Exploring the Global Soy Complex using a Structural Supply-Demand Model and Price Cointegration Analysis

by

Fabian Steenbergen

1015230

MSc Thesis

Agricultural Economics and Rural Policy Group

Wageningen University & Research

Supervised by

Dr. ir. Koos Gardebroek

June, 2023

Colophon

Title	Exploring the Global Soy Complex using a Structural Supply and Demand Model and Price Cointegration Analysis
Author	Fabian Steenbergen (ID 1015230)
Place & Date	Wageningen, June 2023
Supervisor	Dr. ir. Koos Gardebroek
Second reader	Dr. ir. Jack Peerlings
Chair group	Agricultural Economics and Rural Policy (AEP)
Study program	MSc Management, Economics and Consumer Studies
University	Wageningen University & Research

Wageningen University & Research

Agricultural Economics and Rural Policy Group

Cover photo: Burgers, 2019

Preface

At the time of writing, the agricultural sector in the Netherlands is going through turbulent times. Growing up on a dairy farm in the Netherlands makes me fully aware of the impact that governmental policies and societal concerns have on a farmer’s personal life. My brother, 26 years old and about to take over the family farm, has a cloudy and uncertain career path ahead of him, which is a worry for everyone around him.

At the same time, studying agricultural and environmental economics in Wageningen taught me the negative consequences of our current agricultural system, which ought to be taken very seriously. In a debate where emotions and short-term thinking prevail, I hope to contribute constructively. The search for fundamental solutions to solve the impasse between agriculture and the environment formed the core motivation for writing this thesis and will keep driving me for the rest of my professional career.

I would like to thank my supervisor, Koos Gardebroek, for his insightful comments and constructive feedback that kept me on track and made the creation of this MSc thesis possible. Our fortnightly meetings, which were diversified with a history lesson every once in a while, were always a pleasure. I am happy that I have been able to make use of Koos’ expertise in agricultural economics and econometrics during this part of my studies.

Next, I would like to express my gratitude to all my friends and family who have contributed to the completion of this MSc thesis, just by listing to me chattering about soy. Their sympathetic ears and blank faces have been invaluable throughout the process.

Abstract

This research aims to describe the main characteristics and developments in the global soy complex. After identification of the main trade flows in the global soy market, which pointed at Brazil, the US, and Argentina as main exporters, and China and Europe as main importers, this research continued by estimating a structural supply and demand model, using annual 1995-2020 data to find major drivers of global soy supply and demand. Besides, a price cointegration analysis is performed, modelled in a VECM framework and using monthly 1980-2022 data to establish dependencies between soybean, -meal, and -oil prices. In explaining global soy oil demand, the structural model revealed significant roles for the growth in China's GDP per capita and global biodiesel production. Meat consumption is the main determinant for soybean meal demand, a market also significantly affected by the BSE-crisis, occurring in the late 90s, and its resulting policies in the early 2000s. The price cointegration analysis showed a central role for the soy oil price, which Granger caused the price for soybeans. Rooted in soy oil's connection to the world market via the crude oil price level and the soybean's oil/meal ratio, this finding contributes to the main/by-product debate around soy consumption currently going on in society.

Table of Contents

1. Introduction

1.1. Background

Already for a long time, it is well-known that the deforestation of tropical rainforests has devastating effects on the local climate. Studies show, among other things, an increase in temperature, a decrease in precipitation and evapotranspiration, and increased soil erosion (Shukla et al., 1990; Houghton, 1990). Consequences on a more global scale are the loss of biodiversity, the increasing levels of greenhouse gasses, and a precipitation deficit (Houghton, 1990; Werth & Avissar, 2002). The main causes of deforestation are agricultural practices, unsustainable logging, fires, mining practices, and infrastructural projects (Bennet, 2017). One of these agricultural practices is the production of soybeans, of which, according to data from the FAO (2023), the production increased significantly in the last decade in Brazil. This trend is not only visible in Brazil, as figure 1 (FAO, 2023) shows a significant increase in global production in the last decade. Although the linkage between soybean production and deforestation is still a debated topic in literature (Barona et al., 2010), there is a consensus that there is an (indirect) relationship between the two (Gollnow et al., 2018; Fehlenberg et al., 2017). The most common line of thought is that land is deforested for cattle ranching, after which it is used for more profitable soy production. Therefore, soy expansion pushes cattle production into new areas, indirectly causing deforestation (Fraanje & Garnett, 2020).

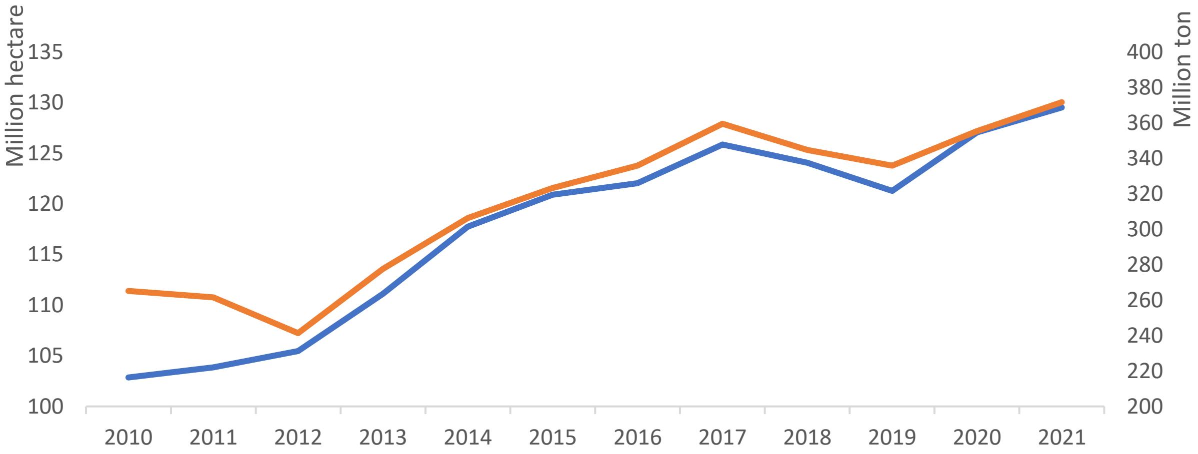

Figure 1: Global production in tonnes (orange line) and hectares (blue line) of soybeans between 2010 and 2021. Based on FAO (2023).

1.2. Problem Statement

Recently, the debate about the sustainability of soy has taken off in Dutch society (de Vries & Smit, 2021; Katan, 2022). Central to this debate is the use of soy meal as animal feed in the agricultural industry in the Netherlands. With 3.8 million cattle, 11.4 million pigs, and 100 million chickens in 2021 (CBS, 2021), the Dutch agricultural sector is an intensive industry. According to Hoste (2014), an average of 13.9 million tonnes of compound feed is produced annually, of which 14% consists of soybean products. For cattle, on average 15.5% of the compound feed is soy, while for pigs it is 8% and for poultry 26%. In absolute numbers, this means that the Dutch animal feed industry uses 2.1 million tonnes of soy, for which 780,000 hectares of land are needed (Hoste, 2014). Often, the soybean as a whole is not mixed into the compound feed. Most of it is soybean meal, which is a protein-rich part of the soybean, containing an excellent set of amino acids and a high level of digestibility, making it a perfect source for the animal feed industry (Willis, 2003). According to Hoste (2014), 78.5% of the soybean consists of soybean meal and it makes up for 79% of the total soy in animal feed in the Netherlands. Hoste (2014) also states that the other main ingredient in animal feed is soy peel, a fibre rich substance that makes up 1% of the bean and 15% of the total soy in animal feed. Next to these two components, soybeans consist for 18.5% of soy oil, which is scarcely used in animal feed (3%). Soy oil is processed in food for human consumption, is used in the cosmetic industry (Hoste, 2014), and as a biofuel (Rodionova et al., 2017).

On one side, the soy debate is fed by NGOs, like WWF in a report about the worldwide effects of the European Union's (EU) consumption of agricultural products (WWF, 2021). They claim the Dutch agricultural sector indirectly causes deforestation in the Amazon rainforest. More specifically, the report claims that the EU-wide soy use in agriculture causes 89,047 hectares of deforestation per year, which makes it top of the list among all agricultural commodities imported by the EU. The Dutch agricultural sector in particular is responsible for 10,000 hectares of deforestation due to its soy use (WWF, 2021). On the opposite side of the debate there is the agricultural sector, which states that the soy used for animal feed is not suited for human consumption. According to Nevedi (2019), an organization in the Netherlands that looks after the interests of the animal feed industry, the soy fed to animals is the by-product that results from producing food from soy for human consumption. This makes it a sustainable practice, as residual products, not suited for human consumption, are turned into high quality products like meat, dairy, and eggs (Nevedi, 2019). However, calling soybean meal a by-product is heavily debated, given soybean meal's large share in the processor's revenues and therefore decision-making. In a study performed by Cheng and Rosentrater (2019), soybean meal made up for 75% of the revenues, compared to 25% from soy oil.

In another argument in favour of the use of soy in animal feed, the industry states that only soy is imported that complies with certain sustainability goals, both environmentally and socially (Nevedi, 2019). This means no illegal deforestation, forced labour, and prohibited pesticides, according to the animal feed advocate. 40% of the imported soy in the Netherlands complies to the EU Soy Sourcing Guidelines of FEFAC and 60% to the guidelines setup by the Round Table on Responsible Soy (RTRS) (Nevedi, 2019). The first-mentioned is debatable, as the FEFAC guidelines only state that producers have to comply to forest and eco-system legislation (FEFAC, 2021). This means it does allow for illegal deforestation, which is, given the political climate in the largest soy exporting country, Brazil, in recent years, an important sidenote.

1.3. Research Objective

In short, it is safe to say that the debate around the use of soy products in animal feed is fully going on, and the (level of) truth in the various claims are far from clear. Therefore, a scientific approach towards these claims is needed to help the debate forward, stimulate the implementation of appropriate policies, and create public awareness about the topic. This research contributes to this by creating a clear overview of which factors and developments are relevant in the global soybean market. The main research question that follows from this objective is the following: What are the main characteristics of the global soy complex and what are its current developments? The *global soy* *complex* is defined as the total production and trade of soybeans and processed soy products (Moschini et al., 2000). The main research question is split into a set of sub research questions:

- What are the main trade flows in the global soybean complex?
- What is the current role of the agricultural industry in the global soybean complex?
- How do biofuel policies and therefore the increasing demand for soy oil affect the soybean complex?
- Is soybean meal or soy oil the main driver for the global demand for soybeans?

These sub research questions aim to give a thorough description of the current status, developments, and future trends in the global soy market. Wherever it is relevant and possible, the role of the Netherlands in particular is highlighted. Together the main and sub research questions will contribute to literature by grasping the complexity of the global soy market, filling up research gaps about several relationships and developments in this market.

1.4. Research Methodology

This research starts with a descriptive section, in which literature research is used to get an in-depth understanding of the global soy market. The main trade flows, the role of the agricultural and energy industry, and global changes in consumption patterns, are qualitatively analysed in this section. After this, a more quantitative approach is used to further explore the soybean complex. This requires some explanation on data collection methods and econometric techniques used. Data is collected from open-source databanks like the FAO (2023), IMF (2022), and World Bank (2023a). Two econometric analyses are done in this research, based on economic theory. The first analysis aims at identifying and quantifying developments and trends in external markets that heavily influence the soy market, using a structural model. Variables representing the main supply and demand factors in the global soy market are theoretically determined. Annual data is used to quantify the effects of the variables included in the model. The econometric framework used to estimate the structural model depends on whether endogeneity is considered a problem, for which the Hausman specification test is used (OLS vs 2SLS). Besides, the model's supply and demand functions may be related via their residuals, in case a system estimator (SUR or 3SLS) is more efficient.

Aim of the second analysis is to zoom in on the soybean market and examine which one of its main subproducts, soybean meal or soy oil, is leading in the demand for soybeans. The model is estimated in a VECM framework and contains monthly time-series data for soybean meal and soy oil prices. The price relationship between soybean meal and soy oil is estimated, using a Granger causality perspective. This type of econometric research is often performed for the price relationship between main- and by-product, for example in the metal industry (Shao et al., 2020). The goal of the analysis is to retrieve information about whether the price relationship is a one-way Granger causal relationship, in which the price of the main-product affects the price of the by-product and not vice versa, or a two-way Granger relationship.

1.5. Structure Overview

This chapter wraps up with a short overview of what can be found in the rest of the thesis. Chapter 2 is a description of the global soy market. More details are provided for the soybean industries in the main producing countries and the largest importers of soybeans. Following the demand side of the global soy market, the focus shifts towards the price determination process and public and private governance initiatives that currently exist in the global soy supply chain (Jia et al., 2020). After this general description of the global soy market, the chapter will zoom in on the role of the Netherlands in this global trade (Hoste, 2014).

Chapter 3 sets up the theoretical framework to analyse the global soy market. In this way, the chapter forms the stepping stone for chapter 4 (Methodology) and 6 (Results). The theoretical framework includes the influence of different industries on the supply and demand of soybeans. As the soybean is a crop that provides food, feed, and fuel, the model includes factors, trends, and policies in all three categories.

The methodology process was shortly touched upon in section 1.4, but requires a more in-depth description. This is done in chapter 4. Chapter 5 summarizes the data used for the analyses. In chapter 6, the results are displayed and interpreted. Based on the results, a concise conclusion is written that can be found in chapter 7. The conclusion will answer the set of sub research questions stated in section 1.3. Together, the sub research questions will lead to an answer to the main research question. The conclusion is supplemented by a subparagraph with a critical reflection, which will discuss the limitations and possible improvements of the research. Besides this, some recommendations for future research are described.

2. The Global Soy Market

Chapter 2 starts with an thorough description of the production (section 2.1) and demand (section 2.2) of soy, creating a clear overview of the trade flows in the global soy market. The full complexity of the soybean complex is captured by describing the price determination process in section 2.3 and the current forms of supply chain governance in section 2.4. The chapter is wrapped up with a section describing the role of the Netherlands in the global soy market.

2.1. Soy Production

According to FAO (2023), global soybean production increased from almost 30 million tonnes in 1961 to more than 370 million tonnes in 2021. This makes it the sixth most gown crop in the world in terms of production volume, and fourth by economic value (Fraanje & Garnett, 2020). Soybean production is highly geographically determined, which means that despite the enormous increase, only a small group of countries is responsible for meeting global demand (Montanía et al., 2021). This is confirmed by figure 2, which shows that the production of soybeans in 2021 mainly took place in three countries, Brazil, the United States (US), and Argentina (FAO, 2023). This is also reflected in export quantities. In 2021, the largest exporters of soybeans were Brazil, the US, and Argentina, together with Paraguay and Canada (figure 2). For soybean oil and soybean meal Argentina is the biggest exporter, followed by Brazil and the US (figure 3).

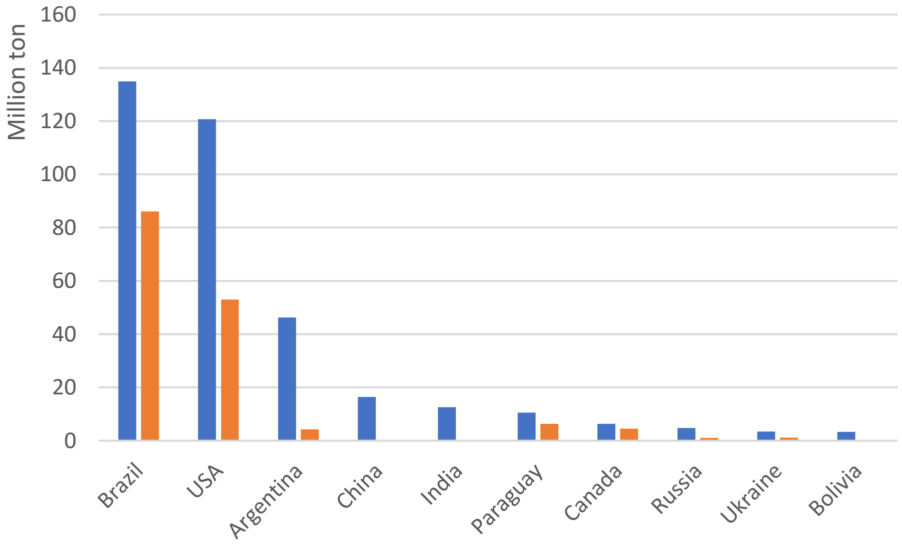

Figure 2: Production (blue) and export (orange) of soy beans for the 10 largest producers in tonnes in 2021. Based on: FAO (2023).

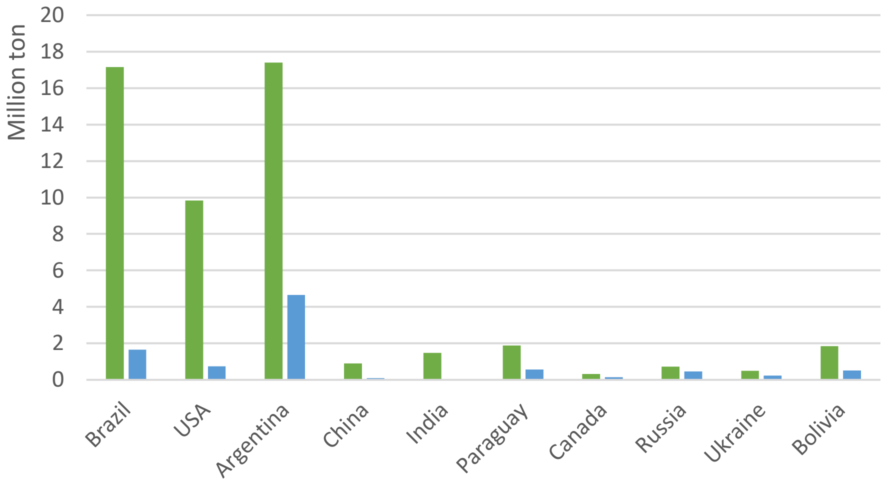

Figure 3: Export quantities for soybean meal (green) and soy oil (blue) in tonnes in 2021.
Based on: FAO (2023).

The growth of soybean production in the last 60 years was not equally spread over its main producing countries. In the 1950s and 1960s, the US was responsible for the majority of the world's soybean production. According to Hart (2017), significant production already took place in the 100 years prior, but remained limited due to the fact that it was only used as a forage crop for grazing animals. After the Second World War, soybean cultivation in the US took a flight. The US started to grow its own edible fats and oils as a result of trade disruptions during the war (Hart, 2017). Together with the need for a protein source for the fast growing poultry industry (Goldsmith & Hirsch, 2006) and soybean's perfect characteristics to function as a rotation crop for corn (Hart, 2017), this is seen as the main reason for the large growth in soybean production in the 1950s and 60s. The US' share in the global soy production has been over 50% until the 1980s (Masuda & Goldsmith, 2009). FAO data shows that the US had an average yearly production of 35 million tonnes from 1961 to 1980, while Brazil produced only 5 million tonnes per year in the same time period. Soybean production in Brazil was almost non-existent in the early 1960s. After new varieties were developed that could grow on lower latitudes, Brazil started to set up its own industry from the late 1960s onwards (Goldsmith & Hirsch, 2006). In the decades that followed Brazil's production increased quickly to its current point. Especially the steep increase from the 1990s onwards contributed to Brazil producing more or less a third of the global soy today (Ritchie & Roser, 2021). Goldsmith and Hirsch (2006) point at several consecutive developments that enabled Brazil to become the world's biggest soy producer in half a century. First, the further genetic adaptation of the soybean in the 1980s, which allowed farmers to grow on even lower latitudes. This made enormous plots of land available for soybean production. Plots of land that, unlike in the US, were not privatized into small plots of land, and could therefore be used for large-scale soy production. Combining this with the mechanical innovations from the US, and high yields and low operating costs in the Brazilian soy sector were the result (Goldsmith & Hirsch, 2006).

While Brazil specializes in producing low cost soybeans, the country lacks the large-scale agro-industrial sector to process the soybeans into the raw commodity's more valuable sub products. Different is the situation for Argentina, a country that started on the same page as its neighbour; little to no soybean production half a century ago. The reasons for Argentina's production growth are similar to that of Brazil, however, according to Regunaga (2010), the difference is that the processing industry grew at the same rate as the country's soybean production. This is the result of a high export tax on whole soybeans, to lower the soybean export and increase the export of the more valuable soybean meal and soy oil (Meade et al., 2016). The result of the export tariffs is visible in figure 3, which shows that Argentina is the biggest exporter in both meal and oil, while taking a third place when it comes to the whole bean (figure 2). Recent tax reforms have made an end to this differentiation, as the export of both whole soybeans and its sub products are currently taxed on an equal rate of 33% (USDA, 2020).

The three main producing countries differ from each other in terms of production costs and export competitiveness. Production costs of soybeans are made up of four components: land costs, crop input costs, machinery costs, and labour costs (Hart, 2017). According to a study conducted by Meade et al. (2016), average production costs per acre are lowest in Argentina and Brazil (lower land and labour costs), while the US has higher yields per acre. Taken together, per bushel, Brazil has the lowest average production costs of the three. In terms of export competitiveness, the US is able to compete with the South-American countries because of the lower transportation costs to the main importing countries (Meade et al., 2016). Policies have a big influence on competitiveness, as is reflected by Argentina's case. The previous Argentinian tax policy put high tariffs on exporting soybeans as a raw commodity. Contrary to Brazil, where the soybeans and soy oil are both free of tax when exported (Jia et al., 2020). Special subsidy programmes that support (soybean) farmers have also been in place in the US, for example the income-support programmes from the 2008 Farm Act (Meade et al., 2016). Another factor worth mentioning is the high input costs in Brazil. Spendings on fertilizers and chemicals, especially in Brazil's largest soy-producing area, Mato Grosso, are high because of the lack of nutrients in the soil. A third interesting component of the competitiveness on the global market are the transportation costs. Meade et al. (2016) show that soybeans from Mato Grosso, the Brazilian state with the lowest farm-level production costs in the world, are exported for a higher price than the US soybeans (+1.5%). This is because the state's inland location leads to high transportation costs (Meade et al., 2016). From this, it is clear that improving infrastructure (Brazil) and policy/tax reforms (Argentina) pave the way for an even more competitive market, in which the pressure, especially on US farmers, rises. These examples, relating to policies, input costs, and transportation costs, are not the only determinants influencing the productions costs and competitiveness. Other factors are climate conditions and resource endowments, which are fixed and therefore a given, and changing exchange rates (Meade et al., 2016). The importance of the exchange rate is underlined by Van Berkum et al. (2006) who point out that the increase in Brazilian soybean production between 2000 and 2004 can be largely attributed to the depreciation of the Brazilian Real. In the years that followed (2005 and 2006), the decrease in soy acreage coincided with an appreciation of the Brazilian currency (Van Berkum et al., 2006).

Another important difference between the producing countries relates to the role of GMO's in their soybean production. The use of genetically modified soybeans differs per country. The largest share of GMO soybeans can be found in Argentina, where almost all soybean production is genetically modified (Milanesi, 2012). The popularity in Argentina is explained by Van Berkum et al. (2006), who state that using GMO soybeans lead to higher production and lower costs for herbicides and tillage activities. Besides this, there is no patent on GMO soy in Argentina (Van Berkum et al., 2006). In Brazil, around 55% of the total acreage consists of GMO soybeans (Milanesi, 2012), compared to 94% in the US (Hart, 2017). The importance of genetic modification in lowering soy's production costs is stressed by Hart (2017). The article states that the modification that made the plant tolerant to the herbicide glyphosate, more commonly known as Roundup, has played an important role in the decreasing input costs for the cultivation of the crop. Overall, it is estimated that 82% of the global soybean acreage consists of GMO soy (De Ridder et al., 2015).

So far, production, export, and import figures where mentioned for countries. However, it is important to mention that not the countries, but private multinational companies are mostly responsible for the actual export, import, and crushing process. Companies like Bunge, Cargill, Dreyfuss, and ADM (partly) own the transportation facilities that bring the commodity from one country to the other (Van Berkum et al., 2006). Processing facilities, both in importing and exporting countries, are often owned by these multinationals. Globally, more than 70% of the soy processing capacity is owned by ADM, Cargill, and Bunge (De Ridder et al., 2015). In the importing countries, animal feed companies are dependent on trading companies for their import of soy, as they lack the economies of scale to import the soy themselves.

2.2. Global Demand for Soybeans

A clear geographical pattern is also visible on the demand side. As can be seen in figure 4, the main importers are China and the European Union (EU). Where China mainly imports whole soy beans, to process them into meal and oil in domestic processing facilities, the EU also directly imports a lot of soybean meal.

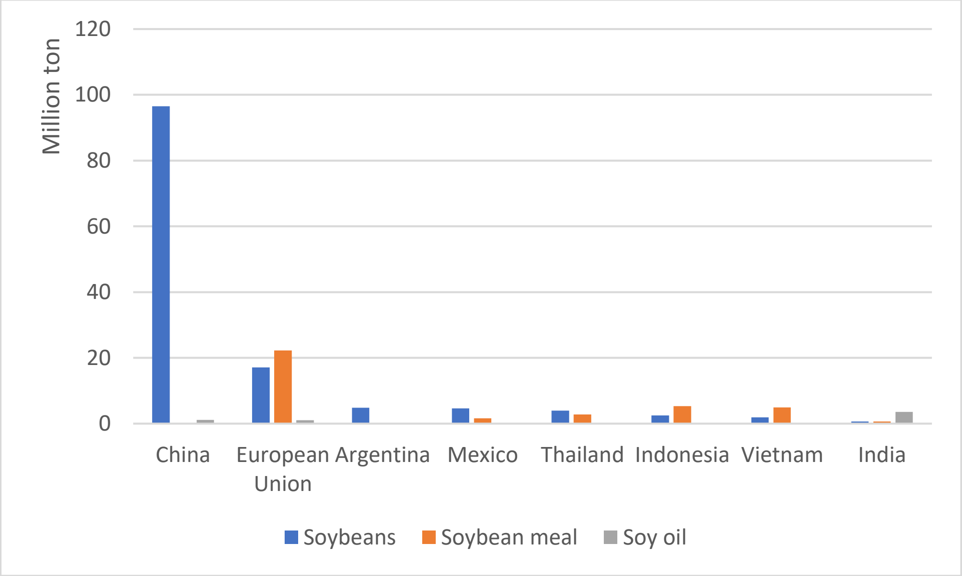

Figure 4: Imports of soybeans (blue), soybean meal (Orange), and soy oil (grey) in tonnes in 2021. Based on: FAO (2023).

The uses of soy, and therefore the reasons for countries to import the commodity, can roughly be split up into three categories: animal feed, human consumption and biofuel (feed, food and fuel). Together, these drivers are responsible for the global demand of soybeans. First, the increasing world population and living standards leading to a higher demand for meat and other animal products. Meat consumption has increased threefold in the last 50 years (Ritchie & Roser, 2021), and is expected to rise with an additional 14% in 2030 (FAO, 2021). FAO (2021) states that poultry is projected to be the meat type with the highest increase in production and consumption. This higher meat consumption has resulted in a rise in the demand for animal feed proteins (Kim et al., 2019). According to Kim et al. (2019), in 2016, it was estimated that globally a total of 1 billion tons of compound feed was produced, of which 44% was used for poultry, 26% for pigs, and 22% for ruminants. 70% of oilseed meals processed in the global compound feed production consists of soybean meal (Kim et al., 2019).

China and the EU are responsible for the majority of demand for soybean meal, making both regions important drivers of the global soybean demand. China produced only 16.4 million tonnes of soybeans, but imported over 100 million tonnes in 2021 (FAO, 2023). Because of increasing feed protein demand, it is estimated that China's self-sufficiency rate will continue to decrease (Kim et al., 2019). In addition to this, Wang and Wei (2021), state that the minimum price for domestic soybean producers, introduced by the Chinese government in 2008, created a gap between the domestic soybean price and the world market price. This resulted in a simultaneous increase in China's soybean imports and government-owned stockpiles (Wang & Wei, 2021), further decreasing China's self-sufficiency rate. The second largest importer of soybeans is the EU. Europe's climate is ill-suited for soybean cultivation, resulting in a feed protein shortage (Kim et al., 2019). The EU's soybean imports totalled 40 million tonnes in 2021 (FAO, 2023).

The influence of policies on the exporting side has been made visible from Argentina's case. When it comes to imports, EU regulation is heavily influencing trade flows. First, the EU's demand for plant-based proteins took a flight from 2003 onwards, after EU legislation was put in place that determined that animal-based proteins (i.e. meat and bone meal (MBM)) were not allowed in animal feed anymore, as a result of the BSE (mad cow disease) crisis (Van Berkum et al., 2006). The same holds for the US, where MBM-restricting legislation was introduced in 2004 (Babula et al., 2004). Prior to these policies, animal-based proteins were a significant component of the world's supply of animal-feed proteins. The fall in demand resulted in prices for blood meals, bones, and ruminant meat used for animal feed drop from $295 per ton before the discovery of BSE to $140 per ton in January 2004 (Babula et al., 2004). A second set of policies heavily influencing the EU's role in the global soy market relate to the large share of GMO varieties in global soy production. The EU allows only a limited amount of soy varieties due to health and environmental concerns. According to De Ridder et al. (2015), the EU's strict regulation on GMO's partly explains the relative low imports from Argentina, where almost all soy is genetically modified. Another clear example is the drop in Dutch imports from the US in the 2000's, in favour of imports from Brazil, where a larger share of the soybeans are non-GMO (De Ridder et al., 2015).

The second main driver of soybean demand is the global increase in biofuel production. According to Sorda et al. (2010), globally, the amount of policies stimulating the use of biofuels, and therefore biofuel production, has increased between 2000 and 2009. The majority of the bioethanol and biodiesel are first-generation biofuels, which means they are made of agricultural commodities, of which soy oil is an important component (Hirani et al., 2018). Currently, only 2.8% of the global soy production is used for biofuels (Ritchie & Roser, 2021), but this is increasing since the 2000s and is expected to keep doing so (Fraanje & Garnett, 2020). The biggest biofuel producers in 2021 were the US, Brazil, Indonesia, and China, respectively (Statista, 2023). The US roughly uses a quarter of its soy oil production to produce biodiesel (Hart, 2017). In terms of consumption of biodiesel, the US ranked first, followed by Germany, Brazil, and France (IndexMundi, 2023). For bioethanol, the top consumers were the US and Brazil, followed by China, Canada and some European countries (IndexMundi, 2023).

The third driver for global soybean demand is the use of soybeans in products for human consumption, i.e. food and cosmetic products. According to Rizzo and Baroni (2018), in Asian countries, soy is used as a food ingredient already for a very long time. In Western society, the use of soybeans is a recent development. Soy is often consumed as a nutritional solution in a vegetarian diet, because of its high level of proteins and its similarity with meat and dairy products (Rizzo & Baroni, 2018). Despite this growing popularity in the Western diet, still only 7% of the soy that is produced is processed into tofu, soy milk, tempeh, and similar products. 13.2% of total soybean production is consumed directly by humans as vegetable oil (Ritchie & Roser, 2021). Next to food, soy is also processed in other products meant for human consumption. The cosmetic and pharmaceutical industry both make use of soybean components (Ali, 2010).

Summarizing, global soy production, expressed in weight, is used for animal feed (75%), human consumption (20%), and industrial purposes (5%) (Fraanje & Garnett, 2020). Based on these numbers, it is easy to state that the growing demand for animal products, and thus the processing of soybeans in animal feed, is the main driver for the large production. However, this does not tell the whole story. First, in terms of economic value, the ratios already change significantly. Fraanje and Garnett (2020) state that currently on average, a third of the revenues are earned with the soy oil and two thirds of the value can be attributed to the soybean meal. Add this to changing consumption patterns, combined with the rise of biofuel policies in the US and EU, and it is safe to say that the demand for soybeans is subject to change in the future.

2.3. Price Determination

This section goes deeper into the economics of the soybean. How the price of soybean comes about and which actors play a role in this market. The type of commodity, the local and global price, and agricultural trading companies form the core of this section.

First, the type of commodity, which highly affects how the price of a commodity comes about. Soybeans are defined as a so-called bulk commodity (Hart, 2017). In essence, this means that soybeans from different farmers or varieties cannot be differentiated from each other, and therefore cannot be priced differently. Exceptions are non-GMO versus GMO and organic versus non-organic soybeans. However, as the majority of produced soy is GMO and less than 0.1% is organic (Hartman et al., 2016), the crop is defined as a homogenous commodity. The result is a competitive market where it is easy for producers to enter and exit. Compared to the high number of sellers (the producers), the amount of actors buying the homogenous soy is much more limited. According to De Ridder et al. (2016), 70% of the global processing capacity is owned by the trading companies Bunge, Cargill, and ADM. A market where many sellers produce a non-differentiable product has important implications for the price development. If profits in the soybean market are high, farmers will switch (partly) from e.g. corn to soy. The higher supply leads to lower prices and profits, and vice versa. In the long-run, prices on a competitive market like this will always be equal to production costs. Important side-note here is that this equilibrium is a long-term phenomenon, with profits and losses on the short-term. This is mainly due to the fact that soybean cultivation takes time, which means that farmers are unable to switch to a different crop from one day to the other. Also, farms are often family business with knowledge on specific crop(s). Switching to a different crop comes with a time investment to get acquainted with the cultivation process, lengthening the market fluctuations.

The paragraph above gives a general idea about the origins of the price of soybeans. Local prices are depending on the global price for soybeans, which is determined on commodity exchanges, of which the Chicago Mercantile Exchange (CME) is the largest. On these exchanges, prices of so-called ‘futures’ are established through global supply and demand, holding all factors into account, like production forecasts and macro-economic trends. Future contracts exist for both the raw commodity (soy beans) and its sub products (soy oil and soybean meal) (Pennings & Leuthold, 2001). The price farmers receive for their soy is based on the future prices, corrected for the transportation costs and the local conditions influencing supply and demand. This difference is called the ‘basis’. For example, on the local market in Mato Grosso, Brazil’s largest soy producing state, supply exceeds demand and transportation costs are low. This creates a situation in which the local (farmers) price is lower than the global (futures) price. As a result, the basis is negative. The basis is positive, i.e. the local price is exceeding the global price, in case of a supply or demand shock, often a drought or flood in a soybean producing area.

2.4. Supply Chain Governance

As the environmental impact of soybean production, primarily in South-America, became clear, the call for sustainable production initiatives became louder, especially in Europe. A number of certificates and schemes have been put in place in the last decades. A public governance initiative that aims at making the soybean supply chain more sustainable is the Forest Code (FC) in Brazil. Liu et al. (2020) state that the FC was set up by the national government and aimed at preventing forested land from being converted into agricultural land. The code is much criticised, as it allowed for a significant amount of legal deforestation and lacked proper enforcement (Liu et al., 2020).

On an international level, several public and private governance initiatives currently exist in or influence the global soy supply chain. The first agreement to be put in place was the Soy Moratorium. Initiated by Greenpeace, the Soy Moratorium is an agreement between (trading) companies, NGOs, and the Brazilian government. Despite being highly successful in reducing the deforestation directly caused by soy production, it has not been able to prevent land from being deforested indirectly (Fraanje & Garnett, 2020). Indirect deforestation is, as explained in the introduction, a significant part of the total deforestation related to soybean cultivation.

Private certificates can be split up in two categories. The majority aims at complying to the EU Soy Sourcing Guidelines of FEFAC. The FEFAC guidelines are debatable as they only state that producers have to comply to forest and eco-system legislation (FEFAC, 2021), meaning that the certificate does allow for legal deforestation. According to Fraanje and Garnett (2020), legal deforestation can lead to 88 mega hectares of deforested land in Brazil per year. Therefore, a new set of certificates was initiated that does not allow for deforestation at all. Examples are the Round Table on Responsible Soy (RTRS) and ProTerra. According to Heron et al. (2018), ProTerra was initiated in 2004 and focusses mainly on excluding GMO material. Next to this, it has standards for sustainable soil and water use, pesticides, and deforestation. RTRS on the other hand accepts both GMO and non-GMO soy, and focusses more on deforestation instead. Other than this, both private certification schemes have a lot of similarities (Heron et al., 2018). So far, these certificates have not been applied on a large scale. According to Heron et al. (2018), the main reason for the limited impact of these schemes is the lack of brand recognition by consumers and the fact that no product differentiation is possible, as there are no differences in quality between the certified and non-certified soy. Criticism added by Jia et al. (2020) relates to the limited scope and geographical coverage, and vagueness regarding definitions, criteria, and indicators. According to Fraanje and Garnett (2020), premiums for soy farmers are too low compared to other commodities. Currently, 22% of the soy imported by the EU complies to any of these certifications, while this only holds for 0.2-6% of soy globally (Fraanje & Garnett, 2020).

2.5. Role of the Netherlands

This section zooms in on the role of the Netherlands in the global soy trade. In 2021, the Netherlands imported over 4 million tonnes of soybeans, which makes the Netherlands, together with Argentina and Mexico, the second largest importer, after China (FAO, 2023). Most soybeans are imported from the US (43.1%) and Brazil (41.9%), while the meal mainly originates from Brazil (54.5%) and Argentina (34.2%) (De Ridder et al., 2015).

The Netherlands' role in the soybean complex is mainly due to trade. Not all the imported soybeans are consumed or processed in the Netherlands. On the contrary, the Netherlands is a transit port for a quarter of the soybeans, which are directly exported after import (FAO, 2023). Half a million tonnes of soy oil is exported after the crushing process, which makes the Netherlands the fourth largest soy oil exporter in the world.

The processed soybeans that remain in the Netherlands are mainly used by the animal feed industry (Hoste, 2014). The four biggest compound feed producers in Europa are all located in the Netherlands, having a market share of over 40% in the EU (Heron et al., 2018). Further down the supply chain, the Netherlands remains an important exporting country, meaning that the soy is indirectly consumed outside the country. Hoste (2014) estimates that 5% of compound found produced in the Netherlands is exported. The meat, eggs and milk, for which the compound feed is an input, is also exported on a large scale. According to Jukema et al. (2022), exports for meat totalled 9.1 billion euros in 2020. For dairy and eggs, export totalled 8.7 billion euros.

The intensive agricultural sector, combined with the Netherlands' role as a transit port to the rest of Europe, explain the prominent Dutch role in the global soybean market. The importance of soy is illustrated by De Ridder et al. (2015), who state that the industries in the Netherlands (food processing industry, animal feed and livestock industry, etc.) that depend on soy imports make up for 7% of the GDP in the Netherlands. In many cases, no substitute of equal quality and quantity is available, making the Netherlands highly dependent on the global soy market. This vulnerability is a risk for the Netherlands, as certain trends, like the global increase in wealth and meat consumption and climate change, can cause shocks on the global soy market (De Ridder et al., 2015).

3. Theoretical Framework

In this chapter economic theory and concepts are used to create a system of functions that together form a structural theoretical model of the soybean complex. The structural model consists of three functions, introduced in section 3.1. The three sections that follow dive deeper into each of the functions, explaining the different variables and proving their economic link to the global soy market. In the final section of this chapter, the theoretical framework of an additional analysis is presented, in which the price relationship between the soy commodities is explored. The presented models are designed in line with the quantitative sub research questions posted in section 1.3.

3.1. The Soybean Complex

Three products play a central role in the soybean complex: whole soybeans, soybean meal, and soy oil. The majority of the global soybean production is processed into soybean meal and soy oil, using a process called crushing. As a result, the supply function describes the supply of the whole soybean, i.e. the raw product. The demand functions describe the demand for the processed products soybean meal and soy oil.

In order to link the supply and demand functions in the model, a transformation function is formulated with one input, the whole soybean, and two outputs, soy oil and soybean meal. Transformation production functions are a convenient way to describe a production set with multiple outputs (Boyd, 2023). Soybeans are used to produce either soy oil or soybean meal, which is captured in the following condition:

$$g_o(qo) + g_m(qm) \leq qb \tag{1}$$

In which $g_o(qo)$ represents the soybeans required to produce soy oil, and $g_m(qm)$ the soybeans required to produce soybean meal. More specifically, these functions imply the oil and meal content of the soybean. The combined output of soy oil and soybean meal can never exceed the total production of soybeans (qb).

The remainder of the chapter goes deeper into the individual supply and demand functions. Including both supply and demand is important, as studies that focus on either the supply or demand side tend to exaggerate effects of shocks in such a limited model (Wang & McPhail, 2014). The supply function contains output and variable input prices, fixed factors, and a set of supply shifters, e.g. technology (Krugman & Wells, 2013). The function is formulated as follows:

$$qsb_t = B(pb_t, w_t, F_t, xb_t) \tag{2}$$

where qsb_t denotes the total quantity of soybeans supplied; pb_t the price of soybeans; w_t the price of variable inputs; F_t the fixed input factors; and xb_t the supply shifters. The subscript t represents the time period.

This supply function follows from the underlying assumption that farmers aim to maximize profits. By differentiating the profit function subject to the producer's technology constraint, a supply function is created which allows producers to use relative input and output prices to determine the optimal level and input demands and output supply. This is called Hotelling's Lemma (Gehrke & Peerlings, 2022).

On the demand side, demand functions for soybean meal and soy oil are set up. The two subproducts have their own demand function, as this allows for the inclusion of more specific variables. Demand is generally determined by prices, income, and a set of other demand shifters, e.g. policies or changes in tastes and expectations (Krugman & Wells, 2013). The demand function for soybean meal looks as follows:

$$qdm_t = M(pm_t, I_t, xm_t) \tag{3}$$

where qdm_t denotes the total quantity of soybean meal demanded; pm_t the price of soybean meal; I_t the income factor; and xm_t the specific demand shifters for soybean meal.

The demand function for soy oil is formulated as:

$$qdo_t = O(po_t, I_t, xo_t) \tag{4}$$

where qdo_t represents the total quantity of soy oil demanded; po_t the price of soy oil; I_t the income factor; and xo_t the specific demand shifters for soy oil.

Both demand functions follow from the assumption that the users of soybean meal and soy oil (animal feed industry, vegetable oil industry, etc.) choose optimal input levels to minimize their costs. This is called Shephard's Lemma (Gehrke & Peerlings, 2022).

The following sections aim at identifying the price variables ($pb_t, pm_t, po_t\ and\ w_t$), the income variable ($I_t$), the fixed factors ($F_t$) and the supply and demand shifters (xm_t, xo_t, and xb_t). If the details of the supply and demand functions are known, the analysis can continue with finding the presence and magnitude of these effects on the global soy market. To keep the model manageable, going into full detail is not possible. Therefore, the model only includes the principal determinants of supply and demand in the soybean complex, leaving out elements that are less relevant for the analysis.

3.2. Soybean Supply

Soybean price

General economic theory suggests that a positive relationship exists between the quantity of soybeans produced and the price of soybeans. This is confirmed by Iqbal and Babcock (2018), who found that soybeans have a long-run own-price elasticity of supply equal to 0.631. Put differently, a 1% price increase leads to a 0.631% increase in quantity supplied on the long-run.

This strong relationship does not hold on the short-run, when the own-price elasticity is 0.213 (Iqbal & Babcock, 2018). This makes sense, as farmers are not able to directly adjust their production to a change in price. This lag in response is further analysed by Babula et al. (2004), who state that soybean production quantity reacts with a six-month lag to a price shock. At time periods beyond 18 months, the price's influence on the supply of soybeans really becomes significant (Babula et al., 2004).

Fixed production factor

The fixed production factors in agriculture can be divided into three categories: land, labour and capital; where technology and management are embedded in capital and labour respectively (Blank, 2015).

The importance of each production factor differs per agricultural practice. For crop cultivation, land availability is generally considered the most limiting factor (Blank, 2015). Therefore, this model includes the total amount of land dedicated to agriculture ($land_t$).

Energy price

Next to fixed inputs, soybean producers use variable inputs. The difference is that the quantity of variable inputs used can be changed on the short term (Krugman & Wells, 2013). For soybeans, and for agricultural commodities in general, holds that energy is an important input. Wang and McPhail (2014), identify the ways energy, and more specifically crude oil, affect the supply of agricultural commodities. First, production costs increase in case of higher energy prices. Both directly, e.g. for farm machinery, and indirectly, e.g. for fertilizers and chemicals. Second, higher energy prices lead to more biofuel production, increasing the competition of using land for food/feed or fuel (Wang & McPhail, 2014). For soybean production in particular, these links with energy as a production input also hold.

Energy prices are represented by the global average crude oil price (pco_t). This is the most important energy source, for which the price is determined by a combination of its physical availability, global business cycles, and expectations (Kilian, 2009).

Corn price

Crop rotation between corn and soybeans is a common practice in the US (Livingston et al. 2008). Consequently, price changes in either one of them has an impact on both supplies. This is confirmed by Livingston et al. (2008), who state that US farmers switch from a corn-soybean rotation to a corn monoculture in case of rising corn prices. More quantitatively, supply responses can be captured in substitution elasticities. Kim and Moschini (2018) link changing corn prices to the soybean acreage response and found cross-price elasticities of -0.50 (short run), -0.32 (long run), and -0.65 (static). Similar elasticities were found by Iqbal and Babcock (2018).

Literature provides conclusive evidence for soybeans and corn being substitutes in production. Therefore, the price of corn (pc_t) is included in the model.

Adding all variables to the general function leaves us with the following supply function for soybeans:

$$qsb_t = S(pb_t, land_t, pco_t, pc_t) \tag{5}$$

3.3. Soy Oil Demand

Soy oil price

A negative relationship is expected between the quantity of soy oil demanded and the price of soy oil (Babula et al., 2004). This means that rising prices lead to declining demand, and vice versa, following standard microeconomic theory (Krugman & Wells, 2013). The soy oil price and demand affect each other without delay (Babula et al., 2004).

The negative relationship is expected to be strong. This is shown by Santeramo (2017), who calculated a price elasticity of demand of -3.22 for soy oil in the EU.

Sunflower oil price

The price of sunflower oil is included in the demand function because it is the most relevant substitute for soy oil. When it comes to human consumption, both oilseeds are direct substitutes for the production of vegetable oil (Ritchie & Roser, 2021). Next to this, soy and sunflower oil are substitutes for the production of biodiesel (Pimentel & Patzek, 2005).

Overall, vegetable oils are easily substitutable, creating a highly competitive market in which small price changes can have large effects. Therefore, the model includes the sunflower oil price, represented with the variable psf_t.

Chinese income growth

The second main component of a Marshallian demand function is income. General theory states that there is a positive relationship between income and demand for most goods. According to Tilman and Clark (2014), this holds for soy oil in particular as an increase in income leads to rising total calorie demand per capita and an increase in 'empty calories', i.e. refined fats and oils, alcohol, and oils.

China's income growth is expected to have the most significant effect on the global soy oil demand. According to Jenkins (2011), China's demand effect is 6.4% between 2002 and 2007, which means that global demand for soy oil was 6.4% higher compared to the scenario in which China's demand had increased on an average rate. This provides a strong basis for including a variable capturing the average Chinese per capita GDP, denoted as ($cgdp_t$).

Biodiesel production

An interesting exogenous factor influencing the soy oil demand are biofuel-related policies, which started in the first decade of the 21st century in the EU and US (Sorda et al., 2010; IEA, 2023). As a result of these policies, soy oil is used for both biodiesel and food, meaning that the food and energy sector compete over soy oil (Enciso et al., 2016). Overall, there is a large body of literature that points at biofuels as a contributing factor to increasing food prices (Tomei & Helliwell, 2016; Hirani et al., 2018).

This factor, aiming to identify and capture the shocks caused by policy interventions relating to biofuels, is represented in the model by the total global biodiesel production (bio_t).

Industrial raw materials

Examples of industrial uses for soybean oil are plastics, paint, and ink (Masuda & Goldsmith, 2009). Next to this, the cosmetic and pharmaceutical industry both make use of soybean components (Ali, 2010). Taken together, 2% of the global soy oil production is used for these industrial purposes, which includes further processed food components like lubricants and emulsifiers (Fraanje & Garnett, 2020). This makes the impact of industrial usage on the total demand for soy oil too small, and is therefore not included in the model.

Adding all variables to the general function leaves us with the following demand function for soy oil:

$$qdo_t = O(po_t, psf_t, cgdp_t, bio_t) \tag{6}$$

3.4. Soybean Meal Demand

Soybean meal price

A negative relationship is expected between the quantity of soybean meal demanded (*qdm_t*) and the price of soybean meal (*pm_t*). The effect of a price change on the quantity of soybean meal demanded starts within the same month (Babula et al., 2004).

Corn price

Soybeans and corn are substitutes when it comes to consumption (Goodwin et al., 2005). Animal-feed producers that use corn as an input, are able to switch to e.g. soybeans if changes in relative prices occur. However, the level of substitution between the two commodities is limited because animal feed needs to meet certain criteria, like protein, fibre, and starch contents (Babula et al., 2004). As corn is primarily used for its high starch-content, while soybeans are the primary protein source in animal feed, the substitution effects between these two commodities is limited.

Despite the limited substitution elasticity, literature suggests a significant relationship between the corn and soybean meal markets (Roberts & Schlenker, 2013). To capture the effects of this relationship on the demand for soybean meal, the price of corn (*pc_t*) is included in the model.

Meat consumption (income)

According to Tilman and Clark (2014), an increase in income is among the most important factors in changing people's diets. Increasing income (i.e. gross domestic product, GDP) leads to a greater demand for meat proteins per capita. Meat consumption on its turn, is positively linked to soybean meal demand. Masuda and Goldsmith (2012) state that for China, a 1% increase in meat and egg production leads to the country demanding 0.909% more soybean meal. The variable included in the model to represent the income component is global meat production, which is denoted as $meat_t$.

Meat and bone meal regulation

An interesting exogenous demand shifter for soybean meal is the BSE-crisis, starting in 1996 and resulting in bans on the use of ruminant meat and bone meal (MBM) in animal-feed in 2001 and 2004 (EU and US respectively). The drop in the price for MBM in the mid-1990s is a direct effect of the crisis. The 2003/2004 soybean meal price spike can mainly be attributed to the increase in demand as a result of the BSE-crisis (Babula et al., 2004).

The effects of this crisis and the resulting policy are included in the model by adding a price variable for MBM (pmb_t). This allows the model to capture and quantify the changes in soybean meal demand due to disruptions in the MBM market.

Adding all variables to the general function leaves us with the following demand function for soybean meal:

$$qdm_t = M(pm_t, pc_t, meat_t, pmb_t) \tag{7}$$

3.5. Price Relationship

Dividing (sub)products into categories based on their economic importance is a common practice in literature. Kim and Heo (2012) provide a product classification for production processes with multiple products. First, the main product, which accounts for a significant part of the revenues, and therefore is a major determinant in decision-making related to the production process. Second, if two or more products are influencing decision-making to a more or less equal extent, they are considered co-products. The third category is when a product has little to no influence on decision-making, making it a by-product. The supply of a by-product depends on how much is produced of the main product (Kim & Heo, 2012).

Section 3.1 described how soybeans as a raw material are crushed into soy oil and soybean meal. This is captured in the following equation, which is based on Baum and Hurn (2021):

$$pb_t + \pi_t = 0.785 * pm_t + 0.185 * po_t \tag{8}$$

Where π_t denotes the profit margin of the soybean processor in year t. All price levels indicate the average price per ton. The equilibrium equation assumes that the average soybean consists of 78.5% meal and 18.5% oil, and that the residual 3% is worthless.

This link has some important implications for the price relationship between these three commodities, and therefore for their classification as main-, by-, or coproduct. Simple reasoning suggests that a price rise in one of the processed commodities, results in more supply of both soybean meal and oil. Before diving into the analysis of how the prices relate to each other, first some theoretical and empirical evidence of the existence of such a relationship is provided.

From a theoretical perspective, the existence of a price relationship between soybeans, soy oil, and soybean meal makes sense. The Law of One Price should apply on the global soybean market, given the fact that soybeans, soybean meal, and soy oil are homogenous products. Therefore, prices should equalize at local markets around the globe in the long-run, driven by the global process of arbitraging. Proof of the existence of the Law of One Price on the long-run in the global soybean market and the presence of arbitrage is given by Margarido et al. (2007). If the Law of One Price holds for the sub-products as well, this is a very strong indication of a long-term relationship between all three commodities. If all would equalize towards a certain level in the long-run according to the Law of One Price, the difference between the three commodities is very likely to have a clear pattern.

Empirical evidence of a price relationship between the three soy commodities is provided in figure 5. The graph shows how a price spike in one market is often followed by similar spikes in the other markets, or that the price spikes happen simultaneously. More empirical evidence for a price relationship between soybeans, soybean meal, and soy oil is provided by (Babula et al., 2004).

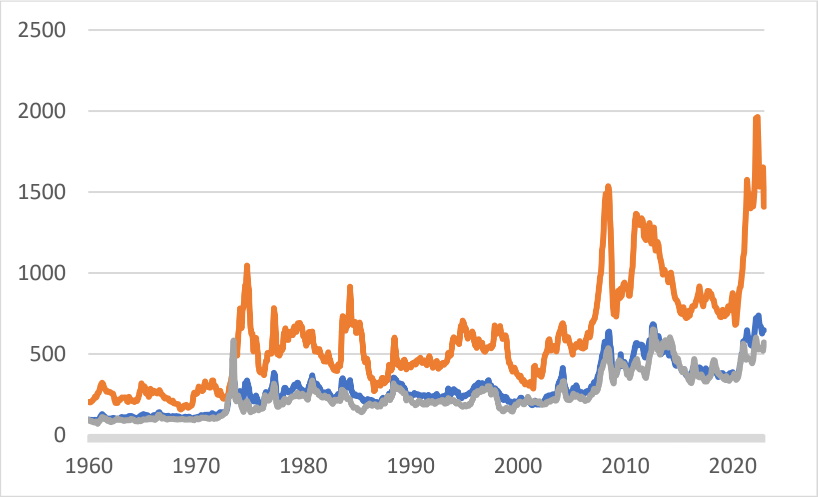

Figure 5: Price development over time of soy beans (blue), soy oil (orange), and soybean meal (grey). Based on: World Bank (2023a).

Theory and empirical evidence give a strong indication of a long-term price relationship between the three price variables in the structural model.

In order to wrap-up this chapter, a schematic overview of the soybean complex is given in figure 6. The overview does not reflect the complexity of the real situation, but is a helpful tool to grasp the soybean complex. Note that the overview does not fully match with the model described, as some extra variables and relationships are included in the overview.

Figure 6: Overview of the soybean complex.

4. Methodology

This chapter describes the empirical structural soy model, theoretically described in chapter 3, and the estimation techniques used. Section 4.2 focuses on the empirical approach for the price relationship analysis.

4.1. Structural Supply and Demand Model

The set of equations in the structural model, consisting of a supply function and two demand functions, is simultaneously estimated. The equations form a system according to the following equilibrium condition:

$$lqdo_t + lqdm_t = lqsb_t + v_t \tag{9}$$

where v_t captures the soy peel component of the soybean, the annual net storage, and the unprocessed soybeans used for human consumption and animal feed. The model is not focussing on these components, therefore treating v_t as part of the error term. Note the prefix l, which indicates that the equations are estimated in a logarithmic form. This is done to mitigate the impact of outliers and for matters of interpretation.

Before simultaneously estimating the system, the three equations are estimated separately using OLS:

$$lqsb_t = \beta_{1,0} + \beta_{1,1} \cdot lrpb_t + \beta_{1,2} \cdot lland_t + \beta_{1,3} \cdot lrpco_t + \beta_{1,4} \cdot trend + \varepsilon_{1t} \tag{10}$$

$$lqdo_t = \beta_{2,0} + \beta_{2,1} \cdot lrpo_t + \beta_{2,2} \cdot lrcgdp_t + \beta_{2,3} \cdot lbio_t + \beta_{2,4} \cdot trend + \varepsilon_{2t} \tag{11}$$

$$lqdm_t = \beta_{3,0} + \beta_{3,1} \cdot lrpm_t + \beta_{3,2} \cdot lmeat_t + \beta_{3,3} \cdot lpmb_t + \beta_{3,4} \cdot trend + \varepsilon_{3t} \tag{12}$$

where the β-coefficients denote the effect of each variable on the dependent variable. The error terms ε capture all the variation not explained by the independent variables and $trend$ includes a linear time trend where 1994=0. Note here that, following standard microeconomic theory, the model makes use of relative prices and real income, indicated with the prefix r. Theory states that supply and demand functions satisfy the condition of homogeneity of degree zero. This implies that the level of supply and demand is determined by relative prices instead of absolute prices. In the supply function, the corn price will be used as a numeraire for the soybean and energy prices ($rpb_t = pb_t/pc_t$ and $rpco_t = pco_t/pc_t$). The corn price is chosen as numeraire because it is highly related to both the soybean and energy price, but not directly important for this research. In the demand function for soy oil, the sunflower oil price is used as a numeraire for the soy oil price ($rpo_t = po_t/psf_t$). The real income variable is created by dividing the income factor by the numeraire ($rcgdp_t = cgdp_t/psf_t$). The relative price between soybean meal and corn is used in the demand function for soybean meal ($rpm_t = pm_t/pc_t$).

In case the standard econometric framework of OLS is biased and inconsistent, this could be due to covariance between $lrpb_t$, $lrpo_t$, and $lrpm_t$ and the equation's error terms. Endogeneity is a possibility in these variables in particular because a certain shock, like a bad harvest, is likely to affect both the production and the price level. The simultaneous equations bias is solved by estimating the model in a 2SLS framework. To ensure the strength and validity of the IV's used, external variables are introduced. Sorghum prices for the supply and meal demand functions and palm oil prices for the soy oil demand. These variables are selected because they are likely to correlate with the price variables, but are uncorrelated to shocks in production (residuals). Palm oil and soy oil, despite being direct substitutes in the feed and energy industry (which connects their prices), are not produced in the same geographical area (uncorrelated production). A data overview (table A1) and descriptive statistics (table A2) for the IV's can be found in the appendix. Post-estimation tests for overidentifying restrictions and the strength of the IV's are performed to confirm that the variables used are indeed valid and strong. All equations are over- or exactly-identified (more or equal amount of potential IVs than explanatory endogenous variables), which is a necessary condition for a 2SLS estimation. To determine the appropriate estimation technique, a Hausman specification test on the equations (10)-(12) is performed.

Equations may also be related via their residuals, if shocks affect the various supply functions in the same way (e.g. bad weather or a plant disease). This can be exploited using a system estimator such as SUR or 3SLS to obtain efficiency gains in estimation. Based on theory and common sense, interrelated residuals are likely to be present; the demand for soybean meal and soy oil and the supply of soybeans are affected by common shocks, creating correlation between the residuals of the three equations. As a result, after estimating the system using the single-equation estimator OLS or 2SLS, the analysis continues by estimating the structural model using either a SUR or 3SLS framework, again depending on the outcomes of the Hausman specification tests.

4.2. Price Cointegration Analysis

The second analysis examines the relationship between the three price variables. This analysis follows from section 3.5, where it was theoretically established that a price relationship in the soybean complex exists. The cointegration parameters in the soybean complex are estimated using monthly prices, as opposite to the estimation of the structural model in (4.1) which is based on yearly prices.

Model specification

Theory (section 3.5) and visual inspection (figure 5) provide a strong indication of a long-run equilibrium relationship among the variables, which is formulated as follows:

$$lpb_t = \beta_1 \cdot lpm_t + \beta_2 \cdot \ lpo_t + \varepsilon_t \qquad (13)$$

Note the prefix l, indicating that the analysis is performed using logarithmic price variables. Implicit in this equation is the technical relationship between the prices of soybeans, soy oil, and soybean meal (8). As theory suggests strong long-term price relationships, the parameters are estimated in a VEC framework. Therefore, an unrestricted VAR is reparametrized as a VECM, which is generally formulated as:

$$\Delta z_t = \delta + \sum_{i=1}^{k-1} \Gamma_i \Delta z_{t-i} + \Pi z_{t-1} + \varepsilon_t \qquad (14)$$

where $\Gamma_i = -\sum_{j=i+1}^{k} \Phi_j$ and $\Pi = \sum_{i=1}^{k} \Phi_i - I_k$, i.e. the coefficients representing the short-run dynamics and the rank of the long-run matrix, respectively. This means the model includes the long-run equilibrium and captures short-term adjustments to that equilibrium. As the VECM is a differenced VAR model, z_t, which is a vector of all the variables used in the model, is in first differences (Δ) and the number of lags is identified as $k-1$. Shocks to the model are captured by ε_{it}.

The rank of Π indicates the number of cointegrating relationships. Practically, this means that if the rank is zero, a VAR model in first differences is appropriate. If all variables in z_t are of the same order of integration and are cointegrated with a certain rank, it means that $r < m$ long-run equilibrium relationships exist among the variables that are stationary at *I*(0).

The general VECM in equation (14) is rewritten to a system of equations to be estimated:

$$\Delta lpb_t = \alpha_1 + \sum_{j=1}^{k-1} \beta_{11j} \Delta lpb_{t-j} + \sum_{j=1}^{k-1} \beta_{12j} \Delta lpo_{t-j} + \sum_{j=1}^{k-1} \beta_{13j} \Delta lpm_{t-j} + \lambda_1 ECT_{t-1} + \varepsilon_{1t} \quad (15)$$

$$\Delta lpo_t = \alpha_2 + \sum_{j=1}^{k-1} \beta_{21j} \Delta lpb_{t-j} + \sum_{j=1}^{k-1} \beta_{22j} \Delta lpo_{t-j} + \sum_{j=1}^{k-1} \beta_{23j} \Delta lpm_{t-j} + \lambda_2 ECT_{t-1} + \varepsilon_{2t} \quad (16)$$

$$\Delta lpm_t = \alpha_3 + \sum_{j=1}^{k-1} \beta_{31j} \Delta lpb_{t-j} + \sum_{j=1}^{k-1} \beta_{32j} \Delta lpo_{t-j} + \sum_{j=1}^{k-1} \beta_{33j} \Delta lpm_{t-j} + \lambda_3 ECT_{t-1} + \varepsilon_{3t} \quad (17)$$

where β_{ij} represent the coefficients of the short-run dynamics of the model's long-run equilibrium. α indicates the constant term and ECT_{t-1} the error correction terms. The latter allow the model to deviate from the long-run equilibrium between supply and demand variables. They are denoted as a lagged value of the error term, which makes them correct for the deviation in the equilibrium from the previous time period. The $\lambda's$ indicate the speed of adjustment, i.e. how fast the system adjusts back to the long-run equilibrium.

Estimation procedure

The analysis starts with unit root tests for the stationarity of the data series. If the level values of the three variables appear to be non-stationary, a first-differenced dataset will be used to establish the presence of long-run equilibrium relationships among the variables. First-differenced data is used to prevent a spurious regression. The presence of cointegrating relationships is examined using the Johansen procedure. The trace and maximum eigenvalue test statistics are used, where H_0 is $r = m$, and H_a $r \leq m + 1$ cointegrating relationships. With the rank of Π known, this results in Π being decomposed into two $m \times r$ matrices, α and β:

$$\Pi = \alpha\beta' \qquad (18)$$

where β denotes the long-run cointegrating vectors and α the adjustment coefficients matrix. The latter is a measurement of how changes in the long-run equilibrium (ϵ_t) impact the system:

$$\epsilon_t = \beta' z_t \sim I(0) \qquad (19)$$

The information criteria and common sense are used to establish the optimal lag length. After running the full model, some diagnostic tests are performed to check for the model's stability, and to make sure that the residuals are not autocorrelated. The last step is to retrieve the parameters for the long-run relationships and short-run model dynamics and interpret the results.

Granger causality test

Granger causality indicates whether the lagged values of a variable help in predicting another variable (Stock & Watson, 2001). Put differently, which variable changes first and which variable follows. By doing this, the processed soybean products can be classified as either main-, by- or co-product. This makes the Granger causality test well-suited to achieve the goal of this analysis.

Note that Granger causality tests should not be interpreted as full causal relationships. Full causality requires a set of information containing all the relevant data that exists in the world. As information sets are always restricted, Grosche (2014) reformulates the interpretation of the test to *prima facie causality*. This means that the results should be interpreted as one variable causing the other, as long as the best dataset available contains no opposing information.

The correct estimation of the model reveals potential Granger causality relationships among the price variables. When looking at the Granger causality between the three variables, a distinction can be made between short- and long-run causality. According to Pala (2013), short-run causality is tested by the Wald-statistic for the coefficients of the explanatory variables. Long-run causality is indicated by the F-statistic of the error correction terms (Pala, 2013), which can be retrieved from the VECM estimation results.

In performing the analysis, special attention should be paid to several steps in the process. According to Grosche (2014), the results are depending on the composition of the dataset, and are therefore sensitive to measurement errors, omitted variables, incorrect specification of the variables, time varying effects in the data set period, and high temporal data aggregation levels (Grosche, 2014).

5. Data

This section presents the data used for both analyses. For each model, a data overview containing all the variables is provided. The variable's source, definition, time period, and unit are also included. A second table is shown containing the data's basic statistics. Both tables contain information on the variables prior to their transformation to a logarithmic form.

5.1. Structural Supply and Demand Model

Table 1 summarizes the variables used in the structural equations for soybean supply, soy oil and soybean meal demands:

Table 1: Structural model data overview

Variable	Abbr.	Definition	Source	Data availability	Unit
Soybean supply	qsb_t	Global production of soybeans	FAO (2023)	1961-2021	Million tonnes (T)
Soybean price	pb_t	Average price of soybeans	World Bank (2023b)	1960-2022	US dollars per metric ton ($/mt)
Land availability	$land_t$	Land used for agricultural production	World Bank (2023d)	1961-2020	Million square kilometres (km^2)
Crude oil price	pco_t	Average price of crude oil	World Bank (2023b)	1960-2022	Nominal US dollar per barrel ($/bbl)
Soy oil demand	qdo_t	Global production of soy oil.	FAO (2023)	1961 – 2020	Million tonnes (T)
Soy oil price	po_t	Average price of soy oil	World Bank (2023b)	1960-2022	US dollars per metric ton ($/mt)
Sunflower oil price	psf_t	Average price of sunflower oil	World Bank (2023a)	Jan 1990 -Feb 2023 (averaged to annual)	US dollars per metric ton ($/mt)
China's GDP per capita (income)	$cgdp_t$	Chinese average GDP per capita	World Bank (2023c)	1960-2021	Constant 2015 US dollars ($)
Biodiesel production	bio_t	Global biodiesel production	Undata (2023)	1990-2021	Thousand metric tonnes (mt)
Soybean meal demand	qdm_t	Global production of soybean meal	USDA (2023)	1995-2021	Million metric tonnes (mt)
Soybean meal price	pm_t	Average price of soybean meal	World Bank (2023b)	1960-2022	Nominal US dollars per metric ton ($/mt)

Corn price	pc_t	Average price of corn	World Bank (2023b)	1960-2022	Nominal US dollars per metric ton ($/mt)
Meat consumption (income)	$meat_t$	Global meat production	Ritchie et al. (2019)	1961-2020	Million tonnes (T)
Meat and bone meal price	pmb_t	Producer price index of meat and bone meal	U.S. Bureau of Labor Statistics (2023).	Dec 1982- Feb 2023 (averaged to annual)	Index Dec 1982=100

Note that for most variables, data is available for the years 1961-2020. Exceptions are biodiesel and soybean meal production and sunflower and MBM prices. Data is most limited for soybean meal production, for which the data starts at 1995. In estimating the model, soybean meal production is the limiting factor. Therefore, based on the data limitations, the structural model is estimated using 1995-2020 data. Table 2 gives the descriptive statistics of the variables in this time range:

Table 2: Descriptive statistics structural model variables

Variable	**Abbr.**	**Obs.**	**Mean**	**Std. Dev.**	**Min**	**Max**
Soybean supply	qsb_t	26	236.62	73.94	126.92	359.51
Soybean price	pb_t	26	359.40	116.55	195.83	595.51
Land availability	$land_t$	26	48.00	0.48	47.23	48.73
Crude oil price	pco_t	26	53.28	30.21	13.06	105.01
Soy oil demand	qdo_t	26	37.82	12.38	20.11	59.89
Soy oil price	po_t	26	752.55	265.24	338.14	1297.66
Sunflower oil price	psf_t	26	816.10	266.82	419.90	1495.40
China's GDP per capita (income)	$cgdp_t$	26	5074.23	2934.37	1520.03	10358.26
Biodiesel production	bio_t	26	17750.77	17474.47	213.00	48914.00
Soybean meal demand	qdm_t	26	163.40	49.91	89.08	248.25
Soybean meal price	pm_t	26	326.52	115.20	165.17	543.46
Corn price	pc_t	26	156.83	59.29	88.53	298.42

Meat production	$meat_t$	26	276.79	44.25	206.88	342.10
MBM price	pmb_t	26	125.01	38.12	75.11	212.48

5.2. Price Cointegration Analysis

Table 3 summarizes the variables used in the price relationship analysis for soybean, soy oil, and soybean meal prices. All data is retrieved from World Bank (2023a).

Table 3: Data overview for the price relationship analysis

Variable	**Abbreviation**	**Definition**	**Monthly time period**	**Unit**
Soybean meal price	pm_t	Global average price level of soybean meal	Jan 1960 – Dec 2022	Nominal US dollar per metric ton ($/mt)
Soy oil price	po_t	Global average price level of soy oil	Jan 1960 – Dec 2022	Nominal US dollar per metric ton ($/mt)
Soybean price	pb_t	Global average price level of soybeans	Jan 1960 – Dec 2022	Nominal US dollar per metric ton ($/mt)

Note that monthly data is available for the 1960-2022 time period. For this research, the 1960-1980 price data is deemed irrelevant, as it does not represent the current state of the global soy market. Therefore, table 4 gives the descriptive statistics of the price variables used in the price relationship analysis, which is between January 1980 and December 2022.

Table 4: Descriptive statistics soybean, soy oil, and soybean meal prices

Variable	**Abbr.**	**Obs.**	**Mean**	**Std. Dev.**	**Min**	**Max**
Soybeans	pb_t	516	336.09	126.06	183.00	737.06
Soy oil	po_t	516	697.46	316.93	271.00	1962.88
Soybean meal	pm_t	516	296.50	118.99	141.00	651.35

Note that the number of observations for the price relationship analysis far exceeds those of the structural model. Higher data availability allows for the estimation of a more accurate price relationship between soybeans, soy oil, and soybean meal.

6. Results

This chapter starts with the estimation results and interpretation of the structural supply and demand model. Specific attention is given to endogeneity of prices by means of a Hausman specification test. Section 6.2 displays the results of the price analysis estimated in a VEC framework, followed by interpretation of the Granger causality tests.

6.1. Structural Supply and Demand Model

The analysis starts by estimating the structural model using the single-equation estimators OLS and 2SLS, for which the results can be found in table 5. Next, a Hausman specification test is performed on both sets of results. No significant endogeneity appears to be present in the equations (table 6), making OLS the preferred estimator since it gives more efficient estimates. Given the potential correlation among the residuals of the different equations, using a simultaneous estimation technique can lead to efficiency gains. Therefore, the system is also estimated in a SUR framework, for which the results can also be found in table 5.

Table 5: Estimation results in OLS, 2SLS and SUR framework

Equation	**Variable**	**OLS**	**2SLS**	**SUR**
$lqsb_t$	$lrpb_t$	0.22 (0.09)**	0.16 (0.11)	0.24 (0.08)***
	$lland_t$	0.03 (0.05)	1.33 (2.21)	-0.01 (1.79)
	$lrpco_t$	0.02 (0.03)	0.03 (0.03)	0.03 (0.02)
	$trend$	0.04 (0.00)***	0.04 (0.00)***	0.04 (0.00)***
	$cons$	3.30 (2.23)	-0.38 (8.60)	4.76 (7.00)
$lqdo_t$	$lrpo_t$	-0.01 (0.16)	0.48 (0.27)*	-0.01 (0.14)
	$lrcgdp_t$	0.25 (0.04)***	0.24 (0.05)***	0.25 (0.03)***
	$lbio_t$	0.11 (0.01)***	0.10 (0.01)***	0.11 (0.01)***
	$cons$	2.16 (0.08)***	2.34 (0.12)***	2.16 (0.07)***
$lqdm_t$	$lrpm_t$	0.06 (0.03)*	0.04 (0.04)	0.06 (0.03)**
	$lmeat_t$	0.96 (0.29)***	0.96 (0.29)***	0.76 (0.25)***
	$lpmb_t$	-0.08 (0.02)***	-0.08 (0.02)***	-0.08 (0.02)***
	$trend$	0.02 (0.01)***	0.02 (0.01)***	0.03 (0.01)***
	$cons$	-0.28 (1.53)	-0.27 (1.54)	0.76 (1.30)

*Note: *, **, and *** denote the critical significance levels of 10%, 5%, and 1% respectively. Standard errors in parentheses.*

Table 6: Hausman specification test

Hausman specification test			
	$lqsb_t$	$lqdo_t$	$lqdm_t$
H_0	$Cov(lrpb_t, \varepsilon_{1t}) = 0$	$Cov(lrpo_t, \varepsilon_{2t}) = 0$	$Cov(lrpm_t, \varepsilon_{3t}) = 0$
Test result	1.25 (0.87)	4.92 (0.18)	1.52 (0.82)

Note: p-values in parentheses

Lower standard errors indicate a slight efficiency gain when using a system estimation technique over a single-equation estimator. However, the Breusch-Pagan test does not reject the H_0 of independence among the residuals on a 10% critical level. The correlation matrix of the residuals support the notion of uncorrelated residuals. Therefore, the interpretation of the results is based on the OLS parameters.

The most significant determinant in explaining variation in the global supply of soybeans is the price of soybeans; a 1% increase of the soybean price leads to 0.22% more production. Similar results were found by Iqbal and Babcock (2018), who state that the soybean's short-run own-price elasticity is 0.213. Total land availability is not significantly influencing the soybean supply. In this case, an insignificant parameter is a notable finding, as globally there is a downward trend for total agricultural land since 1995. The insignificant parameter indicates that the soy acreage did not proportionally decrease with total agricultural land. A non-significant parameter is also found for the price of crude oil. As literature suggests that such a relationship does exist (Moss et al., 2010), this is a remarkable finding. A potential explanation arises when the crude oil price and soybean production are analysed graphically. Where crude oil prices fluctuate, the supply of soybeans tends to follow an upward trend. This is confirmed when looking at the $trend$ coefficient, which shows that the linear trend included in this equation is significant. This indicates a general exogenous growth in the soybean production.

The estimation results for the soy oil demand function show that a 1% increase in Chinese GDP per capita, leads to a 0.25% increase in global soy oil demand. This is an indication for China's dominance in the global market for soy oil. Similar results about the impact of China's demand on the global soy oil market has been found by Jenkins (2011). Next, the parameter for biodiesel production shows that a 1% increase leads to 0.11% more demand for soy oil. This result confirms that the soy oil demand for Europe and the US is significantly determined by their biodiesel production. This result is backed by literature, as an increasing demand for vegetable oils due to biodiesel production is a well-established relationship (Tomei & Helliwell, 2016; Hirani et al., 2018). The price coefficient for soy oil appears to be non-significant, which is a counter-intuitive finding given the strong price elasticity suggested in literature (Santeramo, 2017) and standard micro-economic theory. Besides, an indirect relationship between the soy oil price and soy oil demand is likely to exist. The fixed crush ratio between beans and soy oil and soybean meal causes a relationship where the price is indirectly affecting the quantity demanded. The second analysis, in which the relationship between the three price variables is examined, will shine more light on this issue, as it shows whether $lrpo_t$ indeed affects $lrpb_t$. All in all, the parameter's insignificance is remarkable, given the soy oil price variable's direct and indirect connection to the soy oil demand.

Estimating the equation for the soybean meal demand shows that a 1% increase in meat production leads to a 0.96% increase in soybean meal demand. This is an intuitive result and in line with literature. A similar result is found by Masuda and Goldsmith (2012), who state that China's soybean meal demand increases with 0.91% in case of a 1% increase in meat and egg consumption. Next, the positive relationship between the price and demand for soybean meal is counterintuitive. It is potentially an indication of a disequilibrium between supply and demand, which means that total production is a wrong proxy for demand. As a result, an increase in the price of soybean meal leads to more soybeans being crushed, and therefore an increase of total soybean meal production. This matter will be elaborately addressed in chapter 7. The negative relationship between the meat and bonemeal price and the soybean meal demand seems counter-intuitive, because lower prices for one product generally lead to lower demand for the other in case of substitutes. The negative relationship is most likely the result of the BSE-crisis, which resulted in low meat and bonemeal prices and a higher demand for soybean meal. Therefore, the parameter should not be interpreted as a causal relationship, in which a decreasing meat and bonemeal price causes demand for soybean meal to rise. A significant effect of the BSE-crisis on the demand for soybean meal is a finding backed by literature (Babula et al., 2004).

Post-estimation tests

The ‘Extremes’ Stata software, developed by Cox (2003), is used to identify outliers in the data. The datasets used contain no significant outliers that could disturb the results. Next, post-estimations tests prove the absence of first- and higher order serial autocorrelation. All equations are also tested for the presence of a structural break, for which no evidence is found.

6.2. Price Cointegration Analysis

This section aims on quantifying the relationship among the price variables in the soybean complex, theoretically established in section 3.5. The analysis is complementary to the estimation of the structural model in the previous section, as it shines more light on the interaction between the different price variables.

The analysis starts with an ADF test with only an intercept, and with both an intercept and trend (table B1). An ADF test with trend is included because figure 5 does not give a decisive answer on the absence of a time trend in the price variables. From the test results, it is concluded that non-stationarity cannot be rejected on a 5% critical level for the original series, contrary to first differenced variables. The soy price data is non-stationary, but the price changes are stationary. Therefore, the data is $\mathrm{I}(1)$ and the analysis continues with differenced data.

Before estimating the VECM model to test for cointegration, some assumptions about the inclusion of a trend and constant have to be made. Based on theory, it is likely for the cointegrating equations to be stationary around constant means. Besides this, no linear time trends appear to exist in the level values of the soybean, meal, and oil prices (figure 5). Therefore, only a constant is included in the cointegration part of the VECM model.

The information criteria, do not give a decisive indication for the optimal lag length (table B2). Where LR, FPE, and AIC suggest 3 lags, HQIC and SBIC indicate an optimal lag length of 2. Based on the growth cycle of soybean’s, which is between a 100 and 130 days (FAO, 2023), the optimal lag length is determined to be 3.

The test used to establish the presence of cointegration is the Johansen test . The Johansen test results indicate a maximum rank of 1 (table B3). The suggested presence of one cointegrating relationship supports the use of a VECM over a VAR to test for Granger Causality.

With the optimal number of lags and the rank established, the VECM parameters are estimated. The following error correction term (ECT) results from the estimation:

$$ECT = lpb_t - 0.36lpo_t - 0.54lpm_t - 0.39 \qquad (20)$$

The ECT represents the long-run cointegrating relationship among the price variables. The adjustment speed of the variables in case of a disequilibrium in the cointegrating relationship is captured in the adjustment parameters, which are displayed in table 7:

Table 7: Estimated parameters of the adjustment parameters

	Adjustment parameters		
	lpb_t	lpo_t	lpm_t
λ_i	-0.19 (0.04)***	-0.07 (0.05)	-0.01 (0.05)

The results show that only the soybean prices adjust to deviations in the long-run equilibrium between beans, meal and oil. The other two adjustment parameters are not significantly different from zero.

The VECM is completed with the estimation and interpretation of the short-run dynamics, for which the parameters are displayed in Table 8:

Table 8: Estimated parameters of the short-run dynamics

	Short-run dynamics		
	Δlpb_t	Δlpo_t	Δlpm_t
Adj. par.			
Δlpb_{t-1}	0.10 (0.07)	0.10 (0.08)	0.12 (0.08)
Δlpb_{t-2}	0.02 (0.07)	0.14 (0.08)*	0.04 (0.08)
Δlpo_{t-1}	0.14 (0.05)***	0.30 (0.06)***	-0.04 (0.05)
Δlpo_{t-2}	-0.05 (0.05)	-0.23 (0.06)***	0.01 (0.05)
Δlpm_{t-1}	0.05 (0.06)	-0.04 (0.06)	0.27 (0.06)***
Δlpm_{t-2}	0.05 (0.06)	0.00 (0.06)	-0.10 (0.06)*

The first notable finding from the short-run parameters is that changes in the soybean price are only significantly explained by the first lag values of the soy oil price and not even by its own lags. Second, the soy oil price is determined by its own lagged values, and is not significantly influenced by the soybean or soybean meal prices. The same holds for the soybean meal price level, which is not influenced by, and is not influencing, the price levels of the other two commodities.

Granger causality

VEC models allow for distinguishing between short- and long-run causal relationships. The results in table 9 show that the soy oil price Granger causes the soybean price on the short-run. This finding is confirmed by the estimation results of the short-run dynamics, where the lagged value of the soy oil price significantly explains changes in the soybean price level.

Table 9: Results of the Wald test for short-run Granger causality

	Wald test		
	lpb_t	lpo_t	lpm_t
lpb_{t-1}		0.574	0.108
lpo_{t-1}	0.005***		0.430
lpm_{t-1}	0.156	0.900	

Long-run Granger causality is defined as the significance of the error correction term's coefficient, i.e. the adjustment parameter. The results in table 7 indicate that the soybean price level is Granger caused by the other two on the long-run.

A variable is said to strongly Granger cause another variable, if both the short- and long-run Granger causality relationships are significant. Therefore, the VECM estimation results suggest that the price of soybeans (lpb_t) is strongly Granger caused by the soy oil price (lpo_t).

Batista et al. (2023) performed a relationship analysis on the future prices of soybeans, soy oil, and soybean meal. This study showed that there is a Granger relationship between the future prices of soybeans and both soy oil and soybean meal. The authors note however that this relationship is much stronger for soy oil. This makes the results found by Batista et al. (2023) in line with the results found in this empirical study.

VECM diagnostics

Post estimation tests are done to make sure that there is no autocorrelation in the residuals and that the VECM is stable. The H_0 of no autocorrelation is not rejected for all 3 lags (table B4), indicating the absence of autocorrelation among the residuals, and confirming the legitimacy of the findings.

7. Conclusion and Critical Reflection

7.1. Conclusions

Chapter 1 introduced the main research question, which was formulated as follow: “What are the main characteristics of the global soy complex and what are its current developments?” This chapter forms a direct answer to this question and in the process answers the sub research questions that were formulated.

From diving into the sub question “What are the main trade flows in the global soybean complex?”, it became clear that the global trade flows go from the US, Brazil, and Argentina (major producers) to China and Europe (major consumers). The most notable findings relates to the role of China in the global soy market, by far the biggest importer. China’s dominance in the global soybean market was confirmed in the structural model analysis, where Chinese GDP growth per capita appeared to be a significant contributor to global demand for soy oil. This makes involvement of the country in sustainability practices crucial. The various forms of supply chain governance, described in section 2.4, mainly apply to European soy imports and therefore leave the majority of soybean production open to unsustainable practices like deforestation and unfair wages.

In answering the sub question “What is the current role of the agricultural industry in the global soybean complex?”, the agricultural industry is identified as a main pillar for the soybean meal demand. This notion was confirmed in the structural model, which proved meat consumption to be a major contributor to soybean meal demand. Soybean meal’s role as the dominant protein component in animal feed is not without risk. This research found that a shock in an external market, in this case the BSE-crisis, has a significant impact on the soybean meal demand. Future shocks can lead to demand and price peaks in the global soybean market, putting global food security at risk. This makes investments in viable alternatives for soybean meal needed.

The third sub question, formulated as: “How do biofuel policies and therefore the increasing demand for soy oil affect the soybean complex?”, related to the role of biodiesel production in the demand for soy oil. The increase in biodiesel production in the last two decades, mainly due to biofuel policies in the EU and the US, had a significant impact on the demand for soy oil. Given soy oil’s role as food, feed, and fuel input, it is important for policymakers to hold this finding into account when creating biofuel related policies. With the relationship between biodiesel production and soy oil demand confirmed, and given the importance of the food and feed component of soy oil for global food security, it is important for policy makers to thoroughly investigate the effects that future biofuel policies have on the global soy oil market.

The fourth sub question is formulated as follow: “Is soybean meal or soy oil the main driver for the global demand for soybeans?”. Modelling the price variables in a VECM framework proved the central role of soy oil in the global soybean market. The soy oil price appears to be significantly Granger causing the soybeans price, contrary to the soybean meal price. This is the result of soy oil prices being connected to the world economy due to its increasing use as biodiesel input. This makes soy oil a substitute for crude oil, connecting soy oil prices to global oil prices (Batista et al., 2023). Besides, Batista et al. (2023) state that if demand for soy oil increases, the quantity of soybeans crushed increases disproportionately due to the soybean’s relatively lower oil content. As a result, changes in soy oil demand have a greater impact on the soybean market, compared to soybean meal.

7.2. Critical Reflection

In this chapter, limitations of the research and some potential improvements are listed. This chapter forms a starting point for future researchers diving into the global soy market.

First, data availability formed a limitation in this research. Significant data was available for a broad range of variables. However, for a handful of variables, most notably the demand for soybean meal, limited data was available, which shortened the model's data range to 26 annual observations. Despite retrieving numerous significant coefficients, the conclusions of this research can be strengthened if the model is estimated over a longer time period or with a higher data frequency.

Second, some questionable assumptions about the equilibrium between supply and demand in the soybean complex have been made. As mentioned earlier, not all soybeans are crushed into meal and oil. A small but significant amount (7%) is directly consumed as either food or feed. In this research, total demand for soy oil and soybean meal are assumed equal to the total supply of soybeans. This assumption might have led to flawed results. Therefore, future researchers are recommended to include a fourth equation, denoting demand for whole soybeans, to create a system that more accurately reflects the equilibrium in the soybean complex. Besides, it allows the model to capture the shift in the Western diet from meat-based to plant-based proteins, as the latter is often made from whole soybeans. This change in consumption pattern is currently limited but is projected to become more important, which makes it a useful and interesting factor to include.

Third, a critical reflection on the proxies used is necessary for improving future research. Because an equilibrium situation between supply and demand was assumed in this research, total production of soy oil and soybean meal were used as proxies for demand. However, total production is not necessarily equal to total demand in the same year. This is illustrated with a hypothetical situation in which high prices exist in year *t*. As the soybean supply function suggested, this leads to more production in year *t+1*. This causes over-production if demand falls or if production increases too rapidly. Over-production of oil or meal leads to storage of these commodities, creating surpluses that are consumed in future years when demand exceeds supply. The storage component, which corrects for the disequilibrium between production and consumption, would improve the structural model developed in this research. Therefore, future models of the soybean complex should incorporate annual net storage, in order for the model to get an accurate proxy of demand.

Fourth, the structural model could be further improved by integrating the long-run cointegration variables, found in the second analysis. The price cointegration analysis made clear that the price variables in the soybean complex are related. In the structural model, explicitly incorporating this relationship improves the model's power to explain changes in demand and supply. It is acknowledged that this is complicated, given the different data frequencies used for both analyses. However, a variety of techniques exist to handle mixed-frequency data (Foroni & Marcellino, 2013), allowing future researchers to link both analyses.

Future research in this field could focus on improving the structural model, based on the limitations as presented above. Despite the importance of structural models for policymakers and researchers alike, construction of these models has become a thing of the past. With most models dating back to the early 2000s (Westcott & Hoffman, 1999; Babula et al., 2004; Goodwin et al., 2005). This research, and the follow-up studies that follow, are of great importance for the literature on agricultural commodities and soy in particular.

References

Ali, N. (2010). Soybean processing and utilization. In Singh, G., *The Soybean: Botany, Production and Uses*, 345–374. https://doi.org/10.1079/9781845936440.0345

Babula, R. A., Bessler, D. A., Reeder, J. C., & Somwaru, A. (2004). Modeling U.S. Soy-Based Markets with Directed Acyclic Graphs and Bernanke Structural VAR Methods: The Impacts of High Soy Meal and Soybean Prices. *Journal of Food Distribution Research, 35*(3), 29–52. https://doi.org/10.22004/ag.econ.27559

Barona, E., Ramankutty, N., Hyman, G., & Coomes, O. T. (2010). The role of pasture and soybean in deforestation of the Brazilian Amazon. *Environmental Research Letters, 5*(2), 024002. https://doi.org/10.1088/1748-9326/5/2/024002

Batista, G., Brum, A. L., Baggio, D. K., & Souza, F. M. (2023). The establishment of soybean price and its interrelation with the prices of its derivatives: soybean meal and soybean oil. *Ciência e Natura, 44*, e56. https://doi.org/10.5902/2179460x68858

Baum, C. F., & Hurn, S. (2021). *Environmental Econometrics Using Stata.* Stata Press.

Bennet, L. (2017). *Deforestation and Climate Change.* The Climate Institute. Retrieved from http://climate.org/deforestation-and-climate-change/

Blank, S. C. (2015). Chapter 8: Economic Development and the Irony of American Agriculture's Technical Progress. In *The Economics of American Agriculture: Evolution and Global Development.* Routledge.

Boyd, J. H. (2023). Chapter 13: Production Sets and Profit. In *Microeconomics of Competitive Markets.* Department of Economics, Florida International University, Miami.

Burgers, R. (2019, April). Veehouder niet verantwoordelijk kap regenwoud door verhoogde import Amerikaanse soja. *Pig Business*. Accessed on 12-6-2023 from: https://www.pigbusiness.nl/artikel/193833-geen-kaalslag-regenwoud-door-verhoogde-import-amerikaanse-soja/

CBS. (2021, December). Hoeveel landbouwdieren telt ons land? In *Nederland in cijfers 2021*. Retrieved from https://longreads.cbs.nl/nederland-in-cijfers-2021/hoeveel-landbouwdieren-telt-ons-land/

Cheng, M. & Rosentrater, K. (2019). Techno-Economic Analysis of Extruding-Expelling of Soybeans to Produce Oil and Meal. *Agriculture, 9*(5), 87. https://doi.org/10.3390/agriculture9050087

Cox, N. J. (2003). EXTREMES: Stata module to list extreme values of a variable. *Statistical Software Components* S430801, Boston College Department of Economics, revised 16 Jun 2017.

De Ridder, M., Ten Haaf, P., Olah, N. & Bolscher, H. (2015). Soy Supply Security for the Netherlands. Anticipating Future Global Challenges through Strategic Responses. *Centre of Expertise on Resources.* Retrieved from https://trinomics.eu/wp-content/uploads/2018/07/Soy-supply-security-for-the-Netherlands.pdf

FAO. (2023). *FAOSTAT* [Dataset]. Accessed on 16-1-2023 from https://www.fao.org/faostat/en/#data/QCL/visualize

FEFAC. (2021, February). *Soy Sourcing Guidelines 2021. Towards a Mainstream Market Transition for Responsible Soy.* Retrieved from https://fefac.eu/resources/professional-guidelines/

Fehlenberg, V., Baumann, M., Gasparri, N. I., Piquer-Rodriguez, M., Gavier-Pizarro, G. & Kuemmerle, T. (2017). The role of soybean production as an underlying driver of deforestation in the South American Chaco. *Global Environmental Change, 45*, 24–34. https://doi.org/10.1016/j.gloenvcha.2017.05.001

Foroni, C., & Marcellino, M. (2013). A Survey of Econometric Methods for Mixed-Frequency Data. *Social Science Research Network*. https://doi.org/10.2139/ssrn.2268912

Fraanje, W., & Garnett, T. (2020). Soy: food, feed, and land use change. *Food Climate Research Network*, University of Oxford. https://doi.org/10.56661/47e58c32

Gehrke, E., Peerlings, J. (2022). *Economics of Agribusiness.* Agricultural Economics and Rural Policy group, Wageningen University. 203p.

Goldsmith, P. D., & Hirsch, R. (2006). The Brazilian Soybean Complex. *Choices. The Magazine of Food, Farm, and Resources Issues, 21*(2), 1–8. Retrieved from https://www.researchgate.net/publication/227364759_The_Brazilian_Soybean_Complex

Gollnow, F., Hissa, L. D. B. V., Rufin, P. & Lakes, T. (2018). Property-level direct and indirect deforestation for soybean production in the Amazon region of Mato Grosso, Brazil. *Land Use Policy, 78,* 377–385. https://doi.org/10.1016/j.landusepol.2018.07.010

Goodwin, B. K., Schnepf, R., & Dohlman, E. (2005). Modelling soybean prices in a changing policy environment. *Applied Economics, 37*(3), 253–263. https://doi.org/10.1080/00036840412331315060

Hart, C. (2017). The Economic Evolution of the Soybean Industry. *The Soybean Genome, 1*–9. https://doi.org/10.1007/978-3-319-64198-0_1

Hartman, G., Pawlowski, M., Herman, T., & Eastburn, D. (2016). Organically Grown Soybean Production in the USA: Constraints and Management of Pathogens and Insect Pests. *Agronomy, 6*(1), 16. https://doi.org/10.3390/agronomy6010016

Heron, T., Prado, P., & West, C. (2018). Global Value Chains and the Governance of 'Embedded' Food Commodities: The Case of Soy. *Global Policy, 9,* 29–37. https://doi.org/10.1111/1758-5899.12611

Hirani, A. H., Javed, N., Asif, M., Basu, S. K., & Kumar, A. (2018). A Review on First- and Second-Generation Biofuel Productions. *Biofuels: Greenhouse Gas Mitigation and Global Warming*, 141–154. https://doi.org/10.1007/978-81-322-3763-1_8

Hoste, R. (2014). *Sojaverbruik in de Nederlandse diervoederindustrie 2011-2013; Inventarisatie in opdracht van Stichting Ketentransitie verantwoorde soja.* Wageningen, LEI Wageningen UR (University & Research centre). Retrieved from https://edepot.wur.nl/316027

Houghton, R. A. (1990). The global effects of tropical deforestation. *Environmental Science & Technology, 24*(4), 414–422. https://doi.org/10.1021/es00074a001

IEA. (2023). *Policy database - Data & Statistics.* Accessed on 4-4-2023 from https://prod.iea.org/policies?sector%5B0%5D=Biofuel%20production&page=1

IMF. (2022). *IMF Data* [Dataset]. Accessed on 23-1-2023 from https://www.imf.org/en/Data

IndexMundi. (2023). *World Biodiesel Consumption by Year (Thousand Barrels per Day)* [Dataset]. Accessed on 27-1-2023 from https://www.indexmundi.com/energy/?product=biodiesel.

Iqbal, Z., & Babcock, B. A. (2018). Global growing-area elasticities of key agricultural crops estimated using dynamic heterogeneous panel methods. *Agricultural Economics, 49*(6), 681–690. https://doi.org/10.1111/agec.12452

Jenkins, R. (2011). The “China effect” on commodity prices and Latin American export earnings. *Cepal Review, 2011(103),* 73–87. https://doi.org/10.18356/ceb8a981-en

Jia, F., Peng, S., Green, J., Koh, L., & Chen, X. (2020). Soybean supply chain management and sustainability: A systematic literature review. *Journal of Cleaner Production, 255,* 120254. https://doi.org/10.1016/j.jclepro.2020.120254

Jukema, G., Ramaekers, P. & Berkhout, P. (2022). *De Nederlandse agrarische sector in internationaal verband – editie 2022.* Wageningen Economic Research en Centraal Bureau voor de Statistiek, Rapport 2021-001. 124 blz.; 40 fig.; 38 tab.; 129 ref. https://doi.org/10.18174/561610

Katan, M. (2022). Minder boeren? Eén ding is zeker: eten blijft er genoeg. *NRC Handelsblad.* Accessed on 10-1-2023 from https://www.nrc.nl/nieuws/2022/09/16/er-blijven-genoeg-karbonades-over-a4141880

Kim, H., & Heo, E. (2012). *Causality between main product and byproduct prices of metals used for thin-film PV cells.* In Proceedings of the 3rd IAEE Asian Conference, Kyoto, Japan.

Kim, S. W., Less, J. F., Wang, L., Yan, T., Kiron, V., Kaushik, S. J., & Lei, X. G. (2019). Meeting Global Feed Protein Demand: Challenge, Opportunity, and Strategy. *Annual Review of Animal Biosciences, 7*(1), 221–243. https://doi.org/10.1146/annurev-animal-030117-014838

Kim, H., & Moschini, G. (2018). The Dynamics of Supply: U.S. Corn and Soybeans in the Biofuel Era. *Land Economics, 94*(4), 593–613. https://doi.org/10.3368/le.94.4.593

Krugman, P., & Wells, R. (2013). Chapter 3: Supply and Demand. In *Microeconomics* (third edition). Worth Publishers.

Livingston, M., Roberts, M. S., & Rust, J. (2008). *Optimal Corn and Soybean Rotations.* 2008 Annual Meeting, July 27-29, 2008, Orlando, Florida. https://ideas.repec.org/p/ags/aaea08/6213.html

Margarido, M. A., Turolla, F. A., & Júnior, C. R. B. (2007). The world market for soybeans: price transmission into Brazil and effects from the timing of crop and trade. *Nova Economia, 17*(2), 241–270. https://doi.org/10.1590/s0103-63512007000200002

Masuda, T., & Goldsmith, P. D. (2009). World Soybean Production: Area Harvested, Yield, and Long-Term Projections. *The International Food and Agribusiness Management Review, 12*(4), 143–161. https://doi.org/10.22004/ag.econ.92573

Masuda, T., & Goldsmith, P. D. (2012). China’s Meat and Egg Production and Soybean Meal Demand for Feed: An Elasticity Analysis and Long-Term Projections. *The International Food and Agribusiness Management Review, 15*(3), 35–53. https://doi.org/10.22004/ag.econ.132788

Meade, B., Puricelli, E., McBride, W. D., Valdes, C., Hoffman, L. A., Foreman, L. F., & Dohlman, E. (2016). Corn and Soybean Production Costs and Export Competitiveness in Argentina, Brazil, and the United States. *Economic Information Bulletin.* https://doi.org/10.22004/ag.econ.262143

Milanesi, J. (2012). Current and Future Availability of Non-Genetically Modified Soybean Seeds in the USA, Brazil and Argentina. *Genetically Modified and Non-Genetically Modified Food Supply Chains: Co-Existence and Traceability,* 89–112. https://doi.org/10.1002/9781118373781.ch7

Montanía, C. V., Fernández-Núñez, T., & Márquez, M. A. (2021). The role of the leading exporters in the global soybean trade. *Agricultural Economics*, *67*(No. 7), 277–285. https://doi.org/10.17221/433/2020-agricecon

Moschini, G., Lapan, H. E., & Sobolevsky, A. (2000). Roundup ready soybeans and welfare effects in the soybean complex. *Agribusiness, 16*(1), 33–55. https://doi.org/10.1002/(sici)1520-6297(200024)16:1

Moss, C. B., Livanis, G., & Schmitz, A. (2010). The Effect of Increased Energy Prices on Agriculture: A Differential Supply Approach. *Journal of Agricultural and Applied Economics, 42*(4), 711–718. https://doi.org/10.1017/s1074070800003904

Nevedi. (2019, May). *Grondstoffenwijzer* (3th edition). Retrieved from https://nevedi.nl/publicaties/

Pala, A. (2013). Structural Breaks, Cointegration, and Causality by VECM Analysis of Crude Oil and Food Price. *International Journal of Energy Economics and Policy, 3*(3), 238–246. Retrieved from https://dergipark.org.tr/tr/download/article-file/361230

Pennings, J. M., & M. Leuthold, R. (2001). Introducing new futures contracts: reinforcement versus cannibalism. *Journal of International Money and Finance, 20*(5), 659–675. https://doi.org/10.1016/s0261-5606(01)00013-4

Pimentel, D., & Patzek, T. W. (2005). Ethanol Production Using Corn, Switchgrass, and Wood; Biodiesel Production Using Soybean and Sunflower. *Natural resources research, 14*(1), 65–76. https://doi.org/10.1007/s11053-005-4679-8

Regunaga, M. (2010). *Implications of the Organization of the Commodity Production and Processing Industry : The Soybean Chain in Argentina.* LCSSD occasional paper series on food prices. World Bank, Washington, DC. Retrieved from http://hdl.handle.net/10986/18710

Ritchie, H., Rosado, P. & Roser, M. (2019). *Meat and Dairy Production.* Our World in Data. Retrieved from https://ourworldindata.org/meat-production

Ritchie, H. and Roser, M. (2021). Forests and Deforestation - Soy. Published online at *OurWorldInData.org*. Retrieved from https://ourworldindata.org/forests-and-deforestation [Online Resource]

Rizzo, G., & Baroni, L. (2018). Soy, Soy Foods and Their Role in Vegetarian Diets. *Nutrients, 10*(1), 43. https://doi.org/10.3390/nu10010043

Rodionova, M., Poudyal, R., Tiwari, I., Voloshin, R., Zharmukhamedov, S., Nam, H., Zayadan, B., Bruce, B., Hou, H. & Allakhverdiev, S. (2017). Biofuel production: Challenges and opportunities. *International Journal of Hydrogen Energy, 42*(12), 8450–8461. https://doi.org/10.1016/j.ijhydene.2016.11.125

Santeramo, F. (2017). *Cross-Price Elasticities for Oils and Fats in the US and the EU.* Consultant report. Retrieved from https://theicct.org/publication/cross-price-elasticities-for-oils-and-fats-in-the-u-s-and-the-eu/

Shao, L., Hu, W., & Yang, D. (2020). The price relationship between main-byproduct metals from a multiscale nonlinear Granger causality perspective. *Resources Policy, 69,* 101846. https://doi.org/10.1016/j.resourpol.2020.101846

Shukla, J., Nobre, C., & Sellers, P. (1990). Amazon Deforestation and Climate Change. *Science, 247*(4948), 1322–1325. https://doi.org/10.1126/science.247.4948.1322

Statista. (2023). *Production of soybean meal worldwide from 2013/14 to 2022/23, by country* [Dataset]. Accessed on 2-3-2023 from https://www.statista.com/statistics/612537/soybean-meal-production-worldwide-by-country/

Stock, J. H., & Watson, M. W. (2001). Vector Autoregressions. *Journal of Economic Perspectives, 15*(4), 101–115. https://doi.org/10.1257/jep.15.4.101

Sorda, G., Banse, M., & Kemfert, C. (2010). An overview of biofuel policies across the world. *Energy Policy, 38*(11), 6977–6988. https://doi.org/10.1016/j.enpol.2010.06.066

Tilman, D., & Clark, M. (2014). Global diets link environmental sustainability and human health. *Nature, 515*(7528), 518–522. https://doi.org/10.1038/nature13959

Tomei, J., & Helliwell, R. (2016). Food versus fuel? Going beyond biofuels. *Land Use Policy, 56,* 320–326. https://doi.org/10.1016/j.landusepol.2015.11.015

UNdata. (2023). *Biodiesel* [Dataset]. Accessed on 2-3-2023 from http://data.un.org/Data.aspx?d=EDATA&f=cmID%3ABD#EDATA

U.S. Bureau of Labor Statistics. (2023). Producer *Price Index by Industry: Rendering and Meat Byproduct Processing: Meat and Bone Meal* [Dataset]. Retrieved from FRED, Federal Reserve Bank of St. Louis. Accessed on 29-3-2023 from https://fred.stlouisfed.org/series/PCU31161331161351

USDA. (2020). Argentina: New Export Tax Rates in Argentina. Retrieved from https://www.fas.usda.gov/data/argentina-new-export-tax-rates-argentina

USDA. (2023). *Oil Crops Yearbook* [Dataset]. Accessed on 27-4-2023 from https://usda.library.cornell.edu/concern/publications/5x21tf41f?locale=en&page=2#release-items

Van Berkum, S., Roza, P. & Pronk, B. (2006). *Sojahandel- en ketenrelaties. Sojaketens in Brazilië, Argentinië en Nederland.* LEI Wageningen. Retrieved from https://library.wur.nl/WebQuery/wurpubs/352665

Vries, J. de. & Smit, P. H. (2021, April). Eet de Nederlandse koe gaten in het Braziliaanse Amazonegebied? *De Volkskrant.* Accessed on 10-1-2023 from https://www.volkskrant.nl/nieuws-achtergrond/eet-de-nederlandse-koe-gaten-in-het-braziliaanse-amazonegebied~b6f3f1a7/

Wang, S. L., & McPhail, L. (2014). Impacts of energy shocks on US agricultural productivity growth and commodity prices—A structural VAR analysis. *Energy Economics, 46,* 435–444. https://doi.org/10.1016/j.eneco.2014.05.006

Wang, W., & Wei, L. (2021). Impacts of agricultural price support policy on price variability and welfare: Evidence from China’s soybean market. *Agricultural Economics, 52*(1), 3–17. https://doi.org/10.1111/agec.12603

Werth, D. & Avissar, R. (2002). The local and global effects of Amazon deforestation. *Journal of Geophysical Research, 107*(D20). https://doi.org/10.1029/2001jd000717

Westcott, P. C., & Hoffman, L. A. (1999). Price Determination for Corn and Wheat: The Role of Market Factors and Government Programs. *Technical Bulletins.* https://doi.org/10.22004/ag.econ.33581

Willis, S. (2003). *The use of Soybean Meal and Full Fat Soybean Meal by the Animal Feed Industry*. 12th Australian Soybean Conference. Retrieved from https://www.researchgate.net/publication/267242105_The_use_of_Soybean_Meal_and_Full_Fat_Soybean_Meal_by_the_Animal_Feed_Industry

World Bank. (2023a). *World Bank Monthly Commodity Price Data (The Pink Sheet)* [Dataset]. Accessed on 23-1-2023 from https://www.worldbank.org/en/research/commodity-markets

World Bank. (2023b). *World Bank Annual Commodity Price Data (The Pink Sheet)* [Dataset]. Accessed on 15-2-2023 from https://www.worldbank.org/en/research/commodity-markets

World Bank. (2023c). *GDP per capita (constant 2015 US$)* [Dataset]. Accessed on 15-2-2023 from https://data.worldbank.org/indicator/NY.GDP.PCAP.KD

World Bank. (2023d). *Agricultural land (% of land area)* [Dataset]. Accessed on 15-2-2023 from https://data.worldbank.org/indicator/AG.LND.AGRI.ZS?end=2020&start=1961&view=chart

WWF. (2021). *Stepping Up? The continuing impact of EU consumption on nature worldwide.* https://www.wwf.nl/globalassets/pdf/stepping-up-the-continuing-impact-of-eu-consumption-on-nature-worldwide.pdf

Appendices

Appendix A

Table A1: Data overview for the IV's used in the 2SLS estimation

Variable	Abbreviation	Definition	Annual time period	Unit
Palm oil price	$palm$	Global average price level of palm oil	1995 - 2020	Nominal US dollar per metric ton ($/mt)
Sorghum price	$sorghum$	Global average price level of sorghum	1995 - 2020	Nominal US dollar per metric ton ($/mt)

Note: All data is retrieved from World Bank (2023a).

Table A2: Descriptive statistics for the price variables of palm oil and sorghum

Variable	Abbr.	Obs.	Mean	Std. Dev.	Min	Max
Palm oil price	$palm$	26	668.31	229.32	287.46	1193.37
Sorghum price	$sorghum$	26	153.26	54.59	84.39	271.93

Appendix B

Table B1: Results of the ADF test (price cointegration analysis):

	Variable	Test statistic
Intercept	pb_t	-1.077
	po_t	-1.029
	pm_t	-1.097
Intercept and trend	pb_t	-2.337
	po_t	-2.408
	pm_t	-2.571
Intercept	Δpb_t	-18.473 ***
	Δpo_t	-16.458***
	Δpm_t	-16.470***
Intercept and trend	Δpb_t	-18.476***
	Δpo_t	-16.460***
	Δpm_t	-16.472***

Note: *, **, and *** denote the critical significance levels of 10%, 5%, and 1% respectively. Δ means that the variable is differenced.

Table B2: Information criteria for optimal lag length (price cointegration analysis):

Lag	**LR**	**FPE**	**AIC**	**HQIC**	**SBIC**
0		.000023	-2.168	-2.158	-2.143
1	4056.7	8.6e-09	-10.056	-10.017	-9.957
2	110.2	7.2e-09	-10.236	-10.168*	-10.062*
3	29.7*	7.0e-09*	-10.259*	-10.162	-10.011
4	12	7.1e-09	-10.247	-10.121	-9.924

Table B3: Johansen cointegration test (price cointegration analysis):

H0	**Trace test**		**Maximum eigenvalue test**	
Maximum rank	Statistic	5% critical Value	Statistic	Critical value
0	53.894	34.91	38.313	22.00
1	15.581*	19.96	13.748*	15.67
2	1.833	9.42	1.833	9.24

Table B4: VECM diagnostics test results of residuals (price cointegration analysis):

Autocorrelation (H0: no residual autocorrelation)			
Lagrange-multiplier test	*Lag*	*Test statistic*	*P-value*
	1	6.851	0.653
	2	7.924	0.542
	3	10.582	0.305