Norwegian School of Economics
Bergen, Fall 2019

The Effect of Commodity Prices on Deforestation in Brazil

An Empirical Analysis Investigating the Effect of Futures and Spot Prices for Soy and Maize on Deforestation in Brazilian municipalities

Ingrid Emilie Flessum Ringstad & Julie Lie Tandberg
Supervisor: Torfinn Harding

Master Thesis, MSc in Economics and Business Administration

Majors: Energy, Natural Resources and the Environment, Finance

Acknowledgements

We would like to start by thanking our supervisor Torfinn Harding for providing us with constructive feedback and input on our work, in relation to the econometric analysis and encouraging us to think outside the box. Throughout the process, his knowledge of the topics studied has been invaluable.

Furthermore, we would to extend our gratitude to Vilde Larsen Nakkim for her informative comments and excellent guidance with LaTeX and stata. Last but not least, we want to thank our family and friends for great support and always being there when needed.

Abstract

In this thesis we aim to provide empirical evidence of the effect of spot and futures prices on municipal land transition in Brazil. We take advantage of newly developed, detailed, municipality-level data on land transition, for the whole of Brazil. Additionally, we exploit data on potential yield for soy and maize in our analysis, and we present two simple models for two different kinds of land transition to guide our empirical work. We mainly utilize a First Difference approach to study the effects of both spot and future prices for soy and maize. We also estimate the effects of our commodity prices for different levels of potential yield, and investigate the long-term effects of our prices.

Our main findings are that the spot-future spread has a seemingly larger effect on deforestation than the spot price alone for soy. This contributes to the extensive research on deforestation in Brazil by suggesting that the price expectations for the future should be included in further research on Brazilian deforestation. Maize on the other hand, show more ambiguous effects. Finally, municipalities with higher levels of potential yield seems to have a more price sensitivity towards land transition. Further, the long-term effects are larger for deforestation than regrowth of forests.

Contents

Page

1 Introduction

Today global societies are paying increasingly more attention to climate change and the potential impacts and consequences. Scientists, world leaders and international organizations are coming together to find ways to deal with the challenges linked to climate change. In order to develop a resilient strategy to deal with climate change, both adaptive and mitigation strategies are imperative. One such strategy, which has been proposed is reducing emissions from deforestation (Stern, 2006). According to the IPCC the world's forests play a significant role in the global carbon cycle (International Panel on Climate Change, 2018) as they work as carbon sinks. Several estimates have been made in order to show the global biophysical sequestering potential, where estimates range between 3300 MtCO2/year to 5380 MtCO2/year on average up until 2050(menon2007couplings) (International Panel on Climate Change, 2018). Despite this fact, deforestation has increased over the decades, with agricultural production as one of the leading drivers (benhin2006agriculture). In the period 2000-2012, large scale agriculture accounted for about two-thirds of deforestation in Latin America and one-third in Africa and Asia (Kissinger, Herold, & De Sy, 2012).

Currently, Brazil is the world's leading exporter of soybeans, followed by the United States and Argentina (United States Department of Agriculture, 2018). Brazil's farming regions host a large variety of field crops. Among those are corn, soybeans, wheat, rice and cotton, and they compete with each other, livestock and other agricultural crops for land area (Schnepf, Bolling, Dohlman, et al., 2001). Schnepf et al. (2001) describe the rise of export-oriented soy production as an outcome of macroeconomic conditions stabilizing in Brazil, national agricultural policies becoming more export-oriented in conjunction with trade liberalization gradually removing barriers to trade in the 1990s. Assuming Brazil is a price taker, the country's agricultural commodities are now subject to global commodity prices. This leads to our thesis theme; *how commodity prices affect deforestation in Brazil.*

Commodity prices have had a more steady and slow growth until mid 2000s before a higher increase in commodity prices led to a generally higher price in the world market reaching its top in 2011 (International Monetary Fund, 2019). Some of this price increase can be attributed to rapid population growth and growing economies in Asia, over the last 10-15 years. Mentioned in section 3.4, China imports the majority of soy produced in Brazil today, highlighting the importance of the country's growing economy's effect on export-oriented soy industry in Brazil. The world experienced a price drop in commodities in 2014 which has kept the prices low since then (International Monetary Fund, 2019). We also observe this drop for soy and maize in our price graphs in section 3.5.

In this thesis we aim to *provide empirical evidence of the effect of spot and futures prices on municipal land transition in Brazil.* In order to obtain cross-sectional differences we incorporate the potential yield for soy and maize for each municipality in our analysis. From previous research we are lead to assume that *an increase in the commodity price will increase the transition of land from forest to agriculture*, in other words increase deforestation. We take advantage of newly developed, detailed, municipality-level data on land transition, for the whole of Brazil. This data is developed by the MapBiomas project (2018), and was released April 2019. More specifically, we exploit the data to construct two outcome variables measuring different types of land use; *Transition of land from forest to agriculture* and *Transition of land from agriculture to forest.* To the best of our knowledge this data has never been used to investigate the relationship between *both* spot and future prices on deforestation in Brazil. Our full dataset is constructed by linking these variables, with data assembled by Bustos et al.(2016) from the Brazilian Agricultural Census (IBGE, 1996, 2006), and FAO (2018b). Our price data have been obtained from The World Bank (The World Bank, 2019) and Bloomberg (2019). Moreover, we have supplemented the dataset with additional data obtained from the Brazilian Agricultural Census (1996), (2006) and the Produção Agrícola Municipal (PAM) database (2016).

We utilize an estimation strategy based on First Difference estimation, inspired the paper *Supply Flexibility in the Shale Patch: Evidence from North Dakota* by Hilde C. Bjørnland et al. (2019). In our empirical work, we present two simple models for each outcome variable in order to investigate short term and long term effects of commodity prices.

We add two yearly lagged values of potential value and spot-future spread along with one year lag of our variables on land transition due to the long time span of our data and the possible long term effect of prices. We find that the long run propensities are statistically significant in our models with transition from forest to agriculture as dependent variable, and have the same signs as the associated non-lagged variables. Looking at the estimation using transition from agriculture to forest, this correlation shows less significance in the estimates of the coefficients. From this we can imply that long term price changes are more important for the removal of forest, than transforming agricultural land to forest land.

In sum our results point towards that our results for the spot price is in line with previous literature. What is interesting is that our results indicate that the spot-future spread has a larger impact on transition from forest to agriculture than the spot price, especially for soy. This is a new finding for research investigating deforestation in Brazil. Our First Difference strategy seems to lack some statistical significance when estimating the price effects on transition from agriculture to forest. What is clear is that our results suggest that maize has more unclear effects on land transition than soy. This result also lead us to think that maize might not be the best commodity to use when trying to estimate the effects of spot and future prices together. Reflecting on our results and how complex the challenge of deforestation is, we believe our simplified models lack the power to infer any causal relationships. Still, we have found that statistically significant correlations of several of our futures variables, indicating that there might be an interesting relationship with deforestation and futures.

1.1 Implications for our Study

The topic studied is of particular interest due to several reasons. First, deforestation in Brazil is a massive challenge for combating the effects of climate change in today's society. Second, to the best of our knowledge, this is the first study of its kind to exploit both spot prices and futures data on land transition for the whole of Brazil on a municipal level. Finally, the latest MapBiomas datasets were published in April 2019, includes data dating back to 1985, giving a unique opportunity to investigate the price effects on land over a period of more than 30 years. Consequently, as we are using a novel approach to estimate price effects on Brazilian deforestation we do not expect to find results directly comparable to the results of previous studies, as these primarily focus on the Amazonia biome and do not include both spot prices and future prices, such as (Harding, Herzberg, & Kuralbayeva, 2006), (Richards, Myers, Swinton, & Walker, 2012), and (Faria & Almeida, 2016) among others described in Section 2. Additionally other papers investigating the effects of commodity prices on deforestation in Brazil have faced criticism for not including futures.

The remainder of the thesis continues as follows. In Section 2, we provide a review of the previous literature relevant to our thesis. We present background information on forest and the agricultural industry, as well as information on Brazilian trade before we expand on the development of our chosen commodity prices in Section 3. In Section 4 we describe and elaborate on our data, followed by an overview of our empirical strategy in Section 5. Our analysis and results are presented in Section 6, with a discussion of possible limitations and shortcoming in Section 7. Finally, we conclude in Section 8.

2 Related Literature

In the following subsections, we review studies that analyze the different topics of interest. Over the last decades significant amounts of time have been put into researching the topic of deforestation. Deforestation is one of the biggest threats to saving the climate, and particular focus has been set on tropical forests in regions such as Brazil, the Congo Basin and Indonesia. Considerable amount of research has been dedicated to establish and identify drivers of deforestation. One of these drivers is trade, which has received significant attention from researchers all over the world. Among these we find (Bolling, Somwaru, & Kruse, 2001, (Robalino & Herrera, 2010, (Baldi, Peri, & Vandone, 2011), (Richards et al., 2012), all focusing on trade and commodity prices in relation to deforestation.

P. Richard et al (2012) put particular emphasis on the "impact of currency devaluations on area of production. The results suggest that approximately 80,000 km^2, or 31% of the current extent of soybean production in Bolivia, Paraguay and Brazil, emerged as a supply area response to the devaluation of local currencies in the late 1990's. Their research shows that "amids an increasingly neoliberal economic environment, where barriers to trade are jettisoned in favor of free flow of commodities, relative currency values will occupy an important role in the future sourcing of both agricultural expansion and environmental degradation" (Richards et al., 2012). Brander and Taylor (1998) discuss the pros and cons of trade for exporters of renewable resources, and the role of trade on deforestation is discussed in a theoretical framework by Copeland and Taylor (2009).

Faria and Almeida (2016) investigate the effect of openness to trade on deforestation, also estimate the effect of crop and pasture expansion on deforestation in Amazonia. Harding et al.(2006) find that high international prices on agricultural commodities adds pressure to tropical forests, soy being one of central agricultural commodities since the study also focuses on deforestation in Amazonia. Furthermore, Boerema et al. (2016) investigate the effect of soybean trade in terms of land use displacement using macro data from 1961 to 2008, mainly emphasizing European imports. They find that soy expansion largely replaced other agricultural crops and pastures, but that especially pastures were replacing forests, in other words that the soybean industry indirectly contributed to deforestation in the relevant period. The global soy market, specifically from the viewpoint of the United States as a major exporter, is researched by Bolling et al. (2001) who state that the market can be seen as an example of oligopolistic competition among exporting nations. This indicates that trade policies in the soy market can be interpreted through the classical theories on imperfect competition in international trade as proposed by Dixit (1984).

Empirical findings generally support the price discovery role of futures markets, in other words, spot prices are usually discovered in the future markets. According to Silvapulle and

Moosa (1999) “spot and futures prices on the same commodity have the same fundamentals and change if new information emerges that causes market participants to revise their estimates of physical supply and/or demand. Since contracts sold on futures markets generally do not require the delivery of the commodity but can be implemented immediately with little up-front cash, futures markets generally react more quickly than spot markets.” (Silvapulle & Moosa, 1999). Garbade and Silber (1983), in particular analyze the price discovery for four storable commodities including soybean and corn and conclude that futures markets generally dictate spot markets in recording and disseminating information. Furthermore, Crain and Lee(1996) also find that changes in wheat futures prices lead changes in spot prices, confirming that futures markets dominate spot markets in the price discovery process. In more recent years, Yang et al. (2001) confirmed that futures markets play the dominant role in the price discovery process for storable commodities. Frank Asche and Atle G. Guttormsen (2002) find indications in their empirical study of gas oil prices that future prices do lead spot prices. Additionally they find indications that future contracts with longer expiration time leads short expiration time contracts. Moreover, Hernandez and Torrero (2010) show the relationship between spot and future prices for soft and hard wheat, soybeans, and corn, supporting the evidence of future prices leading spot prices for these commodities. Testing for Granger-causality, they find evidence that future prices Granger-cause spot prices more frequently than the opposite, especially for wheat and corn. Additionally, they find that this relationship is remarkably stronger than in the past. They suggest that this result is due to the increasing importance of electronic trading of futures contracts, yielding more transparent and accessible prices, Hernandez and Torrero (2010). In line with the main findings that emerge in the literature investigating the relationship between spot and futures prices in food commodity markets, futures prices play a major role in price discovery. “That is in registering and transferring information from the related real market; due to the greater transparency and, often, greater liquidity of commodity futures over physical commodities, futures markets react more quickly to new or unexpected information than the underlying spot market. However, in times of crisis and in particular in phases of strong price increases, the cash market also becomes an important actor in the price discovery process.” (Baldi, Peri, & Vandone, 2011). In contrast to this, other studies have contested these results and find that spot prices lead futures prices, among those are Quan (1992), Kuiper et al. (2002), Mohan and Love (2004) (Baldi, Peri, & Vandone, 2011). Meaning one cannot know for sure if there is absolute certainty that futures lead spot prices.

3 Background

Today there is an extensive body of literature both on deforestation and commodity prices in Brazil. As mentioned in Section 2 there is also a considerable amount of literature discussing the relationship between futures and spot prices in the commodity market. Still, most studies on deforestation have focused on Amazonia's primary forests exclusively. As of yet there are no studies combining future and spot prices to investigate their effect on municipal land transition data, stretching over 30 years, to the best of our knowledge.

3.1 Global Deforestation

According to the IPCC the world's forests play a significant role in the global carbon cycle (International Panel on Climate Change, 2018) as they work as carbon sinks. Several estimates have been made in order to show the global biophysical sequestering potential, where estimates range between 3300 MtCO2/year to 5380 MtCO2/year on average up until 2050 (Denman et al., 2007, Table 7.1, Kauppi et al., 2001, linked in (International Panel on Climate Change, 2018)). Despite this fact, deforestation has increased over the decades, with agricultural production as one of the leading drivers (Benhin, 2006). Around half of such deforestation can be attributed to the cultivation of crops for export markets like the EU, China and North America. In the period 2000-2012, large scale agriculture accounted for about two-thirds of deforestation in Latin America and one-third in Africa and Asia (Kissinger et al.(2012). Geographically, the largest share of forest loss has occurred in the tropics (FAO, (2016)). Tropical deforestation represents one of the biggest threats to some of the world's most diverse local and specialized ecosystems, whilst globally it is responsible for close to one-fifth of overall greenhouse gas emissions (Burgess, Hansen, Olken, Potapov, Sieber, 2012). According to the latest Global Forest Resources Assessment, published by the Food and Agriculture Organization of the United Nations (FAO, 2016), just about 3,7 billion hectares of forest remain. This is stark decline from the estimate that the earth's forest cover once measured around 6 billion hectares (Bryant, Nielson, & Tangley, 1997). Deforestation is in itself a classical example of the tragedy of the commons (Harstad and Liski, (2013)). One can describe it as a strategic interaction between multiple individual actors who choose their own extraction rate without fully taking negative externalities into account. The end result is that the sum of individual actors will extract considerably more than what is socially optimal (Harstad and Liski (2013)), leading to the tragedy of the commons where non-excludable resources are subject to over-exploitation and in extreme cases extinction. Consequently deforestation represents one of the essential environmental issues of our time.

3.2 Brazil Opening Up To Trade

Brazil has had a long history of volatile trade policies, which have been strongly influenced by historical events, such as the military coup in 1964, the "economic miracle" between 1967 and 1973, then followed by the 1973 oil crisis. This oil crisis negatively affected Brazil's recent moves towards opening up to trade. Brazil's trade balance hit an unprecedented deficit in 1974 resulting in even tighter import policies. Despite these measures and actions taken to enhance exports, Brazil's trade balance remained in deficit through the 1970s (Hudson, 1997). Following a set of great macroeconomic policies in the early 1980's Brazil turned their trade deficit to a surplus in 1983 (Hudson, 1997). The years that followed provided more policy volatility as the measures taken to turn around the trade balance had negative effects in the domestic market. Schnepf et al.(2001) describe the rise of export-oriented soy production as an outcome of macroeconomic conditions stabilizing in Brazil, national agricultural policies becoming more export-friendly in conjunction with trade liberalization gradually removing barriers to trade in the 1990s. Currently Brazil is the world's leading exporter of soybeans, followed by the United States and Argentina (FAO, 2018a). Brazil's farming regions host a great variety of field crops. Among them are dominant crops like corn, soybeans, wheat, rice and cotton. These crops compete with each other, other agricultural crops, and livestock for land area (Schnepf et al., 2001). As Brazil has positioned itself as a major player in the global commodity market, the country is also more exposed to global politics. An example of a coordinated international political effort potentially affecting Brazil is the EU parliament voting to ban the use of palm oil in biofuels in 2017(Ghani, 2019) (Guardian, 2017). The goal is to prevent the EU's renewable transport targets from contributing to deforestation. Another example is Norway, which more recently announced that the government will reduce money transfers from the REDD+ program as the rate of deforestation has increased from 2018 to 2019 (Government, 2019).

3.3 Forest Cover, Deforestation and Agriculture in Brazil

Deforestation is defined as the conversion from forest to non-forest lands (Achard et al., 2002). Thus, this definition hinges on what is defined to be a forest. According to the United Nations Framework on Climate Change (UNFCCC) a forest has at least 10% cover (N. Sasaki & Putz, 2009). By this definition the central Brazilian Savanna, namely the Cerrado, is considered to be a forest, hence clearing of the cerrado is defined as deforestation (Nophea Sasaki & Putz, 2017). Brazil is home to a complex rich ecosystem, with a wide range of animals, plants, and insects (L. E. Andersen, 1996)(Andersen, 1996). However, for several decades Brazil has been ranging as the country with the highest deforestation rates (2018). This is quite alarming due to the fact that close to 70% of the world's largest rainforest, the Amazon, is located within Brazilian borders. Consequently Brazil sustains 40% of the world's remaining tropical forest (Kirby et al.(2006)). The 1970's is seen as the beginning of the "modern" era of deforestation, catalyzed with the opening of the Transamazon highway (Fearnside, 2005). Brazil's Amazon forest had remained largely untouched up until this point. Approximately 2.04 million hectares of forest was lost annually from the mid 1970's and throughout the 1980's. Substantial rates of deforestation continued into the 1990's and the 2000's, albeit slightly lower than in the past two decades (Fearnside, 2005). Cattaneo (2002) find that between 1980 and 1995 Brazilian forest loss was equivalent to 20% of total tropical forest loss during that period. Netpstad et al. (2009) find that the average clearing from 1996 to 2005 amounted to 1.95 million hectares per year. Increased demand for cropland is one of the major drivers of deforestation (Ehui & Hertel, 1989). During the second half of the 20th century, devotion of land to agriculture in Brazil saw an extensive expansion. Farmland grew from covering 29% to 44% of the country's territory from 1960 to 1985 (Assuncao, Lipscomb, Mobarak, & Szerman, 2016). The Brazilian economy is heavily dependent on the agricultural sector, with soybeans, maize, sugar, and coffee, as some of the main export products (Martinelli, Naylor, Vitousek, & Moutinho, 2010). Furthermore, Brazil is the world's largest exporter of beef in 2018, providing almost 20% of total global beef exports according to the USDA (2018). Barona et al. (2010) show that there is a correlation between cattle price and deforestation, basing the research on the legal Amazon. They find that the correlation between cattle prices and deforestation was 0.86 during the 1995-2003 period, then dropped down between 0.5-0.6m before ascending again to 0.71 during the period of 1999-2007. For soy prices the correlation was 0.27 during the 1995-2003 period, then increased to 0.6 and 0.7 in subsequent periods, before reaching 0.87 during the period of 1999-2007. Furthermore, (Barona, Ramankutty, Hyman, & Coomes, 2010) find that soy price has become increasingly related to deforestation over time, and suggest that it is likely that soy cultivation could be a major underlying cause of deforestation. Example, Fearnside (2005) suggests that while pasture occupies vast areas of land, soybean cultivation carries the political weight necessary to induce infrastructure improvements, which in turn stimulates crop expansion. Nepstad et al. (2006) suggest that the growth of the Brazilian soy industry may have indirectly led to the expansion of the cattle herd. According to Nepstad et al.(2006), soy has driven up land prices in the Amazon, allowing many cattle ranchers to sell valuable holdings at enormous capital gains and purchase new land further north and expand their herd further. These studies elucidate the fact that the rise of agricultural sector has its disadvantages, as there is a close relationship between the expansion of land use for agricultural production and deforestation (Barona et al., 2010), (D. C. Nepstad, Stickler, & Almeida, 2006). Thus the Brazilian rainforest and ecosystems have paid a high price for the success of the country's agricultural sector, as agricultural expansion has been followed by extensive deforestation (Martinelli et al., 2010). Approximately 284 million hectares of land is devoted to agriculture in Brazil FAO (2018a). In addition, it is estimated that approximately 80% of of the deforestation happening in the Amazon is a result of illegal activities (Reuters, 2016). The amount of tropical forest loss in 2016 was estimated at roughly 2 soccer fileds per minute. In a scenario where there is no cooperation, loggers are assumed to behave in an individualistic manner only taking their own costs and welfare into account. Resulting form this behavior is logging decisions optimized by taking the logging of other loggers as given (Eskeland, 2019). In a one-shot, non-cooperative game, one will say this leads to Nash equilibrium, where no single player will wish to deviate form their optimal choice of logging, because these players see other players choices as given. In a given game there may be one, zero, or several Nash equilibrium, but for simplicity it can be assumed there will only be one. A consequence of ending up in a Nash equilibrium is that each logger will clear more forest than what is socially optimal for climate change mitigation and sustainable forest management. This consequence illustrates clearly that without cooperation between loggers, farmers and society, free-riding will ensure that Brazil will experience under-provision of sustainable forest management (Eskeland, 2019). A critical result of this may lead to a tragedy of the commons situation in which tropical forest loss is so severe that it will affect the climate irreversibly. However, in the real world, there are rarely strictly one-shot games with zero means of communication, and in repeated games with communication options cooperation tends to be stronger.

Mapbiomas also illustrate clearly that a large part of the initial forest from 1985 has been transformed into agricultural land for farming purposes by 2018, as can be seen to the left in the illustration below. In 1985 the forest cover measured 590 million hectare, while in 2018 that number had decreased to 514 million hectare.

Development for land areas in Brazil from 1985 to 2018

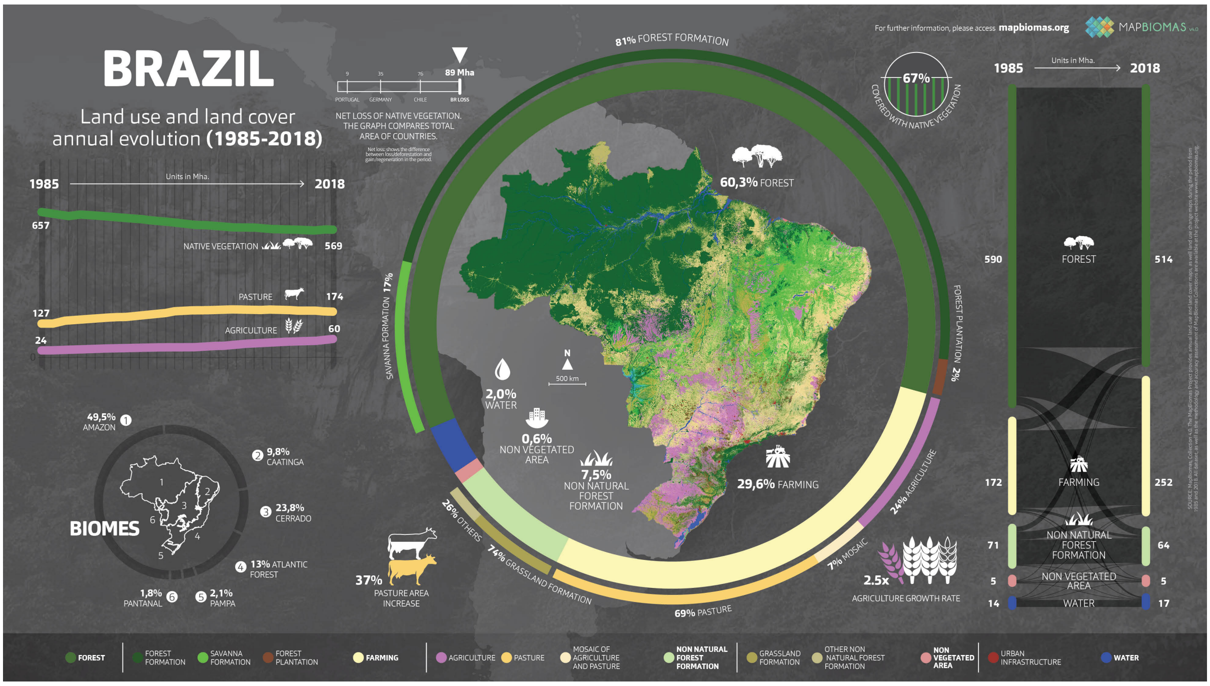


Figure 1: Forest Cover and Agricultural Land Cover in Brazil

Notes: Total area of forest cover and agricultural land cover in Brazil measured in hectares based on data from the MapBiomas Project map statistics database (2019). Forest cover includes forest formations, savanna formations and mangrove. Agricultural land cover includes annual and perennial agriculture and semi-perennial agriculture.

3.4 Commodity Market for Soy and Maize

Today 90% of the world's cultivation of soybeans can be allocated to four main countries, namely Brazil, Argentina, USA and China (Bank, 2018). Additionally, China is today the world's biggest importer of soy, and imports more than the rest of the world combined. 75% of Brazilian soybean exports were imported by China according to recent data from the United States Department of Agriculture (2018). This highlights the importance of China and the country's growing economy's effect on export-oriented soy industry in Brazil, which is briefly discussed by Halvard Sandvik Jansen in his (2018) master thesis. Furthermore soy is an important factor in the production of beef, and is widely used in animal feed (Fund, 2016). According to WWF beef production is the number one contributor to deforestation, followed by soy production. A large extent of the world's livestock production is dependent on the soy production in South-America, and about 75% of global soy is used in animal feed for the world's livestock population, with China and Europe being leading importers of soy. Consequently, most of the world's soy is consumed indirectly through meat. About 40% of Brazilian soy stays in Brazil, where among 90% is used as animal feed (Mongbay, 2019). When such a substantial amount of soy is being used for livestock production, particularly cattle, there is a possibility that a price increase of soy will lead to increased production costs for livestock, and thus decrease the demand for beef resulting in reduced deforestation.

As maize is being utilized for feed, industrial uses, food, and ethanol production it also plays an important role in the global commodity market (Rattray, 2012). Brazil is the third largest maize producer in the world as of 2016 (Pires et al., 2016). Extensive adoption of double cropping systems have been one of the key drivers for the growth in production of maize over the last decades. In the early 1980s some farmers in South-East Brazil introduced a second season of maize production (CONAB, 2012), also called milho safrinha. This second season of maize is cultivated after the summer, between March and July. Even though the technique was only adopted gradually, the use increased significantly from 2008 (Allen & Valdes, 2016).

3.5 Price Development for Soy and Maize

Economic theory states that prices are determined in the interaction between two economic forces, demand and supply. In an open economy of perfect competition, the equilibrium price emerges through a simultaneity process where, in the intersection of the curves, the price and quantity level will be at the point where the demands are indifferent between buying one more unit or not, and the suppliers are indifferent between supplying one more unit or not. With this consideration, economic theory states that a higher price of a good will induce a higher incentive of production. In our case the goods are maize and soy, the suppliers are farmers and the prices are assumed to be determined in an open and efficient world economy with given prices. We would therefore expect increased prices to increase production, and further lead to increased deforestation, and hence decreased prices the opposite.

Below is a graphical illustration of the price development for both soy and maize, moreover it includes both spot prices and futures. For all graphs the prices have been measured in dollars per metric tons.

First is the development of the spot price for soy and maize from 1985 until 2018.

Development of spot prices for soy and maize, 1985-2018

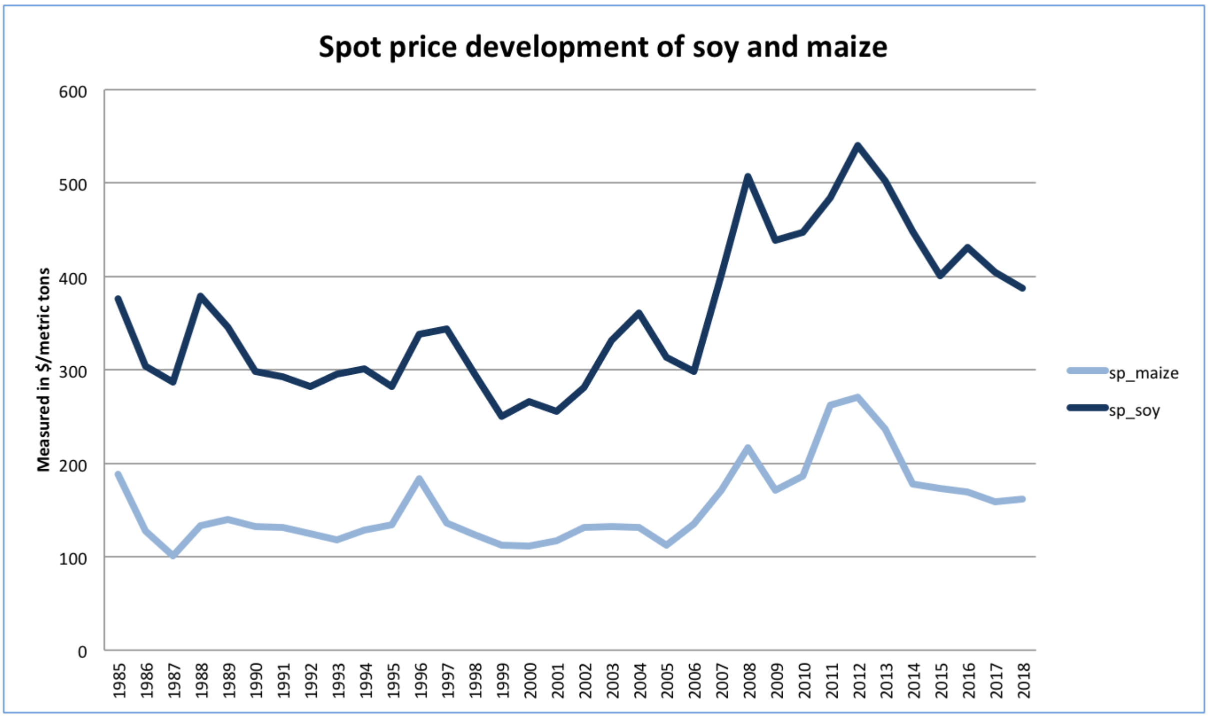


Figure 2: Forest Cover and Agricultural Land Cover in Brazil

Notes: Prices are obtained from The World Bank's Pink sheet (**2019**)

Next, there is the development of the futures for soy and maize for the same period.

Development of futures for soy and maize, 1985-2018

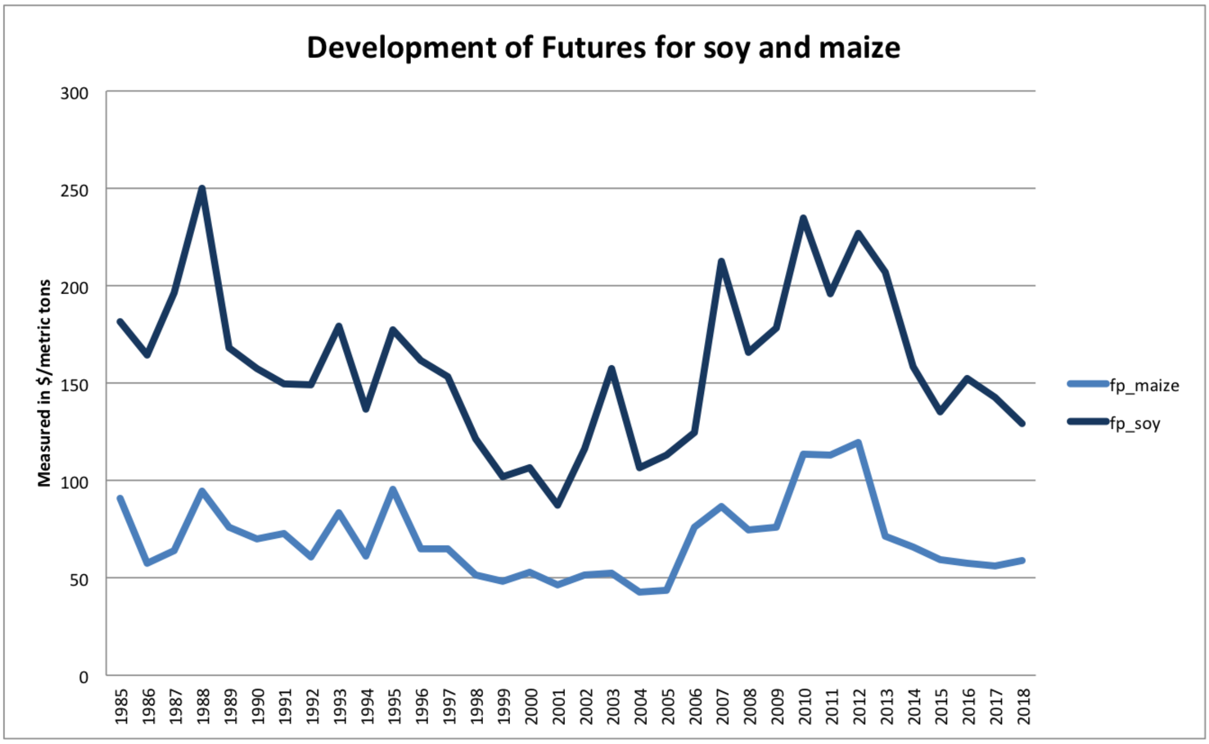


Figure 3: Forest Cover and Agricultural Land Cover in Brazil

Notes: Prices are obtained from the Bloomberg terminal

Finally here we have a graphical illustration combining both spot and future prices for both commodities, over the period 1985 until 2018.

Development of futures for soy and maize, 1985-2018

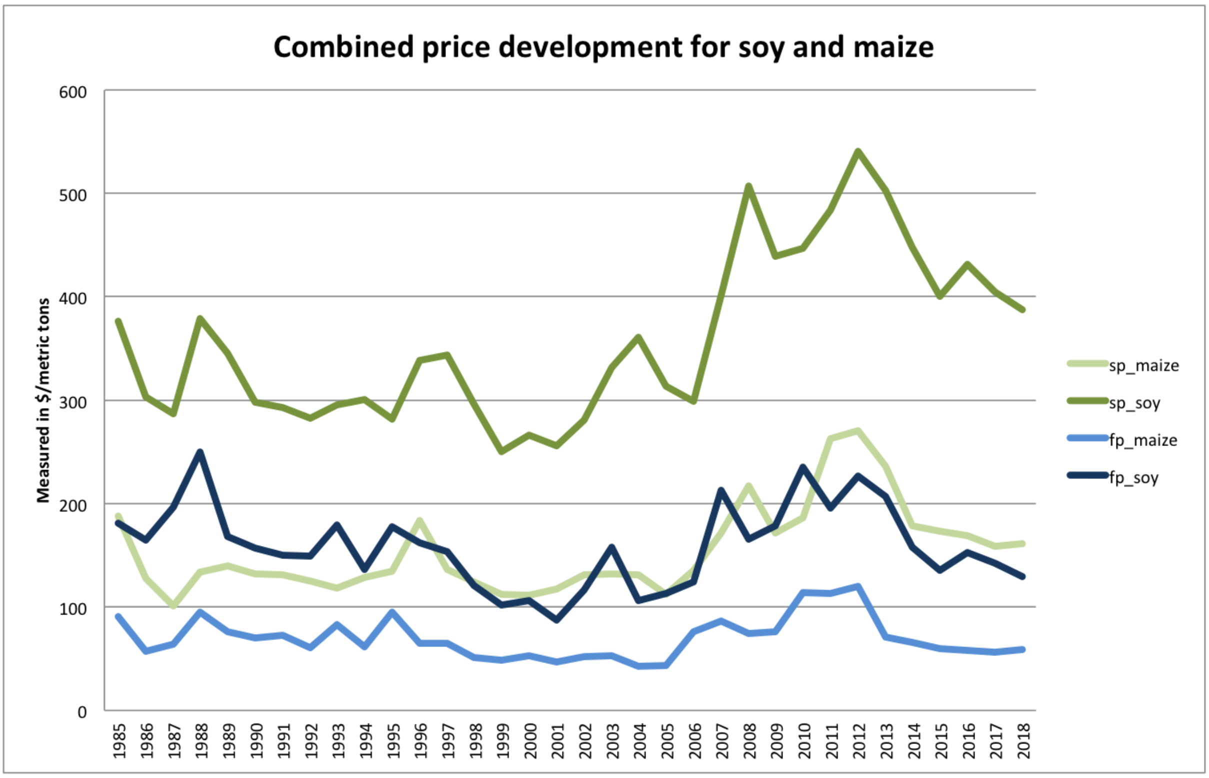


Figure 4: Forest Cover and Agricultural Land Cover in Brazil

Notes: Prices are obtained from the World Bank's pink sheet (The World Bank, 2019)and the Bloomberg terminal

It becomes clear that both the commodities correlate, also there is a clear correlation between the spot and future prices. More precisely the correlation coefficient between spot price soy and spot price maize is 0,9 and between future price soy and maize the correlation coefficient is 0,82, signaling a strong correlation between the two commodities. The relationship between the future and spot price for each commodity is also strongly correlated with a coefficient of 0,59 for soy and 0,63 for maize. Additionally all correlation coefficients are significant at the 5% level when using Pearson's correlation code in Stata.

4 Data

This section describes the data we use in our analysis to estimate the effect of commodity prices from soy and maize affect deforestation on a municipality level in Brazil. Our dataset is constructed mainly from four data sources. Data on land cover is collected from MapBiomas Project (2019). In order to obtain cross-sectional variation in our data, we use data constructed by Bustos et al. (2016) on the potential yield for soy and maize depending on technology. In addition we collect our data on commodity prices of soy and maize from The World Bank Pink Sheet and the Bloomberg database. For the analysis we have chosen to clean out AMC's with no data on transition. After cleaning all the data we have transtion data for 2459 AMC's in our dataset observed over 34 years, and our new constructed panel dataset is considered strongly balanced. We describe our data and data sources in detail in the following subsections below.

4.1 Data on Forest and Agricultural Land

Data on forest transition ranges from 1985 until 2018, and is collected from the MapBiomas Project map statistics database (MapBiomas, 2019). Established in 2015 by an initiative of several universities, technology companies, and NGOs, MapBiomas is a multi-institutional collaboration. The purpose of the project is to develop a reliable and low-cost method to produce detailed temporal series of annual land cover and transition maps and data of Brazil. The series are generated using pixel-per-pixel classification applied to satellite images (MapBiomas, 2019). The process is conducted using extensive machine learning algorithms, and the source of the satellite images is the Landsat Data Archive (LDA), available in the Google Earth Engine platform. The project consists of several collections, whereby this thesis take advantage of the last one, Collection 3.1 and 3.2. This collection is similar to collection 3.0, with improvements to classification published April 2019. For the analysis we have reshaped the data from wide to long format, and removed the level 2 and 3 to look at aggregate forest and agricultural data. We also simplified the transition data into two columns for each AMC and each year, transition from forest to agricultural land and agricultural land to forest. Thus, we are able to analyse the data on panel form.

The dataset we use includes data on forest transition on both state and municipality level for the whole of Brazil, and includes all six biomes: Amazonia, Caatinga, Cerrado, Mata Atlatica, Pampa and Pantanal. Furthermore, the dataset is divided into classes at three levels. At the first level the classes are aggregated into five main classes, including; (1) Forest, (2) Non-forest natural formations, (3) Agriculture, (4) Non-vegetated areas, and (5) Aquatic bodies. Our main interest in this dataset is data on forest cover and agricultural cover, hence we cleaned out all other irrelevant data. Moreover this dataset allows up to investigate transition from forest to agriculture and vice versa. Data on transition from forest to agriculture indicates how much forest has transitioned into agriculture from one year to the next per AMC. The same applies to data on transition from agriculture to forest, it indicates how much agricultural land has been transformed into forest land from one year to the next per AMC. This allows us to consider regrowth of forest as a potential effect of price change. In order to achieve a more robust result we analyse the data transition both from forest to agricultural land and from agricultural land to forest.

Our dependent variable(s) *Transition of land from forest to agriculture*footnoteIncluding forest formations, savanna formations and mangrove and *Transition of land from agriculture to forest*[1] are obtained from the MapBiomas dataset.*Transition of land from forest to agriculture* does not include regrowth, and is therefore a gross measure of deforestation, while in contrast *Transition of land from agriculture to forest* is a measure of regrowth in forest. Both dependent variables includes data from 1985 until 2018. According to MapBiomas all the transition data is measured annually in hectare. The justification behind including both *Transition of land from forest to agriculture* and *Transition of land from agriculture to forest*, is to make the analysis more informative. This enables us to investigate how both spot and future prices affect land transition leading to deforestation and land transition inducing regrowth of forests.

4.2 Agricultural Data

4.2.1 The FAO-GAEZ Database

Variables measuring the potential yield in soy and maize are obtained from the dataset developed by Bustos et al. (2016). The estimates of potential yields are obtained from the FAO-GAEZ v3.0 database (2018b), and have been constructed using an index for soil suitability. Potential yield is measured as ton per hectare. We exploit some of the same data as Harsheim and Nakkim (2018) in order to illustrate the concept of potential yield in Brazil.

The soil suitability index, illustrated in Figure 5, is developed by FAO, in cooperation with the International Institute for Applied Systems Analysis (IIASA), and has been produced using the Agro-Ecological Zones (AEZ) model (FAO, 2018b). The goal of the AEZ model is to estimate agricultural resources and potential. The index is constructed by exploiting knowledge on soil conditions, crop requirements, and soil management (FAO & IIASA, 2018c). More precisely, the measures of soil suitability are used to quantify to what extent the soil conditions in a given area match crop requirements, given defined input and

[1]Including annual and perennial agriculture and semi-perennial agriculture

management circumstances. This model has been applied considering the average climate during the 1961-1990 period (Bustos, Caprettini, & Ponticelli, 2018).

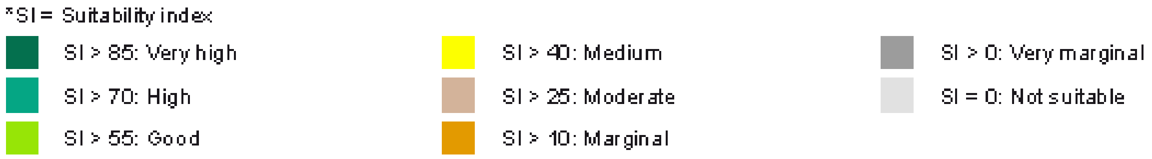


Figure 5: Crop Suitability Index

Source: Obtained from FAO & IIASA, 2018a

In addition, the AEZ model can be used to provide information about potential yields under different technologies, or so-called input combinations. The database separates between low, intermediate and high level inputs, where the different input levels reflect the level of cultivars, labor intensity, nutrient and machinery used in production. The variables on potential soy and maize yields across Brazil have been calculated by exploiting measures of potential yields under different input levels (Bustos et al., 2018). For simplification we assume that it is sufficient to exploit the potential yield measured under high technology inputs, as several technological advances have been adapted across Brazil over the last 15-20 years. Such technological advances include, but is not limited to, genetically engineered soy and second season maize (Harsheim & Nakkim, 2018).

Potential yield, as opposed to actual yield, is considered as exogenous variables, as potential yield is determined based on weather characteristics and soil suitability in each municipality and will not be influenced by any of the independent variables, nor the dependent variable. The paper includes data on potential yield under the assumption of high inputs, and these potential yields are illustrated in Figure 6 and 7. As we assume potential yield to be time constant over the time period examined, a variable containing only potential yield will be omitted in our first difference regressions and fixed effects. Consequently, we use this potential yield variable as an interaction with the price variables in our regressions.

Potential yield for soy under high technology inputs

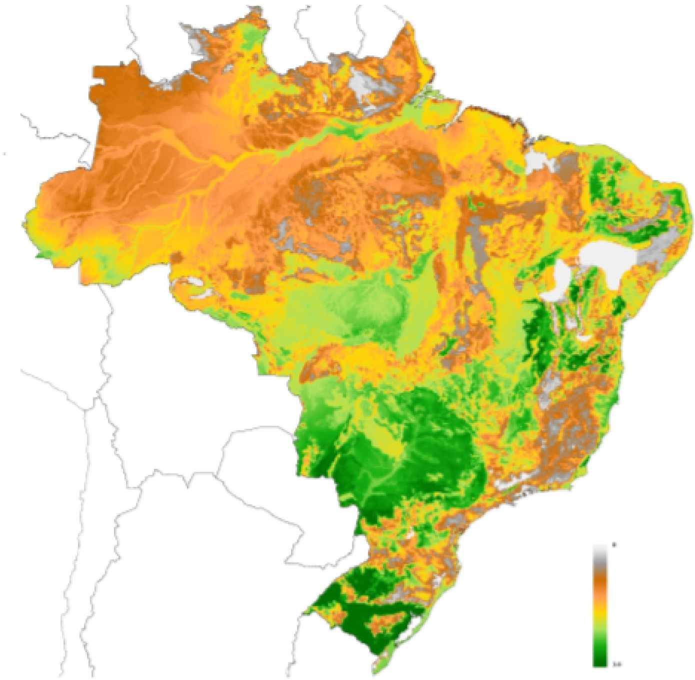


Figure 6: Measures of potential yield for soy

Potential yield for maize under high technology inputs

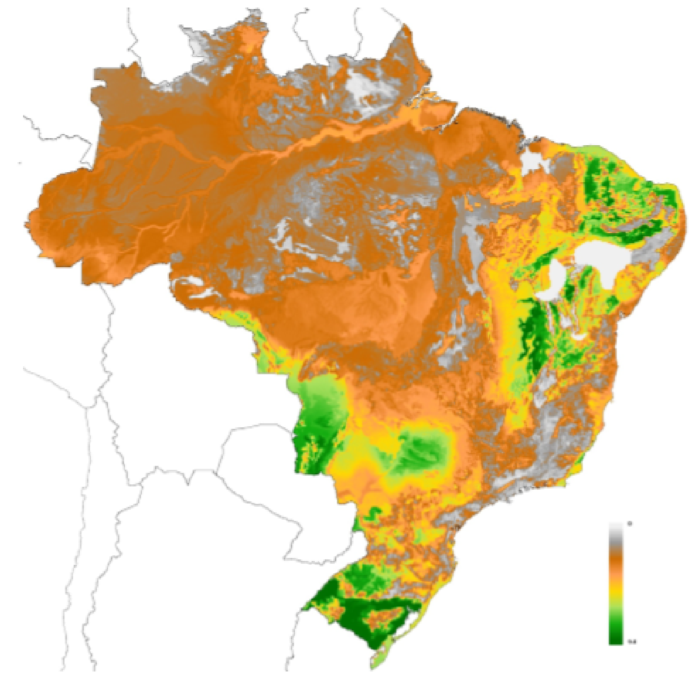


Figure 7: Measures of potential yield for maize

Notes: Figure 6 and 7 was created by Bustos et al. (2016) Source: Bustos, Caprettini, & Ponticelli, 2018, Online Appendix. Based on data from FAO-GAEZ (FAO & IIASA, 2018b).

4.3 Geographical units

As municipalities often experience change in borders as well as merging and divisions with other municipalities, IBGE has defined Área Mínima Comparável (AMC), smallest comparable areas, allowing for comparison over time. We use AMCs as units of observations in our analysis because we are looking at variations over many years (Bustos, Caprettini, & Ponticelli, 2016).

4.4 Spot prices and Futures

Our data on spot and future prices are downloaded from the World Bank's Pink Sheet (2019) and Bloomberg (2019), respectively. World Bank data is open source data on global development, including price development of agricultural commodities. The Pink Sheet is released every second business day of the month and contains multiple commodity prices, both real and nominal, as well as indices of the prices. This is where we obtain our data on spot prices, which are listed yearly. Our data on future prices are obtained from the Bloomberg Terminal, which belongs to Bloomberg L.P., an American private software, data and media company, providing real time and historical information for the global financial sector (Bloomberg L.P., 2019).

Both spot and future prices used in our analysis are in real terms adjusted for American inflation (U.S. Department of Labor Bureau of Labor Statistic, 2019), and are both presented as annual average of the year. The range is from 1985 to 2018 and both price variants are presented as dollars per metric ton, or dollars per 1000 kg. The future prices are one-year contracts and from the Bloomberg source, they were presented as dollars per bushel, where one bushel of soy is 27.2155 kg soy and one bushel of maize is 25,4 kg of maize. Future prices were then converted to dollars per metric tons by dividing the price in dollars per bushel by the number of kg per bushels for each commodity and multiply that with 1000 kg. The price data on spot and futures are used as explanatory variables in our analysis and we will examine the impact of the prices on land transition in interaction with potential yield. As the prices are determined in a global and interactive market, the price variables are assumed to be determined endogenously, despite Brazil being a big international player in the commodity market. Hence, the AMC's in the dataset are equally exposed to world prices, but will experience a slightly difference internally due to characteristics of the individual AMC.

5 Empirical Framework

In this paper, we take an econometric approach in attempting to estimate the effects of future and spot prices for soy and maize on municipal land transition in Brazil. Additionally we exploit the opportunity to estimate the effect of the spread between spot and future prices, inspired by Hilde C Bjørnland (2019). Consequently we set up several models utilizing our dependent variables *Transition of land from forest to agriculture* and *Transition of land from agriculture to forest.*

We attempt to establish a causal relationship between spot and future prices and our outcomes of interest by exploiting exogenous variation in potential yield and global spot and futures prices. Interacting the potential yield and the global prices results in a measure of *potential value* of each commodity for both spot prices and future prices, and results in variables that vary both across time and AMCs.

We mainly focus on one estimation strategy, inspired by the working paper *"Supply Flexibility in the Shale Patch: Evidence from North Dakota"* (Bjørnland, Nordvik, & Rohrer, 2019), which we use to solve the challenge that our prices may follow a random walk. The standard errors are clustered on the AMC level, as this is a way to minimize the risk of biased standard errors (Hansen, 2007). In addition to the estimation strategies, we investigate the directional relationship between the spot and future prices in section 5.1. As mentioned in section 2, there exists a fair amount of research investigating the relationship between spot and futures indicating that future prices tend to lead spot prices. By conducting a Granger-Causality test we are able to estimate a directional relationship between the spot and future prices, to determine if our price data exhibit the same relationship. This brief analysis supports why including future prices in the analysis is a good idea.

This Empirical Framework section is split into four sections, starting with the aforementioned investigation of the relationship the price variables, then we briefly elaborate on our choice of independent variables. The third section elaborates on our main econometric models showing how the different prices affect land transition. In the fourth section we briefly include some theory and assumptions the first difference estimation.

5.1 Investigating the relationship between spot and future prices

Our paper's main contribution to the literature is that we investigate the effect of including both spot and future prices when trying to estimate the effect commodity prices have on municipal land transition in Brazil. From section 3.4 we have already established that there is a significant correlation between the spot prices and future prices. Hence we find it interesting to perform a test, which can indicate whether there is also a clear directional relationship between future prices and spot prices.

The Granger-causality test will check whether one of the variables cause the other variable, in our case we check whether future prices cause the spot prices or vice versa. Inspired by (Asche & Guttormsen, 2002) we set up a simplified model in Stata to test whether current values or spot is correlated with past values of futures, or current values of futures is correlated with past values of spot. From previous empirical work, we may expect that future prices granger-cause spot prices. In our case we test for both directions, thus we have two null-hypothesis for soy and two for maize, which can be found in appendix A.3.

From the results in A.3 we find implications that in our price data, the future price granger-cause the spot price for both commodities, thus we have direction of causality in favor of the future prices. From section 2 it is clear that these results correspond with results that much of the literature have concluded with, yet there are some papers that challenge this which are also mentioned in section 2. These results indicate that it can be quite useful to include future prices when trying to estimate the effect of commodity prices on land transition.

As we aim to infer a causal relationship between prices and land transition we need unbiased estimates of the coefficients. In order to be able to infer a causal relationship between prices and land transition we need unbiased estimates of the coefficients. From the established correlations found in 3.4 and results from the Granger-Causality test, we include both prices and both commodities in hopes that we may be able to avoid potential omitted variable bias in or estimation models by omitting any price variables.

5.2 Choosing our explanatory variables

One of the key assumptions for using panel data is that all variables must vary over both time and units. This is a slight challenge for our price data as we assume all AMC's are price takers and cannot affect the market, and thus the price data varies only over time and not for units. We also have data on potential yield for each AMC as mentioned in section 4.2.1, and this data only varies across AMC's and not over time. In order to get variables that vary across both time and units we construct an interaction term between our price variables and our potential yield variable for high technology inputs. We multiply price with potential yield to get the interaction, which we interpret as the potential value and abbreviated to PV in our models. This is done for both commodities and all price variables.

Inspired by Hilde C Bjørnland's (2019) article using First Differences to estimate the effect of spot and futures on oil production, we use the same strategy as our main estimation strategy. We examine the effect of changes in the spot prices and the change in the spread between spot and future prices. The spread is calculated as spot minus future multiplied with potential yield[2]. If the spot price increases relative to the future price, the spread will increase. An increase in spread is expected to be associated with increased forest transition to agriculture, because we assume that the farmers will maximize production today if the future price will be relatively lower than today. If the analysis shows decrease in forest transition due to increase in spread, this might be associated with farmers having negative expectations for the future and thus decrease production.

For causal interpretation, a linear relationship between the dependent and the independent variables is crucial. If the linearity assumption is violated, the OLS-estimator will aim to fit a linear line through a non-linear relationship. A log-transformation of the variables will normally solve the challenges with non-linearity as well as controlling for skewed normality within the independent variables. A log transformation can be log-log, level-log and log-level. To determine which combinations of log and level relationships to choose, linearity and independence of the residuals, and normality of the interaction terms have been tested informally. As we find the log-level relationship the most intuitive to interpret, we have focused on comparing this with the level-level model. Hence, the tests have been done on level-level equation and log-level and the comparison shows that the log-level-transformation is a more linear relationship and is more normally distributed. Figures of these informal tests can be found the appendix A.4. Therefore, log-transformed variables have been used as a log-level model also gives an intuitive way of interpretation in our models (University of Utah, 2019).

Additionally, we also add lagged variables of spot prices, the spread and the dependent variable for land transition. It is reasonable to assume that changes in land transition of one year is influenced by changes in previous years prices, in addition to current prices. Previous year's land transition may also be assumed to be correlated with current land transition. Prices could be leading variables of land transition due to the time it takes from the commodity price to influence the farmers to expand agriculture and, hence, to

[2]Example; $(Spot^{Soy}_{i,t} - Futures^{Soy}_{i,t}) \cdot PotentialYieldSoy$

clear forest. With price lags, this becomes a dynamic model allowing correlation over time. Thus, we can identify the time delay of the response in deforestation to leading price variables (MathWorks, 2019).

5.3 Our main Econometric Models

The main econometric model for our first difference strategy:

$$LogFA_{i,t} = \alpha_i + \Delta\beta PVSoy^{spot}_{i,t} + \Delta\beta SpreadSoy_{i,t} + \Delta\beta PVMaize^{spot}_{i,t} + \Delta\beta SpreadMaize_{i,t} + u_{jt} \tag{1}$$

In equation 1 $LogFA_{i,t}$ represents the log-transformed *Transition of land from forest to agricultue*, $\Delta\beta PVSoy^{spot}_{i,t}$ and $\Delta\beta SpreadSoy_{i,t}$ represents the change in potential values[3] of soy spot price, and the spot-future spread of soy. For maize, it is the $\Delta\beta PVMaize^{spot}_{i,t}$ and $\Delta\beta SpreadMaize_{i,t}$, which represents the change in potential value and spot-future spread. The α_i is the unobserved time-invariant individual effect and captures any unobserved time-invariant variables on the AMC level. The last term, $u_{i,t}$, contains both fixed errors caused by unobserved time-invariant variables and a random error component. (Wooldridge, 2016). This equation allows us to estimate the effect of an increase in the change of our independent variables, and how this increase affects the land transition from forest to agriculture, also known as deforestation. In our analysis we also include models with a one year lag of the dependent variable along with one and two year lags of our independent variables.

We also estimate the same equation using *Transition of land from agriculture to forest* as the dependent variable, using $LogAF_{i,t}$.

$$LogAF_{i,t} = \alpha_i + \Delta\beta PVsoy^{spot}_{i,t} + \Delta\beta SpreadSoy_{i,t} + \Delta\beta PVmaize^{spot}_{i,t} + \Delta\beta SpreadMaize_{i,t} + u_{jt} \tag{2}$$

This equation lets us estimate how the effect of a change in spot and future prices affect the regrowth of forests. Estimating this equation in addition to equation 1 enables us to investigate whether the estimated effects are consistent. If the estimates of one of our variables indicate an unfavorable effect on forests when we use $LogFA_{i,t}$ as our dependent variable, we expect the same variable to have an unfavorable effect on forests when we use $LogAF_{i,t}$. Hence, we expect our coefficients to exhibit the opposite effect on our two dependent variables. Land transition is measured yearly in hectares for all AMCs, and captures to amount of land that has been transformed from forest to agriculture or from

[3]Potential value is abbreviated to PV

agriculture to forest. It is important to know that in this paper, the definition of forest and agriculture hinges on the definitions made by MapBiomas (2019)

Moreover, we have estimated the effects of our independent variables on both dependent variables using a Fixed Effects strategy in order to investigate the robustness of our results. These estimations are enclosed in appendix A.6.

5.4 Margin Plots of Potential Yield

In our analysis, we are also interested in investigating how the effect of price changes on deforestation may vary across AMCs with different levels of potential yield. The investigation of this relationship is done for all independent variables and both dependent variables, displayed in Figure 8 and 9 as eight margin plots below the regression table in our analysis. For every margin plot, potential yield of high inputs is on the horizontal axis and the average marginal effect of the spot price and the spot-future price is on the vertical axis. We expect the magnitude of the coefficients to increase for AMCs with higher potential yield, due to higher sensitivity for price change.

5.5 First Difference Strategy

Using a First Difference strategy, we control for time-invariant unobserved factors of the AMC's, which may be correlated with the explanatory variables in the model. The mechanism of the first difference approach is illustrated in equation 3 and 4 with multiple time periods

$$y_{it} = \beta_0 + \beta_1 x_{1,it} + ... + \beta_l x_{l,it} + \beta_{l+1} x_{l+1,i} + ... + \beta_k x_{k,i} + \delta_2 1[t=2] + ... + \delta_T 1[t=T] + \alpha_i + v_{i,t} \tag{3}$$

This equation can be rewritten as,

$$\Delta y_i = \alpha_0 + \alpha_3 1[t=3] + ... + \alpha_T 1[t=T] + \beta_1 \Delta x_{1,i} + ... + \beta_l \Delta x_{l,i} + \Delta v_{i,t} \tag{4}$$

4

Where δ represents the time periods, and the base period is t=1. According to Wooldridge (2016), the key assumption is that the idiosyncratic errors are uncorrelated with the explanatory variables in each time period. In other words, the explanatory variables are strictly exogenous after we eliminate i, the unobserved effect. Further, Wooldridge (2016) stresses that this key assumption rules out cases where future explanatory variables

[4]The composite error,$\Delta v_{i,t}$=α_i+u_{it}, and α_i can be eliminated by differing adjacent periods. (Wooldridge, 2016)

answer to current changes in the idiosyncratic errors, as must be the case if x_{itj} is a lagged dependent variable. Moreover, if there is an important time-varying variable that is omitted, then the strict exogeneity assumption is generally violated.

When FE calculates the average of prices, it might be showing a wrong picture because financial prices are sometimes assumed to follow a random walk, in other words, each step is random, stochastic and independent of previous values, and that prices are unpredictable. Assuming commodity prices follow the same pattern as financial assets, their intrinsic value can itself shift to a higher or lower level. This can be the case of the lower commodity prices we have seen across the world in the past years. It could be that the price average themselves has shifted permanently, not only temporarily. Fixed effects strategy would use this average in an estimation, and therefore the change in price is better, which first difference is based on. This behaviour is refereed to as unit root, which can be defined as "a stochastic trend in a time series, sometimes called a "random walk with drift"; If a time series has a unit root, it shows a systematic pattern that is unpredictable" (Statistics-How-To, 2016).

5.5.1 Clustered Standard Errors

Clustered standard errors are also added to the panel data models. Abadie et al. (2017) explain that this is often a good idea in the experimental design situation where clusters are non-randomly sampled and treatment effects are heterogeneous. In this paper, it is clear that AMCs and differences between them are non-random and systematic, indicating that clustered standard errors should be included according to the standards of Abadie et al. (2017). The main point of clustering standard errors is to make the statistical inference about estimators robust to heteroscedasticity and serial correlation (Stock & Watson, 2008). The failure to include clustered standard errors can then lead to biased standard errors, leading to potentially incorrect inference about the statistical significance of estimators (Hansen, 2007).

6 Empirical Analysis

The analysis below is based on the empirical framework discussed in section 5. In section 5.3 we highlight two equations, and our analysis includes a lagged version of both equation 1 and 2, in order to estimate a long term effect of the explanatory variables, namely the long run propensity, exploiting the large time span of our data. Model 1 and 2 are based on equation 1 from section 5.3, while Model 3 and 4 are based on equation 2. The motivation behind using both *Log* $FA_{i,t}$ and Log $AF_{i,t}$ is that we aim to achieve more robustness in our analysis. Furthermore, we have included eight margin plots showing how the coefficient of potential value on land transition changes across different AMCs with different levels of potential yield. This allows us to see how the prices affects land transition given various levels of potential yield.

As mentioned in section 5.3 we have also conducted a Fixed Effects analysis in order to check the robustness of our First Difference analysis. This extra analysis can be found in which can be found in appendix A.6. In the Fixed Effect models, we have replaced spread of spot and future prices with potential value based on future price.

6.1 First Difference Main Estimates

Table 1 shows the results of our First Difference estimations on both dependent variables. Additionally Model 2 and Model 4 incorporates two yearly lags of the independent variables, as well as the one year lagged dependent variable. At the bottom of the table we have included the long-run propensities estimated from Model 2 and 4. These long-run propensities have been calculated for AMC's at different levels of potential yield in order to investigate how the long-term effect changes as potential yield changes.

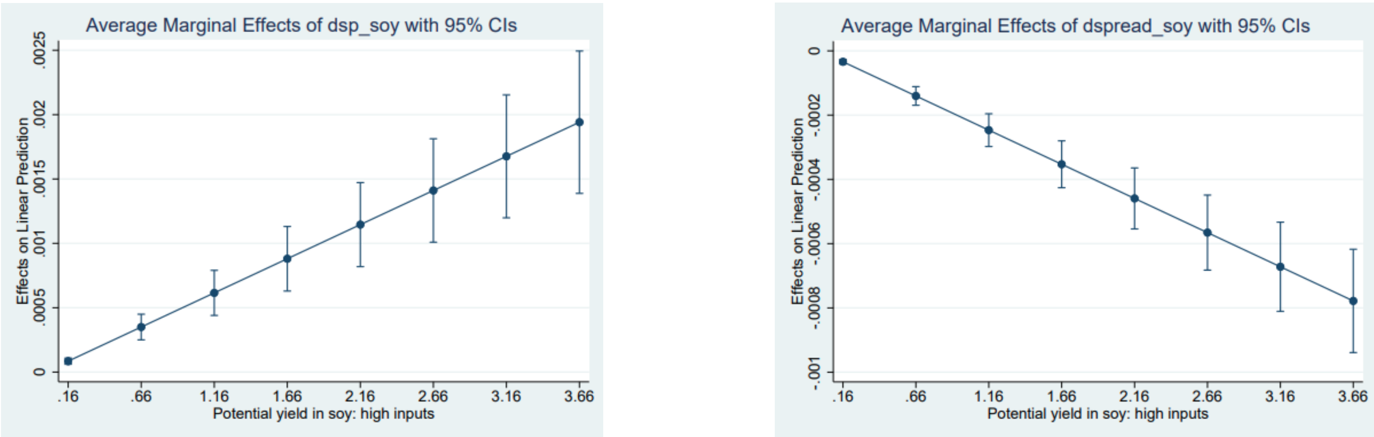
Table 1: First Difference - Commodity Price Effect Transition FA and Transition AF

	(1) Log FA b/se	(2) Log FA b/se	(3) Log AF b/se	(4) Log AF b/se
Δ Potential Value Soy^{Spot}	0.000446***	0.000318***	-0.000087	0.000202**
	(0.000077)	(0.000075)	(0.000084)	(0.000082)
Δ Potential Value Soy$^{Spot}_{t-1}$		0.000140		-0.000593***
		(0.000108)		(0.000114)
Δ Potential Value Soy$^{Spot}_{t-2}$		-0.000155		0.000945***
		(0.000116)		(0.000115)
Δ Spread Spot-Futures Soy	-0.000777***	-0.000649***	0.000712***	0.000906***
	(0.000098)	(0.000124)	(0.000103)	(0.000131)
Δ Spread Spot-Futures Soy$_{t-1}$		-0.000433***		0.000735***
		(0.000151)		(0.000171)
Δ Spread Spot-Futures Soy$_{t-2}$		0.000456***		-0.000276**
		(0.000127)		(0.000134)
Δ Potential Value Maize^{Spot}	0.000128**	-0.000264***	-0.000250***	-0.000323***
	(0.000060)	(0.000071)	(0.000073)	(0.000078)
Δ Potential Value Maize$^{Spot}_{t-1}$		0.000142		0.000054
		(0.000119)		(0.000122)
Δ Potential Value Maize$^{Spot}_{t-2}$		-0.000335***		-0.000098
		(0.000130)		(0.000137)
Δ Spread Spot-Futures Maize	0.000293***	0.000511***	0.000068	0.000068
	(0.000075)	(0.000088)	(0.000081)	(0.000090)
Δ Spread Spot-Futures Maize$_{t-1}$		0.000278*		-0.000355**
		(0.000156)		(0.000165)
Δ Spread Spot-Futures Maize$_{t-2}$		0.000019		0.000058
		(0.000143)		(0.000146)
Log Transition FA$_{t-1}$		-0.496990***		
		(0.003805)		
Log Transition AF$_{t-1}$				-0.514263***
				(0.003288)
Constant	1.161746***	-0.223871***	0.647135***	0.196041***
	(0.020781)	(0.016262)	(0.020544)	(0.018555)
Observations	135985	127478	135982	127469
Adjusted R^2	0.081	0.254	0.045	0.282
Long Run Propensity Soy^{Mean}		0.000426		0.000769
Long Run Propensity Soy$^{25^{th}}$		0.000282		0.000509
Long Run Propensity Soy$^{75^{th}}$		0.000572		0.00103
Long Run Propensity Maize^{Mean}		-0.00124		-0.000982
Long Run Propensity Maize$^{25^{th}}$		-0.000729		-0.000580
Long Run Propensity Maize$^{75^{th}}$		-0.00170		-0.00135
Long Run Propensity Spread Soy^{Mean}		-0.000879		0.00190
Long Run Propensity Spread Soy$^{25^{th}}$		-0.000582		0.00126
Long Run Propensity Spread Soy$^{75^{th}}$		-0.00118		0.00255
Long Run Propensity Spread Maize^{Mean}		0.00218		-0.000611
Long Run Propensity Spread Maize$^{25^{th}}$		0.00129		-0.000361
Long Run Propensity Spread Maize$^{75^{th}}$		0.00300		-0.000841

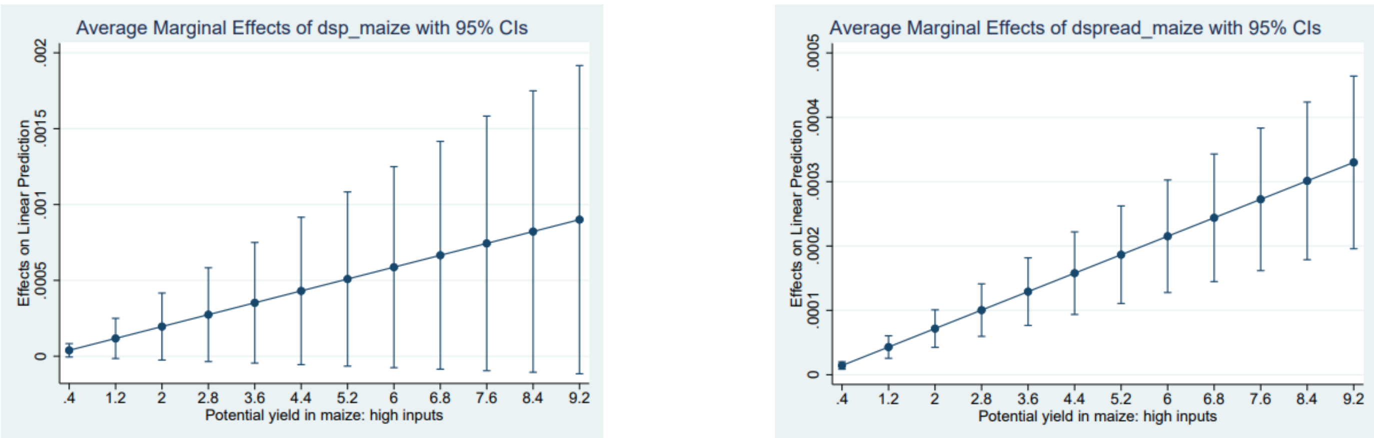
Notes Table 1: Robust standard errors reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Important for the long run propensities. The long run propensities are calculated for three AMC's with different levels of potential yield, namely the *Mean*, 25^{th}, and 75^{th} percentile.

Figure 8: Effect of Prices at Different Levels of Potential Yield on $LogFA_{i,t}$

PANEL A. Margin Plots Soy



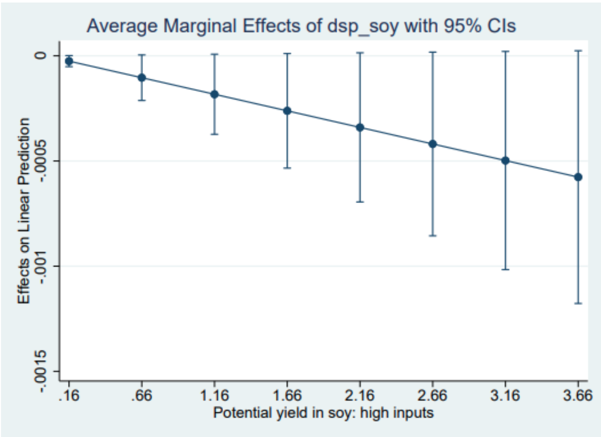
PANEL B. Margin Plots Maize

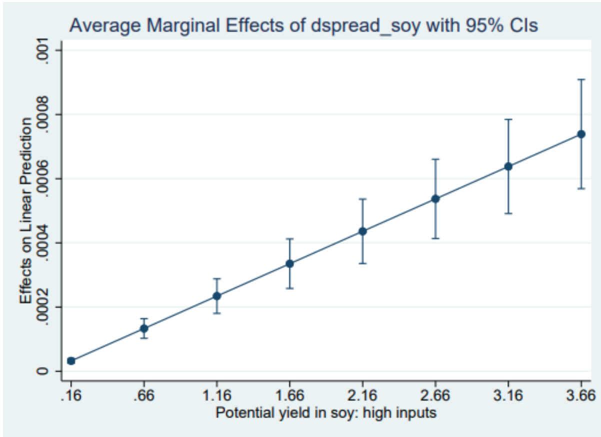


Notes: Panel A shows margin plots for soy. The graph on the left is indicating the effects of the spot depending on the level of potential yield. Potential yield is measured along the x-axis and the y-axis indicate the dependent variable $FA_{i,t}$. The graph on the right is indicating the effects of the spot-future spread for different levels of potential yield. Equal to the graph on the left Potential yield is measured along the x-axis and the y-axis indicate the dependent variable $FA_{i,t}$. Panel B shows margin plots for maize. The graph on the left is indicating the effects of the spot price for maize depending on the level of potential yield. Potential yield is measured along the x-axis and the y-axis indicate the dependent variable $FA_{i,t}$. The graph on the right is indicating the effects of the spot-future spread for different levels of potential yield. Equal to the graph on the left Potential yield is measured along the x-axis and the y-axis indicate the dependent variable $FA_{i,t}$

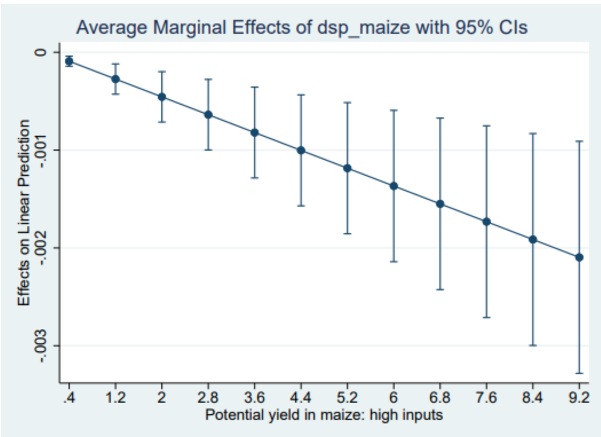
Figure 9: Effect of Prices at Different Levels of Potential Yield on $LogAF_{i,t}$

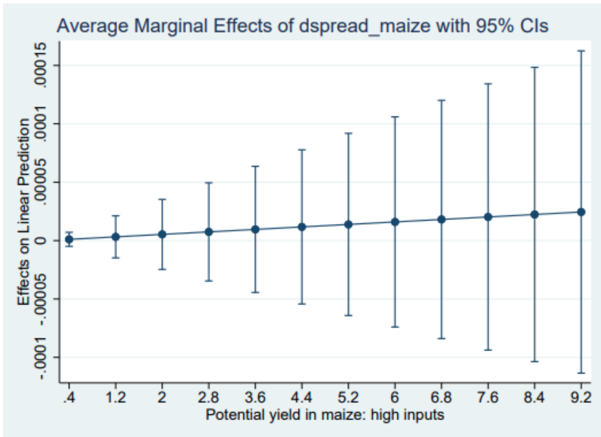
PANEL A. Margin Plots Soy





PANEL B. Margin Plots Maize





Notes: Panel A shows margin plots for soy. The graph on the left is indicating the effects of the spot pirce depending on the level of potential yield. Potential yield is measured along the x-axis and the y-axis indicate the dependent variable $FA_{i,t}$. The graph on the right is indicating the effects of the spot-future spread for different levels of potential yield. Equal to the graph on the left Potential yield is measured along the x-axis and the y-axis indicate the dependent variable $FA_{i,t}$. Panel B shows margin plots for maize. The graph on the left is indicating the effects of the spot price depending on the level of potential yield. Potential yield is measured along the x-axis and the y-axis indicate the dependent variable $FA_{i,t}$. The graph on the right is indicating the effects of the spot-future spread for different levels of potential yield. Equal to the graph on the left Potential yield is measured along the x-axis and the y-axis indicate the dependent variable $FA_{i,t}$

The estimation strategy yields a positive coefficient for the potential value for soy in column 1 and 2, which is in line with our hypothesis that an increase in price leads to an increase in deforestation. Estimates from column 1 indicate that an increase in the change of potential value is associated with an increase in transition from forest to agriculture, and this coefficient is significant at the 1% level. More specifically, it indicates that a 10 dollar increase in the change of potential value is associated with a 0,45% increase in deforestation, which is equivalent to roughly 13,5 soccer fields[5], when using the average transition across all years and all AMC's. The variable of spread between spot and future price of soy indicates that an increase of 10 dollars in the change of the spread is correlated with a decrease of 0,78% in transition from forest to agriculture, which is equal to about 23,5 soccer fields[6]. As the farmers see that the future price is lower relative to the current price, they might expect lower income in the future, potentially leading to decisions that might decrease deforestation as the future incomes are predicted to be lower. What is more, it seems like the magnitude of the spread variable is higher than for the potential value. This is an interesting finding potentially indicating that the variable including future price has a bigger impact on transition of land from forests to agriculture. When looking at maize, estimates indicate that a 10 dollar increase in the change of potential value for maize can be associated with a 0,13% increase in transition from forest to agriculture. Even though this is statistically significant at the 5% level, the economic significance can be questioned in this model, due to the low coefficient value. Using the same comparison of soccer fields as above, this result is associated with approximately 4 soccer fields increase in deforestation, which substantially lower than for soy. Results from the spot-future spread of maize variable also exhibit lower magnitudes than for soy, but it is also substantially higher than the magnitude of potential value variable for maize. There are consistently higher magnitudes of the spread variables, which is an interesting finding.

[5] $(2150 \text{ hectares} \cdot 0,0045) = 9,7 hectars.1 hectare is equal to about 1,4 soccer fields.So we take 9,7 ha \cdot 1,4, which is equal to 13,58.$

[6] Here we use the same calculation strategy as above, $(2150 \cdot 0,0078) \cdot 1,4 = 23,5, using the same mean of transition of land from forest to agriculture across all year and AMC's)$

We also want to estimate the effects our independent variables on transition from agriculture to forest land, in order to determine whether our model shows consistent effects of deforestation or not. In model 3 the potential value of soy is negative as expected, but statistically insignificant. The coefficient also has a rather small magnitude, and should be interpreted with vigilance. The spot-future spread of soy is is very significant and a 10 dollar increase in the change of this spread is associated with a 0,71% increase in transition from agriculture to forest, which can be interpreted as a slowing down of the deforestation and regrowth of forest land. Expanding crop land requires a long term perspective and is assumed expensive investment. As the future price is expected to be lower one might assume that farmers will clear less forest as their income from farming is predicted to decrease in the future, thus there is less incentive to spend time and resources to expand their cropland. For maize the estimates indicate a negative statistically significant relationship between an increase in the change of potential value and transition from agriculture to forest. This is in line with our hypothesis that an increase in commodity prices will be associated with an increase in deforestation, or a decrease in the regrowth of forests. Estimating the change in the spot-future spread in model 3 indicates that the effect of this variable is insignificant on transition from agriculture to forest land. Model 1 and 3 indicate consistent results of the effects on deforestation, with the exception of the spread variable for maize, which is positive in both models and also insignificant in Model 3. In contrast to maize, the coefficients for soy indicate that it is only the spread that is significant for soy in model 3. Moreover, the potential value for soy has substantially lower magnitude in model 3. This result could be indicating that the spot price of soy is less important for farmers in model 3 than the future prices, due to the possibility that transforming land from agriculture to forest requires more time than clearing forests for agricultural expansion. Farmers producing soy and farmers producing maize may exhibit different behavior and decision making strategies. We assume that soy farmers are generally more professional and large scale producers than maize farmers. Thus, soy farmers may act differently in terms of having a more long-term perspectives than maize farmers. Maize farmers may put more emphasis on the current spot price, as they might have more of a short-term perspective and tend to have more small scale production.

Even though some estimates are significant, model 3 seems to lack statistical significance to be a good model for estimating the effects of potential value of our commodities on transition from agriculture to forest. A possible explanation for this could be that there are a multitude of factors affecting transition from agricultural land to forests, i.e politics or demographics. It will require more time transforming land from agricultural land to forest land than the other way around.

In the following panels below Figure 8 and 9 shows margin plots displaying how the effect of of potential value and spread is increasing or decreasing depending on the change in potential yield. These plots are displayed for both for soy and maize and both dependent variables.

According to the margin plot to the left in Panel A in figure 8 AMCs with a higher level of potential yield will experience a bigger effect of a change in potential value on transition from forest to agriculture, in contrast to AMCs with lower levels of potential yield. This is in line with what we may expect as farmers within AMCs with high levels of potential yield of soy are assumed to be more likely to clear forest for production when the price of soy increases. The plot to the right in Panel A indicates that a higher level of potential yield within an AMC, is correlated with a stronger negative effect of the coefficient for the change in spot-future spread. More specifically this shows a similar sensitivity to the change in prices for AMCs with high potential yield.

The margin plot to the left in Panel B in Figure 8, indicates that the effect of increased levels of potential yield of maize, increases the effect of potential value of maize on transition from forest to agriculture. Similar with the results from Panel A, the relationship also indicates a higher price sensitivity for AMCs with higher potential yield. To the right in panel B, the plot indicates that the transition from forest to agriculture in AMCs with high potential yield are seemingly more sensitive to a change in the spread than the spot price. Moreover, the plots in panel B show that the estimates of the spot price and spot-future spread are more uncertain for AMC's with higher potential yield, as the confidence intervals increases. Compared to plots for soy, estimates for maize seems to indicate bigger uncertainty linked to the price estimations.

We also construct the same margin plots in 9 and substitute the transition of forest to agriculture variable with transition from agriculture to forest. Our plots shows that the uncertainty of our estimates are quite similar for potential value of soy and maize by looking at the 95% intervals. For the spot-future spread, on the other hand soy seems to have less uncertainty than maize.

In model 2 and 4 we add two yearly lags on model 1 and 3 in order to estimate the long term effects of the independent variables. The coefficient for change in potential value of soy changes slightly in magnitude when adding lags. An increase of 10 dollars in the change of potential value is now associated with a 0,32% increase in the yearly transition from forest to agriculture. The estimates of the long-run propensities displayed the bottom of the table indicate that the long term change in potential value with spot price is also associated with an increase in transition from forest to agriculture. Moreover, the estimates of long-run propensities of the change in potential value of soy indicate that long term effect of price changes have a larger effect when potential yield is high. A 10 dollar increase in the long term effect of change in potential value is associated with a 0,43% increase in transition from forest to agriculture in the AMC where potential yield is at mean, while the AMC with potential yield at the 25^{th} percentile is associated with a 0,28% increase, and a 0,57% increase at the 75^{th}. The long run propensity for potential value of spot is significant at 5% level, see A.1. When looking at the estimates of the spot-future spread for soy, one can see the same pattern, with the exception that the lagged variables are significant at the 1% level. According to the estimates of the long run propensities of the spread variable for soy, a 10 dollar increase in the long term effect of change in spread is correlated with a 0,88% decrease in transition from forest to agriculture in the AMC with a mean level of potential yield, equivalent to approximately 26,5 soccer fields. The AMCs with potential yield at the 25% and 75% percentile experience a weaker and stronger negative effect, respectively. The long run propensities are only significant at 10% level for the spot-future spread, see A.1. This indicates that a long term effect of increase in spread will lead to a more favourable effect for forests in AMCs with lower potential yield. This may be due to AMCs with higher levels of potential yield for soy being more sensitive to changes in spread of soy in the long term, as being similarly observed in the plots above. Hence, if the spread increases and price of the future is expected to be relatively lower than now, this will have a larger economic impact for AMCs with high potential yield in the long term.

In model 2 the coefficient for potential value of maize changes from positive to negative, and significant at the 1% level. This is somewhat contrary to our hypothesis that an increase in price increases the transition of land from forest to agriculture. Model 2 indicates that an increase in the change of potential value of maize is associated with a decrease in the transition from forest to agriculture. Even though this coefficient is statistically significant, it is slightly hard to interpret economically. One possible explanation may be that some farmers may choose to double crop if the spot price of maize increases, in other words they plan a second season maize. It is also possible that soy farmers may choose to plant an extra crop of second season maize in order to maximize the income from their crop land (Harsheim & Nakkim, 2018). The long run effect of potential value of maize is estimated at -0,1247% and is significant at 5% level, see A.1. The spot-future spread of maize increases slightly in magnitude in model 2, and stays highly significant. Here an increase in the change of this spread is associated with a 0,51% increase in deforestation compared to 0,29% in model 1. Whether this increase in deforestation can be attributed to expansion of maize crops is not easy to determine, but the results are in line with our hypothesis that an increase in commodity prices increases deforestation. Similar to the results from soy, we can see that the coefficient of the spread variable has a larger magnitude than the potential value variable. The long run effect of change in spot-future spread of maize is positive and significant at 5%. It is higher for AMCs with higher levels of potential yield, which can be explained by AMCs with higher levels of potential yield being more adaptive to long term effects of changes in spread. For both commodities, AMCs with high levels of potential yield experience a higher magnitude of the long term effect of spread than AMCs with lower potential yield.

In model 4 estimates show that the coefficient for change in potential value of soy changes from insignificant in model 3 to positive and significant at the 5% level. This now indicated that a 10 dollar increase in the spot price is associated with a 0,20% increase in transition from agriculture to forest, in other words an increase in the potential value of soy is correlated with a favorable effect for forests. On the other hand, the one year lagged variable is correlated with an unfavorable effect for forests, before the two year lag indicates the same effect as today's potential value of soy. It is quite strange that a 10 dollar increase in the potential value of soy can be associated with an increase in yearly regrowth equivalent to roughly 5 soccer fields[7]. The long run propensity for potential value of soy is not significant. Estimates for the spot-future spread for soy show a slight increase in magnitude when we add two yearly lags. The coefficient is consistent with model 3, as opposed to the coefficient for potential value of soy. In other words, these estimates indicates a slowing down of the deforestation, and potentially regrowth of forest land, similar to model 3. The long run propensity of spot-future spread is significant at 1% level, and shows increasingly higher long term effect on transition from agriculture to forest for AMCs with higher potential yield. This is in line with what we expect. When looking at the estimates for maize, it is clear that the coefficients do not change much in model 4 from model 3. The potential value of maize is still indicating a significant negative effect on transition from agriculture to forest, and he long term effect for is significant at 5% level. This long term effect shows a stronger negative magnitude for AMCs with higher levels of potential yield, which is also as expected. The spot-future spread of today is still insignificant, and the long run propensity is far from significant.

[7]This is calculated by taking the average transition for all AMC's across all years, and calculating percentage change to soccer fields: $(1723{\cdot}0,0020) \cdot 1,4 = 4,82$

6.2 Summary of the Results

Mostly, our analysis is showing us that our hypothesis is right; increase of potential value is speeding up the deforestation, and expected fall in future price relatively to the spot price, is slowing down this process. Exceptions are the positive coefficient of spot-future spread in maize, which indicates that lower future price relatively to the spot price, is speeding up deforestation.

In general our results indicate that the spot-future spread for both our commodities exhibit a stronger effect on land transition, both in terms of transition from forest to agriculture and from agriculture to forest. In other words that the expected price development of the future affects land transition more than the market price of today, showing a larger magnitude in our models. In appendix A.6 we find indications of the same pattern. We investigate the robustness of our analysis of the effect of commodity prices on land transition from forest to agriculture by including estimates of the same commodity prices on land transition from agriculture to forest. Results of the effect on agriculture to forest transition are interpreted as the effect on regrowth of forests. The models using $LogAF_{i,t}$ as dependent variable seems to lack some statistical significance in comparison to models using $LogFA_{i,t}$. Furthermore, we conducted a fixed effects analysis in A.6 to further check the robustness of our results. Even though the fixed effects model show the same pattern of future prices having a larger magnitude than spot prices, this fixed effects model show some rather surprising results. Especially since the estimations in the fixed effects model indicate that an increase in the potential value of soy generally can be associated with a decrease in transition from forest to agriculture. This result is counter intuitive and contradicts established research, thus we do not find the estimates trustworthy. Furthermore, our results for maize are somewhat more ambiguous than our results for soy, especially in model 3 and 4. From these two checks for robustness we are not able to be confident in the robustness of our main results.

Our analysis also show that increased potential yield do affect land transition. Our margin plots in figure 8 and 9 illustrate this, with the effect of our commodity price variables on land transition[8] being more profound for AMC's with high potential yield. From the maps in 4.2.1 is is clear that some areas in Brazil are much better suited for cultivation of soy, and some for maize, so in order to get better estimates of the effects of spot and future prices on municipal deforestation, it may be better to examine these areas more in detail. Alternatively expand on the number of commodities to examine, i.e other crops, minerals or livestock. Still, we have found results indicating that there is an interesting relationship between the spot-future spread and deforestation in Brazilian municipalities.

[8]For both $LogFA_{i,t}$ and $LogAF_{i,t}$

Especially the indications that an expected lower commodity price in the future is slowing down deforestation in Brazil.

7 Discussion

In the following section, we discuss possible shortcomings of the dataset and the limitations imposed on our study as a result of our chosen estimation strategy. Finally, we conduct a short discussion on the external validity of our results.

7.1 Limitations to the Dataset

One potential limitation in our data is that we have chosen to only exploit the potential yield using high technology input as mentioned in section 4.2.1. As we cannot be absolutely certain that all AMCs in Brazil exploit high technology inputs this could potentially alter the actual effects of prices on municipal deforestation. Furthermore, there might be a chance that one of the crops, either maize or soy, mainly exploit high technology inputs while the other crops depends more on low technology inputs, which could also affect the actual price effects. In our analysis it is then assumed that AMCs with high technology inputs for soy also have high technology inputs for maize. Furthermore, figure 6 and 7 indicate that Brazil is relatively better suited for soy production than maize production.

In addition, our choice of average global prices is a simplified approach. Brazil is a large exporter of soy, and it is likely that Brazil may have market power especially in the soy market, which might violate our assumption of exogenous prices. It is also a simplification in terms of the prices the Brazilian farmers relate to. We do not differentiate between farmers, only between crops, and we do not know whether farmers are taking spot prices, or future prices or both prices into account. Additionally, we can not be sure whether all farmers of our two commodities take the average global prices into account, as opposed to their own reference prices. These reference prices may be more important to local small scale farmers. As we try to estimate the effect of commodity prices on municipality level, it could be argued that reference prices could be a better to estimate the effect of commodity prices on municipal deforestation. In the future it could interesting to conduct a qualitative study of how farmers in different Brazilian municipalities relate to spot and future prices. Furthermore, we only include one-year futures contracts in our future price data, and futures contracts for commodities are traded on more long term contracts than just one year. Adding different contract lengths could possibly further exploited the fact that our time period is substantially longer than other similar research papers. The choice of global average prices also gave us a challenge in our panel data estimations, due to the fact that the price only varies over time, but is fixed for all AMCs. Thus we had to complicate our price variables somewhat by generating a measure of potential value by multiplying our different price variables with the potential yield in our different AMCs.

This complicates our analysis.

Also, using transition from agriculture to forest, in other words regrowth may not be the best data to use in order to check for robustness of our analysis. The models investigating the effect of the independent variables on transition from agriculture to forest, generally lacks the same statistical significance as our models using transition from forest to agriculture as dependent variable. Prices may not be the best at determining regrowth of forests or a slowing down of agricultural expansion, such a transition could also be affected by politics, demographics or potentially the changing climate.

7.2 Limitations to the Estimation Strategy

One of the main limitations of our estimation strategy, is that we likely have an omitted variable bias. We cannot be sure that the expected value of the error term in our analysis is equal to zero given any value of the independent variables. Hence, there might be factors in the error term correlated with both our independent and dependent variable and generate bias in our coefficients. It is likely that there are other commodity prices affecting both soy and maize prices as well as land transition in Brazil. Such as beef, sugar, coffee, or other commodity prices. As mentioned in 3.3 Brazil is the world's largest exporter of beef in 2018, providing almost 20% of total global beef exports according to the USDA (2018). Research from Barona et al. (2010) show that there is a correlation between cattle price and deforestation, basing the research on the legal Amazon. Additionally Nepstad et al. (2006) suggest that the growth of the Brazilian soy industry may have indirectly led to the expansion of the cattle herd. According to Nepstad et al. (2006), soy has driven up land prices in the Amazon, allowing many cattle ranchers to sell valuable holdings at enormous capital gains and purchase new land further north and expand their herd further. As mentioned above, our data from section 4.2.1 illustrates that soy has a larger potential yield than maize, indicating that soy would naturally be more preferred than maize cultivation. It could be argued that also other commodities may be more relevant for deforestation than maize. Consequently, including only soy and maize in our model is likely create biased estimates and we are not able to causally infer the effect of spot prices and future prices on municipal deforestation in the whole of Brazil.

This limitation would also apply to our fixed effects estimation in appendix A.6, as it is likely this estimation also suffers from omitted variable bias. This is especially clear when looking at the counterintuitive results obtained for the potential value of maize using spot prices.

7.3 External Validity

Essential to any study is the extent that the results can be generalized, more precisely external validity. As mentioned in Subsection 3.4, Brazil sustains 40% of the world's remaining tropical forest. Hence, it is possible that the estimated effects from Brazil are not directly transferable to other countries. A country with significantly less forest and/or agricultural sector will naturally not be as exposed to deforestation resulting from commodity prices. Consequently, it is possible that the estimated effects found in our analysis are not directly transferable to other countries, sustaining less, or other types, of forests. Still, countries with large forest areas could potentially experience similar effects of increased spot and future prices. Even though our results cannot infer an actual causal relationship between deforestation and our price variables, it is still interesting that the results point in the direction of expected future prices potentially having a significant impact on deforestation.

8 Concluding Remarks

This thesis aims to provide *empirical evidence of the effect of spot and future prices on municipal transition of land from forest to agriculture*, more specifically deforestation, in Brazil. We exploit price data from two different commodities, namely soy and maize as these are some of the most well known crops in Brazil. In order to find evidence in favour of, or against, the predicted effects, we have used two different estimation strategies, with a First Difference strategy as our first and main strategy inspired by Hilde C. Bjørnland (2019). We use the potential value of each commodity and spot-future spread of each commodity as our independent variables, and *Transition of land from forest to agriculture* as our main dependent variable. We also estimate the effects of our independent variables on *Transition of land from agriculture to forest* in order to check for robustness of our results. Additionally, we use a Fixed Effects strategy as an extra analysis in appendix A.6 to further check the robustness of our analysis. We exploit newly digitized data on land transition from MapBiomas for the whole of Brazil, in an effort to estimate changes in land transition from forest to agriculture and from agriculture to forest. Both types of land transition are included as we try to estimate as much of a holistic effect of spot and futures as possible. We combine this data with an exogenous measure of potential yield for each commodity from the FAO-GAEZ database, in order to get variation across units.

Our main finding is that the spot-future spread has a seemingly larger effect on deforestation than the spot price alone for soy. We illustrate this in the terms of soccer fields, as this is a widely used measure when addressing deforestation. This is contributing to the extensive research on deforestation in Brazil by suggesting that the price expectations for the future should be included in further research on Brazilian deforestation. Maize show more ambiguous effects.

Due to the long time span of our data and the possible long term effect of prices, we added two year lagged values of our all independent variables. We found that the long run propensities are statistically significant in our models with transition from forest to agriculture as dependent variable, and have the same signs as the associated non-lagged variables. This correlation shows less significance in the estimation with transition form agriculture to forest as dependent variable. From this we can imply that long term price changes are more important of removal of forest, than transforming agricultural land. This is rather intuitive as transforming agriculture to forest land requires more time and resources than clearing forests. Incentives for removing or relocating agricultural activities would likely be more affected by politics or environmental changes.

From our analysis we find that AMCs with higher potential yield for our crops will experience a bigger effect of a change in the price variables on land transition. This is likely due to the fact that cultivation of soy and maize is larger in areas with a high potential yield. Thus a change in the commodity price will have a large effect on these AMCs. Especially since many of these AMCs are located outside the legal Amazon. From figure 6 it is clear that there are regions south of the legal Amazon that are even better suited to grow soy than in much of the Amazon. For maize we can see some of the same indications from figure 7. Regions to the south of the legal amazon and regions in the north east display better conditions for maize cultivation. Thus, one of the strengths of our analysis is that we look at AMCs for the whole of Brazil and not just in the Amazon region. Furthermore, these maps indicate that Brazil is relatively better suited for soy production than maize production. For further research it could be interesting to examine AMCs with higher potential yield more closely. Moreover, in our analysis we briefly discuss that farmers of soy and maize may have different decision strategies. Thus, in section 7.1 we discuss that it could be argued that reference prices could be a better to estimate the effect of commodity prices on municipal deforestation. In the future it could interesting to conduct a qualitative study of how farmers in different Brazilian municipalities relate to spot and future prices.

Reflecting on our results and how complex the challenge of deforestation is, we believe our simplified models lack the power to infer any causal relationships. There are a multitude of factors affecting deforestation in Brazil, and prices are just one of these factors. As mentioned in section 3.3 nearly 80% of the deforestation in the Amazon can be attributed to illegal logging, implying that Brazil is suffering from the tragedy of the commons. Even with some established land rights, these rights are clearly not respected. Thus cooperation between farmers, society and the government is crucial in order to fight deforestation. Prices alone, cannot explain the complex issue of deforestation. Even though we cannot provide causal inference of our estimates due to biased estimates[9], we hope this thesis may trigger interest in further research into the extent that both spot and futures affect municipal deforestation in Brazil or other countries experiencing significant deforestation, and especially focus on the the effects of futures.

[9]As discussed in section 7.2

References

Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. (2017). *When should you adjust standard errors for clustering?* National Bureau of Economic Research.

Achard, F., Eva, H. D., Stibig, H., Mayaux, P., Gallego, J., Richards, T., & Malingreau, J. (2002). Determination of Deforestation Rates of the World's Humid Tropical Forests. *Science*, *297*(5583), 999–1002.

Allen, E., & Valdes, C. (2016). Brazil's Corn Industry and the Effect on the Seasonal Pattern of US corn exports. *AES-93 Economic research Service/USDA*.

Andersen, L. E. (1996). The Causes of Deforestation in the Brazilian Amazon. *The Journal of Environment & Development*, *5*(3), 309–328.

Asche, F., & Guttormsen, A. G. (2002). Lead lag relationships between futures and spot prices.

Assuncao, J., Lipscomb, M., Mobarak, A. M., & Szerman, D. (2016). *Agricultural Productivity and Deforestation in Brazil.* Mimeo, Yale University. 37 The methodology is explained by Camara, Valeriano.

Baldi, L., Peri, M. [Massimo], & Vandone, D. (2011). *Price discovery in agricultural commodities: The shifting relationship between spot and future prices*.

Baldi, L., Peri, M. [MassimoM], & Vandone, D. (2011). Spot and futures prices of agricultural commodities: Fundamentals and speculation. *Proceedings in Food System Dynamics*, 110–125.

Bank, W. (2018). Soybean International Commodity Profile. Retrieved December 14, 2019, from http://siteresources.worldbank.org/INTAFRICA/Resources/257994-1215457178567/Soybean_Profile.pdf

Barona, E., Ramankutty, N., Hyman, G., & Coomes, O. T. (2010). The role of pasture and soybean in deforestation of the brazilian amazon. *Environmental Research Letters*, *5*(2), 024002.

Benhin, J. K. (2006). Agriculture and deforestation in the tropics: A critical theoretical and empirical review. *AMBIO: A Journal of the Human Environment*, *35*(1), 9–17.

Bjørnland, H. C., Nordvik, F. M., & Rohrer, M. (2019). Supply flexibility in the shale patch: Evidence from north dakota.

Bloomberg L.P. (2019). Where data, people and ideas intersect. Retrieved December 18, 2019, from https://www.bloomberg.com/company/

Bloomberg Terminal. (2019). Data received from DNB analyst.

Boerema, A., Peeters, A., Swolfs, S., Vandevenne, F., Jacobs, S., Staes, J., & Meire, P. (2016). Soybean Trade: Balancing Environmental and Socio-economic Impacts of an Intercontinental Market. *PloS one*, *11*(5), e0155222.

Bolling, H. C., Somwaru, A., & Kruse, J. B. (2001). *The united states in the global soybean market: Where do we go from here?*

Brander, J. A., & Taylor, M. S. (1998). Open access renewable resources: trade and trade policy in a two-country model. *Journal of International Economics*, *44*(2), 181–209.

Bryant, D., Nielson, D., & Tangley, L. (1997). The Last Frontier Forests. *Issues in Science and Technology*, *14*(2), 85.

Bustos, P., Caprettini, B., & Ponticelli, J. (2016). Agriculture Productivity and Structural Transformation: Evidence from Brazil. *American Economic Review*, *106*(6), 1320–1365.

Bustos, P., Caprettini, B., & Ponticelli, J. (2018). Technical Appendix to Agricultural Productivity and Structural Transformation: Not for Publication. Retrieved from https://www.aeaweb.org/articles?id=10.1257/aer.20131061

Cattaneo, A. (2002). *Balancing Agricultural Development and Deforestation in the Brazilian Amazon*. Intl Food Policy Res Inst.

CONAB. (2012). *Levantamento De Avaliação Da Safra*. Companhia Nacional de Abastecimento.

Copeland, B. R., & Taylor, M. S. (2009). Trade, tragedy, and the commons. *American Economic Review*, *99*(3), 725–49.

Crain, S. J., & Lee, J. H. (1996). Volatility in wheat spot and futures markets, 1950–1993: Government farm programs, seasonality, and causality. *The journal of finance*, *51*(1), 325–343.

Dixit, A. (1984). International trade policy for oligopolistic industries. *The Economic Journal*, *94*(Supplement), 1–16.

Ehui, S. K., & Hertel, T. W. (1989). Deforestation and Agricultural Productivity in the Côte d'Ivoire. *American Journal of Agricultural Economics*, *71*(3), 703–711.

Eskeland, G. (2019). *Combined note on game theory 2019*. NHH, Department of Business Management and Science.

FAO. (2016). GLOBAL FOREST RESOURCES ASSESSMENT 2015, How are the world's forests changing?, Second Edition. Retrieved December 15, 2019, from https://www.ers.usda.gov/webdocs/publications/87459/oce-2018-1.pdf?v=43146.

FAO. (2018a). Cuntry Profiles: Brazil. Food and Agriculture Organization of the United Nations. Retrieved November 3, 2018, from http://www.fao.org/countryprofiles/index/en/?iso3=BRA

FAO. (2018b). GAEZ - Global Agro-Ecological Zones. Retrieved September 25, 2018, from http://www.fao.org/nr/gaez/en/#/

FAO & IIASA. (2018a). Agricultural Suitability and Potential Yields - Global Agro-Ecological Zones (GAEZ). Retrieved November 2, 2018, from http://www.fao.org/nr/gaez/about-data-portal/agricultural-suitability-and-potential-yields/en/

FAO & IIASA. (2018b). Data Portal - Global Agro-Ecological Zones (GAEZ). Retrieved December 8, 2018, from http://gaez.fao.org/

FAO & IIASA. (2018c). Definitions - Global Agro-Ecological Zones (GAEZ). Retrieved November 2, 2018, from http://www.fao.org/fileadmin/user_upload/gaez/docs/Definitions_EN.pdf

Faria, W. R., & Almeida, A. N. (2016). Relationship between openness to trade and deforestation: Empirical evidence from the Brazilian Amazon. *Ecological Economics*, *121*, 85–97.

Fearnside, P. M. (2005). Deforestation in Brazilian Amazonia: History, Rates, and Consequences. *Conservation biology*, *19*(3), 680–688.

Fund, W. W. (2016). Oilseeds and products update 16. Retrieved December 14, 2019, from https://www.worldwildlife.org/stories/the-story-of-soy#

Garbade, K. D., & Silber, W. L. (1983). Price movements and price discovery in futures and cash markets. *The Review of Economics and Statistics*, 289–297.

Ghani, S. S. A. (2019). Could the EU's ban on palm oil in biofuels do more harm than good? Retrieved December 16, 2019, from http://countrystudies.us/brazil/80.htm

Global Forest Watch. (2018). Forest Change. Retrieved October 27, 2018, from https://www.globalforestwatch.org/dashboards/global?category=forest-change&faoDeforest=eyJwZXJpb2QiOjIwMTB9

Government, N. (2019). Brazil sees reduced payments from Norway following increased deforestation in the Amazon. Retrieved December 16, 2019, from https://www.regjeringen.no/en/aktuelt/okt-avskoging-i-brasils-regnskog-gir-redusert-utbetaling-fra-norge/id2581396/

Guardian, T. (2017). The Amazon's new danger: Brazil sets sights on palm oil. Retrieved December 16, 2019, from https://www.theguardian.com/sustainable-business/2017/jun/29/brazil-palm-oil-amazon-rainforest-deforestation-temer-farming-para-cerrado

Hansen, C. B. (2007). Generalized least squares inference in panel and multilevel models with serial correlation and fixed effects. *Journal of econometrics*, *140*(2), 670–694.

Harding, T., Herzberg, J., & Kuralbayeva, K. (2006). Commodity Prices and Robust Environmental Regulation: Evidence from Deforestation in Brazil. (1), 1–35.

Harsheim, I. G., & Nakkim, V. L. (2018). *Agricultural productivity and the fight against deforestation: An empirical analysis of the effect of factor-biased technical change on agricultural land use and forests in brazil* (Master's thesis).

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica: Journal of the econometric society*, 1251–1271.

Hernandez, M., Torero, M. et al. (2010). *Examining the dynamic relationship between spot and future prices of agricultural commodities*. International Food Policy Research Institute (IFPRI).

Hudson, R. A. (1997). Country Studies Brazil: Trade Policies. Retrieved December 16, 2019, from http://countrystudies.us/brazil/80.htm

IBGE. (1996). *Censo Agropecuário 1995-1996*. Instituto Brasilerio de Geografia e Estatística: Rio de Janeiro.

IBGE. (2006). *Censo Agropecuário 2006*. Instituto Brasilerio de Geografia e Estatística: Rio de Janeiro.

IBGE. (2016). Produção Agrícola Municipal. Retrieved December 1, 2018, from https://sidra.ibge.gov.br/pesquisa/pam/tabelas

International Monetary Fund. (2019). Primary Commodity Price System. Retrieved December 17, 2019, from https://data.imf.org/?sk=471DDDF8-D8A7-499A-81BA-5B332C01F8B9

International Panel on Climate Change. (2018). ar4-wg3-chapter9-1 Forestry. Retrieved December 1, 2019, from https://www.ipcc.ch/site/assets/uploads/2018/02/ar4-wg3-chapter9-1.pdf

Kirby, K. R., Laurance, W. F., Albernaz, A. K., Schroth, G., Fearnside, P. M., Bergen, S., ... Da Costa, C. (2006). The Future of Deforestation in the Brazilian Amazon. *Futures*, *38*(4), 432–453.

Kissinger, G., Herold, M., & De Sy, V. (2012). *Drivers of deforestation and forest degradation: A synthesis report for redd+ policymakers*. Lexeme Consulting.

MapBiomas, P. (2018). MapBiomas General “Handbook”. Algorithm Theoretical Basis Document (ATBD). Collection 3 - Version 1.0. Retrieved September 24, 2018, from http://mapbiomas.org/pages/atbd

MapBiomas, P. (2019). Collection 3.1-3.2 of Brazilian Land Transition Map Series. Retrieved from https://mapbiomas.org/en/download_estatisticas?cama_set_language=en

Martinelli, L. A., Naylor, R., Vitousek, P. M., & Moutinho, P. (2010). Agriculture in Brazil: Impacts, Costs, and Opportunities for a Sustainable Future. *Current Opinion in Environmental Sustainability*, *2*(5-6), 431–438.

MathWorks. (2019). Time Series Regression VIII: Lagged Variables and Estimator Bias. Retrieved December 15, 2019, from https://www.mathworks.com/help/econ/examples/time-series-regression-viii-lagged-variables-and-estimator-bias.html

Mongbay. (2019). Brazilian hunger for meat fattened on soy is deforesting the Cerrado. Retrieved December 14, 2019, from https://news.mongabay.com/2019/01/brazilian-hunger-for-meat-fattened-on-soy-is-deforesting-the-cerrado-report/

Nepstad, D., Soares-Filho, B. S., Merry, F., Lima, A., Moutinho, P., Carter, J., … Schwartzman, S., et al. (2009). The End of Deforestation in the Brazilian Amazon. *Science*, *326*(5958), 1350–1351.

Nepstad, D. C., Stickler, C. M., & Almeida, O. T. (2006). Globalization of the amazon soy and beef industries: Opportunities for conservation. *Conservation biology*, *20*(6), 1595–1603.

of Agriculture, U. S. D. (2018). Oilseeds and products update 16. Retrieved December 14, 2019, from https://www.fas.usda.gov/data/%7BBrazil%7D-oilseeds-and-products-update-16.

Pires, G. F., Abrahão, G. M., Brumatti, L. M., Oliveira, L. J., Costa, M. H., Liddicoat, S., … Ladle, R. J. (2016). Increased Climate Risk in Brazilian Double Cropping Agriculture Systems: Implications for Land Use in Northern Brazil. *Agricultural and forest meteorology*, *228*, 286–298.

Rattray, J. (2012). The implications of the increasing global demand for corn. *UW-L Journal of Undergraduate Research XV*.

Reuters. (2016). Brazil losing forest the size of two soccer fields per minute. Retrieved December 19, 2019, from https://www.reuters.com/article/us-latam-environment-brazil-forests/brazil-losing-forest-the-size-of-two-soccer-fields-per-minute-idUSKCN0XQ1MP

Richards, P. D., Myers, R. J., Swinton, S. M., & Walker, R. T. (2012). Exchange rates, soybean supply response, and deforestation in South America. *Global environmental change*, *22*(2), 454–462.

Robalino, J., & Herrera, L. D. (2010). *Trade and deforestation: A literature review*. WTO Staff Working Paper.

Sajwan, R., & Chetty, P. (2018). How to perform Granger causality test in STATA. Retrieved October 31, 2019, from https://www.projectguru.in/publications/granger-causality-test-stata/

Sandvik, H. J. (2018). A Brazilian soy story : how international soy demand affects deforestation and agricultural land use. Retrieved December 14, 2019, from https://openaccess.nhh.no/nhh-xmlui/bitstream/handle/11250/2560748/masterthesis.PDF?sequence=1&isAllowed=y

Sasaki, N. [N.], & Putz, F. E. [F. E.]. (2009). Critical Need for New Definitions of "Forest" and "Forest Degradation" in Global Climate Change Agreements. *Conservation Letters*, *2*(5), 226–232.

Sasaki, N. [Nophea], & Putz, F. E. [Francis E]. (2017). Deforestation of the brazilian amazon. *Cxford Research Encyclopedia of Environmental Science*.

Schnepf, R., Bolling, C., Dohlman, E., et al. (2001). *Agriculture in Brazil and Argentina*. United States Department of Agriculture, Economic Research Service.

Shogren, J. (2013). *Encyclopedia of energy, natural resource, and environmental economics*. Newnes.

Silvapulle, P., & Moosa, I. A. (1999). The relationship between spot and futures prices: Evidence from the crude oil market. *Journal of Futures Markets: Futures, Options, and Other Derivative Products*, *19*(2), 175–193.

Statistics-How-To. (2016). Unit Root: Simple Definition, Unit Root Tests. Retrieved December 5, 2019, from https://www.statisticshowto.datasciencecentral.com/unit-root/

Stern, S. N. (2006). Stern Review: The Economics of Climate Change. (1), 5–7.

Stock, J. H., & Watson, M. W. (2008). Heteroskedasticity-robust standard errors for fixed effects panel data regression. *Econometrica*, *76*(1), 155–174.

The World Bank. (2019). World Bank Commodity Price Data (The Pink Sheet), annual prices 1960 to present. Retrieved September 1, 2019, from https://www.worldbank.org/en/research/commodity-markets

U.S. Department of Labor Bureau of Labor Statistic. (2019). Consumer Price Index Data from 1913 to 2019. Retrieved November 12, 2019, from https://www.usinflationcalculator.com/inflation/consumer-price-index-and-annual-percent-changes-from-1913-to-2008/

University of Utah. (2019). Checking Linearity. Retrieved December 15, 2019, from https://campusguides.lib.utah.edu/c.php?g=160853&p=1054160

USDA. (2018). Brazil Once Again Becomes the World's Largest Beef Exporter. Retrieved December 16, 2019, from https://www.ers.usda.gov/amber-waves/2019/july/brazil-once-again-becomes-the-world-s-largest-beef-exporter/

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. The MIT Press.

Wooldridge, J. M. (2016). *Introductory econometrics: A modern approach*. Cengage Learning.

Yang, J., Bessler, D. A., & Leatham, D. J. (2001). Asset storability and price discovery in commodity futures markets: A new look. *Journal of Futures Markets: Futures, Options, and Other Derivative Products*, *21*(3), 279–300.

A Appendix

A.1 Long-Run Propensity Estimations with Corresponding Significance Levels

TABLE A. Long Run Propensity

FA potential value spot soy mean

D.ln_sum_FA	Coef.	Std.Err.	z	P>z	[95%Conf.	Interval]
_nl_1	0.001	0.000	2.470	0.014	0.000	0.001

FA potential value spot soy p25

D.ln_sum_FA	Coef.	Std.Err.	z	P>z	[95%Conf.	Interval]
_nl_1	0.000	0.000	2.470	0.014	0.000	0.001

FA potential value spot soy p75

D.ln_sum_FA	Coef.	Std.Err.	z	P>z	[95%Conf.	Interval]
_nl_1	0.001	0.000	2.470	0.014	0.000	0.002

FA potential value spot maize mean

D.ln_sum_FA	Coef.	Std.Err.	z	P>z	[95%Conf.	Interval]
_nl_1	-0.001	0.001	-2.420	0.016	-0.003	-0.000

FA potential value spot maize p25

D.ln_sum_FA	Coef.	Std.Err.	z	P>z	[95%Conf.	Interval]
_nl_1	-0.001	0.000	-2.420	0.016	-0.001	-0.000

FA potential value spot maize p75

D.ln_sum_FA	Coef.	Std.Err.	z	P>z	[95%Conf.	Interval]
_nl_1	-0.002	0.001	-2.420	0.016	-0.003	-0.000

Figure 10: Forest to Agriculture - Potential Value

TABLE B. Long Run Propensity

FA spread soy mean

D.ln_sum_FA	Coef.	Std.Err.	z	P>z	[95%Conf.	Interval]
_nl_1	-0.001	0.000	-1.940	0.053	-0.002	0.000

FA spread soy p25

D.ln_sum_FA	Coef.	Std.Err.	z	P>z	[95%Conf.	Interval]
_nl_1	-0.001	0.000	-1.940	0.053	-0.001	0.000

FA spread soy p75

D.ln_sum_FA	Coef.	Std.Err.	z	P>z	[95%Conf.	Interval]
_nl_1	-0.001	0.001	-1.940	0.053	-0.002	0.000

FA spread maize mean

D.ln_sum_FA	Coef.	Std.Err.	z	P>z	[95%Conf.	Interval]
_nl_1	0.002	0.001	2.530	0.012	0.000	0.004

FA spread maize p25

D.ln_sum_FA	Coef.	Std.Err.	z	P>z	[95%Conf.	Interval]
_nl_1	0.001	0.001	2.530	0.012	0.000	0.002

FA spread maize p75

D.ln_sum_FA	Coef.	Std.Err.	z	P>z	[95%Conf.	Interval]
_nl_1	0.003	0.001	2.530	0.012	0.001	0.005

Figure 11: Forest to Agriculture - Spot-Future Spread

TABLE C. Long Run Propensity

AF potential value soy mean

D.ln_sum_AF	Coef.	Std.Err.	z	P>z	[95%Conf.	Interval]
_nl_1	0.000	0.000	1.400	0.160	-0.000	0.001

AF potential value soy p25

D.ln_sum_AF	Coef.	Std.Err.	z	P>z	[95%Conf.	Interval]
_nl_1	0.000	0.000	1.400	0.160	-0.000	0.001

AF potential value soy p75

D.ln_sum_AF	Coef.	Std.Err.	z	P>z	[95%Conf.	Interval]
_nl_1	0.001	0.000	1.400	0.160	-0.000	0.002

AF potential value maize mean

D.ln_sum_AF	Coef.	Std.Err.	z	P>z	[95%Conf.	Interval]
_nl_1	-0.001	0.001	-2.190	0.028	-0.003	-0.000

AF potential value maize p25

D.ln_sum_AF	Coef.	Std.Err.	z	P>z	[95%Conf.	Interval]
_nl_1	-0.001	0.000	-2.190	0.028	-0.002	-0.000

AF potential value maize p75

D.ln_sum_AF	Coef.	Std.Err.	z	P>z	[95%Conf.	Interval]
_nl_1	-0.002	0.001	-2.190	0.028	-0.004	-0.000

Figure 12: Agriculture to Forest - Potential Value

TABLE D. Long Run Propensity

AF spread soy mean

D.ln_sum_AF	Coef.	Std.Err.	z	P>z	[95%Conf.	Interval]
_nl_1	0.002	0.001	3.710	0.000	0.001	0.003

AF spread soy p25

D.ln_sum_AF	Coef.	Std.Err.	z	P>z	[95%Conf.	Interval]
_nl_1	0.001	0.000	3.710	0.000	0.001	0.002

AF spread soy p75

D.ln_sum_AF	Coef.	Std.Err.	z	P>z	[95%Conf.	Interval]
_nl_1	0.003	0.001	3.710	0.000	0.001	0.004

AF spread maize mean

D.ln_sum_AF	Coef.	Std.Err.	z	P>z	[95%Conf.	Interval]
_nl_1	-0.001	0.001	-0.690	0.489	-0.002	0.001

AF spread maize p25

D.ln_sum_AF	Coef.	Std.Err.	z	P>z	[95%Conf.	Interval]
_nl_1	-0.000	0.001	-0.690	0.489	-0.001	0.001

AF spread maize p75

D.ln_sum_AF	Coef.	Std.Err.	z	P>z	[95%Conf.	Interval]
_nl_1	-0.001	0.001	-0.690	0.489	-0.003	0.002

Figure 13: Agriculture to Forest - Spot-Future Spread

A.2 Hausman-test

We conduct a Hausman-test in order to be certain Fixed Effects is the right strategy as compared to Random Effects. The results indicate that the null-hypothesis can be rejected, as there are no difference between the random effects estimator and the fixed effects estimator. In line with (Sandvik, 2018) this indicates that the individual-fixed effects, not controlled for fully in the random-effects estimation, affect the system and lead to inconsistent estimators.

Hausman test

Hausman (1978) specification test

	Coef.
Chi-square test value	86566.8
P-value	0

Figure 14: Figure

A.3 Granger Causality test

Granger Causality Test

Granger causality Wald tests

Equation	Excluded	chi2	df	Prob > chi2
fp_soy	sp_soy	2.7366	2	0.255
fp_soy	ALL	2.7366	2	0.255
sp_soy	fp_soy	18.701	2	0.000
sp_soy	ALL	18.701	2	0.000

Figure 15: Results of Granger Causality Test for Soy

Notes: These results are based purely on soy prices. The results includes two equations, the first investigating whether spot prices lead futures, and the second investigating whether futures lead spot prices.

Granger Causality Test

Granger causality Wald tests

Equation	Excluded	chi2	df	Prob > chi2
fp_maize	sp_maize	.58651	2	0.746
fp_maize	ALL	.58651	2	0.746
sp_maize	fp_maize	18.265	2	0.000
sp_maize	ALL	18.265	2	0.000

Figure 16: Results of Granger Causality Test for Maize

Notes: these results are based purely on maize prices. The results includes two equations, the first investigating whether spot prices lead futures, and the second investigating whether futures lead spot prices.

The model used includes past values of "other series to the serie's own history"(Sajwan & Chetty, 2018). Granger causality implies a correlation between current values of one variable and the past values of other variables (Sajwan & Chetty, 2018). In other words, we set up a model to test whether current values or spot is correlated with past values of futures, or current values of futures is correlated with past values of spot. In our case we test for both possible directional relationships, thus we have two null-hypothesis for soy and two for maize. Our first null-hypothesis is that lagged values of spot price of soy *do not* cause future price of soy. Second null-hypothesis is that lagged values of future price of soy do not cause spot price of soy.

Next we have the equivalent for maize, and the third null-hypothesis is that lagged values of spot price for maize do not cause future price of maize. Finally, we have the null-hypothesis that lagged values of future price for maize do not cause spot price for maize. We cannot reject the first and third 0-hypothesis as the p-value is greater than 0,05. On the other hand we can reject the second and fourth null-hypothesis, as these both have p-value less than 0,05, respectively 0,000 in both figure 15 and 16.

A.4 Scatter Plots

This is the results of our linearity test, clearly showing that a log-level model is in favour of a level-level model as the residuals become more normally distributed and random after the log-transformation of the dependent variable.

PANEL A. Level-level test

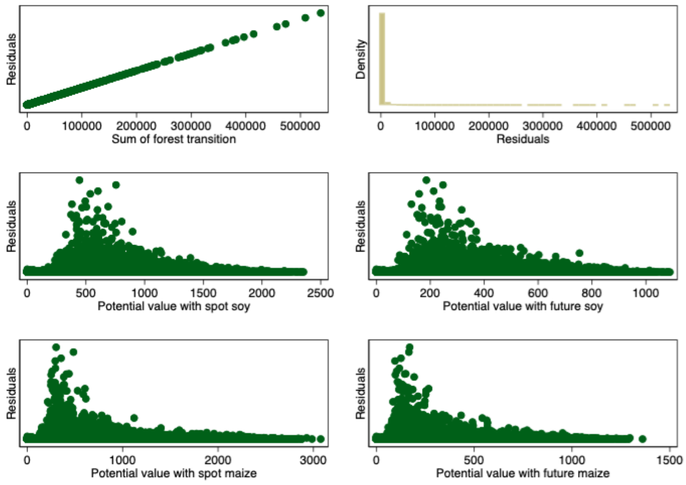


Figure 17: Level-level scatter plots

PANEL B. Log-level test

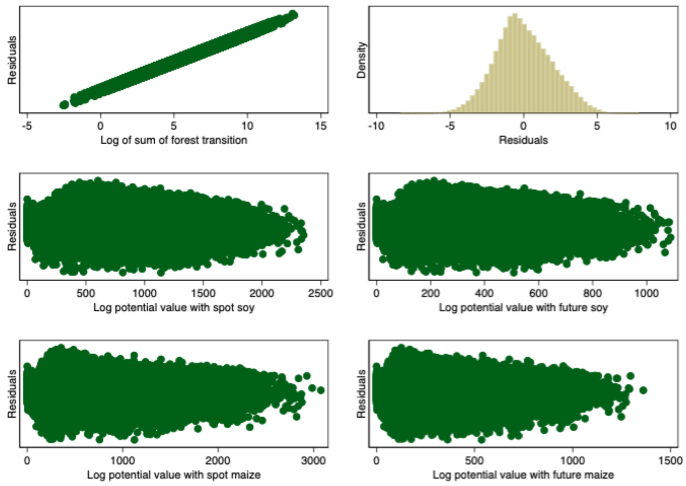


Figure 18: Log-level scatter plots

PANEL A. Level-level test

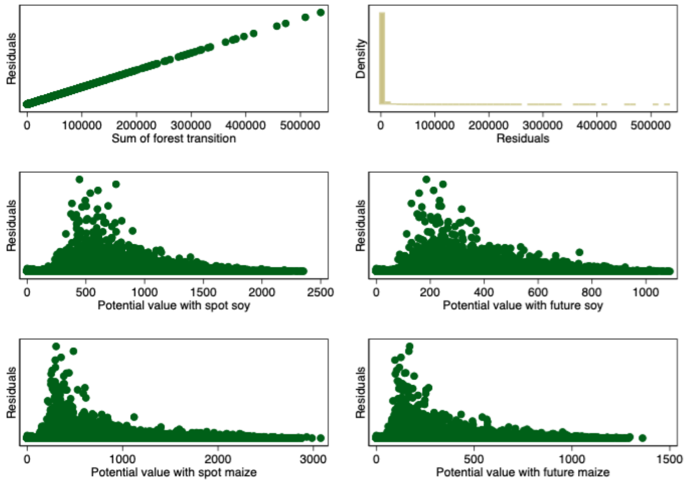


Figure 19: Level-level scatter plots

PANEL B. Log-level test

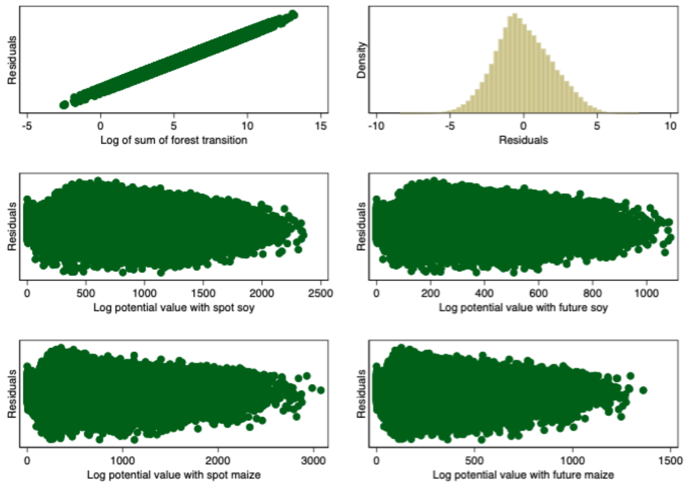


Figure 20: Log-level scatter plots

A.5 Data

This part of the appendix contains a detailed description of the main variables used in the empirical analysis.

Transition of Land From Forest to Agricultural Land: The variable is defined as the amount of area, measured in square kilometres, covered with forest in 1995, which were transformed to areas covered with agriculture land by 2005. To calculate total transition between sub-classes of Natural forest and Agriculture, we use the same approach as described for *Agricultural Land Cover* and *Forest Cover*. All variables come at the municipality level and have been aggregated at the at the level of AMC using the correspondence used by Bustos et al. (2016), proposed by IPEA and IBGE.

Log Transition of Land From Forest to Agriculture: The variable is defined as the logarithm of the variable *Transition of Land From Forest to Agriculture*.

Transition of Land From Agriculture to Forest: The variable is defined as the amount of area, measured in square kilometres, covered with agricultural land in 1995, which were transformed to areas covered with forest by 2005. To calculate total transition between sub-classes of Natural forest and Agriculture, we use the same approach as described for *Agricultural Land Cover* and *Forest Cover*. All variables come at the municipality level and have been aggregated at the at the level of AMC using the correspondence used by Bustos et al. (2016), proposed by IPEA and IBGE.

Log Transition of Land From Forest to Agriculture: The variable is defined as the logarithm of the variable *Transition of Land From Agriculture to Forest*.

Potential Value of Soy using spot prices The Potential Value of Soy using spot prices is one of the main interaction variables in our models. It is constructed by multiplying Potential Yield using high technology inputs for soy with the spot price of soy.

Potential Value of Maize using spot prices The Potential Value of Maize using spot prices is one of the main interaction variables in our models. It is constructed by multiplying Potential Yield using high technology inputs for maize with the spot price of maize.

Potential Value of Soy using futures The Potential Value of Soy using spot prices is used as an interaction variables in our Fixed Effects models. It is constructed by multiplying Potential Yield using high technology inputs for soy with the futures price of soy.

Potential Value of Maize using futures The Potential Value of Maize using spot prices is used as an interaction variables in our Fixed Effects models. It is constructed by multiplying Potential Yield using high technology inputs for maize with the futures price of maize.

Lagged Value of Log Transition of Land from Forest to Agriculture This variable represents last year's logged transition of land from forest to agriculture. We use this variable to investigate to what extent last year's land transition affects current land transition from forest to agriculture. In other words, how previous deforestation affects deforestation today.

Lagged Value of Log Transition of Land from Agriculture to Forest This variable represents last year's logged transition of land from agriculture to forest. We use this variable to investigate to what extent last year's land transition affects current land transition from agriculture to forest. In other words, how previous regrowth affects regrowth today.

Potential Value of Soy using spot future spread The Potential Value of Soy using spot prices is used as an interaction variables in our Fixed Effects models. It is constructed by multiplying Potential Yield using high technology inputs for soy with the spread between spot and future prices for soy.

Potential Value of Maize using spot future spread The Potential Value of Maize using spot prices is used as an interaction variables in our Fixed Effects models. It is constructed by multiplying Potential Yield using high technology inputs for maize with the spread between spot and future prices for maize.

A.6 Fixed Effects Estimations

A.6.1 Main econometric models for Fixed Effects Strategy

The main econometric models for our fixed effects strategy is:

$$LogFA_{i,t} = \alpha_i + \beta PVsoy_{i,t}^{spot} + \beta PVsoy_{i,t}^{futures} + \beta PVmaize_{i,t}^{spot} + \beta PVmaize_{i,t}^{futures} + u_{jt} \tag{5}$$

In order to investigate whether our results from the First Differences approach are robust, we have conducted an extra analysis using Fixed Effects to investigate the effects of our independent variables on deforestation and regrowth of forest. One change from our Fixed Effects strategy is that we substitube the spot-future spread variables with a potential value varible using future prices.

This model is also a log-level model trying to estimate the percentage effect the coefficients of the independent variables have on $LogFA_{i,t}$. More specifically this change is represented by the β-coefficient, which represents the percentage effect of potential value of each commodity[10], dependent on whether it is constructed with spot prices or future prices. We assume that the only thing affecting change in potential value is the commodity price. If the β-coefficient is statistically significant, the null hypothesis is that there is no link between the different commodity prices and the transition of land from forest to agriculture can be rejected. The dependent variable is the *Log Transition of land from Forest to Agriculture* for AMC i at time t, written as $FA_{i,t}$. Transition from forest to agriculture is measured yearly in hectares for all AMCs, and captures to amount of forest that have been transformed into agricultural land. Important to know that in this paper, the definition of forest and agriculture hinges on the definitions made by MapBiomas (2019). The α_i is the unobserved time-invariant individual effect and captures any unobserved time-invariant variables on the AMC level. The last term, $u_{i,t}$, contains both fixed errors caused by unobserved time-invariant variables and a random error component (Wooldridge, 2016).

Next we estimate same equation as equation 5 just with added lags on the variables. In order to be able to estimate the long-term effects of the independent variables, we incorporate two yearly lagged values of each independent variable constructed with the spot price. In addition, we add the one year lagged dependent variable to estimate how big of an effect last years land transition have on this years transition.

In equation 6 the dependent variable $FA_{i,t}$ is substituted with $AF_{i,t}$, more specifically

[10] Potential value is abbriviated to PV

Transition of land from agriculture to forest.

$$LogAF_{i,t} = \alpha_i + \beta PVsoy^{spot}_{i,t} + \beta PVsoy^{futures}_{i,t} + \beta PVmaize^{spot}_{i,t} + \beta PVmaize^{futures}_{i,t} + u_{jt} \tag{6}$$

Additionally we want to run the same regression to estimate the effects on our second dependent variable, $LogTransitionoflandfromAgriculturetoForest$. We do this in order to investigate whether our results are consistent. Similar to the first dependent variable, this one measures yearly land transition, except now we investigate the transition from agriculture to forest. Such transition may be interpreted as a slowing down of deforestation, or to some extent regrowth of forest land. The same conditions for the definition applies to this dependent variable, the definition of agricultural land and forest hinges on the definitions made by the MapBiomas project.

Similar to equation 5, we estimate the long-term effect for equation 6. We incorporate two yearly lagged values of each independent variable constructed with the spot price in order to be able to estimate the long-term effects of the independent variables. We also add one year lagged dependent variable to estimate the effect of last year's transition from agriculture to forest affects this years transition.

A.7 Fixed Effects Estimation and Clustered Standard Error

In this paper, we utilize fixed effects estimation to control for systematic and time-invariant differences between AMC's.The fixed-effects model controls for all time-invariant differences between individuals, so the estimated coefficients of the fixed-effects models cannot be biased because of omitted time-invariant characteristics, such as culture, religion, gender, race, etc (Kohler, Ulrich, Frauke Kreuter, Data Analysis Using Stata, 2nd ed., p.245). In order to see what this method involves, one can start with a general unit-effects model, such as in the equation below.

$$y_{it} = (\sum_j(\beta_1 x_{it})) + \alpha_i + u_{it} \tag{7}$$

According to Wooldridge 2016, the fixed-effects model is attained by time-demeaning the unobserved effects mode

$$\ddot{y}_{it} = (y_{it} - \bar{y}_{it}) = (\alpha_i - \alpha_i + (\sum_j \beta_j \cdot (x_{it} - \bar{x}_{it})) + (u_{it} - \bar{x}_{it}) = (\sum_j \beta_j \cdot \ddot{x}_{it}) + \ddot{u}_{it} \tag{8}$$

When time-demeaning the variables, one will eliminate all individual unobserved time-invariant effects i and any fixed components present in the error term. Wooldridge

(2016) explains that the estimates from the fixed effects method will be the same as if one introduced dummies for every individual unit, which in this paper would be municipalities. The main argument for why a fixed effects model is a suitable tool to use in this paper, is that the groups of units in this paper, namely the AMC's, face different and non-random treatments and time-invariant factors that affect the soy and/or maize production. By using a fixed effects model one can control for or time-demean away all fixed differences between municipalities and more directly be able to pinpoint the effect of our explanatory variables over time, as AMC's are systematically different. Consequently we focus exclusively on the variance within AMC's as compared to the variance between them (Wooldridge, 2016).

In order to formally test whether using a fixed effects strategy is correct we perform a Hausman test. This test originates from an acclaimed article on the Specification Tests in Econometrics from 1978 (Hausman, 1978) and tests whether a random effects estimator is inconsistent and a fixed effects estimator is preferred. According to Wooldridge (2016) the fixed effects model allows capricious correlation to happen between i and x_{itj} while that is not the case for random effects. Hence, "fixed effects is generally thought to be a more convincing tool for estimating ceteris paribus effects" (Wooldridge, 2010. Wooldridge (2016) demonstrate that the Hausman test is based on evaluating the difference between the random and fixed effects estimators. If the test estimates a statistically significant difference, it is taken as evidence that the null-hypothesis of no significant difference between the two estimators can be rejected, and be taken to mean that key random effects assumption is false, then fixed effects is normally used. Furthermore, Sandvik(2018) highlights that "the fixed effects model is consistent even in the event that the Hausman test does not reject the null hypothesis of there not being any systematic differences between the panel data groups". Using random effects models in that case can be argued for by the fact that they are known to be more efficient when there are no systematic differences between units. Meaning that fewer observations are needed for unbiased estimation and that the random effects estimator in that sense is a stronger tool due to less variance being differenced away (Wooldridge, 2016).

When using the Hausman test on our regression models, we find that the null-hypothesis can be rejected. The results from the Hausman test can be found in appendix A.2. In line with (Sandvik, 2018) this indicates that the individual-fixed effects, not controlled for fully in the random-effects estimation, affect the system and lead to inconsistent estimators (Wooldridge, 2016). If an estimator is inconsistent, its expected value will not converge on the true population parameter even for large numbers of observations (Wooldridge, 2016). Since there are both empirical indications of systematic differences between AMC's we have chosen to use fixed effects estimation as our extra estimation strategy for this paper.

A.8 Fixed Effects Model Estimates

In Table 1 column 1 and 2 the dependent variable is the *Transition of land from forest to agriculture*, while the dependent variable in column 3 and 4 is *Transition of land from agriculture to forest*. In the regressions the independent variables are interaction variables between potential yield and both spot and future prices for both commodities, and we also include the relevant lagged dependent variable.

All models control for time fixed effects and AMC fixed effects. The standard errors are clustered on AMC level for both models to get the standard errors more robust to heteroscedasticity, and deal with serial correlation within AMCs. According to Hansen (2007) clustering is necessary due to group-level arbitrary shocks provoking correlation between observations within the groups for all time periods, in other words that the observations become serially correlated. In the two first models we try to estimate the effect the independent variables have on transition from forest to agriculture before we run the same regressions on transition from agriculture to forest.

In column 1 the regression estimates that the interaction variable for soy using spot prices is negative and significant at the 1% level, indicating that an increase in potential value by 10 dollars due to an increase in spot price of soy, is associated with a decrease of 0,22% in transition from forest to agriculture measured in hectares, a surprising result taking previous literature into account. When looking at the interaction variable incorporating futures instead of spot prices, estimates indicate that an increase of potential value by 10 dollars due to an increase in future price of soy, is associated with an increase of 0,25% in transition from forest to agriculture and is significant at the 5% level. Consequently this model suggests that for soy there may be more of a relationship between futures and deforestation, than between spot prices and deforestation.

For maize the estimates for the interaction variable using spot prices show that an increase of potential value by 10 dollars due to an increase in spot price of maize, is associated with an increase of 0,2% in transition from forest to agriculture, and is significant at the 1% level. An increase of potential value by 10 dollars due to an increase in future price of maize, is associated with a decrease of 0,04% in the dependent variable. In contrast to all the other coefficients in column 1, this is insignificant. An interesting finding in this model is that the estimates for soy and maize tend to be opposite of each other. Furthermore we see that the coefficients for maize using futures are much lower than the others, in addition to being insignificant.

In column 2 we add two yearly lagged values of the potential value spot for both com-

Table 2: Fixed Effects - Commodity Price Effect Transition FA and Transition AF

	(1) Log FA b/se	(2) Log FA b/se	(3) Log AF b/se	(4) Log AF b/se
Potential Value Soy^{Spot}	-0.000220***	-0.000210***	-0.000050	-0.000234***
	(0.000067)	(0.000078)	(0.000060)	(0.000070)
Potential Value $\text{Soy}^{Futures}$	0.000246**	0.000435***	0.000458***	0.000373***
	(0.000098)	(0.000094)	(0.000092)	(0.000090)
Potential Value Maize^{Spot}	0.000216***	0.000328***	0.000135***	0.000217***
	(0.000047)	(0.000050)	(0.000044)	(0.000048)
Potential Value $\text{Maize}^{Futures}$	-0.000042	-0.000379***	-0.000689***	-0.000629***
	(0.000073)	(0.000074)	(0.000079)	(0.000077)
Log Transition FA_{t-1}		0.174315***		
		(0.007053)		
Potential Value Soy^{Spot}_{t-1}		-0.000177***		0.000184***
		(0.000067)		(0.000065)
Potential Value Soy^{Spot}_{t-2}		0.000254***		0.000054
		(0.000049)		(0.000050)
Potential Value $\text{Maize}^{Spot}_{t-1}$		0.000048		-0.000153***
		(0.000043)		(0.000045)
Potential Value $\text{Maize}^{Spot}_{t-2}$		-0.000261***		0.000084**
		(0.000035)		(0.000037)
Log Transition AF_{t-1}				0.130087***
				(0.007114)
Constant	5.106105***	5.175155***	4.883873***	4.840210***
	(0.031715)	(0.049902)	(0.031815)	(0.051597)
Observations	140261	131740	140260	131738
Adjusted R^2	0.116	0.109	0.092	0.055
FE	YES	YES	YES	YES
Long Run Propensity Soy^{Mean}		-0.000340		0.0000109
Long Run Propensity $\text{Soy}^{25^{th}}$		-0.000225		0.00000720
Long Run Propensity $\text{Soy}^{75^{th}}$		-0.000458		0.0000146
Long Run Propensity Maize^{Mean}		0.000562		0.000689
Long Run Propensity $\text{Maize}^{25^{th}}$		0.000562		0.000689
Long Run Propensity $\text{Maize}^{75^{th}}$		0.000562		0.000689

Notes: Robust standard errors reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Important for the long run propensities. The long run propensities are calculated for three AMC's with different levels of potential yield, namely the *Mean*, 25^{th}, and 25^{th} percentile.

modities, the motivation for doing this is to be able to estimate a potential long-term effect of both commodities on the dependent variable. When looking at the potential value of soy using spot prices, it is clear that the coefficient barely changes when we add the lagged values. What is interesting when studying the coefficients related to soy is that the one year lagged value has a seemingly negative effect, while the two year lag has a seemingly positive effect and they are both statistically significant. Calculating the long-run propensity from the variables related to spot price of soy, yields a negative long-run propensity. The coefficientt is implying that a 10 dollar increase in the long run propensity is associated with a 0,34% decrease in deforestation , lending doubt to whether it is economically significant in this model. One potential explanation of the generally unexpected negative effect of soy spot price on transition from forest to agriculture may have something to do with the fact that about 90% of domestic soy is used as animal feed[11], so as the spot price for soy increases, the cost of animal feed increases. This again may lead to an increased cost and decreased production of livestock. As mentioned in section 3.4, beef is one of the main contributors to deforestation in Brazil. For maize, column 2 shows that when adding lagged values for the potential value of maize using spot price, the coefficient for today increases from 0,000216 to 0,000328, and stays significant. In other words, in this model a 10 dollar increase in the potential value of soy using spot prices of today is associated with a 0,32% increase in the transition from forest to agriculture. Studying the coefficients for maize, one can find the same pattern as for soy just opposite. The one year lagged variable is positive, but insignificant for maize, while the two year lagged variable is negative and statistically significant. When calculating the long-run propensity for maize we get a positive long term effect, indicating that a 10 dollar increase in the potential value of maize for spot prices is associated with a 0,6% increase in deforestation. This seems quite odd, especially when comparing to the estimates of soy and the technological advances in maize cultivation (Harsheim & Nakkim, 2018). Thus it is unclear whether there is any meaningful economic significance. The coefficient for potential value of maize using futures changes from insignificant to significant at the 1% level when the lagged variables are added. This is an interesting finding, indicating that in model 2 the associated effect of futures increases when the model incorporates the lagged effects of potential value using spot. We find this result quite hard to explain. Compared to previous research, it is surprising that these regressions indicate that soy is associated with a negative effect on deforestation, while maize is associated with a positive effect. This is contrary to our hypothesis that both commodities have a positive relationship on deforestation, in other words that an increase in the price of both would be associated with an increase in the transition from forest to agriculture. Moreover, the fact that futures and spot are opposite of each other for both commodities is an interesting finding,

[11] As mentioned in 3.4

and could potentially be due to different decision making strategies for farmers.

In model 2 we add two yearly lagged values of the potential value spot for both commodities, the motivation for doing this is to be able to estimate a potential long-term effect of both commodities on the dependent variable. When looking at the potential value of soy using spot prices, it is clear that the coefficient barely changes when we add the lagged values. What is interesting when studying the coefficients related to soy is that the one year lagged value has a seemingly negative effect, while the two year lag has a seemingly positive effect and they are both statistically significant. Calculating the long-run propensity from the variables related to spot price of soy, yields a negative long-run propensity. The coefficient is quite small, lending doubt to whether it is economically significant in this model. One potential explanation of the generally unexpected negative effect of soy spot price on transition from forest to agriculture may have something to do with the fact that about 90% of domestic soy is used as animal feed, as mentioned above, so as the spot price for soy increases, the cost of animal feed increases. This again may lead to an increased cost and decreased production of livestock. As mentioned in section, beef is one of the main contributors to deforestation in Brazil.