo.

8. The group of five states leading in dairy consumption: Santa Catarina, Rio Grande do Sul, Minas Gerais, Mato Grosso do Sul, and São Paulo.

9. Two leading consumption with three leading production: Santa Catarina, Rio Grande do Sul, Minas Gerais, Goiás, and Paraná.

Some other combinations will be tried in order to test the market integration in the Brazilian milk market. However, only the positive results of the cointegration test will be presented in this chapter. Nevertheless, the author can show all the results if requested. Table 10 shows the results of the cointegration test for the first six groups.

Table 10 – Results of the cointegration test for group 1 to 5.

Group	rank	LR	Critical values 10%	5%	1%
1. Region South: PR-SC-RS	$r = 0$	62.35	39.73	42.77	48.87
	$r \leq 1$	27.18	23.32	25.73	30.67
	$r \leq 2$	5.86	10.68	12.45	16.22
2. Region Southeast: MG-SP-RJ	$r = 0$	44.49	39.73	42.77	48.87
	$r \leq 1$	24.97	23.32	25.73	30.67
	$r \leq 2$	8.28	10.68	12.45	16.22
3. Region Center-West: GO-MS-MT	$r = 0$	68.89	39.73	42.77	48.87
	$r \leq 1$	24.96	23.32	25.73	30.67
	$r \leq 2$	6.82	10.68	12.45	16.22
4. Center-West and South: GO-MS-PR-SC-RS	$r = 0$	172.52	84.27	88.55	96.97
	$r \leq 1$	84.61	60.00	63.66	70.91
	$r \leq 2$	51.19	39.73	42.77	48.87
	$r \leq 3$	24.72	23.32	25.73	30.67
	$r \leq 4$	6.53	10.68	12.45	16.22
5. South and Southeast: SC-PR-MG-SP-RJ	$r = 0$	183.98	84.27	88.55	96.97
	$r \leq 1$	110.63	60.00	63.66	70.91
	$r \leq 2$	62.99	39.73	42.77	48.87
	$r \leq 3$	25.86	23.32	25.73	30.67
	$r \leq 4$	5.99	10.68	12.45	16.22
6. Center-West and Southeast: GO-MS-MG-SP-RJ	$r = 0$	169.41	84.27	88.55	96.97
	$r \leq 1$	81.54	60.00	63.66	70.91
	$r \leq 2$	50.93	39.73	42.77	48.87
	$r \leq 3$	26.72	23.32	25.73	30.67
	$r \leq 4$	4.59	10.68	12.45	16.22

Source: Results of the research

The interpretation of Table 10 is made following the sequential procedure suggested by Johansen (1992), i.e., we begin testing for zero cointegrating vectors ($r = 0$). If we reject the first test, i.e., if the test statistic (LR) were higher than the critical values, we move on to test $r \leq 1$. We continue until we fail to reject the null hypothesis, i.e., until the value of the test statistic is smaller than the critical values. Following this procedure, Table 10 shows us that all the states in South and Center-West are cointegrated. However, in the Southeast region, only Minas Gerais, São Paulo, and Rio de Janeiro are cointegrated. We tested the cointegration in the Southeast including Espírito Santo, but, in this case, the states presented 2 cointegrating relationships, which indicate that one state should had been excluded.

We also found out that the states in Center-West and Southeast are cointegrated with each other when we included Rio Grande do Sul, instead of Mato Grosso. The same thing happened between Center-West and Southeast. They are cointegrated if we include

Mato Grosso, but not cointegrated if we use Rio de Janeiro. Similarly, between South and Southeast, there is cointegration if we replace Rio Grande do Sul for Rio de Janeiro.

By this analysis, we can conclude that most of the states in South, Southeast, and Center-West of Brazil are cointegrated. It means that their prices move together in the long-run, and they represent a single market. There is a flow of milk and information among these states. Nevertheless, there are some problematic states in this analysis. They are Espírito Santo, Rio de Janeiro, Rio Grande do Sul, and Mato Grosso. Sometimes, these states show cointegrating relationships with the states in other regions, and other times they are the cause of the non-cointegration. Table 11 shows the results of the cointegration test between the dairy leading production and consumption states.

Table 11 – Results of the cointegration test for group 7 to 9.

Group	rank	LR	Critical values 10%	5%	1%
7. Leading production and processing: MG-GO-PR-RS-SP	$r = 0$	137.51	84.27	88.55	96.97
	$r \leq 1$	88.74	60.00	63.66	70.91
	$r \leq 2$	54.38	39.73	42.77	48.87
	$r \leq 3$	23.62	23.32	25.73	30.67
	$r \leq 4$	3.64	10.68	12.45	16.22
8. Leading consumption: SC-RS-MG-SP	$r = 0$	112.14	60.00	63.66	70.91
	$r \leq 1$	52.01	39.73	42.77	48.87
	$r \leq 2$	23.43	23.32	25.73	30.67
	$r \leq 3$	4.47	10.68	12.45	16.22
9. Combination of leading production and consumption: SC-RS-MG-GO-PR	$r = 0$	130.02	84.27	88.55	96.97
	$r \leq 1$	89.68	60.00	63.66	70.91
	$r \leq 2$	51.44	39.73	42.77	48.87
	$r \leq 3$	25.76	23.32	25.73	30.67
	$r \leq 4$	4.76	10.68	12.45	16.22

Source: Results of the research

The cointegration test for the dairy leading production, processing, and consumption groups denotes that they are cointegrated with *k-1* cointegrating relationships. It means that these states present a common trend in the long-run and are in the same milk market.

Therefore, Tables 10 and 11 gives us important information about the extension of the market in Brazil. Although we cannot test the cointegration between more than 5 states, the tests among groups indicate that the milk market in Brazil seems to concentrate in the Center-West, Southeast, and South. The states in these regions are cointegrated with the other states within region, and also are cointegrated with states in the other two regions.

The different groups tested demonstrated that geographical location and similarities in production, processing, and consumption are determining characteristics for the Brazilian milk market. We found out that the states geographically close and/or sharing the same pattern (level) of production, processing, and consumption have a common trend between the milk prices and represent the Brazilian milk market. In other words, the extension of the Brazilian milk market is likely given by Mato Grosso, Goiás, Mato Grosso do Sul, Minas Gerais, Rio de Janeiro, São Paulo, Paraná, Santa Catarina, and Rio Grande do Sul. As explained before, the limitations of the methodology does not allow us to assure it. However, the results of the test strongly indicate it. Future researchers should give particular attention to Rio de Janeiro, Mato Grosso, and Rio Grande do Sul. They demonstrated to be cointegrated in certain cases, but not in others. It is also an interesting result, because these three states represent the extreme boundaries of the Brazilian milk market, which reaffirms that the geographically proximity are a very important feature of a market. Mato Grosso is the northwest extreme; Rio de Janeiro is the east; and Rio Grande do Sul is the south extreme. It likely happens because of the transportation cost. According to the market integration literature, the distance as well as the transportation cost corresponding are the most significant points to determine the extension of an integrated market.

These results make sense according to the literature, and also according to Pereira's (2005) results. She found similar results for live cattle in Brazil. This study showed that eleven Brazilian states (Rondônia, Mato Grosso, Goiás, Mato Grosso do Sul, Bahia, Minas Gerais, Espírito Santo, São Paulo, Paraná, Santa Catarina, and Rio Grande do Sul) were cointegrated and presented a common trend in the long-run. These states also coincide to the cattle leading states in production, processing, and consumption, as occurred in our analysis.

Figure 5 illustrates the Brazilian milk market.

Figure 5 – Brazilian milk market

Source: Developed by the author.

It is important to note that, although the white areas are not include in the Brazilian milk market by our analysis, some of these states may be part of the milk market in Brazil. Especially Bahia, Pernambuco, and Rondônia demonstrated to be very integrated with other states of Brazil in the cointegration in pairs. However, in the group tests, they did not show cointegration. On the other hand, Amazonas, Roraima, Amapá, Tocantins, Maranhão, Piauí, and Alagoas are states that we did not have enough data to use in our research. Because of this, they were excluded from this analysis. Nevertheless, it does not mean that they are not in the national milk market.

6.2. Pattern of integration

Since the pattern of interdependence can be analyzed either for VEC or for DAG, we are going to present both results together in order to facilitate the comprehension. In recent studies, DAG has seemed to be more appropriate to analyze contemporaneous causal relationship, identifying the restrictions for structural VAR models (Awokuse & Duke, 2006). In the DAG analysis, we are going to use PC algorithm, and a p-value of 0.01, as recommended for a sample size between 100 and 300.

As the objective of the VEC and the DAG is to analyze the pattern of interdependence between the states, it does not make sense to use a bivariate model. Therefore, we are going to present the results of the pattern of interdependence using the same groups of states established previously.

Group 1 - Region South: Paraná, Rio Grande do Sul, and Santa Catarina

Table 12 – Matrix of long-run coefficients of the VEC model for the region South.

	Paraná	Rio Grande do Sul	Santa Catarina
Paraná		-0.224*** {0.000}	0.143*** {0.003}
Rio Grande do Sul	0.027 {0.482}		-0.085*** {0.008}
Santa Catarina	0.066 {0.265}	0.019 {0.690}	

*** means statistically significant at 1% level

The p-values are presented in brackets.

Source: Results of the research.

As affirmed by Harris (1995), cited by Barbosa et al. (2002), the matrix of long-run coefficients provides two important information, depending on the significance and the magnitude of the coefficients. A significant coefficient means that the price is not weak exogenous, i.e., a significant coefficient reacts to shocks in the long-run equilibrium relationship. On the other hand, the magnitude of the coefficient informs how fast it will adjust in the long-run.

Looking at the p-values at Table 12, we can notice that there are only three coefficients statistically significant at 1% level. In this case, it means that the milk price in Paraná is dependent on the milk price in Rio Grande do Sul and in Santa Catarina, and the milk price in Rio Grande do Sul is dependent on the price in Santa Catarina. It evidences certain level of leadership of Santa Catarina. It makes sense, since Santa Catarina is the leading consumption in the South of Brazil. However, the small values (<0.5) of the adjustment coefficients indicate that if a price shock occurs in one state, the other states will adjust slowly (Table 12).

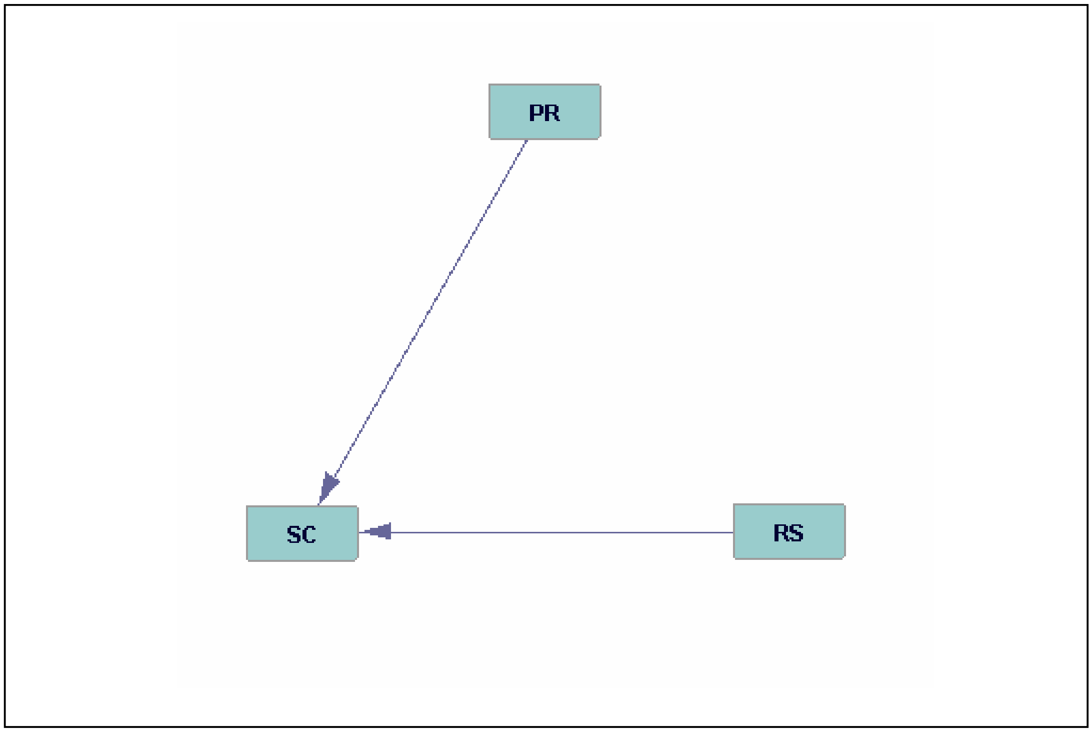


Figure 6 – Patterns from PC Algorithm on innovations from VAR on the South of Brazil.
Source: Results of the research.

Figure 6 shows the results of the DAG analysis for the region South. It demonstrates that the milk price in Rio Grande do Sul and Paraná cause the milk price in Santa Catarina, which is contradictory with the results of the VEC model. In other words, the VEC and the DAG disagree. However, the DAG's results state that the milk prices in Rio Grande do Sul, as well as Paraná, are caused by an exogenous state or power, i.e., the price in Rio Grande do Sul and Paraná are caused by the price in another state out of the region South or by other market elements (i.e., exports, processors, etc). The results of the DAG for Rio Grande do Sul seems reasonable, because this state is the leading milk production state in the region South of Brazil during all the period analyzed. The p-value for this model is 0.49589855. In the DAG analysis, p-values higher than 1% are reasonable, while higher than 5% are great. Therefore, the p-value of 49.58% is quite significant.

It is important to note that VEC and DAG analysis evidenced different leaders, based on different criterion. The leader supported by VEC is the leading consumption state in the South, while the leader supported by DAG is the leading production state in the South. Both results have a good reason to be, and we cannot choose one of them.

Group 2 - Region Southeast: Minas Gerais, São Paulo, and Rio de Janeiro

Table 13 – Matrix of long-run coefficients of the VEC model for the region Southeast.

	Minas Gerais	São Paulo	Rio de Janeiro
Minas Gerais		-0.080 {0.248}	0.375*** {0.008}
São Paulo	0.132*** {0.001}		-0.114 {0.177}
Rio de Janeiro	0.052** {0.034}	0.056 {0.278}	

*** means statistically significant at 1% level, and ** is at 5% level.

The p-values are presented in brackets.

Source: Results of the research.

As happened with region South, half of the coefficients in the Southeast are significant at 5% level, which suggests that the milk price in Minas Gerais is dependent on the price in Rio de Janeiro; and the price in São Paulo and Rio de Janeiro depends on prices in Minas Gerais. It indicates that there is not a leader in milk price formation in the Southeast of Brazil. On the other hand, the prices adjust very slowly, since the coefficients are all smaller than 0.5 (Table 13).

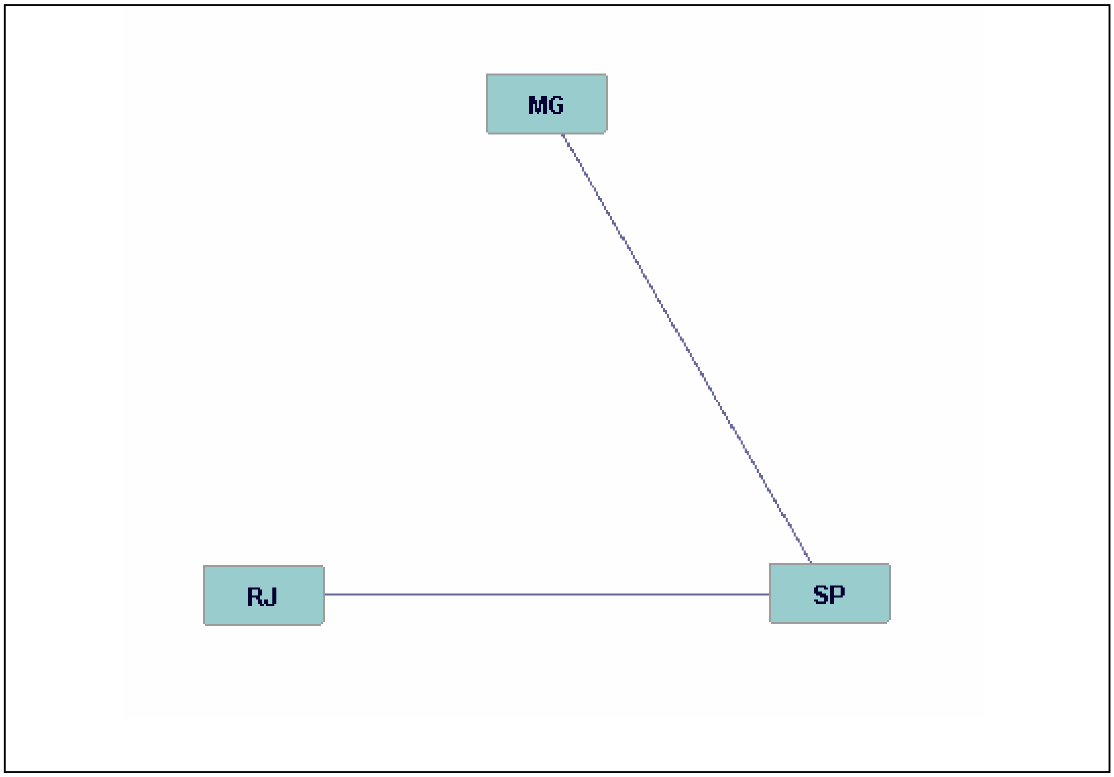


Figure 7 – Patterns from PC Algorithm on innovations from VAR on the Southeast of Brazil.
Source: Results of the research.

Basically, this figure shows that there are certain interdependencies between Minas Gerais, São Paulo, and Rio de Janeiro. However, there is not a cause-effect relationship in this region. It means that the prices in these three states are connected and correlated, but they are not caused by any other state in the Southeast of Brazil. There is not a leader in milk price formation in the region Southeast, which is quite understandable because all the states in this region are economically developed, and have a great level of production and consumption of milk. Nevertheless, we expected that Minas Gerais would be the national leader in milk price formation in Brazil, since this state contains most of the milk production and processing of Brazil. Nonetheless, if it is not the regional leader, it will not be the national leader.

Group 3 - Region Center-West: Goiás, Mato Grosso do Sul, and Mato Grosso

Table 14 – Matrix of long-run coefficients of the VEC model for the region Center-West

	Goiás	Mato Grosso do Sul	Mato Grosso
Goiás		-0.268*** (0.000)	0.183*** (0.004)
Mato Grosso do Sul	0.338*** (0.000)		-0.338*** (0.000)
Mato Grosso	0.248* (0.081)	0.389*** (0.002)	

*** means statistically significant at 1% level, and * means at 10% level.

The p-values are presented in brackets.

Source: Results of the research.

The p-values in Table 14 show that all of the coefficients are statistically significant at 10% level. It is an extreme situation and means that each state reacts to every single disequilibrium of the others. In other words, the milk prices in the Center-West of Brazil are very integrated and interdependent. This table also demonstrates that the speed of adjustment of the prices at the region Center-West is faster than in the other regions, but it is still slow (i.e., they are closer to 0.5, but is still smaller). Especially Mato Grosso do Sul presents higher coefficients, which indicates speed of adjustment faster than the other two states.

The results of the region Center-West are quite different from the South and Southeast's results. Both South and Southeast presented considerable number of non-significant coefficients, while in Center-West the coefficients were all significant. In addition, the coefficients in Center-West were higher than in the other two regions, which indicate that the speed of adjustment in this region is faster. Center-West, as well as South, and Southeast, produces a lot of milk. However, what differentiates this region from the other two is its geographical position. It is located in the center of the country, which allows direct contact with all the other regions. Moreover, Mato Grosso do Sul, the state that presented higher speed of adjustment, is located in a strategic position. It is connected with the two states in the Center-West (Mato Grosso and Goiás), two states in the Southeast (Minas Gerais and São Paulo), and one state in the South (Paraná). Therefore, it is likely that Mato Grosso do Sul are receiving influence of these three regions, which may be the cause of its higher coefficient of adjustment (i.e., the fact that Mato Grosso do Sul are connected with five states and three regions could make it a more dynamic state, and easier to adjust to shocks).

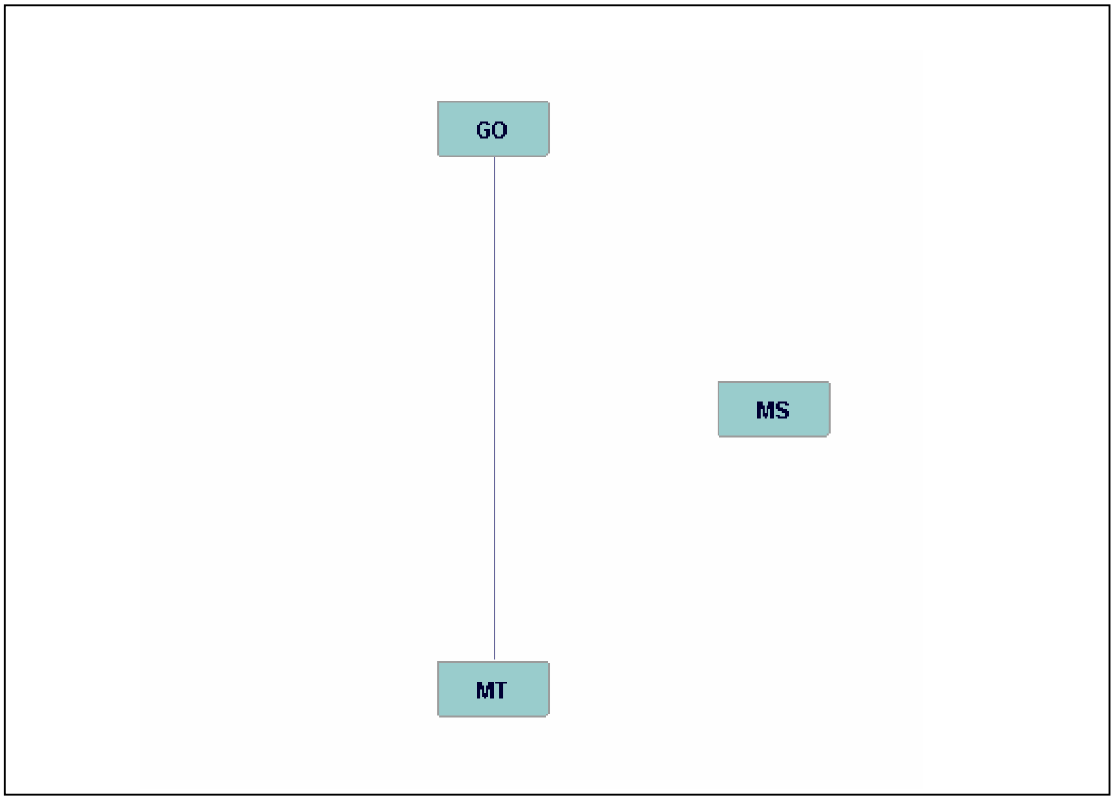


Figure 8 – Patterns from PC Algorithm on innovations from VAR on the Center-West of Brazil.
Source: Results of the research.

This figure confirms part of the results of the VEC model, since the VEC said that these three states of the Center-West are interconnected, and DAG says that Goiás and Mato Grosso are connected. Figure 8 shows that the milk price in these two states are interdependent, but they do not cause each other. However, Mato Grosso do Sul is presented as independent of the other. As explained before, it might be that Mato Grosso do Sul have been influenced by the states out of the region Center-West (for example, Minas Gerais, São Paulo, or Paraná, which share borders with Mato Grosso do Sul).

Since our main objective is to find the leader of milk price formation at the national level, we will still analyze the pattern of interdependence between the states in different regions.

Group 4 - Region Center-West and South: Goiás, Mato Grosso do Sul, Paraná, Santa Catarina, Rio Grande do Sul

Table 15 – Matrix of long-run coefficients of the VEC model for the region Center-West and South

	Goiás	Mato Grosso do Sul	Paraná	Santa Catarina	Rio Grande do Sul
Goiás		-0.540***	0.143***	0.076	0.252***
		{0.000}	{0.009}	{0.242}	{0.001}
Mato Grosso do Sul	0.353***		-0.350***	0.037	-0.062
	{0.003}		{0.000}	{0.637}	{0.509}
Paraná	0.140	0.051		-0.288***	-0.028
	{0.177}	{0.384}		{0.000}	{0.732}
Santa Catarina	-0.183*	0.191***	0.053		-0.115
	{0.084}	{0.001}	{0.455}		{0.172}
Rio Grande do Sul	-0.174**	0.058	0.077	0.077***	
	{0.018}	{0.160}	{0.177}	{0.001}	

*** means statistically significant at 1%, ** is at 5% level, and * is at 10% level

The p-values are presented in brackets.

Source: Results of the research.

Table 15, as well as Tables 12 and 13, has half of the coefficients statistically significant at 10% level. We can notice some tendencies in this table. Rio Grande do Sul seems to have influence on the milk prices in Goiás; Santa Catarina influences prices in Paraná and Rio Grande do Sul; Paraná influences Mato Grosso do Sul and Goiás; Mato Grosso do Sul influences Santa Catarina and Goiás; and Goiás influences Mato Grosso do Sul, Santa Catarina and Rio Grande do Sul. Goiás is the state with more causal relationship (three). However, Rio Grande do Sul and Mato Grosso do Sul seems to have more influence under Goiás, because their coefficients are higher in absolute value. In addition, Table 15 shows that Goiás and Mato Grosso do Sul (in both directions) and Mato Grosso do Sul and Paraná have higher speed of adjustment. Especially Goiás-Mato Grosso do Sul presented a coefficient of adjustment higher than 0.5, which characterizes fast response to economic shocks.

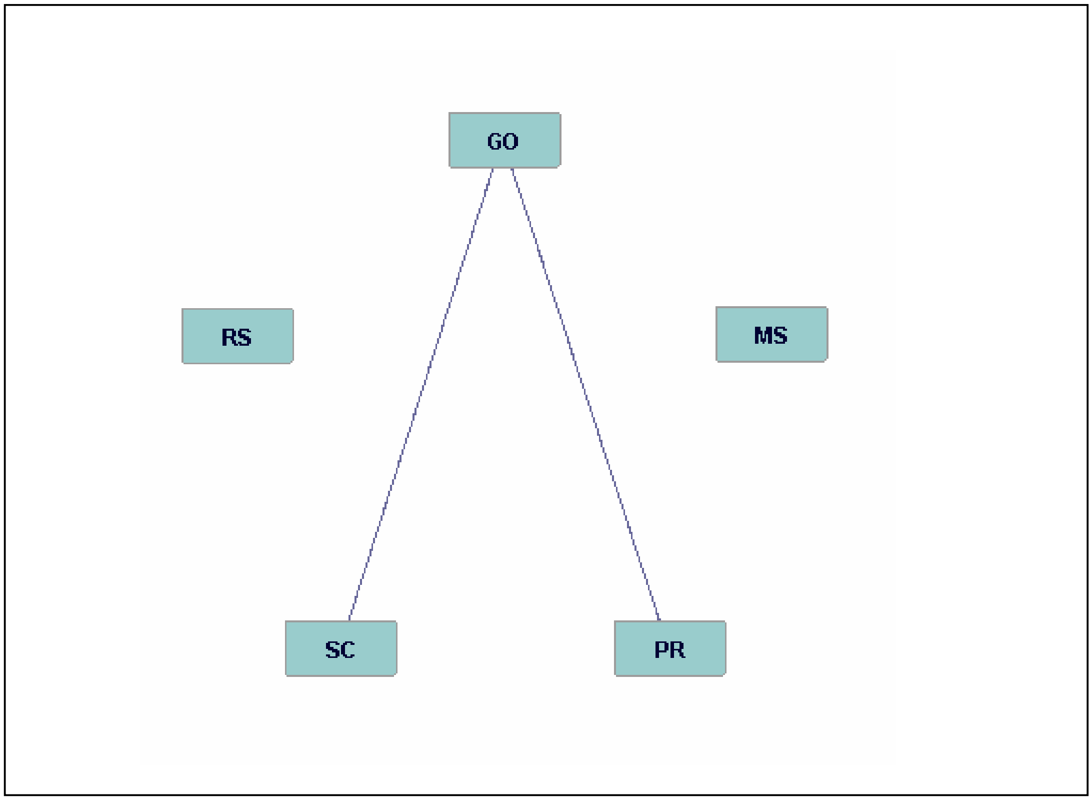


Figure 9 – Patterns from PC Algorithm on innovations from VAR on the Center-West and South of Brazil.
Source: Results of the research.

As indicated by the VEC, Figure 9 shows that there is some level of independence among the states in Center-West and South of Brazil. Goiás, Santa Catarina, and Paraná are interconnected, but the cause-effect relationship between them is not well defined. On the other hand, Mato Grosso do Sul and Rio Grande do Sul are completely independent of the other states in these analysis. It is intriguing, since Rio Grande do Sul was showed as the cause of the price in Santa Catarina (Figure 6), and the cause of the price in Goiás (Table 15). Instead of its location, Mato Grosso do Sul is still not correlated with the others, as defined in Figure 8.

Group 5 - South and Southeast: Santa Catarina, Paraná, Minas Gerais, São Paulo, and Rio de Janeiro

Table 16 – Matrix of long-run coefficients of the VEC model for the region South and Southeast

	Santa Catarina	Paraná	Minas Gerais	São Paulo	Rio de Janeiro
Santa Catarina		-0.375***	0.174***	0.087**	0.174
		{0.000}	{0.006}	{0.011}	{0.101}
Paraná	0.115		-0.179***	-0.009	0.234**
	{0.234}		{0.004}	{0.796}	{0.024}
Minas Gerais	-0.091	0.302**		-0.085	0.256
	{0.629}	{0.013}		{0.193}	{0.205}
São Paulo	0.343***	0.110	0.113***		-0.442***
	{0.001}	{0.102}	{0.002}		{0.000}
Rio de Janeiro	0.115*	0.042	0.051**	-0.001	
	{0.084}	{0.332}	{0.026}	{0.989}	

** means statistically significant at 5% level

The p-values are presented in brackets.

Source: Results of the research.

Table 16 presents similar results to the previous tables. Part of the coefficients is significant at 10% level, which represents dependence between the milk prices in these states. Minas Gerais is the one that affect the price in all the other states, while Santa Catarina is affected by all the others. Also, Table 16 shows that Santa Catarina influences the milk price in São Paulo and Rio de Janeiro, Paraná influences in Santa Catarina and Minas Gerais; São Paulo influences Santa Catarina; and Rio de Janeiro influences Paraná and São Paulo. Moreover, all the coefficients are smaller than 0.5, indicating that the milk prices adjust slowly to return to the long-run equilibrium. There is only three values higher than 0.3 (São Paulo-Santa Catarina, São Paulo-Rio de Janeiro, and Santa Catarina-Paraná).

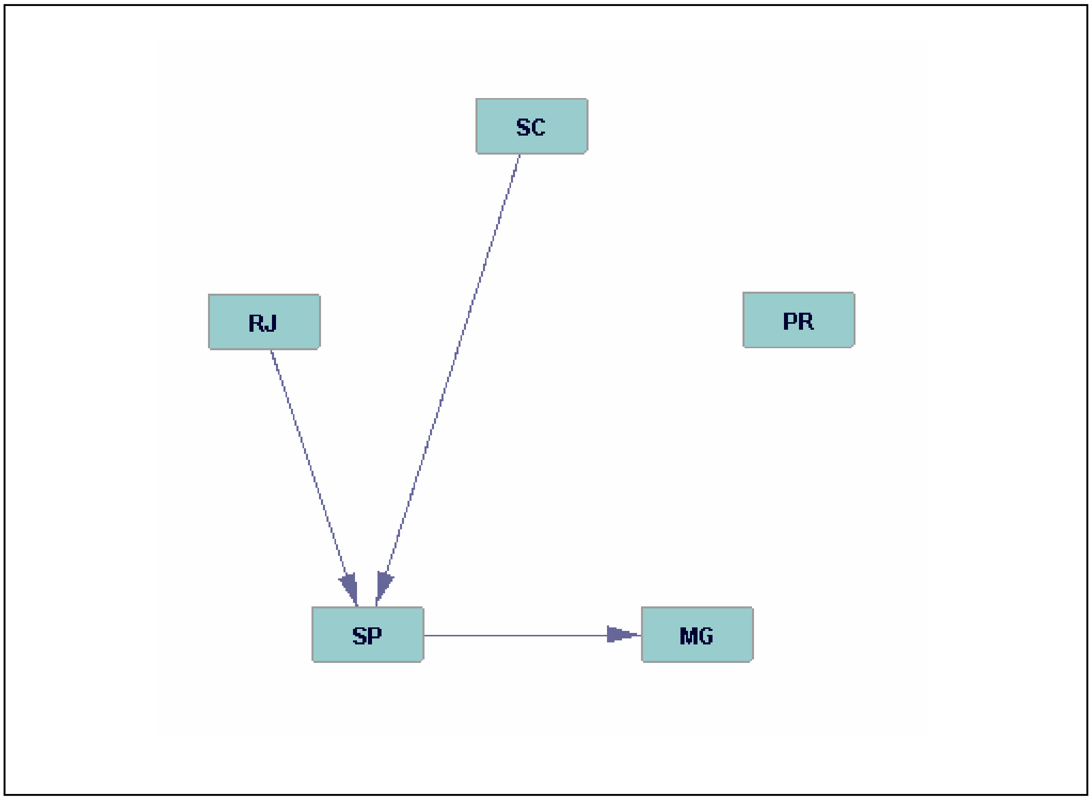


Figure 10 – Patterns from PC Algorithm on innovations from VAR on the South and Southeast of Brazil.
Source: Results of the research.

In this case, the results of the VEC and the DAG analysis are similar, i.e., there is interdependence among the states. Both analysis point out that Santa Catarina and Rio de Janeiro causes São Paulo's milk price. Nevertheless, only DAG shows that São Paulo causes Minas Gerais' prices, and Paraná is independent of the other states. Rio de Janeiro causing São Paulo does not make sense, because the second one is more important in milk production, processing, and consumption than the first one. However, DAG and VEC evidence it. In this case, the DAG analysis should not be taken in consideration, since the p-value associated with it is 0.00003625, which is too small. In the DAG analysis, p-values higher than 1% are good.

Group 6 - Center-West and Southeast: Goiás, Mato Grosso do Sul, Minas Gerais, São Paulo, and Rio de Janeiro.

Table 17 – Matrix of long-run coefficients of the VEC model for the region Center-West and Southeast

	Goiás	Mato Grosso do Sul	Mato Grosso	Minas Gerais	São Paulo
Goiás		-0.418*** {0.000}	0.107** {0.033}	0.019 {0.455}	0.076** {0.014}
Mato Grosso do Sul	0.397*** {0.000}		-0.403*** {0.000}	0.016 {0.618}	0.028 {0.478}
Mato Grosso	-0.130 {0.478}	0.218* {0.058}		-0.215*** {0.000}	0.036 {0.612}
Minas Gerais	-0.306* {0.077}	0.223** {0.041}	0.077 {0.156}		-0.153** {0.023}
São Paulo	0.059 {0.541}	0.051 {0.403}	0.031 {0.309}	0.092** {0.014}	

*** means statistically significant at 1% level, ** is at 5% level, and * is 10% level.

The p-values are presented in brackets.

Source: Results of the research.

Once more, half of the coefficients are significant at the level of 10%. However, we can notice that the milk prices in Goiás and Minas Gerais are influenced by the prices in three states, while in São Paulo, the milk prices are influenced only by Minas Gerais. In addition, the coefficients of São Paulo are very small, which evidences a slow adjustment to shocks. On the other hand, two coefficients of Mato Grosso do Sul, one of Goiás, and one of Minas Gerais are higher than 0.3. It does not represent a fast adjusment to disequilibrium, but it is faster than the other adjustments in this group.

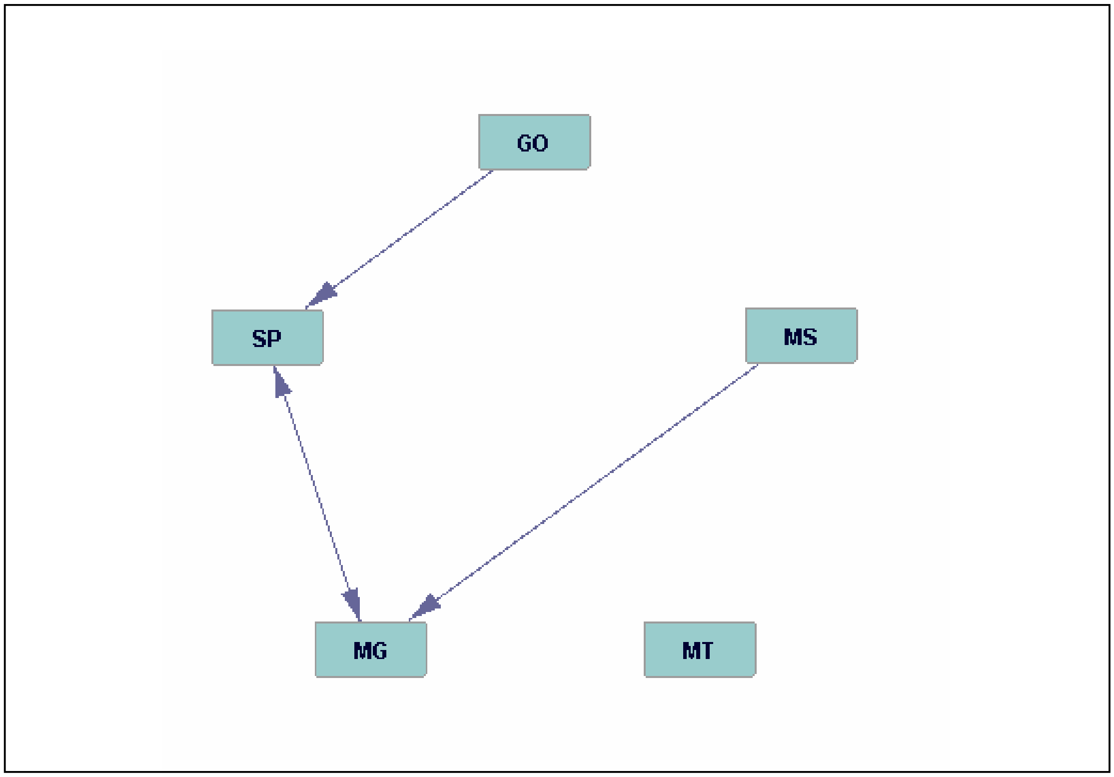


Figure 11 – Patterns from PC Algorithm on innovations from VAR on the Center-West and Southeast of Brazil.
Source: Results of the research.

This graph shows a leadership of Mato Grosso do Sul and Goiás; with Goiás causing São Paulo; São Paulo and Minas Gerais causing each other; Mato Grosso do Sul causing Minas Gerais and Mato Grosso is not caused or cause any other state. In certain ways this results agree with the VEC results, since Mato Grosso causes Minas Gerais, and Minas Gerais and São Paulo cause each other. However, it is intriguing that there is no relationship among the states of Center-West when they are associated with other regions. Although the VEC identifies higher coefficients among these states, which suggests that prices are interdependent, the DAG shows strong independence among them.

Group 7 - Five states leading in milk production and processing

Table 18 – Matrix of long-run coefficients of the VEC model for the five states leading in milk production

	Minas Gerais	Goiás	Paraná	Rio Grande do Sul	São Paulo
Minas Gerais		-0.122*	-0.153	0.089	0.153
		{0.067}	{0.374}	{0.485}	{0.146}
Goiás	0.095***		-0.399***	0.143**	-0.016
	{0.002}		{0.000}	{0.014}	{0.746}
Paraná	-0.022	0.067		-0.260***	0.066
	{0.505}	{0.435}		{0.000}	{0.209}
Rio Grande do Sul	0.042	-0.120*	0.114**		-0.162***
	{0.102}	{0.066}	{0.019}		{0.000}
São Paulo	0.098***	0.100	-0.001	0.120**	
	{0.007}	{0.286}	{0.985}	{0.036}	

*** means statistically significant at 1% level, ** is at 5% level, and * is at 10% level.

The p-values are presented in brackets.

Source: Results of the research.

It is difficult to affirm something about to the interdependence or independence between the milk prices in these states, since part of the coefficients are significant at 10% level and others are not. Table 7 shows that Minas Gerais is weakly dependent on Goiás' prices; Goiás is dependent on Minas Gerais, Paraná, and Rio Grande do Sul; Paraná is on Rio Grande do Sul; Rio Grande do Sul is on Goiás, Paraná, and São Paulo; and São Paulo is on Minas Gerais and Rio Grande do Sul. Except for the coefficient Goiás – Paraná, all the other coefficients are small; hence, the speed of adjustment in these states will be slow.

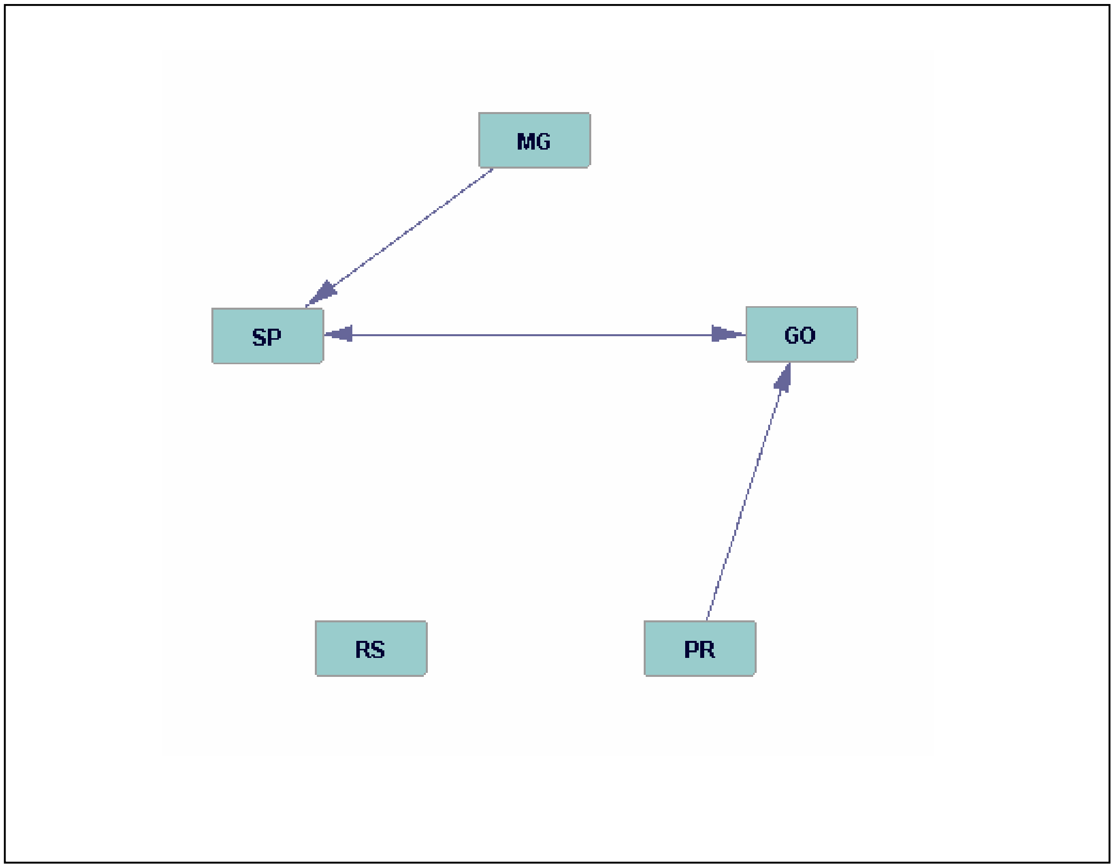


Figure 12 – Patterns from PC Algorithm on innovations from VAR on the five states leading in milk production in Brazil.
Source: Results of the research.

Figure 12 evidences that Minas Gerais causes São Paulo, which causes, and at the same time, is caused by Goiás. In this group, Paraná seems to cause Goiás as well, and Rio Grande do Sul is independent of the others. It is a very interesting scenario, since evolve all the greatest milk production in Brazil. At this scenario, Rio Grande do Sul, as occured at Figure 9, is also independent of the others. Again, as shown in Figure 9, Paraná seems to be correlated with Goiás, with the difference that in Figure 12, Paraná is the cause of Goiás. On the other hand, the relationships between São Paulo and Minas Gerais, and São Paulo and Goiás are different in the other figures. Although Minas Gerais appears as cause of São Paulo's milk price, this association seems to be ambiguous, since Figure 10 shows the opposite, and Figure 11 shows a bidirectional cause. The same is true between São Paulo and Goiás. In this case, the VEC analysis does not help, since neither São Paulo affects Goiás' prices nor Goiás affects São Paulo's prices. Nevertheless, the DAG indicates that the prices in São Paulo are dependent on the prices in Minas Gerais, as well as prices in Goiás are dependent on prices in Paraná. In this sense, we can see a leadership of Paraná and Minas Gerais in this group.

Group 8 - Five states leading in milk consumption: Santa Catarina, Rio Grande do Sul, Minas Gerais, and São Paulo

Table 20 – Matrix of long-run coefficients of the VEC model for the four states leading in milk consumption

	Santa Catarina	Rio Grande do Sul	Minas Gerais	São Paulo
Santa Catarina		-0.314***	-0.052	0.085**
		{0.001}	{0.223}	{0.015}
Rio Grande do Sul	0.082		-0.095***	0.027
	{0.232}		{0.001}	{0.270}
Minas Gerais	0.007	0.200**		-0.132**
	{0.971}	{0.012}		{0.042}
São Paulo	0.358***	0.111**	0.110***	
	{0.000}	{0.011}	{0.002}	

*** means statistically significant at 1% level and ** is at 5% level

The p-values are presented in brackets.

Source: Results of the research.

In this case, most of the coefficients are significant at the level of 5%, which shows that the milk price in the leading states in milk consumption are interdependent, especially in Rio Grande do Sul (as a cause) and in Santa Catarina (as the effect). As occurred before, the coefficients in Table 20 are also small, evidencing slow adjustment to shocks. It means that whether disequilibrium happen, the milk price in these states will take a long time to adjust. It might be an effect of the economic instability of the country, which affect all the agriculture segments.

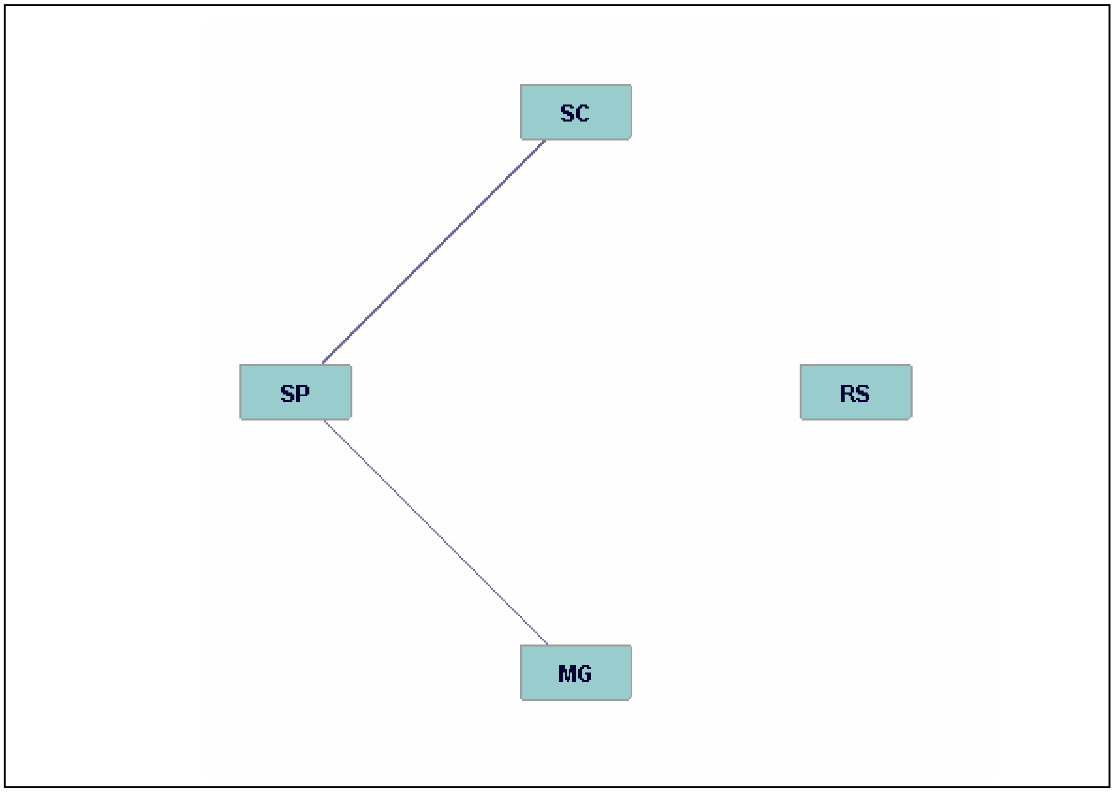


Figure 13 – Patterns from PC Algorithm on innovations from VAR on the five states leading in milk consumption in Brazil.
Source: Results of the research.

Figure 13 shows that Santa Catarina, São Paulo, and Minas Gerais are correlated, but they do not have a cause-effect relationship. On the other hand, Rio Grande do Sul are once again evidenced as independent states in milk price formation in Brazil. The relationship between São Paulo and Santa Catarina, as well as Minas Gerais and São Paulo is also evidenced by VEC. It shows that there is bidirectional influence in both cases. However, as occurred in the other groups, VEC shows Rio Grande do Sul very integrated with the other states and DAG shows Rio Grande do Sul as an independent states in milk price formation.

Group 9 - Combination of the states leading in milk production and consumption: Santa Catarina, Rio Grande do Sul, Minas Gerais, Goiás, and Paraná.

Table 21 – Matrix of long-run coefficients of the VEC model for a combination of states leading in milk production and consumption

	Santa Catarina	Rio Grande do Sul	Minas Gerais	Goiás	Paraná
Santa Catarina		-0.160* {0.057}	-0.075 {0.168}	0.100*** {0.002}	-0.109 {0.256}
Rio Grande do Sul	0.185*** {0.002}		-0.141*** {0.000}	0.036 {0.118}	-0.165** {0.014}
Minas Gerais	0.111 {0.499}	0.192* {0.069}		-0.132** {0.042}	-0.141 {0.454}
Goiás	0.188** {0.014}	-0.019 {0.703}	0.099*** {0.001}		-0.449*** {0.000}
Paraná	-0.031 {0.706}	0.137*** {0.009}	0.005 {0.875}	0.198** {0.034}	

*** means statistically significant at 1% level, ** is at 5% level, and * is at 10% level.
The p-values are presented in brackets.
Source: Results of the research.

Once more, the coefficients of the VEC analysis are small and some of them are significant at the level of 5%, while others are not. Table 21 shows that the milk prices in Santa Catarina are dependent on the milk prices in Rio Grande do Sul and Goiás; prices in Rio Grande do Sul are dependent on Santa Catarina, Minas Gerais, and Paraná; Minas Gerais is dependent on Rio Grande do Sul and Goiás; Goiás is dependent on Santa Catarina and Minas Gerais, and São Paulo; and Paraná's price is influenced by Rio Grande do Sul and Goiás. In addition, the small value of the coefficients evidences slow adjustment to shocks.

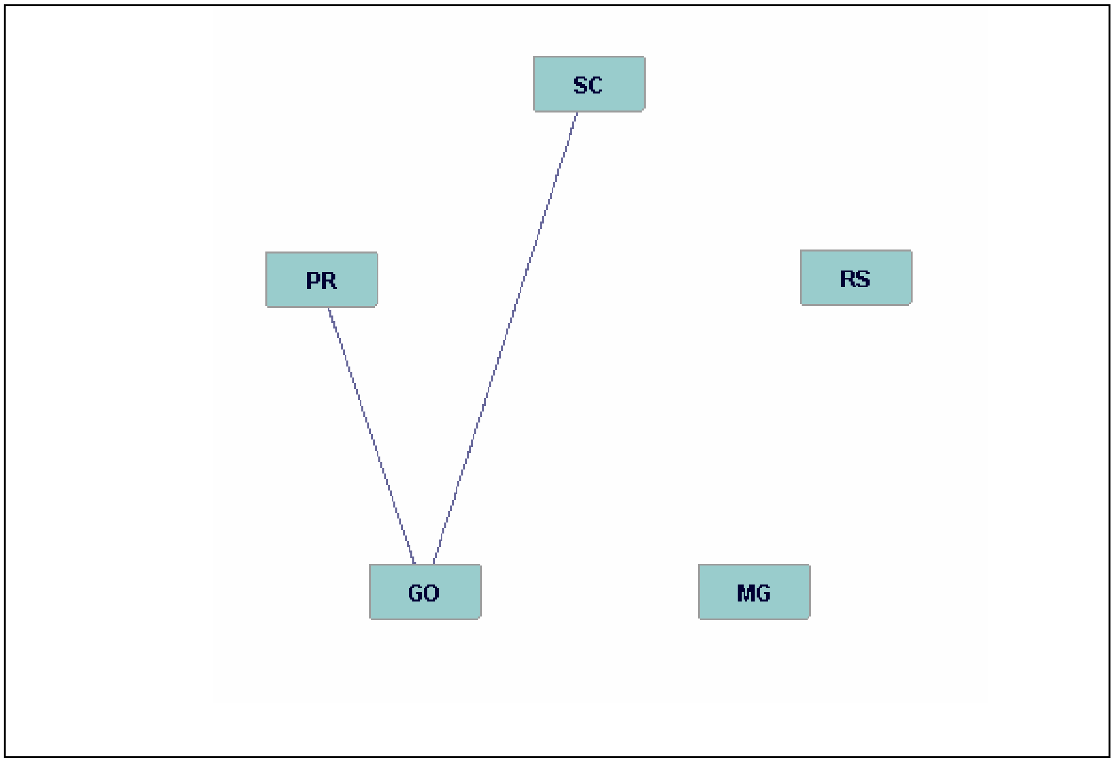


Figure 14 – Patterns from PC Algorithm on innovations from VAR on a combination of states leading in milk production and consumption.
Source: Results of the research.

Figure 14 does not indicate exactly the direction of the causation between Santa Catarina, Paraná, and Goiás. However, it shows that Minas Gerais and Rio Grande do Sul have their milk price formed independently of the other states in this analysis, which is against the VEC's results.

As a conclusion for all this graphs and analysis, we can say that the speed of adjusment to economic shocks is slow for all the states of the Brazilian milk market. This result may be a reflection of the instability of the economy in Brazil. The milk prices, as well as other commodity prices, suffer the consequences of the instability of the economy, and hence, they had difficulty to adjust to shocks.

Moreover, we can notice from the analysis that the definition of the causal relationship between the milk prices in Brazil is not an easy work. Both VEC and DAG analysis presented ambiguous results. It is intriguing, because the DAG uses the VEC/VAR covariances to find the causal relationship among the variables. Recent studies indicate that DAG is a more efficient tool to identify causal relationships, but it showed mixed results for the milk market in Brazil. It signifies that the relationships between the states in the milk market still undetermined. 15 years after the deregulation, this market is still developing and trying to adapt to the new reality. Another reason for this would be that the data used is by state, while the recommended data would be by processor or by production zone. Since the multinational processors exert market power, affecting the milk price, it may influence the performance of the VEC and DAG analysis.

6.3. Degree of integration

6.3.1. Impulse response functions

Since the region South was the only one that presented unidirected acyclic graphs, it is possible to run the impulse response functions for the South of Brazil. It is presented forward.

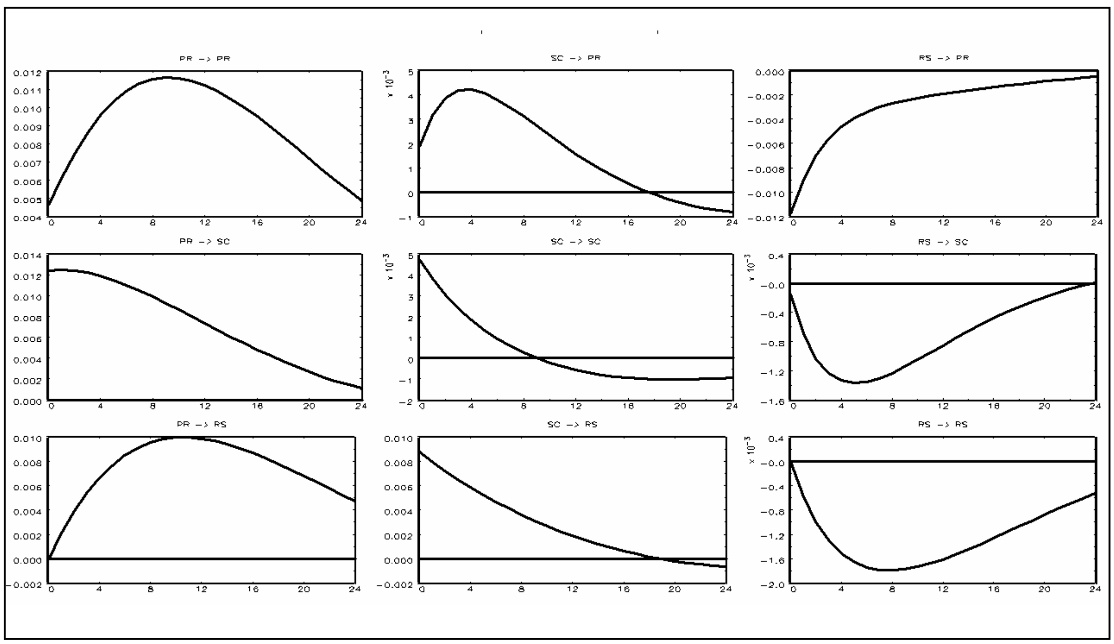


Figure 15 – Set of impulse response functions for the region South.
Search: Results of the research.

The first point we can observe from the Figure 15 is that the effects of the shocks seem to be almost instantaneous, i.e., we can see the response on the prices even in the first month. It indicates that the three states from the region South are very integrated with each other.

A shock in Paraná's price has a positive effect in Rio Grande do Sul and Santa Catarina, since the values on the plots are positive. However, Santa Catarina responds with a continuous falling in the prices, and it seems to return to the equilibrium after two years. On the other hand, Rio Grande do Sul have peak of price in twelve months, dropping it after that. It does not converge to the equilibrium in two years. Moreover, shocks in Paraná's price seems to have higher effects on the neighbors than shocks in Santa Catarina and Rio Grande do Sul, since the values on the graphs of Paraná are higher.

Impacts of the Santa Catarina's shocks start with positive responses in Rio Grande do Sul and Paraná's prices. However, they start to decline in the first month in Paraná, and the fourth month in Rio Grande do Sul. In both states, the prices decline under zero, evidencing a cyclic behavior.

In addition, a positive one-time price shock in Rio Grande do Sul promotes a continuous increase in the prices in Paraná, but a decrease in the prices in Santa Catarina, in the first months, followed by a return to the equilibrium in two years.

The effect of shocks on the same state presents divergent behaviors in the three cases. A shock in Paraná cause an inverse U effect in Paraná, returning to the initial equilibrium in two years. The effect of Santa Catarina in Santa Catarina is a falling in the prices during the following twenty months. However, the prices cross the axle X, evidencing a cyclic behavior. On the other hand, Rio Grande do Sul have a decrease, followed by an increase in the prices, due a shoch in Rio Grande do Sul.

In order to see how significant the responses to the shocks are, we used the accumulated impulse response functions (Figure 16).

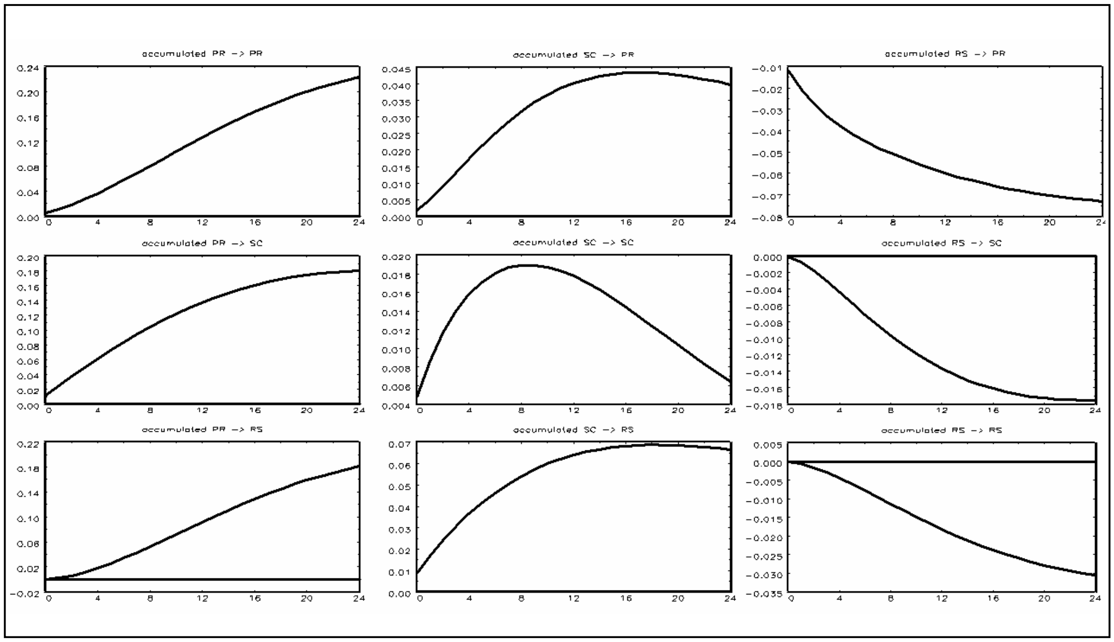


Figure 16 – Set of accumulated impulse response functions for the region South.
Search: Developed by the author.

Figure 16 indicates that the effects of these shocks are very small, since the values at the axle Y at the plot are small. In sum, it means that the states in the South of Brazil are very integrated with each other (since they react immediately to the shocks), but the real effect in the prices is not very significant. From this table, one can notice that the effect of the shock in the first month is really small and increasing for shocks in Paraná and Santa Catarina and decreasing for shocks in Rio Grande do Sul.

However, one should notice that the impulse response functions present only the effects of a shock in one state. It does not mean that a global shock at the economy or in more than one state will have the same impacts.

As explained previously, the small adjustment to shocks and return to equilibrium may be caused by the instability of the Brazilian economy. Another fact that may influence this result is that the milk production, as well as the whole agriculture and livestock production, is a seasonal or annual activity. It means that the producers plan the production 6-12 months before, and once planned (i.e., once buy the input and equipments) it is hard to change the production quickly. The farmers are not able to change their planning quickly, and hence it takes at least one year or two to the activity adjust to disequilibria.

Moreover, it evidences price inefficiency in the Brazilian milk market, which may be reflex of the oligopsony situation. It indicates that processors in each state adjust price slowly. The information is not flowing correctly in this market. One alternative for this would be the futures markets for milk. According to the literature, one of the functions of the futures markets is to allow a better flow of information among markets and, hence, come the system price more efficient.

7. CONCLUSION

Dairy is a highly relevant sector of the Brazilian agribusiness economy. However, this segment has changed significantly after the deregulation (1991). Changes happened in the production regions, consumption, and techniques of production.

Thus, it is worthwhile knowing about the spatial integration of the market and milk price formation at the farm level after deregulation. This problem has not been studied in Brazil. The general objective of this work is to analyze the dynamics of milk price formation in Brazil by identifying where the milk price is formed and what the relationship is between milk prices in different Brazilian states.

The theory adopted is a version of Faminow & Benson (1990), applied for an oligopsony market in Brazil. Instead of the Takayama and Judge, which is the most well known theory about market integration, we believe that the Faminow and Benson approach is more appropriate for this study, since it utilizes more realistic assumptions (intra-market trade, oligopsony, etc). However, we encourage researchers to integrate both approaches, developing a theory that encloses intra and inter-market trade. We do agree that either the Faminow and Benson theory or the Takayama and Judge Theory are interesting. Nevertheless, future researches needs to develop a theory that incorporates intra and inter-market trade.

We noticed that there are a lot of disagreements and misunderstandings in the market integration field. Tons of papers discuss this subject, and few of them agree with each other. There are some interesting methodologies, but the choice for one of them depends on the problem studied. In this research, we choose a modification of the Gonzalez-Rivera & Helfand (2001) methodology, which is divided into extension of the market, pattern of integration, and degree of integration. However, we suggested that future researchers consider the threshold model, since we found out that the transportation cost might be the cause of non-cointegration among states in Brazil.

One modification that we made in Gonzalez-Rivera & Helfand (2001) methodology was to use nominal prices instead of real prices. Due to the fact that we were interested in analyzing the milk price formation in Brazil, we thought that the nominal price would be more approppriate. Previous tests and the literature also indicated that deflating the price creates a trend between milk price series, which is not desirable in a cointegration test.

The extension of the market was determined through the measure of a self-sufficiency index, unit root test, and Johansen procedure. The last one focused on searching for a common trend between the time series. The pattern of integration was studied using the VEC/VAR analysis in association with the DAG. Lastly, the degree of integration was measured by the impulse response functions derived from the Bernanke decomposition.

The Johansen procedure is simply a cointegration test. It is a very useful tool for time series analysis. Nevertheless, it demands extreme attention to detail, such as number of lags, presence of structural breaks, intercept, trend, level of significance, etc. All of these factors can affect the results of the cointegration test. Moreover, as a challenge for future studies, we recommend additional investigation for methods that allow the researcher to work with more than five time series, since it is a shortcoming of this methodology.

Also about the methodology, we alert the researchers that it is very important to find a decomposition of the VAR matrix that is not subjective and limited to do the impulse response decomposition. We made a great improvement in using Bernanke decomposition, instead of Choleski's. However, it still has some problems.

As a result, the cointegration in pairs showed that states in the Center-West, South, and Southeast have more cointegrating relationships than states in the North and Northeast of the country. The cointegration in groups strengthened the results of the cointegration in pairs, since we found out that the states in the South, Southeast, and Center-West are cointegrated within their regions and among regions. In addition, states leading consumption, production, and processing were cointegrated in groups of five. It suggests that the extension of the milk market in Brazil is composed of states from South, Southeast (except Espírito Santo), and Center-West: Rio Grande do Sul, Santa Catarina, Paraná, Minas Gerais, Rio de Janeiro, São Paulo, Mato Grosso, Mato Grosso do Sul, and Goiás. All states seem to follow the same milk price movement in the long-run. On the other hand, states from North and Northeast of Brazil have local milk markets.

Since Rio de Janeiro, Mato Grosso, and Rio Grande do Sul were sometimes cointegrated and other times not, we could conclude that the geographic proximity (as well as the transportation costs) are determining of the extension of the market. In addition, common characteristics among states; as leading consumption, processing, and production; were evidenced as important for the market integration. It is also a useful information, because allow the government to group integrated states according common characteristics to apply policies. In addition, since the policymakers know that some states are not integrated because of the transportation costs, they can act on the transportation costs to improve the integration among the states in Brazil.

As an implication, we can say that this analysis will be very useful especially for policymakers on the milk market in Brazil. They will need different policies for different markets, i.e, one policy for the national market (South, Southeast, and Center-West) and another one for the local markets (North, Northeast, and Espírito Santo).

Although we identified the milk market extension in Brazil, recent important states, for example Rondônia, does not appear in our list. It indicates that the period analyzed was not appropriate to cover and identify all the recent changes in this market. Future works can possibly have better results incorporating 2000's decade. It seems like a decisive period for the consolidation of the new rearragements at the dairy market. We recommend that future work redo this analysis for different periods of time, as an analysis for 1990's and another for 2000's, because the milk market in Brazil has changed significantly in the period considered for us.

Regarding the pattern of integration among the states, we did an analysis within each group, because of the limitations of the methodology. We found that there is a mix of causation, depending on which group of states we considered. Moreover, the VEC and DAG indicated different causation relationship. It indicates that the milk market in Brazil is still developing, and the period analyzed was not able to catch all the rearrangements on this market. We found that the states in the Brazilian milk market are integrated with each other, but there is not a leader in this market. It rejects our initial hypothesis. It means that several states are important in milk price formation in Brazil. Because of this, the policies need to be applied in all states, instead of one leader state. If we had a leader, the government would be able to work with policies in one leader state, and the effect of this policy would transfer for all the other states.

The results of the pattern of interdependence indicated small values of the long-run coefficients at the VEC. It means slow speed of adjustment, or small reaction to shocks in one state. This was strengthened by the impulse response analysis, which indicated that the states in the South of Brazil are integrated with each other, but a shock in one state causes a slow response in the others. In other words, we noticed that the effects of the shocks are quite small. The adjustments to the shocks is slow, i.e., on average, the states spend two years or more to return to the initial equilibrium or a new equilibrium situation. It can be a reflection of the instability of the economy in Brazil. It also means that the information flow is slow among states. It implies that there are inefficient prices in the Brazilian milk market, which may be reflex of the oligopsony situation. It indicates that processors in each state adjust price slowly. The information is not flowing correctly in this market. One alternative for this would be the futures markets for milk. According to the literature, one of the functions of the futures markets is to allow a better flow of information among markets and, hence, come the system price more efficient. In certain ways, the slow adjustment is positive for the milk market, since in an instable economy as in Brazil, the crisis in one state will be transferred slowly for the other states.

It is also a great finding for policymakers, because it allows them to visualize how integrated the milk price is in Brazil, and allows them to predict the reaction of the market or shocks in the milk production or on the prices. Information on speed of adjustment and integration between the states are quite important in determining policies to develop the dairy sector and applying these policies in a more effective way to improve the welfare of the economic agents. Since they know that the milk market in Brazil is is integrated, but price information moves slowly from one state to another, they need to work with stronger policies in order to improve the speed of adjustment to shocks. Knowing the effect in the prices (through the impulse response functions), they can also choose between investment, storage, or production policies, depending on the effect that they want (increase or decrease prices).

In sum, the government can use this information to improve the integration of the milk market, including the North and Northeast in the national market, to work on the stabilization of this market; to develop the milk trade and flow within the country; to predict the price movement after economic shocks, etc. We also believe that this research provides useful information to build the milk price index to the milk futures contracts in Brazil.

8. REFERENCE

Abreu, M.P., M.C. Medeiros, and R.L.F. Werneck. 2003. "Formação de Preços de Commodities: Padrões de Vinculação dos Preços Internos aos Externos." Texto para Discussão no. 474. Departamento de Economia, PUC, Rio de Janeiro.

Ardeni, P.G. 1989. "Does the law of one price really hold for commodity prices." *American Journal of Agricultural Economics*, 71: 661– 669.

Arnade, C. 2006. Shock Absorbing Prices: A Look at Cattle and Feed. *AAEA Meetings*. Long Beach, California.

Asche, F., D.V. Gordon, and R. Hannesson. 2004. "Tests for Market Integration and the Law of One Price: The Market or Whitefish in France." *Marine Resource Economics* 19: 195–210.

Awokuse, T.O., and J.M. Duke. 2006. "The causal structure of Land Price Determinants." *Canadian Journal of Agricultural Economics*. 54: 227-245.

Awokuse, T. O., and D. A. Bessler. 2003. "VAR, Policy Analysis, and Directed Graphs: An Application to the US Economy." *Journal of Applied Economics.* 5: 1–24.

Babula, R. A., D. A. Bessler, and W. S. Payne. 2004. "Dynamic Relationships among US Wheat-Related Markets: Applying Directed Acyclic Graphs to a Time Series Model." *Journal of Agricultural and Applied Economics*. 36: 1–22.

Baek, J., and W.W. Koo. 2005. Price dinamics in the North American Wheat Market. *AAEA Meetings*, Providence, Rhode Island.

Baffes, J. 1991. "Some Further Evidence on the Law of One Price —The Law of One Price Still Holds." *American Journal of Agricultural Economics*. 73, 1264–1273.

Baffes, J., and G. Gohou. 2003. The Co-movement Between Cotton and Polyester Prices. World Bank Policy Research Working Paper 3534. *Presented at the FAO/ESC Commodity Symposium*, Rome.

Barbosa, M.F., M.A. Margarido, and S.N. Junior. 2002. "Análise da Elasticidade de Transmissão de Preços no Mercado Brasileiro de Algodão." *Nova Economia*. 12 (2): 79-108.

Barrett, C.B. 2001. "Measuring integration and efficiency in international agricultural markets." *Review of Agricultural Economics*, 23, 19–32.

Baulch, B. "Transfer Costs, Spatial Arbitrage, and Testing for Food Market Integration." *American Journal Agricultural Economics*. 79 (May 1997):477-487.

Bernanke, B. S. 1986. Alternative explanations of the money-income correlation. *Carnegie-Rochester Conference Series on Public Policy*. 25: 49–99.

Bessler, D.A. "Unpublished Lecture Notes: Frontiers in Markets and Information," Department of Agricultural Economics, Texas A&M University, Fall 2006.

Bessler, D. and D. Akleman (1998). "Farm Prices, Retail Prices, and Directed Graphs: Results for Pork and Beef." *American Journal of Agricultural Economics* 80 (5) 1144-1149.

Cassel, G., 1918. Abnormal Deviations in International Exchanges. *Economic Journal* 28:413–15.

Chickering, D.M. 2002. "Optimal Structure Identification with Greedy Search" *Journal of Machine Learning Research* 3 :507-554.

Coelho, A.B. 2002. "A Cultura do Algodão e a Questão da Integração entre Preços Internos e Externos." São Paulo. MS thesis, Universidade de São Paulo.

Dahlgran, R.A. 2000. An examination of futures price determination though the lens of market integration. *Western Agricultural Economics Association Annual Meeting*, Vancouver – British Columbia.

Davies, N., P. Newbold. Some power studies of a portmanteau test of time series model specification. *Biometrika*. 1979, 66 (1):153-155.

Dickey, D.A, and W.A. Fuller. 1979. Distribution of estimates for autoregressive time series with unit root. *Journal of American Statistics Association*, 74: 427– 431

Empresa Brasileira de Pesquisa Agropecuária - Gado de Leite (Embrapa). *Banco de dados econômicos*. 2006. Available at www.cnpgl.embrapa.br

___. 2005. Available at www.cnpgl.embrapa.br

___ 2003. Available at www.cnpgl.embrapa.br

___ 2002. Available at www.cnpgl.embrapa.br

Enders, W. 2005. *Applied Econometric Time Series*. New York: John Wiley.

___ 1995. *Applied Econometric Time Series*. New York: John Wiley.

Engle, R.F., and C.W.J. Granger. 1991. *Long-Run Economic Relationship: Readings in Cointegration.* Advanced texts in econometrics. New York: Oxford University Press.

Engle, R.F, and C.W.J. Granger. 1987 Cointegration and error correction: Representation, estimation and testing. *Econometrica*, 55: 251–276.

Fabiosa, J.F. 2000. "Impact of GATT in the Functioning of Agricultural Markets: An Examination of Market Integration and Efficiency in the World Beef and Wheat Market Under the Pre-GATT and Pos-GATT Regimes." *AAEA Annual Meeting*, Tampa, Florida.

Fackler, P.L., and B.K. Goodwin. 2000. Spatial price analysis. *Handbook of Agricultural Economics.* pp. 1-59.

Faminow, M.D.; Benson, B.L. 1940. Spatial Market Integration. *American Agricultural Economics*. 72 (1): 49-62.

FAO – Food and Agriculture Organization. 2006. Available at www.fao.org.

Fardmanesh, M., and S. Douglas. 2003. Foreign Exchange Controls, Fiscal and Monetary Policy, and the Black Market Premium. Yale University Economic Growth Center *Discussion Paper No. 876*. Available at ssrn.com/abstract=487485

Fassio, L.H., R.P. Reis, L.C.T. Yamaguchi, and A.J. Reis. 2006. "Custos e Shut-Down Point da Atividade Leiteira em Minas Gerais." Available at www.scielo.br/pdf/resr/v43n4/27754.pdf.

Franken, J.R.V., and J.L. Parcell. 2003. "Market Integration: Case Studies of Structural Change." *NCR-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management.* St. Louis, Missouri, April 21-22,

Goletti, F., and E. Christina-Tsigas. 1995. Analyzing market integration. In: Scott, G.J., ed. *Prices, products and people: analyzing agricultural markets in developing countries*. Boulder: Lynne Riener.

Gomes, S.T. 2006. *Diagnóstico da pecuária leiteira do estado de Minas Gerais em 2005: relatório de pesquisa*. Belo Horizonte: FAEMG.

Gomes, A.L. 2003. "O comportamento do preço do leite recebido pelo produtor: uma abordagem de curto e longo prazo." MS thesis, Universidade Federal de Viçosa.

Gomes, A.T., J.L.B. Leite, A.V. Carneiro, A.L. Gomes. 2002. "Formas e mecanismos de relacionamento entre indústrias e produtores de leite: casos de sucesso." *Congresso Brasileiro de Economia e Sociologia Rural*, Passo Fundo, RS.

Gomes, S.T. 1997. “Cadeia agroindustrial do leite no Mercosul.” In: W.C. Vieira; F.M.A. Carvalho (eds.). *Mercosul: agronegócios e desenvolvimento econômico*. Imprensa Universitária: Universidade Federal de Viçosa, pp. 155-176.

González-Rivera, G., S.M. Helfand. 2001. The extent, the pattern and degree of market integration: a multivariate approach for the Brazilian Rice Market. *American Journal of Agricultural Economics* 83 (3): 576-592.

Gonzalo, J., and C.W.J. Granger. 1995. “Estimation of Common Long-Memory Components in Cointegrated Systems. *J. Buss. Econ. Stat.* 13(January 1995): 27-35.

Goodwin, B.K., H. Djunaidi. 2000. “A necessary condition for market integration test: the case of U.S. broiler industry.” *Southern Agricultural Economics Association. Annual Meeting.* Lexington – KE.

Goodwin, B.K., T.J. Grennes, and M.K. Wohlgenant. 1990. Testing the Law of One Price When Trade Takes Time. *Journal of International Money and Finance* 5:682–93.

Goodwin, B.K., N.E. Piggotti. 1999. “Spatial market integration in the presence of threshold effects.” *AAEA Meetings*, Nashville.

Granger, C.W.J. 1988. Causality, cointegration, and control. *Journal of Economic Dynamics and Control*, North-Holland, 12, p. 551-559.

Greenhut, M.L. 1971. *A Theory f the Firm in Economic Space*. Austin TX: Lone Star Publishers.

Gujarati, D. N. 2000. *Econometria básica*. São Paulo: Makron Books, 846 p.

Haigh, M.S., N.K. Nomikos, and D.A. Bessler. 2001. Integration and Causality in International Freight Markets – Modeling with Error Correction and Directed Acyclic Graphs. *Working Paper 02-11*. Department of Agricultural and Resource Economics.University of Maryland, College Park.

Haigh, M.S., and D.A. Bessler. 2004. "Causality and Price Discovery: An Application of Directed Acyclic Graphs," *Journal of Business*, University of Chicago Press, 77(4): 1099-1098.

Hall, S.G., and Milne, A. 1994. The Relevance of P-Star Analysis to UK Monetary Policy. *The Economic Journal* 104, (424): 597-604.

Hamilton, J. 1994. *Time Series Analysis*, Princeton University Press, Princeton New Jersey.

Hemme, T., O. Garcia, and A. Saha. 2003. “A Review of Milk Production in India with Particular Emphasis on Small-scale Producers.” Working paper. Pro-poor livestock policy initiative: FAO.

Holistic Numerical Methods Institute. 2006. Available at numericalmethods.eng.usf.edu

IBGE. 2006. Available at www.ibge.gov.br

Jarvis, L.S., J.E. Bervejillo, and J.P. Cancino. 2005. International beef prices: is there evidence of convergence? *Review of Agricultural Economics* 27 (3): 449-455.

Johansen, S. "Determination of Cointegration Rank in the Presence of A Linear Trend." *Oxford Bulletin of Economics and Statistics* 54 (1992):383-397.

Kawaguchi, T., N. Suzuki, and H.M. Kaiser. 1997. A spatial equilibrium model for imperfectly competitive milk markets. *American Journal of Agricultural Economics.* 79 (3): 851-859.

Krugman, P.R., M. Obstfeld. 2003. *International Economics: Theory and policy*. 4. ed., Massachusetts: Addison Wesley.

Leite Brasil. 2006. Available at www.leitebrasil.org.br

Li, J.R., and C.B. Barrett. 1999. "Distinguishing between equilibrium and integration in markets analysis." Working paper 99-8. International Agricultural Trade Research Consortium. Available at www.umn.edu/iatrc.

Liang, C., D.M. Feuz, and R.G. Taylor. 1997. "Cointegration tests of spatial and variety price linkages in regional dry bean markets." *Western Agricultural Economics Association Annual Meeting*, Reno – Nevada.

Liu, Q., and H.H. Wang. 2003. "Market Integration Test for Pacific Egg Markets." *American Agricultural Economics Association Annual Meeting*, Montreal, Canada, July 27-30.

MacKinnon, J.G. 1991. Critical Values for Co-Integration Tests. *Long-Run Economic Relationships*. In: R. F. Engle and C.W.J. Granger, eds., pp. 267–76. Oxford: Oxford University Press.

Maddala, G.S., and I. Kim. 1998. *Units roots, cointegration, and structural change*. Cambridge: Cambridge University Press, 505 p.

Marshall, A. 1947. *Principles of Economics*. London: Macmillan.

Martins, P.C. 1999. "O sistema agroindustrial do leite: transações e incentivos à adoção de tecnologias." *Congresso Internacional de Economia e Gestão de Negócios Agroalimentares*, Ribeirão Preto, SP.

Mosconi, R., and C. Giannini. 1992. "Non-causality in Cointegrated Systems: representation, estimation and testing." *Oxford Bulletin of Economics and Statistics*, 54 (3): 399-418.

Massey, P. 2000. Market Definition and Market Power in Competition Analysis: Some Practical Issues. *The Economic and Social Review*, v. 31, (4) 309-328.

Mattos, F., and P. Garcia. 2004. Price Discovery in Thinly Traded Markets: Cash and Futures Relationships in Brazilian Agricultural Futures Markets. *NCR-134 Conference on*

Applied Commodity Price Analysis,Forecasting, and Market Risk Management. St. Louis, Missouri, April 19-20.

McNew, K. 1996. Spatial market integration: definition, theory, and evidence. *Agricultural and Resource Economics Review*.

McNew, K., and P.L. Fackler. 1997. Testing Market Equilibrium: Is Cointegration Informative? *Journal of Agricultural and Resource Economics.* 22(2):191-207

Meirelles, A.J. 2003. "Setor lácteo: história recente e construção de um novo tempo." *Congresso Internacional de Laticínios*, Foz do Iguaçu, RS.

Meirelles, A. J. "Pecuária leiteira: um olhar crítico sobre os anos 80." *Revista Balde Branco* 2, n.4 (1989): 18-26.

Meyer, J. 2003. Measuring market integration in the presence of transaction costs - A threshold vector error correction approach. *25th International Conference of Agricultural Economists*, August 16-22, 2003, Durban, South Africa

Miljkovic, D. 1999. Transporting the export-bound grain by rail: a study of market integration. *Research Discussion Paper n. 29*. Available at www.agecon.lib.umn.edu

Milkpoint. Giro Lácteo. 2005. Available at www.milkpoint.com.br

Mohanty et al., W.H. Meyers, and D.B. Smith. 1998. Fractional cointegration and the fractional rejection of the Law of One Price in international commodity markets. *Journal of Agricultural and Applied Economics*, 30, 2(December 1998):267–276.

Moser, C., C.B. Barrett., and B. Minten. 2005. Missed opportunities and missing markets: Spatio-temporal arbitrage of rice in Madagascar. *American Agricultural Economics Association Annual Meeting*, Providence, Rhode Island, July 24-27.

Neves, M.F., and Consoli, M.A. 2006. *Estratégias para o leite no Brasil*. 1st ed. Ribeirão Preto.

Nogueira, M. P. 2006: "O ano da reviravolta para o leite. Federação da Agricultura do Estado do Paraná." Available at www.faep.com.br/comissoes/leite/130206.asp.

Pearl, J. *Causality*, Cambridge: Cambridge University Press, 2000.

Pede, V.O., and A.M. Mckenzie. 2005. Integration In Benin Maize Market: An Application Of Threshold Cointegration Analysis. American *Agricultural Economics Association Annual Meeting*, Providence, Rhode Island, July 24-27.

Pendell, D.L., and T.C. Schroeder. 2004. Spatial market integration in regional cattle markets. *Western Agricultural Economics Association Annual Meeting*. Honolulu, Hawaii,

Pereira, L.R. 2005. PhD Thesis. Universidade Federal de Viçosa.

PlanetMath, 2006. Available at www.planetmath.com

Rashid, S. 2004. *Spatial integration of maize markets in post-liberalized Uganda*. Markets, trade and Institutions Division, International Food Policy Research Institute. Available at www.agecon.lib.umn.edu.

RATS Manual, 2004. Estima.

Ravallion, M. 1986.. Testing market integration. *American Journal of Agricultural Economics.* n. 68, p. 102-109.

Richardson, J.D. 1978. Some Empirical Evidence on Commodity Arbitrage and The Law of One Price. *Journal of International Economics* 8:341–51.

Saghaian, S.H. L.G. Maynard, M.R. Reed. 2006. The Effects of E. Coli 0157:H7, FMD and BSE on Japanese Retail Beef Prices: A Historical Decomposition. *International Association of Agricultural Economists Conference*, Gold Coast, Australia, August 12-18.

Samuelson, P.A., "Spatial Price Equilibrium and Linear Programming." *The American Economic Review*, 42(1952):283-303.

Sbrissia, G.F. 2005. *Sistema Agroindustrial do Leite: Custos de Transferência e Preços Locais*. Piracicaba.

Scheines, R., P. Spirtes, C. Glymour, and C. Mee. (1994). *TETRADII: Tools for Causal Modeling* (Pittsburgh, PA: Carnegie Mellon University).

Schroeder, T.C. Fed Cattle Spatial Transactions Price Relationships. Journal of Agricultural and Applied Economics, 29 (2):347-362.

Seddighi et al., 2000. Time series Analysis.

Sephton, P.S. 2003. Spatial market arbitrage and threshold cointegration. *American Journal Agricultural Economics*, 85 (4), Nov 2003, 1041-1046.

Sexton, R.J., C.L. Kling, and H.F. Carman. 1991. Market Integration, Efficiency of Arbitrage, and Imperfect Competition: Methodology and Application to U.S. Celery. *American Journal of Agricultural Economics* 73 (3): 568-580.

Silveira, A.M. 2004. *A relação entre os preços de açúcar nos mercados domésticos e internacional*. Piracicaba, 74 p. MS thesis. Escola Superior de Agricultura Luiz de Queiroz, Piracicaba.

Souza, D.P.H. 2000. *Análise da estrutura de custo e preço de sobrevivência dos principais sistemas de produção de leite.* MS thesis. Universidade Federal de Viçosa.

Spirtes, P., C. Glymour, and R. Scheines. (1993). Causation, Prediction and Search, New York: Springer-Verlag. Spirtes, P., C. Glymour, R. Scheines, C. Meek, S. Fienberg, and E. Slate, (1999). "Prediction and Experimental Design with Graphical Model," in Clark Glymour and Gregory F. Cooper editors: *Computation, Causation and Discovery*, Cambridge, Massachusetts: MSIT Press, pp. 65-93

Spirtes, P., R. Scheines, C. Meek, T. Richardson, C. Glymour, H. Hoijtink, and A. Boomsma. 2006. *TETRAD 3: Tools for Causal Modeling*. User's Manual. Available at http://www.phil.cmu.edu/projects/tetrad/tet3/master.htm

Stigler, G.J. 1969. *The Theory of Price*. London: Macmillan.

Stigler, G.J., and R.A. Sherwin. 1985. The Extent of a Market. *Journal of Law and Economics* 28:555–85.

Swanson, N.R. and C.W.J. Granger. "Impulse Response Functions Based on A Causal Approach to Residual Orthogonalization in Vector Autoregressions." *Journal of the American Statistical Association* 92, (1997):357-67.

Tannús, S.P. 2001 "Competitividade e importações: a produção leiteira em Minas Gerais." MS thesis, Universidade Federal de Viçosa.

Takayama, T. and G.G. Judge. *Spatial and Temporal Price and Allocation Models.* North-Holland Publishing Company, 1971.

Terraviva. 2006. "Leite – Observações sobre a concentração da produção e da industrialização no Brasil." Available at www.terraviva.com.br.

Tomek, W.G., and K.L. Robinson. 1990. *Agricultural Product Prices*. 3rd ed. Cornell University. 360 p.

USDA. 2006. Available at www.usda.gov

Wikipedia. 2006. Available at www.wikipedia.com

Yu, T., and D.A. Bessler. 2006. Cointegration and Causality Analysis of World Vegetable Oil and Crude Oil Prices. *American Agricultural Economics Association Annual Meetings*, Long Beach, California July 23-26.