CHARLES UNIVERSITY

FACULTY OF SOCIAL SCIENCES

Institute of Economic Studies

Karolína Gajdušková

Identifying driving factors of coffee prices

Bachelor thesis

Prague 2020

Author: Karolína Gajdušková
Supervisor: Petr Pleticha, MSc.

Academic Year: 2019/2020

Bibliographic note

GAJDUŠKOVÁ, Karolína. *Identifying Driving Factors of Coffee Prices.* Prague 2020. 69 p. Bachelor thesis. Charles University, Faculty of Social Sciences, Institute of Economic Studies, Supervisor Petr Pleticha, MSc.

Abstract

This bachelor thesis focuses on the price of coffee and the factors affecting it. The study aims to identify these factors to predict price movement. This thesis also answers the question whether the factors observed have the same impact on two coffee varieties: arabica and robusta. There were defined five groups of the factors: supply, demand, climate, commodity market and financial. The effect is observed for annual coffee price data that were obtained from the International Coffee Organization (ICO). The observed period is between 1965 and 2018. The impact of the factors was detected by the ARDL regression, both in the short run and the long run. Based on the results, the price is affected mostly by supply, demand, and financial factors. Moreover, robusta's price specifically is driven mainly by the factors explaining the power of the US dollar. Arabica's price, on the other hand, is affected by the supply and demand factors: exports, and the growth of GDP. This finding supports the claim that these two coffee varieties should be considered as two different commodities. The results suggest that the establishment of an association of the biggest coffee producers regulating coffee exports would help in control of coffee prices.

Keywords

Coffee price, Arabica, Robusta, ARDL method, ECM model

Range of thesis: 69 400 characters with spaces

Abstrakt

Tato bakalářská práce se zabývá cenou kávy a faktory, které ji ovlivňují. Cílem práce bylo tyto faktory odhalit a zjistit, jak se bude cena kávy na základě jejich změn vyvíjet. Práce se také zabývá otázkou, zda zkoumané faktory působí stejně na dva druhy kávy: arabicu a robustu. Celkem bylo definováno pět skupin faktorů: nabídka, poptávka, klimatické změny, komoditní trh a finanční aspekty. Výsledný efekt byl zkoumán na ročních datech ceny kávy, která poskytla organizace International Coffee Organization (ICO). Zkoumané období bylo mezi lety 1965 až 2018. Regrese ARDL odhalila, že faktory ovlivňují cenu kávy jak krátkodobě, tak i dlouhodobě. Výsledky ukázaly, že na danou proměnnou mají největší vliv faktory zachycující poptávku, nabídku a stabilitu amerického dolaru. Kromě toho se potvrdilo, že arabica a robusta by měly být považovány za samostatné komodity, neboť na ně dané faktory mají jiný vliv. Cena robusty je ovlivněna především finančními faktory, zatímco arabicu nejvíce ovlivňují exporty a růst HDP. Jedním z opatření, které by mohlo pomoci s kontrolou ceny kávy, je založení asociace největších světových producentů kávy, která by regulovala exporty, s jejichž nárustem se snižuje cena.

Klíčová slova

Cena kávy, Arabica, Robusta, metoda ARDL, model ECM

Rozsah práce: 69 400 znaků včetně mezer

Declaration of Authorship

1. The author hereby declares that he compiled this thesis independently, using only the listed resources and literature.

2. The author hereby declares that all the sources and literature used have been properly cited.

3. The author hereby declares that the thesis has not been used to obtain a different or the same degree.

Prague, 29 July 2020 Karolína Gajdušková

Acknowledgments

I would like to express my gratitude to my supervisor Petr Pleticha for his advice, support and patience when discussing this thesis.

I place on record my sincere thanks to the International Coffee Organization, which provided me with non-public historical data regarding coffee.

I am also grateful to all my family and friends for the moral support without which I would not be able to finish this thesis. Finally, I would like to thank all who were involved in getting my daily doses of coffee.

Bachelor's Thesis Proposal

Institute of Economic Studies
Faculty of Social Sciences
Charles University in Prague

Author's name and surname: Karolína Gajdušková

Supervisor's name: Petr Pleticha, MSc.

Proposed Topic:

Identifying driving factors of coffee prices

Preliminary scope of work:

Research question and motivation

The main research question I would like to study and answer is which factors influence coffee prices and what is their estimated relation to the prices.

More than 125 million of people are employed in the coffee industry (coffee farms give livelihood to 25 million of people) and 2,25 billion of coffee cups are consumed around the world every day (Farah, 2019). These numbers indicate that coffee is an important commodity for the whole world's welfare because it provides plenty of jobs and opportunities to earn money.

Coffee as a commodity is important not only for developing countries (such as Ethiopia) but also for developed countries, where the coffee market is significant. For example, in the United States the economic coffee activity made up 1.6 percent of its GDP (Menke; 2018). The fact that coffee consumption increases every year (ICO; 2019) could be for many reasons (the worldwide population increases, drinking coffee is more popular, etc.); however, coffee being a commodity, its price is not given only by demand. There are other factors, such as supply production constraints (e.g., weather), government trade policies (e.g., quotas), changes in exchange rates affect commodity prices or even changes in oil prices, affecting commodity prices (Maurice; 2011) therefore it affects prices of coffee. The process of making coffee is a long and complex one. At every stage, there are several factors that can positively or negatively impact the final price of coffee.

In this thesis I would like to analyze the factors which affect the price of coffee the most and how do they change over time. One of the models which I would use is the one regarding commodity prices in general (Chen, S. L., Jackson, J. D., Kim, H., & Resiandini, P.; 2014), where one of the main independent variables is the nominal exchange rate. Especially for coffee, there are factors such as prices of substitutes for coffee (e.g., tea or cocoa) and components to coffee (e.g., sugar or milk), which are given by demand. Also, there are factors given by coffee supply, such as weather, political issues, quotas or number of countries producing coffee and their productivity. Another meaningful model is the Modelling Coffee Prices model (Russell, B., Mohan, S., & Banerjee, A.; 2012) which reflects the impact of government policies on coffee prices.

Contribution

The main purpose of this thesis is to estimate the model predicting the future price of coffee based on the driving factors and provide information on whether these factors can be controlled and/or rationally and meaningfully regulated. This would be helpful for other commodities' estimation models.

There are a lot of other studies analyzing the price of coffee on the market (Palm, Vogelvang; 1986), its fluctuations and its stabilization mechanism (Malan; 2013). However only a few of them deals with the factors the price depends on and some of these studies are outdated (Hopp, Foote; 1955).

Methodology

The International Coffee Organization (ICO) provides a lot of data regarding coffee. These data show which countries export and import coffee and how much of it is traded, which is an important factor affecting coffee prices (supply and demand part). There are also sets of data about the countries exporting coffee the most (Brazil, Vietnam, Colombia, etc.). Mainly when it comes to the conditions

important for coffee breeding which partly reflects the coffee price as a commodity. Another set of data is that regarding weather conditions in countries exporting coffee. These data are available on the website called Realible Prognosis (rp5.ru) and TuTiempo.net.
All of these will be used in the section the identificators will be analyzed.

In order to estimate the model the independent variables needs to be known. Therefore the driving factors of coffee price will be identified by regression model (Ji, Q.; 2012) and then these identified factors will be used as the independent variables in regression model estimating the coffee price.

Outline

Abstract
Introduction
Theoretical Background
- History of coffee trading – from beginning to 1990
- Coffee crises
- Factors given by supply (exporters)
- Factors given by demand (importers)

Methodology
- Modeling Coffee Prices model (Russell, B., Mohan, S., & Banerjee, A.; 2012)
- Simple factor-based model (Chen, S. L., Jackson, J. D., Kim, H., & Resiandini, P.; 2014)
- Model identifying driving prices for crude oil (Ji, Q.; 2012)

Research
- Model estimation

Conclusion

List of academic literature:

Bibliography

Berhane, T., Shibabaw, N., Shibabaw, A., Adam, M., & Muhamed, A. A. (2018). Forecasting the Ethiopian Coffee Price Using Kalman Filtering Algorithm. *Journal of resources and ecology*, *9*(3), 302-306.
Chen, S. L., Jackson, J. D., Kim, H., & Resiandini, P. (2014). What drives commodity prices? *American Journal of Agricultural Economics*, *96*(5), 1455-1468
Cortazar, G., Millard, C., Ortega, H., & Schwartz, E. S. (2018). Commodity Price Forecasts, Futures Prices, and Pricing Models. *Management Science*.
Durevall, D. (2007). Demand for coffee in Sweden: The role of prices, preferences and market power. *Food Policy*, *32*(5-6), 566-584
FARAH, Adriana. *Coffee: Chemistry, quality and health implications*. Cambridge: Royal Society of Chemistry, [2019]. ISBN 1782620044
Hopp, H., & Foote, R. J. (1955). A statistical analysis of factors that affect prices of coffee. *Journal of Farm Economics*, *37*(3), 429-438
International Coffee Organization (ICO); Trade Statistics Table: Consumption (2019) [online]; Available at: http://www.ico.org/prices/new-consumption-table.pdf
Ji, Q. (2012). System analysis approach for the identification of factors driving crude oil prices. *Computers & Industrial Engineering*, *63*(3), 615-625.
Malan, B. B. (2013). Volatility and stabilization of the price of coffee and cocoa in Côte d'Ivoire. *Agricultural Economics*, *59*(7), 333-340.
Maltsbarger, R. D. (2011). *A structural model of the international coffee sector: an econometric investigation* (Doctoral dissertation, University of Missouri--Columbia)
Maurice, N., & Davis, J. (2011). Unravelling the underlying causes of price volatility in world coffee and cocoa commodity markets.
Menke, Andrew (2018), *The Global Coffee Industry* [online] Available at: https://globaledge.msu.edu/blog/post/55607/the-global-coffee-industry
Milas, C., Otero, J., & Panagiotidis, T. (2004). Forecasting the spot prices of various coffee types using linear and non-linear error correction models. *International Journal of Finance & Economics*, *9*(3), 277-288.
Palm, F. C., & Vogelvang, E. (1986). A short-run econometric analysis of the international coffee market. *European Review of Agricultural Economics*, *13*(4), 451-476
Russell, B., Mohan, S., & Banerjee, A. (2012). Coffee market liberalisation and the implications for producers in Brazil, Guatemala and India. *The World Bank Economic Review*, *26*(3), 514-538.

Ubilava, David. "El Niño, La Niña, and world coffee price dynamics." *Agricultural Economics* 43.1 (2012): 17-26

Vogelvang, E. (1992). Hypotheses testing concerning relationships between spot prices of various types of coffee. *Journal of Applied Econometrics*, *7*(2), 191-201

Contents

List of Figures and Tables

List of Figures

List of Tables

Introduction

Coffee is one of the most consumed beverages worldwide. In coffee year 2018/2019 the total consumption was approximately 168.1 million bags (International Coffee Organization, 2020a), which is about 10.06 billion kilograms of coffee beans. The coffee industry has a significant impact on the country's economy. Based on data of 32 countries exporting coffee, one percent increase of coffee export results in 0.0217 percent increase in country's GDP on average (Murindahabi et al., 2019). For example, in Colombia, the third biggest producer of coffee worldwide, a rise of coffee production by one percent causes Colombia's GDP to rise by 0.34 percent (Ferguson, 2017). Another example is Ethiopia, being the biggest coffee producer in Africa, where one percent increase in coffee exports leads to a rise in GDP per capita by 0.0692 percent (Yifru, 2015).

The coffee tree is grown mostly in the area between Tropics of Capricorn and Cancer, called the Bean Belt where lots of developing countries are situated (NCAUSA). For these countries, the coffee industry is essential because it provides job opportunities for the citizens and forms a considerable part of the nations' GDP. In Ethiopia, the coffee share of GDP was 1.57 percent and of export value was 20.2 percent in 2015 (World Integrated Trade Solution, 2020), and it provides livelihood to 16% of its population which is approximately 15 million people (Tefera, 2014). According to Fairtrade Foundation, around 125 million people's livelihoods are dependent on coffee production.

Moreover, coffee is an important commodity for importers, which are mostly developed countries such as the USA and the members of the EU (International Coffee Organization, 2020b). These countries profit from the subsequent coffee manufacturing, which includes the operation of coffee shops, production and selling of coffee equipment such as coffee machines and coffee cups. The coffee industry is an important economic component of these countries. For example, in 2015, it formed 1.6% of the US's GDP (National Coffee Association USA, NCAUSA, 2016).

All these numbers mentioned indicate that coffee is not only a popular beverage but also a commodity which has an impact on the global economy. The price of such commodity is an important economic factor which is worth to be examined. Some researchers had already concluded that coffee is an inelastic commodity (Yohannes, 2016; Teuber, 2012) which is the reason why its price suffers from big volatility. This makes it challenging to predict the behaviour of the price, which is vital for all farmers, traders and consumers.

The aim of this thesis is to find the factors that have a significant impact on the price of coffee. Based on these findings, we can predict, how the coffee price reacts to movements in these. At the same time, identifying the price factors and applying them to the two coffee types shall show the dependence of the individual coffee types on various factors.

Hopp and Foote (1955) and Akiyama et al. (1982) had already discussed similar topic and provided a statistical analysis of coffee price based on some factors across categories, mainly supply and demand ones. Hopp (1955) introduced one multiple linear regression, including time trend, where supply factors were deflated by appropriate demand factors. On the other hand, Akiyama (1982) provided two models, one for supply and one for demand. Both studies are outdated since they used data at least 40 years old, and the significance of factors could have changed over time. The recent studies (Cuaresma et al., 2018) focus only on one type of coffee – arabica – or only on some countries producing coffee, for example, Ethiopia (Behane et al., 2018) or Côte d'Ivoire (Malan et al., 2011).

This thesis aims to provide an econometric model for prices of both types of coffee, arabica and robusta, and for the price given by the International Coffee Organization (ICO). The ambition is to discover various factors which drive the prices at the global level. For this purpose, an adjusted version of the model introduced by Baffes et al. (2016) is used. Baffes focused on the causes of price movements of agricultural commodities and he used the ARDL regression. Tothmihaly (2017) used the same method for examining the behaviour of cocoa prices. These papers will be fundamental to this thesis.

The hypotheses to be tested are:

H1: Factors driving the price of coffee significantly are the amount of coffee and the real interest rate.

H2: The effect of the same factor differs for the coffee types.

The H1 suggests two factors which might be the ones we search for. Coffee supply is expected to drive the prices because with a change in the amount available, the price changes as well (law of supply). The real interest rate is an important factor when speaking of the US dollar in which the prices are stated. The H2 claims that the factors have a different impact on arabica and robusta prices because these types differ in various characteristics, for example, the optimal weather conditions for growing (supply factor) or taste (demand factor). Thus, it is likely that a change in one factor will not have the same magnitude of impact on both arabica and robusta.

The theoretical section is divided into three parts. Chapter 1 summarizes the findings made by researchers who focused on coffee and other agricultural commodities' prices. Chapter 2 provides a background of coffee and its characteristics, differences between arabica and robusta and mentions the policies and important events which might have caused fluctuations in price. Chapter 3 describes the factors of more categories (supply, demand, financial market, weather conditions and commodity market), which might be the driving features when it comes to movements of coffee price.

The methodology section comprises of three parts as well. Chapter 4 discusses data and their description. Chapter 5 introduces the autoregressive distributed lag (ARDL) regression, cointegration, diagnostics tests and multicollinearity problem. The results of the model estimation for coffee prices are summarized in Chapter 6. The conclusion contains a discussion regarding the hypotheses and the factors which were confirmed to be the driving ones.

1. Literature review

This chapter summarizes the findings from studies which analysed similar topics. The first part deals with those that were focused strictly on the coffee prices and factors affecting them. The second part deals with those that describe the driving factors of other commodities' prices.

1.1. Studies analyzing coffee prices

Estimating agricultural commodity prices is not a trivial process. In the past, there were studies that had already discussed the economic factors affecting coffee prices. Hopp and Foote (1955) concluded that the significant variables are the time, the ratio of the world stocks of coffee and the ratio of the exports from Brazil, which was, and still is, the biggest exporter of coffee.

Akiyama et al. (1982) analysed this topic from both supply and demand sides. Unlike Hopp and Foote's model (1955), these contain lags. The results also provided evidence that income elasticities for the countries with the biggest consumption are low, and so is the price elasticity of demand. These conclusions correspond with the finding that the volatility of coffee price is mainly caused by the price of green coffee beans (Lewin et al., 2004), which primarily given by the supply factors. These describe the countries with the biggest share of coffee production and the harvesting conditions, including weather, political and economic ones (e.g. natural disasters and real prices paid to growers including subsidies).

In the recent years, studies forecasting coffee price in a specific country were published. For example, Behane et al. (2018) used Kalman filtering algorithm and Root Mean Square Error (RMSE) to forecast coffee prices in Ethiopia based on daily coffee prices. Other influencing variables were not included because the authors found it difficult to obtain the significant ones, and the validity of the model would be hard to test. Naveena et al. (2017) analysed robusta coffee prices in India and used Autoregressive Integrated Moving Average (ARIMA) and ANN (Artificial Neural Network) model. None of the studies worked with any explanatory variables driving the price of coffee.

Cuaresma et. al. (2018) analysed the impact of market fundamentals, speculations and macroeconomic conditions on prices of Arabica coffee and explained coffee price's fluctuations over history. A total of 13 independent variables is included in the model.

These are the factors suspected from being the driving ones (climatic, macroeconomic, fundamental and financial). They used Bayesian Model Averaging (BMA) analysis, Posterior Model Probabilities (PMP), Vector Autoregressive (VAR) and Vector Error Correction (VEC) specifications, the coffee prices were in logarithmic form. The result of this study is that the most important factors to describe the fluctuations in history and predict the price for the future are the macroeconomic (output for Brazil and real effective exchange rate) and financial ones (stock market index for the USA).

In their study Lewin et al. (2004) discuss coffee's historical background, its prices and their volatility and supply and demand factors. Aside from other papers, new factors which could drive the coffee prices are introduced, for example, the number of new coffee exporters. Regarding the costs of coffee, the most volatile part is the green bean coffee price, on the other hand, the retail prices are usually stable. This finding supports the main idea of this thesis - focusing on the factors of green coffee bean price. This approach is confirmed in the study by Malan et al. (2011) that analysed the relationship between cocoa and coffee prices in Côte d'Ivoire, whose GDP is dependent on this industry. Malan et al. (2011) focused on the price volatility of these commodities and its impact on the farmers, which is caused by endogenous and exogenous (natural disasters) fluctuations. Zheng et al. (2008) concluded that the price elasticity of demand for coffee is -0.083. This means it is inelastic, therefore the demand for the commodity is not very responsive to changes in the price.

1.2. Studies analyzing prices of other commodities

Chen et al. (2014) in their study focus on 51 fuel and non-fuel commodities, and it answers the question why the theoretical and the empirical approach do not provide the same results regarding the stationarity of the commodity's prices. The theory suggests that the behaviour of the prices is dynamic and mean reverted, which means that only short-time deviations are present, and their levels quickly reach the mean back. However, Augmented Dickey-Fuller test confirms nonstationarity. The authors defined two groups of factors. The first one is closely related to the US dollar nominal exchange rate, which is a unit root process. The second group covers other factors. It was concluded that the first group causes the nonstationarity. The second set of factors together with idiosyncratic components suffers only from small deviation, but the stationarity is not violated. One of the second common factors is the crude oil price, which appeared to be significant. This corresponds with results from other studies (Nazliogu, Sagan, 2011; Maurice, Davis; 2011) proving the causality between the crude oil price and other commodities' prices. The finding is important for global investors because they can predict the fluctuations in prices of agricultural commodities based on the changes in oil prices.

Maurice, Davis (2011) and Natanelov (2011) also concluded that the price of crude oil and futures exchange markets cause volatility of the agricultural commodity prices. Coffee, being a commodity, is not an exception. The costs of production include machines operating (for example picking devices) and transportation expenses which are both dependent on the oil prices. Maurice and Davis also pointed out the fact that coffee price behaviour is correlated with cocoa prices, which is considered together with tea to be a substitute for coffee.

Agricultural commodities, whose prices are studied the most often are soybeans, wheat and corn. In order to forecast future prices, Ahumada et al. (2016) analysed the prices of these commodities in one study since they were suspected from being cross-dependent. The researchers took the quarterly nominal prices of the commodities as a dependent variable. The annual production, ethanol production, CPI, real GDP of China, India and OECD, US exchange rate, real monetary base and flows of funds were the explanatory variables. The methods used for estimation were Equilibrium Correction Model (EqCM) and Vector Autoregression in first differences (DVAR). The researchers concluded that embodying cross-dependence between two or more commodities might improve the estimated model. This leads to the suggestion of including cocoa prices in the model since coffee prices are correlated with them (Maurice and Davis, 2011).

Crude oil is considered to be one of the most important commodities worldwide, and several studies focus on crude oil prices, their forecasting and estimation. Ji (2012) used a system analysis approach to identify factors driving prices of crude oil before and after the financial crisis in 2008. The Partial Least Squares (PLS) regression was used to analyse the significance and power of each factor suspected from being the driving one. Each variable was tested by the Augmented Dickey-Fuller test for unit root test, and subsequently, eight of them were selected for the next analysis. The variables EPPI (US energy producer price index) and ECPI (US energy CPI) appeared to have a considerable impact, but they suffered from contemporaneous causality. Therefore, the researcher used technique DAG (directed acyclic graph) to reduce it. The finding of the study was that factors causing the price volatility the most are crude oil stock level, US Dollar index and the investment substitute effect of other markets. Miao (2017) had a similar aim in his study, which was focused on six categories of possible factors influencing crude oil prices (supply, demand, financial market, commodities market, speculative and geopolitical). The researcher used the Least Absolute Shrinkage and Selection Operator (LASSO) method to forecast future prices. Despite crude oil not being an agricultural commodity, its price suffers from similar difficulties causing the volatility.

Keatinge et al. (2015) focused on determinants influencing annual wheat price in the United States. These were split into five groups (supply, demand, climate index, macroeconomic and natural resource) and the model was estimated by Ordinary Least Squares (OLS) regression. The authors concluded that the model explained the variance of prices properly, but there are possibilities of improvements, such as adding financial variables and climate conditions (temperature and rainfall) as well.

There are two fundamental studies for this thesis. Baffes et al. (2016a, 2016b) focused on the factors causing price movements of agricultural commodities, specifically maize, soybeans, wheat, rice, palm oil and cotton. The researchers observed the short run and long run impact of five factors: GDP, real crude oil prices, real interest rate, real effective exchange rate and stock to use ratio of the observed commodity. All these variables appeared to explain price movements of at least one researched commodity. The authors suggested to include variables describing climate conditions to improve the model. Tothmihaly (2017) explained low elasticity of cocoa prices by supply and demand factors. Among supply ones were coffee price, as a substitute for cocoa, yield and lagged values. On the demand side were oil price, GDP and palm oil price. The study concluded that coffee is a weak substitute for cocoa. Both studies used the same ARDL regression, which is specified in Chapter 4.

2. Introducing coffee and its characteristics

This chapter introduces coffee as a plant and its process of growing and events in the last 60 years that might have had an impact on coffee prices. It contains two subchapters, each describing one of the mentioned topics. All of these are important to better understand why the price of coffee is that volatile and why the two coffee types behave differently. The findings are essential for investigating factors the driving factors.

2.1. Varieties of coffee beans and the process of growing

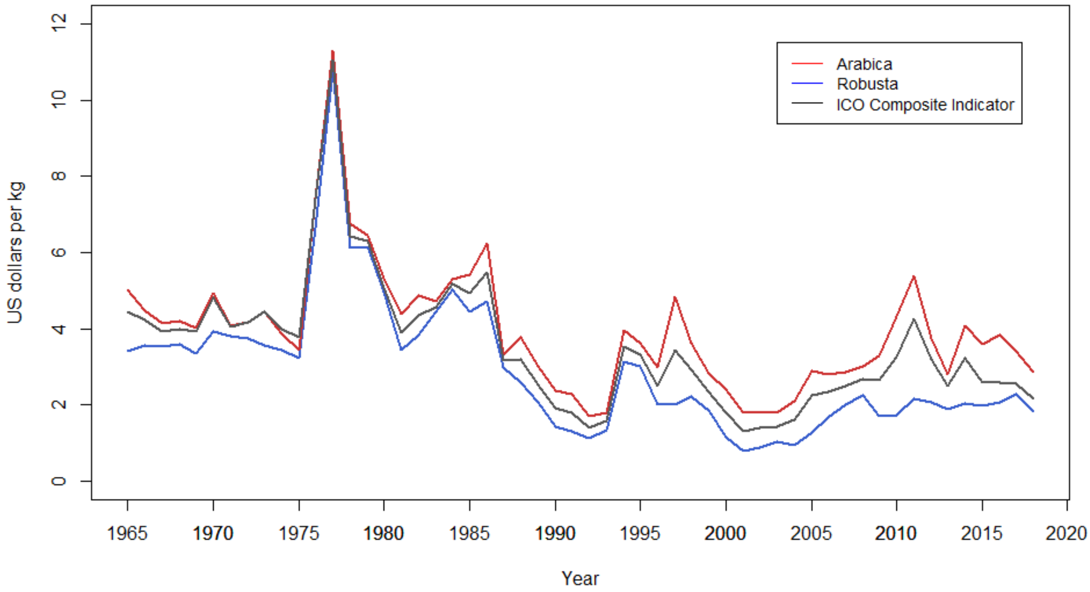
This subchapter summarizes the facts about coffee as a plant, specifically the types of coffee and the process of its growing. It is crucial for determining the factors that influence the price significantly.

2.1.1. Comparison

There are two main varieties of coffee: arabica and robusta. Arabica is further subdivided in three groups established by the ICO – Brazilian Naturals, Colombian Milds and Other Milds. This distribution is based on the countries of origin of coffee. However, some institutions, for example the World Bank's commodity market, considers the price of arabica as an average of the prices of these three types. Since we speak about the global market situation, we will also compute with this value. It should not cause deviations, as the evolution of the prices of arabica's types is similar (see Figure 5 in Appendix A). The arabica's share on the market is around 60 percent (International Coffee Organization, 2020c), and the prices are higher than the robusta's mainly because of the different quality. This relates to the fact that arabica is preferred by most of the big coffee companies. Robusta's share is about the remaining 40 percent. The process of robusta's manufacturing is cheaper, which is reflected in the price and the taste of the beans, that are less flavourful (Mounts, 2018), thus less demanded. The two different types prefer different weather conditions: arabica is grown mostly in Latin America, while robusta in Africa and Asia. This is reflected in the price as well because countries in Latin America are usually more developed than African and Asian countries. Therefore, the technology is more advanced, and people employed in the coffee industry demand higher wages.

The different behaviour of prices can be seen in Figure 1, which shows the evolution of prices coffee types and the ICO composite indicator. All prices were adjusted by the Manufactures Unit Value (MUV) index to obtain the real ones. Arabica's price peaks are in years 1977, 1986, 1997 and 2011. In 1977 the real price for arabica was 11.3 US dollars which is the biggest in recent history. Compared to arabica's prices, the prices of robusta followed similar evolution, but there are some differences, for example, the peak in years 1997 and 2011. The possible causes will be described in detail in the next subchapter. The black line represents the evolution of the ICO composite indicator which combines prices of the two types and is defined by the ICO. It copies the behaviour of the arabica, which is reasonable because it has a major share in the world market. Another fact to be mentioned is that the real price slightly decreases over time. The finding that robusta's and arabica's prices do not evolve the same supports the H2 claiming that the same factors have a different impact on the prices.

Figure 1 - Real prices of coffee types and the ICO composite indicator from 1965 to 2018

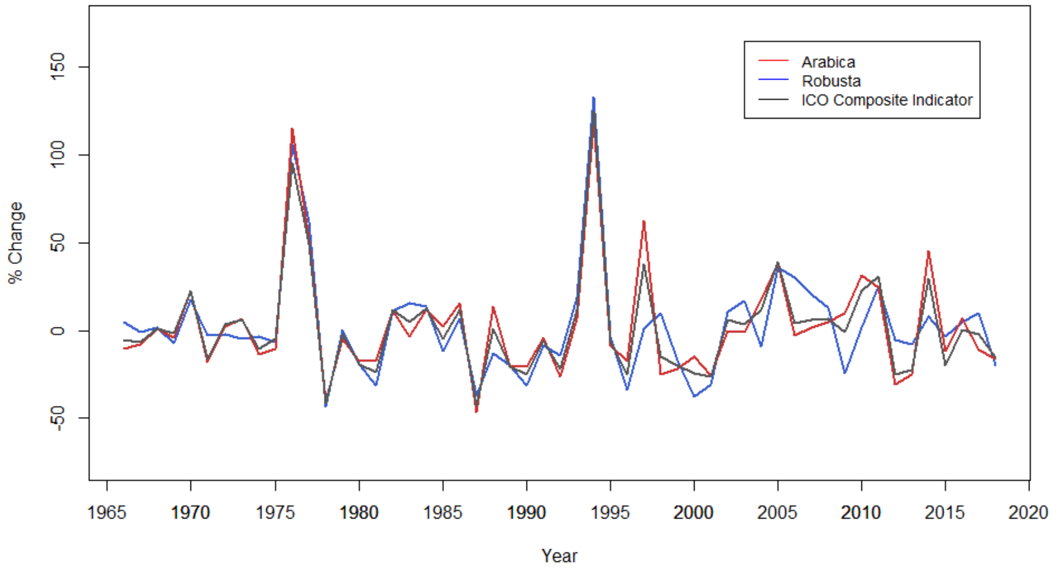


Data source: ICO

The previous graph provides an overview of how the prices evolved. However, to understand the price movements, it is relevant to focus on the annual changes in the prices shown in Figure 2. In this graph, the unit is the percentage change in price in comparison to the previous year. It can be concluded that the price is not stable for neither type since there are some nonnegligible fluctuations. The biggest change came in the year 1994 when the price increased by more than 100%. This event was not visible in the previous graph; however, it is important, as it points out that in that year the price increased the most for the past 50 years. The second biggest increase of price was in the year 1976. The price was rising, and it peaked in 1977 which is visible in the Figure 1. However, the change in price between years 1975 and 1976 was bigger than the one next year.

As in the previous graph, the ICO composite indicator is driven by the arabica behavior, which is expected. Moreover, the price of robusta and arabica changes in a different way in some cases, for example, in years 1982 and 2009. This again supports H2 and the reason, why these two types should be considered as two separate commodities.

Figure 2 - Price volatility of coffee types and the ICO composite indicator from 1965 to 2018



Data source: ICO

2.1.2. Coffee bean journey: The process of growing

There are five major steps of a green coffee bean journey from planting to export. Each additional step means an additional cost of production, which is reflected in the final price. The process is almost the same for both sorts of coffee, and they just prefer different conditions.

Planting the seed

In the beginning, there is a coffee seed to be planted. Geographically, the coffee plant is grown mostly in the Bean Belt, an area between the Tropics of Capricorn and Cancer. The varieties, arabica and robusta, prefer different weather conditions (see summary in Figure 3). Robusta is more resistant to higher temperatures, and it grows in the lower areas. This corresponds with the fact, that it is planted mostly in the countries in Africa, for example Côte d'Ivoire, Ghana and Togo. These are very poor countries, with very cheap labour force, which might be one of the reasons why robusta is cheaper. Arabica, on the other hand, prefers colder temperatures, less of rainfall and higher situated planting areas. It is mostly grown in the countries of Latin America (Colombia and Venezuela) but also Ethiopia produces this type of coffee. Asian countries (Vietnam, India and Indonesia) and Brazil, the biggest coffee producer in the world, provide conditions where both types can be grown (International Coffee Organization, 2020c).

The optimal weather conditions for coffee plant growing are described in Figure 3 might be useful for modifying factors explaining climate conditions.

Figure 3 Optimal weather conditions for both types of coffee

	Arabica	Robusta
Temperature (ºC)	18-21	22-30
Rainfall (mm)	1200-1800	1200-2500
Altitude (m)	600-2200	0-800

Data source: DaMatta (2006)

Harvesting

For the first 3-4 years, the coffee plant breeds no coffee cherries, and then it blossoms once or twice per year. After the coffee cherries are ripe, they are stripped or selectively picked. The strip picking is a process, done by a machine or a human, when the coffee branches are stripped off, so all cherries are harvested, including the unripe and rotten. This method is time efficient and relatively cheap, but this is reflected in the quality. The selective picking is used mainly for the finest types of arabica coffee and can be done only by hand because the picker chooses only the appropriate cherries. This method is costly and not as quick as the strip picking, but the quality is higher. (NCAUSA, Coffee and Health). This is one of the reasons why Arabica coffee is more expensive on the average.

Processing the cherries

In this phase, the coffee bean is removed from the coffee cherry. There are two methods. The first is the dry method, which is the traditional one. It is used for almost all robusta and in some countries (Brazil, Ethiopia and Paraguay) for Arabica. The wet method is more advanced and costly since water and special equipment are needed. Then the beans are being dried until the level of its moisture is around 11 % (NCAUSA, 2020; Coffee and Health).

Milling

In the final phase before exporting, redundant layers are removed from the coffee bean in two stages: hulling and polishing. Both are machine-run processes. Then the defected beans are sorted out, and the rest is sorted based on size and colour. (NCAUSA, 2020; Coffee and Health)

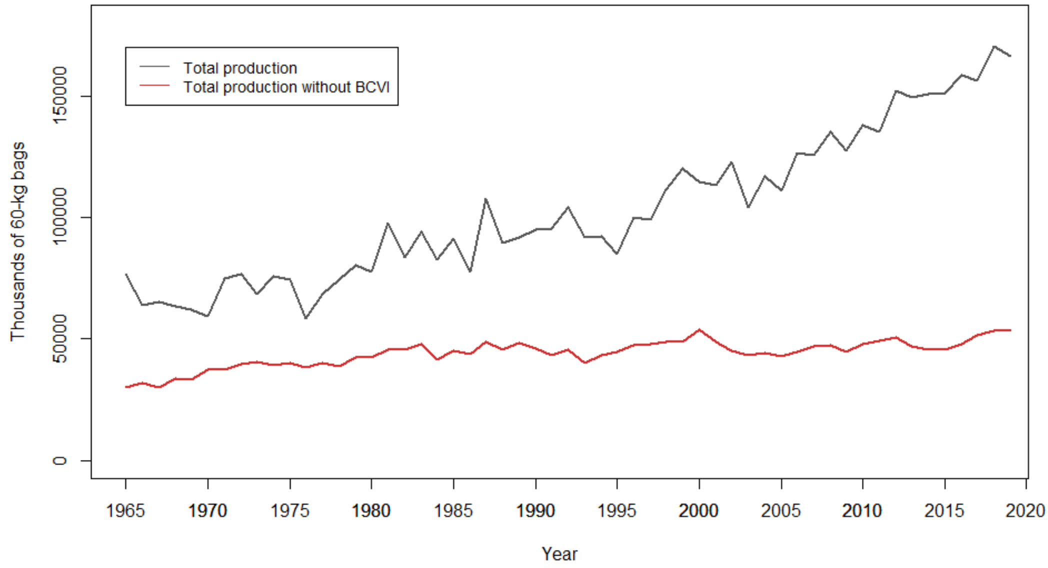
Exporting

After milling, the green beans are packed in 60 kg jute bags and distributed to countries around the world for the next processes which are tasting, roasting, grinding and brewing (NCAUSA, 2020). The export costs (e.g. price of oil) are reflected in the final price of a green coffee bean.

Countries producing coffee

Brazil is the world’s biggest producer of arabica and coffee in general. In 2019 its output was 47.578 million of 60 kilograms bags which is approximately 34% of the worldwide production. Until the 1990´s one of the biggest producers of robusta used to be Indonesia but then Vietnam joined the coffee trade as an exporter. Since then, Vietnam is the second biggest coffee producer and Indonesia is the fourth. The third biggest producer of coffee is Colombia, producing arabica only. Together, these countries cover more than 50% of the worldwide coffee supply (International Coffee Organization, 2020c), as visible in Figure 4. The black line is the total production of coffee, and it increases over time significantly. The red line represents the total production without the four biggest producers. This line is almost constant, which means the amount of coffee produced is in total the same over time. It suggests that the impact of the four biggest producers increases. Therefore, weather and political events in these countries should be considered as well, because they affect total coffee production, thus coffee price and volatility too.

Figure 4 - Total coffee production with and without the biggest producers



Data source: ICO

There are several important facts to be mentioned. Firstly, higher prices of arabica are caused by the area where it grows (richer vs. poorer countries) and by the system of picking and processing, that is costly. These factors partially explain the reason, why arabica's price is slightly more volatile than robusta's. Another finding is that crude oil price might influence the final price of green coffee beans since it is a part of transportation costs.

2.2. Natural disasters and trade policies in 1965 – 2018

In this chapter we summarize important events that might have affected global coffee trade, its prices and volatility. Some of these may explain the peaks recorded in Figure 1 and Figure 2. Moreover, this chapter also mentions the impacts of policies affecting coffee prices in the past. These facts are important for the understanding the coffee price development in the observed era.

The first attempt to regulate the coffee market raised in 1902 and since then it has been discussed topic. There are two levels at which the market can be controlled: domestic and international. The domestic policies are difficult to set as they tend to be sensitive to international market situation, so the responses are unpredictable (Baffes, 2005). The policies on the international level are more relevant for this thesis, since we focus on the global level.

In 1963 the ICO was established to support coffee producers and connect suppliers and consumers. Some policies were specified by the International Coffee Agreements (ICA) which kept the prices higher. In 1973 the organization managed to stabilize the price for coffee globally by levying export quotas. Besides the consequences of frosts in Brazil in 1975, which resulted in a rapid increase in prices for both types, the "quota period" worked until 1989 when it the ICA collapsed mainly due to lack of support of the United States. Since then, the coffee market is driven by supply and demand factors (International Coffee Organization, 2014). By now the ICO has 42 exporting members (19 of these belong to least developed countries in the world) and seven importing members (the EU is taken as one member). In 2007 its members agreed on the 2007 Agreement empowering ICO's position regarding the international trade, and it entered into force in 2011.

In 1976 there was a rapid increase in coffee price, which more than doubled for both coffee types. The leading cause was the frost in Brazil, the biggest coffee producer, which damaged almost 75% of crops a year before, and the total harvest decreased by more than half (Markgraf, 2001). This significant decrease in supply lead to a fast and rapid increase in price, which is visible in both figures. All prices almost doubled due to this event. The price kept increasing, and in the year 1977, the price for arabica was the highest in history, which is visible in Figure 1. However, the change of prices between 1976 and 1977 was lower than the year before which can be seen in Figure 2. In 1978 Brazil recovered, and the production increased resulting to a significant decrease in coffee prices.

Aside from 1977, there was another price peak in 1986 which was a result of a drought in Brazil a year before. The price decreased year later after Brazil recovered, and coffee production increased again.

Another frost causing drop in coffee output occurred in 1994, and the change in ICO Composite indicator increased from 1.57 (1993) to 3.54 (1994) US dollars. Despite the change was rapid – by 125%, Figure 1 did not capture this event, because the price did not seem to be that high comparing to the rest of observations. However, Figure 2 provides insight on changes, and this was the biggest one in the observed time period. To protect the production from weather extremes, Brazil moved coffee crops to areas which are frost-prone, so that the output can be better controlled. (Baffes, 2005)

In 1993 the Association of Coffee Producing Countries was formed, and all big coffee producers were members except for Vietnam. Its intention was to increase coffee prices by regulating coffee exports. However, it disintegrated in 2002 because of several reasons. One of them was for example free rider problem and the absence of the institutional structure which was essential to ensure compliance. (Baffes, 2005)

In 1995 the World Trade Organization (WTO) was established replacing the GATT (General Agreement on Tariffs and Trade) and today it has 164 member countries. Its purpose is to set an open market, make trading cheaper and reduce trade barriers. Unlike GATT, which made agreements between states, WTO is an international organization with common rules to all members. Thus, after 1995 some policies went through changes which probably affected the prices and the amount of goods and services traded (WTO). Almeida et al. (2012) found in their study that TBT (Technical Barriers to Trade) have a negative impact on coffee exports.

One of coffee crisis which was endangering farmers occurred in 2001. The problem was an oversupply of coffee because Vietnam relatively newly entered the coffee market, and Brazil produced more coffee beans than years before. This led to a rapid decrease in green coffee beans price, which fell by approximately 50 percent to 41 cents per pound, the lowest real price in the last 100 years (Boydell, 2018).

In 2011 there was a rise in coffee prices due to both the increase in demand by emerging countries and inclement weather which caused decrease in arabica supply (Harrington, 2011). Three years later, 2014, a drought in Brazil caused lack of supply and increase in price (Yang, 2014). When we look at the Figure 1, we can see, that these two events affected only prices of arabica.

In the previous subchapter, Figure 1 shows how the price of coffee fluctuated over the past 60 years. The main peaks were in 1976, 1994, 2011 and 2014. In the same years, some natural disasters occurred in countries producing coffee the most, which led to a decrease in coffee production and export. This finding supports H1 claiming that one of the driving factors of coffee prices is the amount of coffee produced and exported.

3. Groups of possible factors driving the price of coffee

The number of factors influencing the prices of commodities is substantial. In similar studies, the authors divided them into a couple of groups (Baffes, 2016a, 2016b; Cuaresma, 2018). Based on these, for this study, we define five groups: supply, demand, climate, commodity market and financial. Each subchapter discusses one group.

3.1. Supply factors

The quantity of commodity supplied is one of the most important determinants for explaining the behaviour of the commodity price. There are more approaches to its measurement. Baffes et al. (2016a, 2016b) used the stock to use ratio to explain the quantity supplied. The ratio is defined as total stocks of the commodity over total commodity consumption. Cuaresma et al. (2018) point out that output of coffee is one of the most important variables, when describing coffee price movements. This claim is supported by findings by Tothmihaly (2017) who focused on cocoa price. The researcher included cocoa yield, which is tightly correlated with production, and it appeared to have a significant impact. Also, in 2018 Food and Agricultural Organization (FAO) published analysis mentioning the dependency between coffee price movements and coffee exports (Amrouk; 2018). This variable signifies how much of the commodity is traded in the global market. This is the reason why it is expected to be the best indicator. Coffee is a crop which grows in some specific areas where the demand is not significant in terms of aggregate demand (International Coffee Organization, 2020b). This is driven by EU members and USA who produce almost no coffee so they are dependent on global exports.

All these factors are expected to have a negative impact on coffee prices in both short run and long run. This relates to one of the key thoughts of economics - with an increase in quantity supplied the price of this commodity decreases.

3.2. Demand factors

The demand factors are not as important as the supply ones since we speak about the green coffee bean price. It should be affected by demand only a little because consumers' preferences are not much significant at this point yet. However, the peak of arabica price in 2011 was caused among others by a higher demand. The idea of including the demand factors (e.g. the growth of emerging countries) is supported by Ahumada et al. study (2016), which considers the prices of agricultural commodities. Miao et al. (2017) focused on the crude oil prices and added GDP of the countries with the biggest consumption. Baffes et al. (2016a, 2016b) considered the GDP of the world when investigating price movements of agricultural commodities.

Since coffee is a worldwide consumed beverage, we use the same approach as Baffes did. There are two possibilities. The first one is the real GDP in terms of constant US dollars. The second one is the GDP growth which used Harvey (2017), who concluded that this factor has a significant and negative impact on commodity prices. Both factors reflect the world economic situation and capture technological progress. They are expected to have a negative impact on the price of coffee in both the short run and the long run. One of the reasons for this expectation is the Engel's law claiming that with an increase in income, the marginal propensity to consume (MPC) decreases.

3.3. Climate factors

One of the main reasons why the price volatility cannot be under perfect control is the fact that the production of agricultural commodities is dependent on weather conditions. This claim supported Karthikeyan et al. (2014), who provided a model predicting the price of soybean crops in which the climate factors had a significant role. However, it is complicated to provide climate variable at the global level. The average temperature or amount of rainfall would not be sufficient, because coffee is planted only in some areas in the world. However, there are two options how climate variable can be proxied and become sufficient. The first are indices regarding weather, for example, Southern Oscillation Index (SOI) and Sea Surface Temperature (SST) which were used for estimating prices of wheat (Algieri, 2014). Another index to describe weather conditions is the El Nino index. However, none of these indices captures the complete weather situation in the world. The second possibility is to include a variable reflecting the number of natural disasters in the world that have an impact on crops, specifically extreme temperatures, floods and frosts (Baffes et al., 2016). This would be probably a better option because it explains the situation around the world. These variables are tightly connected to supply of coffee. The impact of natural disasters should be positive because it causes drops of the commodity available for trading, thus decrease in quantity supplied, which leads to a higher price.

3.4. Commodity market factors

The first group are complements, prices of which are expected to have a negative impact on the price of the given commodity. Intuitively, sugar and milk could be considered as complements to coffee. However, there is no study which would confirm the relationship between these commodities. Therefore, none of this group will be considered as a possible driving factor.

The second group are substitutes which are expected to have a positive impact on the price movements of coffee. There are two similar products' prices which should be considered overtime – cocoa and tea. Some studies found a correlation between coffee and these variables, but the causality has not been proved. The agricultural commodities behave likewise and depend on similar factors; thus, their prices evolve in the same way, which might lead to misinterpretation of the effect. Tothmihaly (2017), who used a similar approach to the one in this thesis, concluded, that coffee is a weak substitute for cocoa. This implies the same result when considering cocoa as a substitute for coffee. Moreover, Doan (2014) concluded that there is no evidence that people would buy less coffee when its price increases.

Another commodity, whose price should be considered is the crude oil. Maurice et al. (2011) and Natanelov (2011) provided an evidence that the price of coffee is related to the crude oil price not only due to production costs but also because crude oil price significantly influences the commodity stock market and can reflect the economic activity. Thus, the crude oil price is expected to have a positive impact on price of coffee in both short and long run.

3.5. Financial variables

The data for prices of coffee and crude oil are expressed in US dollars. These variables will be deflated by Manufactures Unit Value (MUV)[1] index (Baffes, 2016, and Tothmihaly, 2017, did the same in their studies). After applying this index, we obtain real prices, so inflation has no impact. Even after this transformation, there are two financial factors to be observed. The first one is US real interest rate. This variable explains the value of holding money (whether it is better to consume or save them). Researchers mostly agree that the real interest rate has a negative impact on commodity prices in the

[1] Real price = (nominal price/MUV) * 100

long run (Akram, 2009, and Frankel, 2006). Regarding the short run effects, the results are not the same. Some studies concluded that there is a positive relationship because the interest rate reflects shock in the business cycle (Akram, 2009).

Another factor regarding currency is the United States' real efficient exchange rate (REER). It is the exchange rate of a currency (US dollar) against a basket of multiple currencies adjusted by a measure of relative prices (Bank for International Settlements, 2018). This variable captures the strength of the US dollar in terms of other currencies. Its impact in the short run is expected to be negative. However, in the long run, it should be positive because, from this perspective, the trade becomes more favourable to exporters to the United States, which is one of the biggest coffee importers (International Coffee Organization, 2020b). This claim is supported by Ayres (2020), who concluded there is a positive relationship between primary commodity prices and exchange rates of developed countries.

4. Data and descriptive statistics

Data regarding coffee, which means coffee prices, stock to use ratio, exports and production, were provided by the International Coffee Organization (ICO). The ICO composite indicator is a weighted average of the prices of four types of coffee (three arabicas, one robusta). The price of arabica is the average of the three types, as we already noted. Units of these variables were US cents per pound which we adjusted to US dollars per kilogram for a better interpretation. Stock to use ratio is computed as total stocks divided by the total consumption in a given year. The unit is a percentage. Units of production and exports are for both thousands of 60 kilograms bags. This is the reason why prices were adjusted from currency per pound to currency per kilogram.

Crude oil prices are obtained from the World Bank, and as there are more varieties, we compute with the average price. The units are US dollars per barrel. All commodity prices we use are adjusted by the Manufactures Unit Value (MUV) index, which deflated nominal prices to real ones. The base year is 2010.

The world's GDP is measured in trillions of real US dollars, and the annual growth of the world's GDP is in percentage. Data for both were provided by the World Bank.

Data for the real effective exchange rate (REER) were provided by Bruegel. It is the US dollar against a basket of foreign currencies which considers 171 trading partners. The unit is an index and the base year is 2010.

The real interest rate is the lending interest rate for the United States adjusted for inflation. It is measured in percentage, and the data are obtained from the World Bank.

The variable called "natural disasters" is the number of floods, droughts, and extreme temperatures per a given year. The data source is Emergency Events Database (EMDAT), and it is measured in units.

All data are annual, therefore there is no problem with seasonality. The period is from 1965 to 2018, which means that there are 54 observations. The year 2019 is not included because data for some variables are not been available yet. The descriptive statistics of all variables are summarized in Table 2.

Table 1 - Description of variables

Variable	*Name*	*Description*	*Unit*	*Notes*	*Source*
ICO	ICO composite indicator	Weighted average of prices of four types of coffee, real price	US dollars per kilogram	Nominal ICO indicator adjusted by monetary unit value	ICO
Price	Price of arabica/robusta		US dollars per kilogram	Nominal price adjusted by monetary unit value	ICO
GDP	World GDP	real GDP	trillion US dollars		World Bank
GDP growth	Annual growth of GDP		Percentage		World Bank
Interest rate	Real interest rate		Percentage		World Bank
REER	Real effective exchange rate	of US dollar	Index, 2010 = 100	US dollar real exchange rate against a broad basket of currencies	Bruegel
Oil price	Crude oil price	Real price of crude oil	US dollars per barrel	Nominal oil price adjusted by monetary unit value	World Bank
StUR	Stock to use ratio	Total stocks divided by total consumption	Percentage		ICO
Exports	Exports	Total exports of coffee worldwide	Thousands of 60 kg bags		ICO
Production	Production	Total production of coffee worldwide	Thousands of 60 kg bags		ICO
Disasters	Natural disasters	Number of cases of floods, droughts, and extreme temperatures	Units		EMDAT

All data are annual and time period for variables is 1965-2018

Table 2 - Descriptive statistics of variables in the original form

Variable	*Minimum*	*1st quartile*	*Median*	*Mean*	*3rd quartile*	*Maximum*	*Standard deviation*
ICO composite indicator	1.313	2.485	3.232	3.522	4.274	11.048	1.72
GDP	14.84	26.52	38.75	42.38	57.53	82.71	19,50
Growth of GDP	-1.678	2.519	3.311	3.347	4.337	6.505	1.54
Real effective exchange rate	91.83	97.34	103.44	105.63	114.86	125.65	9.83
Crude oil price	5.212	19.858	28.416	38.263	53.763	95.266	26.40
Real interest rate	-0.6943	1.198	1.9202	1.9672	2.6802	5.1579	1.28
Stock to use ratio	0.1902	0.3457	0.5537	0.5558	0.6662	1.37	0.28
Total exports	43 276	57 426	71 422	73 043	85 283	111 806	18 118
Total production	58 235	77 126	95 199	101 681	122 204	170 385	29 627
Number of disasters	21	56	98	111.6	184	259	73.10
Price of arabica	1.692	2.876	3.814	3.96	4.661	11.30	1.66
Price of robusta	0.7929	1.8382	2.2723	2.9363	3.594	10.784	1.78
Exports of arabica	31 263	42 206	50 329	49 794	56 168	72 027	9 948
Exports of robusta	12 013	16 231	20 727	23 250	28 702	41 787	8 532

Despite the fact that we will compute with the data in the logarithmic form[2] (the reason is explained in the next chapter), the data are presented in their original form because it can be better explained. Moreover, this format gives us an overview of how they behave and whether there are some big fluctuations, which should be interpreted. For example, we can see that all values of ICO composite indicator are between values of price of arabica and robusta, which confirms it is a weighted mean of these. As was already mentioned, the stock to use ratio is in percentages, however, its maximum is 137 %. There are two more values which exceed 100 %, for the years 1965 – 1968. The cause of this was a rapid increase of production as in the 1960s when the ICO was established

[2] Real interest rate and GDP growth remain in original form because their values are negative in some cases

leading to a support of coffee farmers. However, people did not react immediately, and the consumption did not grow as fast as the output, which is the reason why this ratio was bigger than 1. Another thing, which should be mentioned is that all coffee prices have a maximum which deviates a lot and there is a big difference between the third quartile and the maximum (approximately the value of standard deviation multiplied by 4). For all the coffee prices, this is the observation in the year 1977. The possible causes were already mentioned. This indicates that an outlier is present, and we need to provide a solution on how to deal with it.

Another important finding is that the minimum values of the interest rate and annual growth of GDP are negative, therefore we cannot put it in the logarithmic form. It causes no problem to our model, but the final interpretation will be different.

Table 3 shows the correlation between the variables in the form in which they will be used for our estimation. Some of them are highly correlated – the value is lower than -0.8 or higher than 0.8, which means we should be cautious about multicollinearity problem. However, the autoregressive distributed lag method (ARDL) model supposes the correlation is present in the long run between variables and controls for it. The model will be described in the next chapter.

Table 3 - Correlation matrix, all variables in logarithmic from except from interest rate and growth of GDP

	GDP	*GDP growth*	*REER*	*Oil price*	*Interest rate*	*Stock to use ratio*	*Exports*	*Production*	*Disasters*
GDP		-0.407**	-0.114	0.779***	-0.408**	-0.88***	0.965***	0.945***	0.920***
GDP growth	-0.407**		0.252	-0.49***	0.261	0.371**	-0.350**	-0.397**	-0.367**
REER	-0.114	0.252		-0.29*	0.308*	0.217	0.002	-0.012	-0.076
Oil price	0.779***	-0.49***	-0.29*		-0.427**	-0.833***	0.718***	0.748***	0.713***
Interest rate	-0.408**	0.261	0.308*	-0.427**		0.592***	-0.425**	-0.388**	-0.254
Stock to use ratio	-0.88***	0.371**	0.217	-0.83***	0.592***		-0.83***	-0.802***	-0.728***
Exports	0.965***	-0.350**	0.002	0.718***	-0.425**	-0.833***		0.947***	0.855***
Production	0.945***	-0.397**	-0.012	0.748***	-0.388**	-0.802***	0.947***		0.870***
Disasters	0.920***	-0.367**	-0.076	0.713***	-0.254	-0.728***	0.855***	0.870***	

Computed correlation used Pearson-method with listwise-deletion.

5. ARDL method

5.1. ARDL model and its specification

The ARDL was firstly introduced by Pesaran (2001). Its purpose is to capture the short and the long run impacts of variables. ARDL is commonly used for estimating both impacts of factors in small sample sizes. It is adjusted for variables suffering from the unit root and spurious correlation. The only condition which all variables must satisfy is to be integrated of order 0 or 1 (I(0) or I(1)) which is tested by the Augmented Dickey-Fuller (ADF) test.

Baffes et al. (2016a, 2016b) used the ARDL method to explain price movements of agricultural commodities by five independent variables – GDP, real effective exchange rate, crude oil price, real interest rate and stock to use ratio. This is the original model which we will adjust to identify the driving factors of coffee price:

$$lnY_t = \beta_0 + \beta_1 lnGDP_t + \beta_2 lnREER_t + \beta_3 lnOil_t + \beta_4 IR_t + \beta_5 lnStUR_t + \mu_t$$

Equation 1

The estimated parameters β_i's can be explained as elasticities except from β_4 because the real interest rate cannot be transformed to logarithmic form. The ARDL specification of the model is than defined as:

$$\Delta lnY_t = \beta_0 + \sum_{i=1}^{n} \theta_i \Delta lnY_{t-i} + \sum_{i=0}^{m} \varepsilon_i \Delta lnGDP_{t-i} + \sum_{i=0}^{q} \pi_i \Delta lnREER_{t-i}$$
$$+ \sum_{i=0}^{p} \rho_i \Delta lnOil_{t-i} + \sum_{i=0}^{o} \vartheta_i \Delta IR_{t-i} + \sum_{i=0}^{r} \sigma_i \Delta lnStUR_{t-i} + \beta_1 lnY_{t-1}$$
$$+ \beta_2 lnGDP_{t-1} + \beta_3 lnREER_{t-1} + \beta_4 lnOil_{t-1} + \beta_5 IR_{t-1}$$
$$+ \beta_6 lnStUR_{t-1} + \mu_t$$

Equation 2

The final ARDL model adjusted for the coffee prices is specified here:

$$\Delta lnY_t = \beta_0 + \sum_{i=1}^{n} \theta_i \Delta lnY_{t-i} + \sum_{i=0}^{m} \varepsilon_i \Delta lnGDP_{t-i} + \sum_{i=0}^{o} \vartheta_i \Delta IR_{t-i} + \sum_{i=0}^{p} \rho_i \Delta lnOil_{t-i}$$

$$+ \sum_{i=0}^{q} \pi_i \Delta lnREER_{t-i} + \sum\nolimits_{i=0}^{r} \sigma_i \Delta lnAM_{t-i} + \sum_{i=0}^{t} \tau_i \Delta lnDis_{t-i}$$

$$+ \sum_{j=1}^{s} \alpha_j DM_j + \beta_1 lnY_{t-1} + \beta_2 lnGDP_{t-1} + \beta_3 IR_{t-1} + \beta_4 lnOil_{t-1}$$

$$+ \beta_5 lnREER_{t-1} + \beta_6 lnAM_{t-1} + \beta_7 lnDis_{t-1} + \mu_t$$

Equation 3

The sign Δ means we take the first difference of the variable. Coefficients *θ, ε, π, ρ, υ, σ, τ* represent short run parameters and β_i's long run parameters. The impact of dummy variables for outliers (DM) is explained by *α*'s, and error terms are represented by *μ*.

The dependent variable is Y_t, which is the coffee price. The factors are GDP, interest rate, real oil price, real effective exchange rate, disasters and AM - variable describing the amount of coffee available (stock to use ratio, exports and production). Please, see Table 1 which provides more details. The explanatory variables are lagged price of coffee, set of factors, their lags and dummy variables. Except for interest rate and dummies, all variables are in logarithmic form. Length of lags is chosen based on the Akaike information criterion (AIC), which is designed for model selection[3].

The next step is to test the model for cointegration, which means variables are correlated in the long run. This is important because the long run relationships might affect short run estimates. This is called ARDL bound test, and the null hypothesis says that long run parameters are all equal to zero so cointegration is not present:

H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = 0$

[3] Firstly introduced by Akaike (1974), defined as: AIC = 2k – 2ln($\mathcal{L}$), where k is number of parameters and $\mathcal{L}$ is the maximum value of the likelihood function of the model. The lower the value, the better the model.

We obtain the F statistics and compare it to critical values. This approach was firstly introduced by Pesaran (2001), who defined the critical values. Narayan (2005) later specified new critical values which suit smaller sample sizes (30-80 observations) better. These are the critical values we will use since the number of observations is around 50. The null hypothesis of no cointegration is not rejected if the F statistics is lower than I(0), lower bound, and rejected if it is higher than the I(1), upper bound. If it is between these values, the model is not sufficiently specified, and the conclusion cannot be made.

If there is no cointegration, there are only the short run effects, and the ARDL model is modified to a form without the long run estimates. On the other hand, if the hypothesis is rejected, cointegration is present, and the Error Correction Model (ECM) is defined to capture the short run effects:

$$\Delta lnY_t = \beta_0 + \sum_{i=1}^{n} \theta_i \Delta lnY_{t-i} + \sum_{i=0}^{m} \varepsilon_i \Delta lnGDP_{t-i} + \sum_{i=0}^{o} \vartheta_i \Delta IR_{t-i} + \sum_{i=0}^{p} \rho_i \Delta lnOil_{t-i} + \sum_{i=0}^{q} \pi_i \Delta lnREER_{t-i} + \sum\nolimits_{i=0}^{r} \sigma_i \Delta lnAM_{t-i} + \sum_{j=1}^{s} \alpha_j DM_j + \varphi ECT_{t-1} + \mu_t$$

Equation 4

It is called "Error correction" because it is adjusted to the impact that the variables have in the long run. These are explained by a new variable - the Error Correction Term (ECT) which is a vector of residuals obtained from the long run model (Equation 5) lagged by one. This term explains how fast the dependent variable converges to the equilibrium in the long run (i.e., it measures the speed of adjustment against the long run equilibrium). Its impact is supposed to be significant, and the value of coefficient φ should be between 0 and -1. If it is positive, it diverges from the equilibrium in the long run.

$$lnY_t = \gamma_0 + \gamma_1 lnGDP_t + \gamma_2 IR_t + \gamma_3 lnOil_t + \gamma_4 lnREER_t + \gamma_5 lnAM_t + \gamma_6 lnDis_{t-1} + \sum_{j=1}^{s} \alpha_j DM_j + \epsilon_t$$

Equation 5

If the cointegration is present, the long run estimates are relevant and obtained from the original ARDL model (Equation 3). For independent variable *i*, the long run estimate *λ* is computed in this way:

$$\lambda_i = -\frac{\beta_i}{\beta_1}$$

Equation 6

5.2. Diagnostics test and multicollinearity

The final model must satisfy conditions before interpreting results to ensure the estimates are relevant. Together there are three diagnostics tests regarding residuals, one test for specification and one test for multicollinearity. All tests are made at 5% level of significance.

The first condition is no autocorrelation of residuals which is tested by the Durbin-Watson test. The null hypothesis is that autocorrelation is not present. Thus, if the p-value is greater than 0.05, the hypothesis cannot be rejected. The homoskedasticity of residuals is tested by Breusch Pagan test. The null hypothesis of homoskedasticity is rejected if the p-value is lower than 0.05. The Shapiro-Wilk test tells us whether the residuals are normally distributed. The null hypothesis of normality is rejected if the p-value is lower than 0.05. Another test to be made is the Ramsey Regression Equation Specification Error Test (RESET) which tells us whether there is a problem with the specification of the model. The null hypothesis of no misspecification is rejected it the p-value is lower than 0.05.

The last condition to check is whether there is a problem with multicollinearity. ARDL model supposes cointegration between variables, thus is less sensitive to multicollinearity. To check whether it is not a serious problem, the Variance inflation factor (VIF) is calculated for each explanatory variable in the model. It measures how big is the impact of collinearity of variables. VIF value for the i-th variable is calculated as:

$$VIF_i = \frac{1}{1 - R_i^2}$$

Equation 7

According to Gujarati (2004), the multicollinearity is not a problem in ARDL model as long as for no variable the VIF value exceeds 10, otherwise, the model has to be adjusted by removing lags or the whole variable.

6. Results

6.1. ICO composite indicator

The first step of estimating the model is to check all that all variables are integrated of order 0 or 1. If this condition is violated, the ARDL method cannot be used. The results of ADF test checking for unit root is displayed in Table 10 in Appendix B. Some of the variables are not stationary at the level form, the corresponding p-value is bigger than 0.05 and the null hypothesis of unit root is not rejected. However, all first differences are stationary. None of the variable is integrated of order 2 or higher, so the condition is satisfied.

Firstly, we specified ARDL models Equation 2 based on findings of researchers who dealt with prices of agricultural commodities. The results are in Table 11 in Appendix B. Model 1 uses the same variables as model by Baffes et. al. (2016) who studied causes of price movements of agricultural commodities. This paper did not deal with outliers, but for coffee prices two outliers are identified – years 1976 and 1994. In both situations the coffee price increased rapidly which might cause bias of coefficients. Since we have time series data, these variables should not be omitted. The solution is to add dummy variables for these cases (Hyndman; 2018) and the result is Model 2. The variable yr76 is equal to 1 for year 1976 and 0 for other observations, yr94 is defined analogically. Since the coefficient of variable yr94 is significant, this one should be included. Model 3 is restricted version of Model 2 where the most insignificant variables, those with the highest p-value, were removed to avoid overfitting the model. The F statistics for joint significance of the dropped variables was low, thus the null hypothesis was not rejected and removing these variables does not cause problems with the model estimation. Moreover, the AIC value is lower for the restricted model which means it is better (Aikaike; 1974).

The next step is to check whether models satisfy given assumptions. Neither of homoskedasticity, normality and no autocorrelation of residuals is violated; however, results of VIF indicate complications. VIF value for stock to use ratio (sur1) and GDP (gdp1) is between 16-21 for all three models so these variables suffer from multicollinearity. One possible solution is to remove more lags, but it cannot be applied here, since the all lags are significant and excluding them would cause violation of autocorrelation. To solve this problem, we introduce similar model with two different variables. The first one is GDP growth which describes economic well-being and how fast the economy is growing, negative GDP growth suggests the economy suffers from recession. The second is the amount of coffee exported which substitutes stock to use ratio. Both variables were discussed in Chapter 3. There is enough of evidences in previous studies, that these can describe the behaviour of commodities' prices.

The results of the new model, with GDP growth and exports, are summarized in Table 4. As well as the real interest rate, GDP growth is in the level form and not in logarithmic, because it was negative for some observations.

Table 4 - ARDL model with exports

	Dependent variable:
	Δln(ICO composite indicator)_t
	ICO indicator model
Number of lags	4, 4, 4, 1, 2, 2
Δln(ICO)_1	-0.127
	(0.085)
Δln(ICO)_2	0.262^{***}
	(0.081)
Δln(ICO)_3	0.483^{***}
	(0.083)
ΔGDP growth_1	0.043^{***}
	(0.015)
ΔGDP growth_3	-0.043^{***}
	(0.013)
Δln(REER)_t	0.713^{*}
	(0.411)
Δln(REER)_3	-1.595^{***}
	(0.533)
Δln(Oil)_t	-0.003
	(0.069)
ΔInterest rate_1	0.076^{**}
	(0.032)
Δln(Exports)_t	-1.689^{***}
	(0.299)
yr76_t	0.484^{***}

	(0.134)
yr94_t	0.589***
	(0.135)
ln(ICO)_1	-0.297***
	(0.069)
GDP growth_1	-0.046**
	(0.019)
ln(REER)_1	1.390***
	(0.285)
ln(Oil)_1	0.071*
	(0.041)
Interest rate_1	-0.063***
	(0.018)
ln(Exports)_1	-0.659***
	(0.172)
Constant	1.276
	(2.179)
Observations	50
R^2	0.870
Adjusted R^2	0.794
AIC (original)	-52.314
AIC (restricted)	-56.026
Residual Std. Error	0.118 (df = 31)
F Statistic	11.501*** (df = 18; 31)
Note:	*p**p***p<0.01

Results of the diagnostics test for the model are summarized in Table 13 in Appendix B. There is no problem with characteristics of autocorrelation and multicollinearity. Regarding results for cointegration, the calculated F statistics were compared to critical values introduced by Narayan (2005), see Table 12 in Appendix B. We consider values for 5% significance. The results it the cointegration is present. This means that the variables have a significant impact in the long run and the original model must be adjusted to obtain significant short run estimates.

Now we specify the Error Correction Model (ECM) to see the short run effects. The results are shown in Table 5. The coefficient of ECT1 is significant and negative which confirms presence of cointegration in the long term. Another important point is, that impacts of dummy variables (outliers) are significant, which makes the rest of estimates more accurate.

Table 5 - Error Correction Model for ICO price

	Dependent variable
	Δln(ICO composite indicator)_t
	ICO indicator model
Δln(ICO)_1	-0.134
	(0.114)
Δln(ICO)_2	0.143
	(0.105)
Δln(ICO)_3	0.352***
	(0.108)
ΔGDP growth_1	0.008
	(0.018)
ΔGDP growth_3	-0.036**
	(0.018)
Δln(REER)_t	0.093
	(0.508)
Δln(REER)_3	-0.442
	(0.541)
Δln(Oil)_t	-0.053
	(0.086)
ΔInterest rate_1	0.061
	(0.040)
Δln(xports)_t	-1.501***
	(0.411)
yr76_t	0.708***
	(0.173)
yr94_t	0.628***
	(0.173)
ECT_1	-0.309***
	(0.094)
Constant	-0.008
	(0.025)
Observations	50
R^2	0.708

Adjusted R^2	0.603
Residual Std. Error	0.163 (df = 36)
F Statistic	6.718*** (df = 13; 36)

Note:

We compute the long run estimates from the ARDL model in Table 4. Long run effects of variables are computed as was previously defined (Equation 6). Results are shown in Table 6. All variables have a significant impact in the long run.

Table 6 - Long run coefficients

	ICO indicator model
GDP growth	-0.1548
REER	4.6801
Oil price	0.2390
Interest rate	-0.2121
Exports	-2.2188

Now we interpret the results of the estimated model. Variables, which have a significant impact in the short run on coffee price are GDP growth and coffee exports (Table 5). Annual growth of GDP has negative impact in the short run, specifically one unit change decreases the coffee price by 3.6%. Exports have negative impact as well – change in one unit leads to decrease by 1.501 point. These results correspond with our expectations. Higher income usually results in lower marginal propensity to consume (MPC)[4] and lower exports indicate higher prices.

Regarding long run effects, the coefficient of lagged ECT by one is 0.309. This tells us the system corrects the previous period disequilibrium at a speed of 30.9% per year to reach the long run equilibrium. Both growth of GDP and exports have the impact

[4] Engel's law – as income increases, the proportion of money spent on consumption decreases

on prices as they have in the short run, a negative one, and the absolute value of the coefficients is bigger than in the short term. This means that these variables are not only significant, but also their power is noticeable over time. The reason why the financial variables have no impact in the short run is, that the commodity's price is not able to react quickly to changes in the currency's power and stability.

When we focus on the long run effects, there is a positive relationship between crude oil price and agricultural commodities' prices as expected and was proven already in previous studies (Nazliogu, Sagan; 2011; Maurice, Davis; 2011). It is relevant because crude oil price reflects costs of production which drives coffee price. The elasticity estimate of real effective exchange rate equals 4.6801 and it has explanation. This variable signifies economic situation of the United States which are important part of the global trade. When the REER increases, it means that dollar is stronger and the exports to the US are favourable for exporters. In the long term this results in broader and more balanced global trade. This claim is supported by Ayres (2020) who concluded there is a positive relationship between primary commodity prices and the exchange rate of developed countries, which the United States is. The interest rate affects price negatively in the long run which corresponds with previous findings by Akram (2009) and Frankel (2006). This effect is common in the long run, because with lower interest rate, the real value of holding money is lower, thus investments decreases and consumption increases.

Alongside with this model, we also estimated two alternative models with other variables describing coffee supply. The first is stock to use ratio which was used in previous studies and models (Model1, Model2, Model 3 in Table 11) and the second is the total production of coffee. All results are in Appendix B (Table 14, Table 15, Table 16 and Table 17). Both models passed diagnostics and multicollinearity tests and the cointegration is present. The reason why we estimated the one with the stock to use ratio is to determine, whether this variable is sufficient despite absence of the real GDP in the model. When we focus on short run, the only variable which appears to be significant is the real interest rate which affects the dependent variable positively. This is explained by the fact, that this variable reflects the development of business cycle. Regarding the long run, the stock to use ratio has a positive impact on the dependent variable, which is opposite to the expectations and findings made in similar papers supported by economic theory. One of the possible reasons might be very high correlation between stock to use ratio and crude oil prices which exceeds value 0.8 (Table 3), which might cause the misspecification. Except for crude oil, the rest of variables have significant impact on the price in the long run and signs of all coefficients are as expected. In conclusion, despite all conditions were satisfied, this model did not provide relevant results.

The second alternative model used production as the variable describing quantity of coffee supplied. For this one, in the short run, all variables were significant except from real effective exchange rate (Table 16). An increase in GDP growth by a unit lead to a decrease of the dependent variable by 4-5% which is comparable to the 3.6% from the first model. Both crude oil prices and real interest rate affect the price positively. Although the production has a positive impact on the price, which is against the supposition, the impact is negligible, because its elasticity coefficient is 0.00001. This probably happened because of high correlation between crude oil prices and production (see Table 3). The error correction term indicates, the price corrects by 34.9% per year which is a little faster than in the model we interpreted. When we look at the long run effects (Table 17), we can see, that the GDP growth is not significant. Other variables are significant, and their impact is as expected. Since this thesis is focused on the global coffee prices, we believe, that the variable exports is the best option for proxying coffee supplied. The model which uses production as an explanatory variable might be more suitable in a situation when focusing on the behaviour of coffee price in one country.

Baffes et. al. (2016) suggested that the model they used could be improved by adding variables describing climate conditions, specifically, number of cases of floods, droughts and extreme temperatures over year. Table 18 in Appendix B presents ARDL model with same variables as our model with exports with an additional variable: the number of natural disasters. This model did not pass multicollinearity test because the new indicator is correlated with the GDP growth and exports of coffee. The value of VIF for variable disasters is 11.6572 which is higher than 10.

6.2. *Price of arabica and robusta*

In this part we discuss how the effects of the factors found vary for two coffee types, arabica and robusta. Then we compare them with findings from the previous subchapter. We will use the same model as before. The data are the same as for the previous part, summarized in Table 1.

Firstly, we need to assure that all variables are integrated of order 0 or 1. Results of ADF test are in Table 10 in Appendix B. Again, there is a mixture of variables which are I(0) and I(1). The condition is satisfied so we can use the ARDL method.

The results of models for both types are in Table 7. Number of lags was chosen by AIC and then some insignificant lags were removed to improve the model and avoid overfitting. The AIC values of restricted models are lower for both coffee types, which suggests these models to be better than the original ones.

As in the previous case, there are some outliers. For arabica prices these are the same as for the ICO composite price index – years 1976 and 1994. Robusta prices suffer from an additional outlier – year 1978, when the prices fell rapidly. Thus, we have two, respectively three dummy variables and they all appeared to be significant.

Table 7 - ARDL model for arabica and robusta

	Dependent variable: Δln(Price)_t	
	Arabica	Robusta
Number of lags	4, 4, 1, 2, 2, 1	4, 3, 1, 4, 4, 1
Δln(Price)_1	-0.005	0.087
	(0.093)	(0.093)
Δln(Price)_2	0.275***	0.004
	(0.095)	(0.088)
Δln(Price)_3	0.405***	0.106
	(0.092)	(0.090)
ΔGDP growth_2	-0.050***	0.027
	(0.015)	(0.017)
ΔGDP growth_3	-0.051***	
	(0.014)	
Δln(REER)_t	0.586	-0.138
	(0.484)	(0.453)
Δln(Oil)_1	-0.217**	
	(0.092)	

Δln(Oil)_2		-0.209^{**}
		(0.102)
Δln(Oil)_3		0.377^{***}
		(0.084)
ΔInterest rate_t	0.067^{**}	0.066^{**}
	(0.031)	(0.029)
ΔInterest rate_1		0.126^{***}
		(0.033)
ΔInterest rate_3		0.070^{**}
		(0.033)
Δln(Exports)_t	-1.375^{***}	-0.323
	(0.295)	(0.270)
yr76_t	0.441^{**}	1.002^{***}
	(0.166)	(0.197)
yr78_t		-0.418^{**}
		(0.164)
yr94_t	0.532^{***}	0.744^{***}
	(0.168)	(0.148)
ln(Price)_1	-0.419^{***}	-0.290^{***}
	(0.089)	(0.071)
GDP growth_1	0.025	0.019
	(0.021)	(0.019)
ln(REER)_1	0.705^{**}	0.582^{**}
	(0.279)	(0.257)
ln(Oil)_1	0.172^{***}	0.058
	(0.053)	(0.040)
Interest rate_1	-0.036^{*}	-0.090^{***}
	(0.020)	(0.025)
ln(Exports)_1	-0.974^{***}	-0.464^{***}
	(0.249)	(0.148)
Constant	7.182^{***}	2.088
	(2.579)	(1.569)
Observations	50	50
R^2	0.826	0.871
Adjusted R^2	0.734	0.781
AIC (original)	-31.82	-44.92
AIC (restricted)	-37.41	-47.83
Residual Std. Error	0.142 (df = 32)	0.127 (df = 29)
F Statistic	8.939^{***} (df = 17; 32)	9.751^{***} (df = 20; 29)
Note:		*p**p***p<0.01

Before interpretation of results, we need to check diagnostics test, multicollinearity and cointegration. Table 19 in Appendix B provides results of all of these. Both models satisfy all tests and there is no problem with multicollinearity. The values of F-statistics are compared to critical values (Table 12). Both are higher than the upper bound value, which means the cointegration is present.

After saving the residuals from long run model, the short run model is estimated (Table 8). The coefficient of the lagged value of ECT is negative and significant for both types, which confirms presence of cointegration in the long run. The price of arabica converges to equilibrium faster, its speeds of adjustment is 41.8% per year and Robusta's 24.4% per year. As in the previous case, all dummy variables are significant.

There are only two variables, which are significant for both types in the short run, price of crude oil and the interest rate. The interest rate has positive impact on both between 6-10%. Despite this result seeming to be incorrect, Akram (2009) acknowledged the effect of interest rate on commodity prices is not unambiguous in the short run. Pedersen (2015) concluded, that the interest rate reaction was positively related to copper price shocks. In this case, the positive relationship between these two variables occurs because of business cycle shocks which reflects in the short run. This result supports finding by Baffes et. al. (2006) who found a positive relationship in the short run between real interest rate and commodity prices as well. The crude oil price affects price of both types significantly however the impact differs. Robusta is affected as expected, there a positive impact of third lag of crude oil prices on the dependent variable. On the other hand, the relationship between first lag of oil prices and arabica price is negative. The reason for this might be that it is almost an immediate effect and because crude oil price reflects market situation, thus is partly related to growth of the GDP, which has negative impact.

The growth of GDP and amount of exports have a negative significant impact only on arabica prices in the short run. In comparison to arabica, robusta is less demanded by big coffee companies, which might be the reason why neither of these is significant to its price movements. Overall, effects on arabica price are captured better than on robusta price.

Table 8 - ECM for price of arabica and robusta

	Dependent variable: Δln(Price)_t	
	Arabica	Robusta
Δln(Price)_1	-0.051	0.106

	(0.086)	(0.100)
Δln(Price)_2	0.248**	0.020
	(0.092)	(0.099)
Δln(Price)_3	0.368***	0.140
	(0.085)	(0.089)
ΔGDP growth_2	-0.049***	0.013
	(0.014)	(0.017)
ΔGDP growth_3	-0.051***	
	(0.014)	
Δln(REER)_t	0.309	-0.343
	(0.442)	(0.487)
Δln(Oil)_1	-0.205**	
	(0.082)	
Δln(Oil)_2		-0.137
		(0.106)
Δln(Oil)_3		0.304***
		(0.093)
ΔInterest rate_t	0.096***	0.101***
	(0.028)	(0.028)
ΔInterest rate_1		0.066**
		(0.031)
ΔInterest rate_3		0.020
		(0.031)
Δln(Exports)_t	-1.108***	-0.330
	(0.258)	(0.283)
yr76_t	0.496***	0.899***
	(0.157)	(0.221)
yr78_t		-0.416**
		(0.182)
yr94_t	0.486***	0.705***
	(0.156)	(0.157)
ECT_1	-0.418***	-0.244***
	(0.087)	(0.078)
Constant	-0.005	-0.028
	(0.022)	(0.023)
Observations	50	50
R^2	0.798	0.801
Adjusted R^2	0.732	0.713
Residual Std. Error	0.143 (df = 37)	0.145 (df = 34)
F Statistic	12.171*** (df = 12; 37)	9.111*** (df = 15; 34)
Note:		*p**p***p<0.01

In Table 9 are long run effects of factors on the prices (Equation 6). The growth of GDP has no significant impact on both types. The real effective exchange rate affects both positively and has bigger impact on robusta. As was already mentioned, higher REER results in an advantage for those who export to the United States, because dollar is stronger. This support the world trade balance and prices adjustments. The effect of interest rate is, as expected, negative for both types and it has bigger impact on robusta prices as well. Regarding exports, there is a negative relationship and bigger impact on prices of arabica, the elasticity estimate is -2.3246 compared to robusta's -1.6. The estimate of oil price elasticity is significant only for arabica and the impact is positive as expected.

Table 9 - Long run estimates for models for prices of arabica and robusta

	Arabica	Robusta
GDP growth_1	x	x
REER_1	1.6826	2.007
Oil price_1	0.4105	x
Interest rate_1	-0.0859	-0.3103
Exports_1	-2.3246	-1.6

Note: The "x" means, that the variable is not significant

Overall, we can conclude, that robusta's price is more affected by financial variables – those, which are related to the power and stability of the US dollar. On the other hand, arabica is more driven by exports and GDP growth, which explain the amount of coffee available for trade, the demand and economic situation. The crude oil price affect price of robusta only in the short run with a lag. In this case it is probably captured as the cost of production. However, arabica is affected in the long run, here the variable probably reflects the economic activity, as we explained in the Chapter 3.

Conclusion

This thesis aimed to identify the driving factors of coffee price and to suggest policies that would regulate them to control the price. Some studies dealt with coffee price, but most of them discussed only one type of coffee or the average coffee price. This thesis describes the difference between the results for the ICO composite indicator (the average coffee price) and the results for prices of the two coffee varieties – arabica and robusta – defined separately. In the theoretical part, we introduced coffee as a plant and the difference between the two coffee types. Then we defined five groups of factors which may affect the price movements. Then we used the ARDL method to estimate a model to detect which of the factors have an impact on the price for both the short and long run. The dependent variable was the change of real coffee price between two years, and there were five explanatory variables, GDP growth, real effective exchange rate, real crude oil price, real interest rate and exports. We also included dummy variables for outliers, i.e. years, when the change in price was the biggest. The dummies appeared to have a considerable impact on the model specification.

We discovered that in all cases the cointegration was present, which means the factors have an impact in both the short run and the long run. Arabica's price was driven mainly by the supply and demand factors: exports, GDP growth and crude oil price. On the other hand, price of robusta is mostly affected by financial variables which capture the power of the US dollar. Moreover, the results suggested, that the price of arabica adjusts to the long run equilibrium faster than the price of robusta. This means that arabica's price is more sensitive to long run effects of the factors. These findings confirm the H2 we specified in the introduction. Regarding the impact of the variables on the ICO composite indicator, it is similar to the effects of the variables on arabica prices. This is reasonable, as the share of arabica on the global is larger. All the mentioned factors have a significant impact and results we obtained correspond with conclusions delivered by the previous studies and intuition, described in Chapter 3. They also support the H1 stated in the introduction, which says that coffee price is driven mainly by the coffee supply (exports) and by the real interest rate of the US dollar.

In the beginning, we considered the possibility of adding a variable describing weather conditions – number of specific natural disasters. Unfortunately, this appeared to be highly correlated to the GDP growth and to exports, which made the model insufficient. We also concluded that an alternative model, where exports are substituted by production, provided relevant results. However, it would be more suitable for the situation when observing coffee price in a specific country only and not at the global level.

Based on these findings, we propose policies which would regulate some factors, thus make the coffee price more predictable. One factor which affects the price significantly is exports, which have a negative impact. The solution for this would be forming a similar pact as the Association of Coffee Producing Countries which collapsed due to the absence of Vietnam and due to the problem with free riders. The new association would gather all the biggest coffee producers and set policies which would regulate coffee exports. Because exports are highly dependent on production, another solution would be stabilising the price by implementing buffer stocks scheme in the biggest coffee producers. The idea is buying stocks when the supply of the commodity is high and releasing them when the supply is low to prevent prices from falling, increasing respectively.

The other factors we discovered, have significant economic influence and therefore cannot be modified only for the sake of controlling the coffee price. However, to some degree, we can predict how would the coffee prices behave after a change in these factors. For example, due to the pandemic COVID-19 in the first half of 2020, the global welfare decreased because of the arrangements implemented to reduce the risk of spreading the virus that inevitably lead to global economic lock-down. There is an expected major fall in the GDP growth comparing to the previous years. According to the model, this event causes the real price of coffee to increase. Generally, if the producers understand the coffee price factors properly, they may better control the price through the export controls and cooperation.

Regarding the model specification, we are aware that there are more variables that affect the price of coffee significantly. For example, these are the variables explaining climate conditions and other financial variables, e.g. stocks or other market indices. On the other hand, this model provides the basic concept of how the price behaves and can be applied for similar agricultural commodities, for example, tea or cocoa. Moreover, unlike previous researches on price movements, this model deals with the outliers, which make the results more precise.

Bibliography

Ahumada, Hildegart, and Magdalena Cornejo, (2016). "Forecasting food prices: The case of corn, soybeans and wheat." International Journal of Forecasting 32.3: 838-848.

Akaike, Hirotugu, (1974). A new look at the statistical model identification." *IEEE transactions on automatic control* 19.6: 716-723.

Akiyama, T; Duncan, R. C; Akiyama, T*Duncan, R. C. 1982. Analysis of the world coffee market (English). World Bank staff commodity paper ; no. SCP 7. Washington, DC : The World Bank. http://documents.worldbank.org/curated/en/763401468739518610/Analysis-of-the-world-coffee-market

Akram, Q. Farooq, (2009). Commodity prices, interest rates and the dollar. *Energy economics* 31.6 (2009): 838-851.

Algieri, Bernardina, (2014). A roller coaster ride: an empirical investigation of the main drivers of the international wheat price. *Agricultural economics* 45.4: 459-475.

Almeida, Fernanda Maria de, Marilia Fernandes Maciel Gomes, and Orlando Monteiro da Silva, (2012). Non-tariff measures in international coffee trade. No. 1007-2016-79480.

Amrouk, El Mamoun, (2018). Depressed international coffee prices: Insights into the nature of the price decline. Accessed: 23.05.2020 Available at: http://www.fao.org/fileadmin/templates/est/COMM_MARKETS_MONITORING/Coffee_Cocoa/Documents/coffee_prices_2018.pdf

Ayres, Joao, Constantino Hevia, and Juan Pablo Nicolini. "Real exchange rates and primary commodity prices." Journal of International Economics 122 (2020): 103261.

Baffes, John, Bryan Lewin, and Panos Varangis, (2005). "Coffee: market setting and policies." Global agricultural trade and developing countries. 297-310.

Baffes, John, and Tassos Haniotis, (2016a). "What explains agricultural price movements?." *Journal of Agricultural Economics* 67.3: 706-721.

Baffes, John, and Xiaoli L. Etienne, (2016b). "Analysing food price trends in the context of Engel's Law and the Prebisch-Singer hypothesis." *Oxford Economic Papers* 68.3: 688-713.

Bank of International Settlements. (2018). Exchange rates statistics. Available at: https://www.bis.org/statistics/tables_i_eer.pdf [Accessed 2020-06-23]

BBC News, (2011). Brazil floods: 'Never seen anything like it' Available at: https://www.bbc.com/news/world-latin-america-12189712 [Accessed 2020-03-14]

Berhane, Tesfahun, et al., (2018). Forecasting the Ethiopian Coffee Price Using Kalman Filtering Algorithm. Journal of Resources and Ecology 9.3: 302-305.

Boydell, H., (2018). What Should We Have Learned From The 2001 Coffee Price Crisis? Available at: https://www.perfectdailygrind.com/2018/09/what-should-we-have-learned-from-the-2001-coffee-price-crisis/ [Accessed 2020-03-16]

Chen, J., (2019). Coffee, Sugar and Cocoa Exchange – CSCE . Investopedia. Available at: https://www.investopedia.com/terms/c/csce.asp [Accessed 2020-03-16]

Coffee and Health. Where Coffee Grows. Institute for scientific information on coffee. Available at: https://www.coffeeandhealth.org/all-about-coffee/where-coffee-grows/ [Accessed 2020-06-26]

Cortazar, Gonzalo, et al., (2019). "Commodity price forecasts, futures prices, and pricing models." Management Science 65.9: 4141-4155.

Cuaresma, Jesus Crespo, Jaroslava Hlouskova, and Michael Obersteiner, (2018). Fundamentals, speculation or macroeconomic conditions? Modelling and forecasting Arabica coffee prices. European Review of Agricultural Economics 45.4: 583-615.

DaMatta, Fábio M., and José D. Cochicho Ramalho, (2006). Impacts of drought and temperature stress on coffee physiology and production: a review. *Brazilian Journal of Plant Physiology* 18.1: 55-81.

Doan, Lynn a Marvin Perez, (2014). Coffee Cravers Ignoring Bean-Price Surge for Caffeine Fix. *Bloomberg*. Available at: https://www.bloomberg.com/news/articles/2014-03-12/coffee-cravers-ignoring-bean-price-surge-for-caffeine-fix [Accessed 2020-06-09]

Fairtrade Foundation. Coffee Farmers. Available at: https://www.fairtrade.org.uk/Farmers-and-Workers/Coffee [Accessed 2020-03-14]

Frankel, Jeffrey A., (2006). *The effect of monetary policy on real commodity prices*. No. w12713. National Bureau of Economic Research

Gujarati, D., (2004). *Basic Econometrics*, 4th edition, MacMillan, New York

Harrington, Dan (2011). Coffee Prices 2011 – The Rising Price of Coffee. Available at: http://www.gourmetcoffeelovers.com/coffee-prices-2011-the-rising-price-of-coffee/ [Accessed 2020-07-03]

Harvey, David I., et al., (2017). Long-run commodity prices, economic growth, and interest rates: 17th century to the present day. *World Development* 89: 57-70.

Hopp, H., & Foote, R. J., (1955). A statistical analysis of factors that affect prices of coffee. Journal of Farm Economics, 37(3), 429-438

Hyndman, Rob J., and George Athanasopoulos, (2018). *Forecasting: principles and practice*. Otexts

International Coffee Organization, (2014). World coffee trade (1963 – 2013): A review of the markets, challenges and opportunities facing the sector. ICC (International Coffee Council) Available at: http://www.ico.org/news/icc-111-5-r1e-world-coffee-outlook.pdf [Accessed 2020-03-16]

International Coffee Organization, (2020a). ICO Annual Review 2018/19. Available at: http://www.ico.org/documents/cy2019-20/annual-review-2018-19-e.pdf [Accessed 2020-06-26]

International Coffee Organization, (2020b). Imports by selected importing countries. Available at: http://www.ico.org/historical/1990%20onwards/PDF/2b-imports.pdf [Accessed 2020-06-26]

International Coffee Organization, (2020c). Total production by all exporting countries. Available at: http://www.ico.org/historical/1990%20onwards/PDF/1a-total-production.pdf [Accessed 2020-06-26]

International Trade Centre, ICO Indicator Prices, Available at: http://www.intracen.org/coffee-guide/world-coffee-trade/ICO-indicator-prices/ [Accessed 2020-04-05]

Ji, Qiang, (2012). System analysis approach for the identification of factors driving crude oil prices. Computers & Industrial Engineering 63.3: 615-625.

Karthikeyan, K. and Harlalka, A., (2014). Robust Regression Model for Prediction of Soybean Crop Price Based on Various Factors. International Journal of Emerging Technologies in Computational and Applied Sciences, 14, 258- 263

Keatinge, Fergus JD, (2015). Influential Factors in the Econometric Modeling of the Price of Wheat in the United States of America. *Agricultural Sciences* 6.08 (2015): 758.

Lewin, Bryan, Daniele Giovannucci, and Panos Varangis, (2004). "Coffee markets: new paradigms in global supply and demand." World Bank Agriculture and Rural Development Discussion Paper 3.

Malan B.B., (2013). Volatility and stabilization of the price of coffee and cocoa in Côte d'Ivoire. Agric. Econ. – Czech, 59: 333-340.

Maltsbarger, Robert Dále, (2011). A structural model of the international coffee sector: an econometric investigation. Diss. University of Missouri—Columbia.

Markgraf, Vera, (2001) ed. *Interhemispheric climate linkages*. Elsevier

Miao, Hong, et al., (2017). Influential factors in crude oil price forecasting. Energy Economics 68: 77-88.

Mounts Brian, 2018. Arabica vs Robusta Coffee: Why Are Arabica Beans Better Anyway [online] Available at: https://ggccoffee.com/arabica-vs-robusta/ [Accessed 2020-03-12]

Narayan, P. K., (2005). The saving and investment nexus for China: evidence from cointegration tests. *Applied economics*, *37*(17), 1979-1990.

Natanelov, Valeri, et al., (2011). Is there co-movement of agricultural commodities futures prices and crude oil?. Energy Policy 39.9: 4971-4984.

National Coffee Association. Coffee Around the World. NCAUSA. Available at: https://www.ncausa.org/About-Coffee/Coffee-Around-the-World [Accessed 2020-02-19]

National Coffee Association, (2016). The NCA USA Economic Impact Report. NCAUSA. Available at: https://www.ncausa.org/Industry-Resources/Economic-Impact [Accessed 2020-06-26]

National Coffee Association, (2020). 10 Steps from Seed to Cup. NCAUSA Available at: https://www.ncausa.org/About-Coffee/10-Steps-from-Seed-to-Cup [Accessed 2020-06-26]

Naveena, K., et al., (2017). Hybrid ARIMA-ANN modelling for forecasting the price of Robusta coffee in India. Int. J. Curr. Microbiol. Appl. Sci 6.7: 1721-1726.

Nazlioglu, Saban, (2011). World oil and agricultural commodity prices: Evidence from nonlinear causality. Energy policy 39.5: 2935-2943.

Pedersen, Michael, (2015). The impact of commodity price shocks in a major producing economy. The case of copper and Chile. No. 753. Central Bank of Chile.

Pesaran, M. Hashem, Yongcheol Shin, and Richard J. Smith, (2001). Bounds testing approaches to the analysis of level relationships. *Journal of applied econometrics* 16.3: 289-326.

Russell, Bill, Sushil Mohan, and Anindya Banerjee, (2012). "Coffee market liberalisation and the implications for producers in Brazil, Guatemala and India." The World Bank Economic Review 26.3: 514-538.

Teuber, Ramona, and Roland Herrmann, (2012). Towards a differentiated 48odelling of origin effects in hedonic analysis: An application to auction prices of specialty coffee. *Food Policy* 37.6: 732-740.

Tothmihaly, Andras, (2017). How low is the price elasticity in the global cocoa market?. No. 102. GlobalFood Discussion Papers.

Wooldridge, Jeffrey M., (2016). Introductory econometrics: A modern approach. Nelson Education

World Integrated Trade Solution, (2020). Trade Summary for Ethiopia (Exludes Eritrea) 2015. World Bank. Available at: https://wits.worldbank.org/CountryProfile/en/Country/ETH/Year/2015/Summarytext [Accessed 2020-06-26]

World Trade Organization, History of the multilateral trading system Available at: https://www.wto.org/english/thewto_e/history_e/history_e.htm [Accessed 2020-03-15]

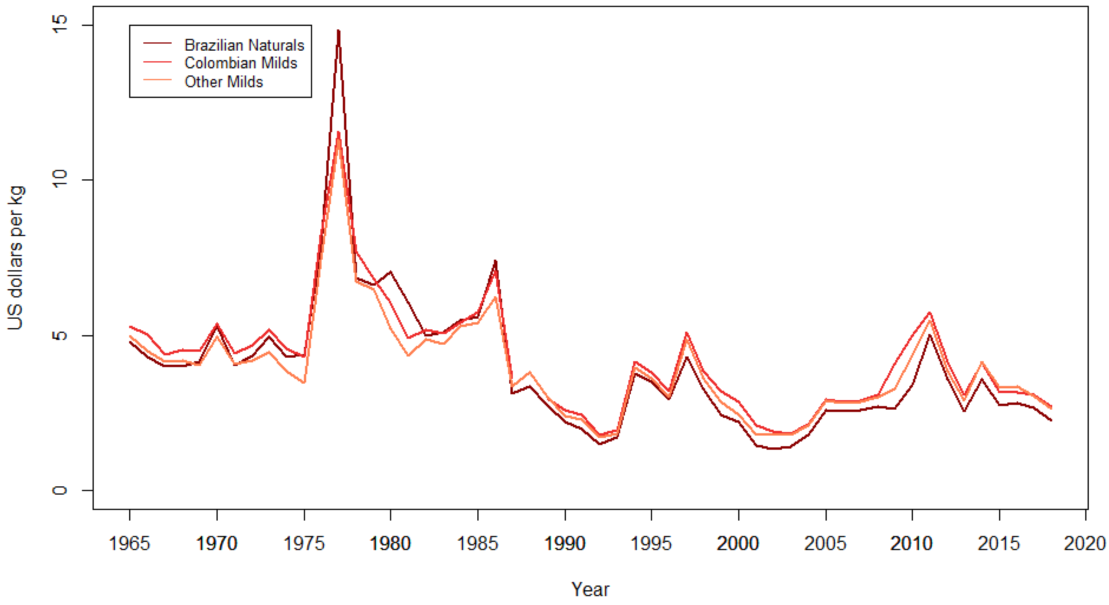
Yang Jia Lynn (2014). Worst drought in decades hits Brazil coffee belt as buyers brace for price rise. Available at: https://www.theguardian.com/world/2014/feb/25/brazil-drought-threatens-coffee-crops [Accessed 2020-07-03]

Yifru, T., (2015). Impact of Agricultural Exports on Economic Growth in Ethiopia: The Case of Coffee, Oilseed and Pulses (No. 243473). Collaborative Master's Program in Agricultural and Applied Economics.

Yohannes, Michael Fesseha, Toshinobu Matsuda, and Naoko Sato, (2016). "Substitution in Consumer Demand for Coffee Product Categories in Japan." *Journal of Agricultural Science* 8.4: 50.

Appendix A

Figure 5 – Evolution of arabica's types real prices



Data source: ICO

Appendix B

Table 10 - Results of ADF test

	Level form	First difference	
Variable	*P-value*	*P-value*	*Result*
ln(ICO)	0.2569	0.0219	I(1)
ln(GDP)	0.0398	<0.01	I(0)
GDP growth	0.0332	0.02342	I(0)
ln(REER)	<0.01	0.0140	I(0)
ln(Oil)	0.5565	<0.01	I(1)
Interest rate	0.3712	<0.01	I(1)
ln(Stock to use ratio)	0.5888	<0.01	I(1)
ln(Exports)	0.1694	<0.01	I(1)
ln(Production)	0.4985	<0.01	I(1)
ln(Disasters)	0.9287	<0.01	I(1)
ln(Price_arabica)	0.1656	<0.01	I(1)
ln(Price_robusta)	0.2594	<0.01	I(1)
ln(Exports_arabica)	0.2528	<0.01	I(1)
ln(Exports_robusta)	0.2530	0.03745	I(1)

Note: The lag length of ADF test was selected by AIC

Table 11 - ARDL model with original variables

	Dependent variable:		
	Δln(ICO)_t		
	Model 1	Model 2	Model 3
Number of lags	4,4,4,1,2,1	4,4,4,1,2,1	4,4,4,1,2,1
Δln(ICO)_1	-0.019	-0.017	-0.019
	(0.132)	(0.115)	(0.106)
Δln(ICO)_2	0.079	0.138	0.087
	(0.117)	(0.103)	(0.093)
Δln(ICO)_3	0.306^{**}	0.297^{**}	0.241^{**}
	(0.129)	(0.118)	(0.108)
Δln(GDP)_t	-1.927	-2.987	-2.138
	(2.739)	(2.417)	(2.211)
Δln(GDP)_1	-1.912	0.649	
	(2.415)	(2.485)	
Δln(GDP)_2	-11.070^{***}	-10.395^{***}	-10.573^{***}
	(2.493)	(2.184)	(2.040)
Δln(GDP)_3	-6.424^{**}	-4.295^{*}	-3.686^{*}
	(2.783)	(2.474)	(2.168)
Δln(REER)_t	0.350	0.387	
	(0.693)	(0.625)	

Δln(REER)_1	-0.256	-0.371	
	(0.848)	(0.749)	
Δln(REER)_2	-1.263	-0.811	-0.960
	(0.784)	(0.715)	(0.612)
Δln(REER)_3	-2.320**	-2.278***	-2.119***
	(0.860)	(0.756)	(0.638)
Δln(Oil)_t	-0.070	-0.043	-0.028
	(0.105)	(0.091)	(0.087)
ΔInterest rate_t	0.052	0.051	0.051
	(0.040)	(0.034)	(0.031)
ΔInterest rate_1	-0.048	-0.042	-0.054
	(0.044)	(0.039)	(0.037)
Δln(Stock to use ratio)_t	-0.595**	-0.411	-0.636***
	(0.254)	(0.284)	(0.207)
yr76_t		0.301	
		(0.225)	
yr94_t		0.536***	0.521***
		(0.170)	(0.162)
ln(ICO)_1	-0.612***	-0.495***	-0.491***
	(0.125)	(0.115)	(0.112)
ln(GDP)_1	-1.367***	-1.087***	-1.164***
	(0.247)	(0.233)	(0.212)
ln(REER)_1	1.436**	1.581***	1.326***
	(0.636)	(0.559)	(0.413)
ln(Oil)_1	-0.103	-0.044	-0.046
	(0.083)	(0.074)	(0.064)
Interest rate_1	0.083*	0.052	0.071*
	(0.045)	(0.043)	(0.037)
ln(Stock to use ratio)_1	-1.159***	-0.873***	-0.980***
	(0.283)	(0.274)	(0.230)
Constant	-0.969	-2.884	-1.554
	(2.848)	(2.540)	(2.044)
Observations	50	50	50
R^2	0.786	0.853	0.836
Adjusted R^2	0.626	0.723	0.733
AIC	-25.205	-39.99	-42.605
Residual Std. Error	0.159 (df = 28)	0.136 (df = 26)	0.134 (df = 30)
F Statistic	4.898*** (df = 21; 28)	6.562*** (df = 23; 26)	8.069*** (df = 19; 30)
Note:	*p**p***p<0.01		

Table 12 - Critical values for the bounds test

Significance	I(0)	I(1)
1 %	3.955	5.583
5 %	2.900	4.218
10 %	2.435	3.600

Note: Narayan (2005), Case III (unrestricted intercept and no time trend), k=5, n=50

Table 13 - Diagnostics tests for the model with exports

	ICO indicator model	
Diagnostics Test	P-value	Result
Durbin Watson test	0.468	No autocorrelation
Breusch-Pagan test	0.7154	No heteroskedasticity
Shapiro Wilk test	0.09554	Normal distribution
RESET test	0.8229	No misspecification
Multicollinearity		
highest VIF	5.5359 (exp1)	No problem
Cointegration		
F statistics	9.8888	Cointegration is present

Table 14 - ARDL model with alternative variables

	Dependent variable:	
	Δln(ICO)_t	
	Stock to use ratio	Production
Number of lags	4, 4, 4, 3, 3, 3	4, 4, 1, 2, 3, 4
Δln(ICO)_1	-0.138	0.014
	(0.120)	(0.097)
Δln(ICO)_2	0.174	0.212^{**}
	(0.116)	(0.093)
Δln(ICO)_3	0.410^{***}	0.460^{***}
	(0.107)	(0.108)
ΔGDP growth_t	-0.053^{**}	
	(0.025)	
ΔGDP growth_1	0.100^{***}	
	(0.024)	
ΔGDP growth_2		-0.041^{**}
		(0.015)
ΔGDP growth_3	-0.029^{*}	-0.049^{***}
	(0.017)	(0.014)
Δln(REER)_t		0.199
		(0.458)
Δln(REER)_3	-2.123^{***}	
	(0.673)	
Δln(Oil)_1		-0.151^{*}
		(0.088)
Δln(Oil)_2	0.039	
	(0.125)	
ΔInterest rate_t	0.065^{*}	0.068^{**}
	(0.033)	(0.029)
ΔInterest rate_1		0.048
		(0.034)
ΔInterest rate_2	0.034	
	(0.039)	
Δln(Stock to use ratio)_1	-0.487^{**}	
	(0.217)	
Δln(Stock to use ratio)_2	-0.233	
	(0.211)	

Δln(Production)_t		-0.00001**
		(0.00000)
Δln(Production)_3		0.00000*
		(0.00000)
yr76_t	0.557**	0.379**
	(0.226)	(0.168)
yr94_t	0.664***	0.600***
	(0.174)	(0.160)
ln(ICO)_1	-0.082	-0.376***
	(0.062)	(0.076)
GDP growth_1	-0.137***	0.015
	(0.040)	(0.019)
ln(REER)_1	1.525***	0.741**
	(0.399)	(0.273)
ln(Oil)_1	0.020	0.154***
	(0.065)	(0.050)
Interest rate_1	-0.056*	-0.056**
	(0.028)	(0.022)
ln(Stock to use ratio)_1	0.216*	
	(0.112)	
ln(Production)_1		-0.749***
		(0.175)
Constant	-6.409***	5.110**
	(1.776)	(2.015)
Observations	50	50
R^2	0.799	0.829
Adjusted R^2	0.661	0.721
AIC (original)	-27.031	-36.102
AIC (restricted)	-30.43	-40.503
Residual Std. Error	0.151 (df = 29)	0.137 (df = 30)
F Statistic	5.778*** (df = 20; 29)	7.672*** (df = 19; 30)
Note:	*p**p***p<0.01	

Table 15 - Diagnostics test for models with alternative variables

	Stock to use ratio		*Production*	
Diagnostics Test	P-value	Result	P-value	Result
Durbin Watson test	0.434	No autocorrelation	0.314	No autocorrelation
Breusch-Pagan test	0.329	No heteroskedasticity	0.4348	No heteroskedasticity
Shapiro Wilk test	0.473	Normal distribution	0.202	Normal distribution
RESET test	0.8952	No misspecification	0.8211	No misspecification
Multicollinearity				
highest VIF	7.9867 (gdpa1)	No problem	6.04175 (prod1)	No problem
Cointegration				
F statistics	4.446709	Cointegration is present	5.342707	Cointegration is present

Table 16 - ECM for models with alternative variables

	Dependent variable:	
	Δln(ICO)_t	
	Stock to use ratio	Production
Δln(ICO)_1	-0.163	0.026
	(0.141)	(0.100)
Δln(ICO)_2	-0.013	0.200**
	(0.132)	(0.097)
Δln(ICO)_3	0.201*	0.409***
	(0.117)	(0.104)
ΔGDP growth_t	0.021	
	(0.019)	
ΔGDP growth_1	0.033	
	(0.021)	
ΔGDP growth_2		-0.050***
		(0.015)
ΔGDP growth_3	-0.028	-0.051***

	(0.019)	(0.014)
Δln(REER)_t		-0.117
		(0.449)
Δln(REER)_3	-0.632	
	(0.640)	
Δln(Oil)_1		-0.163^{*}
		(0.086)
Δln(Oil)_2	-0.072	
	(0.138)	
ΔInterest rate_t	0.081^{**}	0.097^{***}
	(0.037)	(0.028)
ΔInterest rate_1		0.010
		(0.031)
ΔInterest rate_2	-0.016	
	(0.041)	
Δln(Stock to use ratiol)_1	-0.262	
	(0.222)	
Δln(Stock to use ratiol)_2	-0.313	
	(0.235)	
Δln(Production)_t		-0.000001
		(0.00000)
Δln(Production)_3		0.000001^{*}
		(0.00000)
yr76_t	0.712^{**}	0.428^{**}
	(0.263)	(0.166)
yr94_t	0.628^{***}	0.555^{***}
	(0.203)	(0.158)
ECT_1	-0.136^{*}	-0.349^{***}
	(0.079)	(0.078)
Constant	-0.056^{*}	-0.023
	(0.031)	(0.023)
Observations	50	50
R^2	0.645	0.775
Adjusted R^2	0.488	0.685
Residual Std. Error	0.185 (df = 34)	0.145 (df = 35)
F Statistic	4.116^{***} (df = 15; 34)	8.623^{***} (df = 14; 35)
Note:	$^{*}p^{**}p^{***}p<0.01$	

Table 17 - Long run estimates for models with alternative variables

	Stock to use ratio	Production
GDP growth	-1.6707	x
REER	18.5975	1.9707
Oil price	x	0.4096
Interest rate	-0.6829	-0.1489
Stock to use ratio	2.6341	-
Production	-	-1.992

Note: The "x" means, that the variable is not significant

Table 18 - ARDL model including disasters

	Dependent variable:
	Δln(ICO)_t
Δln(ICO)_1	-0.069
	(0.103)
Δln(ICO)_2	0.287**
	(0.103)
Δln(ICO)_3	0.494***
	(0.092)
ΔGDP growth_t	-0.014
	(0.020)
ΔGDP growth_1	0.050**
	(0.018)
ΔGDP growth_3	-0.050***
	(0.015)
Δln(REER)_t	0.007
	(0.005)
Δln(REER)_1	-0.004
	(0.006)
Δln(REER)_2	-0.003
	(0.005)
Δln(REER)_3	-0.018***
	(0.006)

Δln(Oil)_t	0.001
	(0.002)
ΔInterest rate_t	0.036
	(0.028)
ΔInterest rate_1	0.068*
	(0.038)
Δln(Exports)_t	-1.607***
	(0.382)
Δln(Exports)_2	-0.458
	(0.342)
Δln(Disasters)_t	-0.075
	(0.084)
yr76_t	0.361**
	(0.161)
yr94_t	0.582***
	(0.137)
ln(ICO)_1	-0.318***
	(0.090)
GDP growth_1	-0.077**
	(0.033)
ln(REER)_1	0.015***
	(0.005)
ln(Oil)_1	0.003**
	(0.002)
Interest rate_1	-0.020
	(0.024)
ln(Exports)_1	-0.363
	(0.229)
ln(Disasters)_1	-0.111
	(0.073)
Constant	3.524
	(2.501)
Observations	50
R^2	0.903
Adjusted R^2	0.803
AIC	-56.93
Residual Std. Error	0.115 (df = 24)
F Statistic	8.972*** (df = 25; 24)
Note:	*p**p***p<0.01

Table 19 - Diagnostics test for price of arabica and robusta

	Arabica		*Robusta*	
Diagnostic Test	P-value	Result	P-value	Result
Durbin Watson test	0.616	No autocorrelation	0.974	No autocorrelation
Breusch-Pagan test	0.73	No heteroskedacity	0.2776	No heteroskedacity
Shapiro Wilk test	0.1039	Normal distribution	0.1998	Normal distribution
RESET test	0.5279	No misspecification	0.1641	No misspecification
Multicollinearity				
highest VIF	4.998382 (exp1)	No problem	7.577272 (exp1)	No problem
Cointegration				
F statistics	4.723021	Cointegration is present	4.74749	Cointegration is present