# Agricultural price distortions in Africa

## Insights into the determinants of price distortions and their effect on land use efficiency

**Mandy Malan**

November 2015

MSc Thesis

Supervised by:

dr. Jeroen Klomp (Development Economics Group, Wageningen UR)
dr. Ezra Berkhout (Duurzame Ontwikkeling, Planbureau voor de Leefomgeving)
dr. Jetske Bouma (Duurzame Ontwikkeling, Planbureau voor de Leefomgeving)

**Abstract**

Price distortions for agricultural commodities are widespread in Africa and have not been reduced as much as in other developing countries. This research looks at the effect of price distortions on land use efficiency. Panel data for 22 countries in Africa is used to look at the impact of agricultural price distortions and the possible intermediating effects. On the other hand, the determinants of price distortions are examined. In addition to aggregate agricultural effects, the research focusses on three exportable cash crops (cocoa, coffee, and cotton) and four import-competing food crops (maize, rice, sugar, and wheat). Different types of models are estimated to test robustness and to take endogeneity into account. A fixed effects model, IV-regression, and a simultaneous equation model show significant effects of price distortions on productivity. Most notably, large positive effects are found for cotton and cocoa, indicating that a decrease of taxation of these crops will lead to productivity increases. For some of the food crops, which are often subsidized, an increase in subsidization will lead to decreases in productivity. In addition, several economic, financial and political variables are found to influence pricing policy. Democratic governments in a stable macroeconomic context generally tax agriculture less.

# Contents

# List of Figures



# List of Tables

## Abbreviations

FAO: Food and Agricultural Organization
GDP: Gross Domestic Product
IMF: International Monetary Fund IV: Instrumental Variable
LDCs: Least Developed Countries
NRA: Nominal Rate of Assistance
SEM: Structural Equation Model
TFP: Total Factor Productivity

# 1 Introduction

Government pricing policy in agriculture has been subject to much debate throughout the past decades. Such pricing policy may influence agricultural incentives and alter production and investment decisions. These pricing policies, or so-called price distortions, can be domestic measures (e.g. output taxes and input subsidies) or trade measures (e.g. import tariffs and export taxes). In addition, agricultural distortions may arise because of multiple exchange rate systems (Anderson and Masters, 2009). These policy measures are termed price distortions because they distort the market price. Within many countries, the market price may also be distorted because of other factors, for instance because of high transaction costs due to market imperfections. However, following the literature on this subject, in this research, the term agricultural price distortions refers to the pricing policy implemented by governments, as described above.

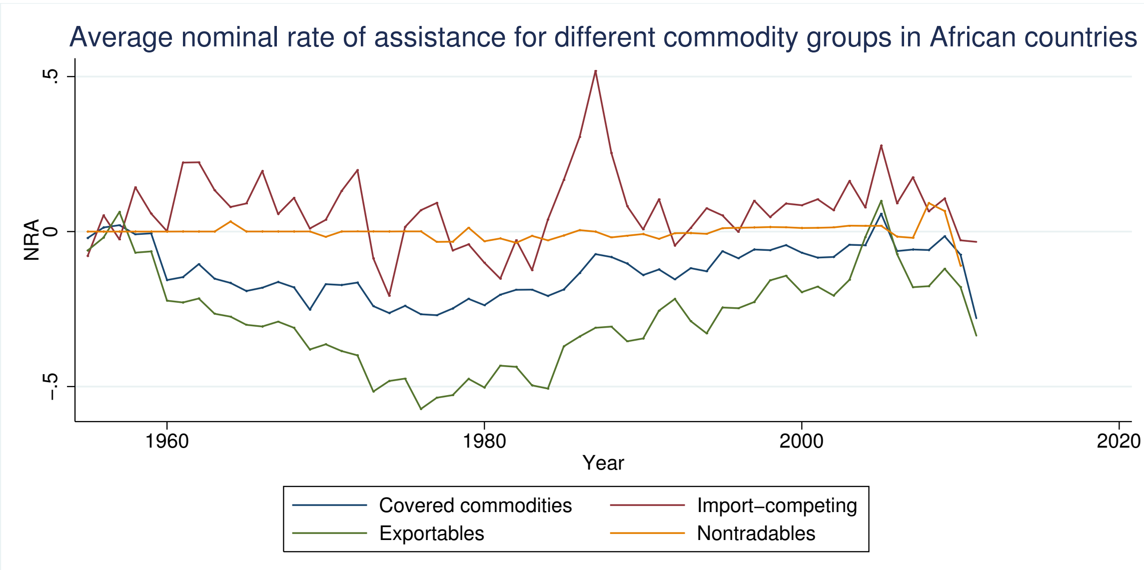


**Figure 1:** average NRA for different commodity groups for a sample of African countries. If NRA is smaller than 0 the commodity group is taxed, if NRA equals 0 there is no pricing policy, if NRA is larger than 0 the commodity group is subsidized. Covered commodities are a selection of commodities that make up around 70% of the total value of production within a country.

Despite (international) efforts to reduce trade barriers and other price distorting measures, agricultural price distortions are still significant worldwide. These agricultural price distortions are particularly important in developing countries, where productivity is low, and employment in agriculture is high. Historically, developing countries have had a policy bias against agriculture, though recently distortions have been reduced significantly. However, most countries in Africa are lagging behind and still have a bias against agriculture (Anderson and Masters, 2009). The motivation for this policy bias against agriculture comes from the desire of governments to enhance industrialization. Figure 1 shows the development of agricultural price distortions in African countries. It shows that import-competing commodities (e.g. maize, rice, wheat) have been mostly subsidized throughout history. Exportables (e.g. cocoa, coffee) have historically been taxed a lot, with taxes reaching their highest point during the late seventies. After that, taxation is decreased though never completely eradicated. Similarly, agriculture as a whole, (measured by a selection of the most important crops of a country), is mostly taxed. Non-tradable crops (e.g. millet, cassave) are hardly taxed or subsidized. Thus, next to an anti-agricultural bias, there is also an anti-trade bias in the African countries. Trade has been traditionally restricted to stimulate domestic food production (Anderson and Masters, 2009).

At the same time agricultural productivity in Africa is also lagging behind. Yet, the need for a rise in productivity is ever-increasing, as food insecurity and the increasing pressure on land are posing to be significant challenges. FAO (2009) predicts that in order to feed the world population in 2050, food production needs to be increased by 70%. Population growth has induced agricultural extensification, increasing land scarcity and threats to biodiversity (Reardon et al., 1999). In order to limit the use of land reserves for agriculture, sustainable intensification and thus productivity increases are necessary (FAO (2009), Reardon et al. (1999)). Existing literature focuses on seeking causes in factors like input use, biophysical indicators, technology adoption, market failure, infrastructure and the institutional context (WorldBank (a), Fulginiti et al. (2004), Reardon et al. (1992) Barrett (2008)). However, there is little focus on the possible role of agricultural price distortions due to government market intervention. When agricultural commodities are being taxed, the farm-gate price becomes lower than the world market price. Thus, farmers receive a lower price than without taxation. This can result in altered production decisions. Farmers may switch to other commodities, leading to less producion of the taxed commodity, and investment in productivity may be reduced. Intensification of agriculture may be inhibited and use of resource allocation is possibly inefficient (Fulginiti and Perrin (1993), Binswanger et al. (1978)). The extent to which farmer decisions are affected by these price changes will depend on whether the farmers market their production. If farmers engage in subsistence farming only, output taxes and trade barriers are unlikely to affect farmers. However, input subsidies or taxes may still be relevant for subsistence farmers. This research is focussed on determining whether agricultural pricing policy in Africa has affected agricultural productivity growth. Moreover, many organizations are proposing subsidizations to agriculture in Africa in order to increase productivity. This research will provide evidence whether such subsidies have productivity-increasing effects.

In this research it is aimed to find evidence on the relationship between agricultural price distortions and land productivity. In addition, it is suspected that this relationship may not be uniform. First off, there are different types of price distortions, ranging from direct domestic policy measures to trade measures and indirect distortions. As research indicating a negative effect of agricultural support on productivity is focussed mostly on subsidies in developed countries, it will be interesting to see how taxation of agriculture affects productivity. Secondly, the magnitude and direction of the effect on land productivity may differ for different commodities. Thirdly, intermediate channels may alter the magnitude and direction of the effect of distortions on land productivity. For instance, if the institutional quality of a country is low, policy may not be effective and farmer incentives may not change. More insight into these differing effects on productivity creates opportunities for productivity increasing intervention where productivity gaps are largest.

Some of the intermediate channels may not only be intermediate, but also affect pricing policy or form the context within which policy choices or farmer decisions are being made. Therefore, this research also aims to look at the other side of price distortions. The main focus being the determinants of agricultural price distortions. Do governments solely use taxes and subsidies on agriculture for economic or budgetary reasons, or are there political economic circumstances influencing policy decisions?

This thesis is structured as follows. After formulation of the research questions, an overview of the literature on these themes is given. This consists of a discussion on the relationship between pricing policy and agricultural productivity, and a discussion of literature on the determinants of price distortions. A complete theoretical framework is presented thereafter. Then, the methodology of the data collection and data analysis are discussed. The data used is panel data and it is analysed using standard panel data analysis techniques and specific econometric techniques to deal with endogeneity issues. The results are presented in the following section and are put into context in the discussion. The thesis is closed with a concluding chapter.

# 2 Research questions

The central question to be answered in this research is:

- How do agricultural price distortions affect land use efficiency in Africa, and what are the determinants of price distortions?

The specific research questions are as follows:

1. How do price distortions on agricultural import-competing commodities affect land use efficiency?
2. How do price distortions on exportable commodities affect land use efficiency?
3. Which conditions mitigate or increase the magnitude of the effect of price distortions on land use efficiency?
4. What are the determinants of agricultural price distortions in Africa?

# 3 Theory

## 3.1 Effect of agricultural price distortions on land use efficiency

The first part of the research consists of determining whether agricultural price distortions have an effect on farmer incentives to adjust productivity. The relationship between agricultural price distortions and productivity is ambiguous. Some evidence and theory suggest that support to agriculture will increase agricultural productivity due to technological innovation and resource use efficiency. A channel through which this mechanism may work is that farmers are more motivated to invest in the productivity of their farm. This may be caused by credit and risk attitudes (Rizov et al., 2013). Subsidies are a form of extra income which could increase access to credit. Farmers attitudes towards risk may also change because of this extra income. Both leading to investment induced productivity gains (*ibid.*). Similarly, the Schmookler-Lucas hypothesis proposes that innovation and product price are positively related. A higher expected output price will increase the innovation of the producer, which leads to productivity growth (Fulginiti and Perrin (1993), Binswanger et al. (1978)). This is supported by evidence of Fulginiti and Perrin (1993), who find that agricultural price increases in Least Developed Countries (LDCs) will lead to substantial productivity growth in agriculture, suggesting that agricultural taxation in these countries has inhibited productivity growth. More recently, Rakotoarisoa (2011) finds evidence that rice subsidies and protection in developed countries and rice taxation in developing countries have jointly widened the rice productivity gap between rich and poor rice countries.

However, there are also arguments against these theories suggesting that support to agriculture will decrease productivity. Subsidies distort the production structure of the farm, which may lead to allocative inefficiency, e.g. overinvestment in subsidized inputs. In addition, technical inefficiency can occur if subsidies lead to slack, lack of effort and limited competition. Kornai (1986) states that budget constraints may become soft due to subsidization. A soft budget constraint means that the strict relationship between income and expenditure is relaxed. Expenditure can then exceed income because it is paid for by another institution, in this case the state (*ibid.*). Subsidizations then work as a form of insurance which may lead to a moral hazard problem. Finally, subsidies may reduce the rate at which resources are reallocated to other (more productive) activities, also leading to inefficiency (Rizov et al., 2013). Rizov et al. (2013) find evidence for a negative relationship between coupled subsidies in the EU and agricultural productivity. Similarly. taxation could also force farmers to produce crops more efficiently to meet income or food needs. In standard economic theory, taxes are treated simply as negative subsidies. Price distortions are measured as either negative (taxation) or positive (subsidization). However, farmers may respond differently to taxes and subsidies, i.e. a subsidy may not have the opposite effect as a tax. This research will provide more insight into this.

Most research focuses on the relationship between distortions and total factor productivity (TFP). However, there are some problems with measuring TFP. TFP is a measure for technological progress and is given by the residual of the production function. All unobserved variables influencing production are included in this measure, and this can lead to inclusion of unobservables that do not necessarily represent productivity growth. Moreover, the increasing pressure on land in Africa motivate us to focus on land use efficiency instead of TFP. Land use efficiency is measured as the yield (production per hectare). There is little literature on the relationship between agricultural price distortions and yield. One study on input subsidies in Malawi shows significant increases in yields due to more efficient use of inputs (Denning et al., 2009). This research will give further insights into the relationship between agricultural price distortions and land use efficiency.

### 3.1.1 Different policy measures

There are different policy measures that distort prices for smallholder farmers in Africa. A distinction can be made between direct and indirect policy measures. Direct policy measures affect prices of commodities directly, whereas indirect policy measures alter agricultural incentives through distortions in other sectors. Examples of direct price distortions are trade distortions, like export taxes or import tariffs, and domestic producer and consumer price-distorting measures like input subsidies or output taxes. In addition, governments can manipulate exchange rate system to distort prices. These different policy instruments may have different effects on the incentives of farmers in Africa. In addition, there are also differences in the way a tax or subsidy is formed. An ad valorem tax reduces or raises the price by a proportion of the price, whereas a lump sum tax or subsidy reduces or raises the price or income by a fixed amount. This research focuses on the ad valorem tariffs, as the focus is crop-specific and lump sum taxes or subsidies are mostly decoupled from specific crops. In this section, an overview of economic theory and prior research on the effects of different price distortions on prices, production and productivity is given. This will give insights into what to expect regarding the relationship being studied in this research.

*Trade distortions*
Trade distortions make up about three-fifths of worldwide agricultural price distortions (Anderson et al., 2013). With respect to trade distortions, economic theory predicts the following:

- An import tariff is a tax on imported commodities. This raises the domestic price of the commodity above the world market price, i.e. producers receive a higher price for the commodity than without an import tariff. Producers will increase production and it is expected that investment in agricultural commodities that are protected through import tariffs will be higher, which will lead to higher productivity. Import quotas, or quantitative import restrictions, have a similar effect, apart from governments not being able to collect tax revenue.

- An export tax lowers the price for farmers that produce for the export market, i.e. the domestic producer price is reduced below the world market price, thereby reducing production. Investments in crops that are taxed are expected to be lower. For export subsidies the effect is the opposite.

*Domestic distortions*
In terms of domestic distortions economic theory predicts the following:

- An output tax will reduce producer prices below consumer prices and thus reduce production. It is expected that farmers will invest less in a crop if output is being taxed. For an output subsidy the effect is the opposite.

- Input subsidies will increase producer prices below consumer prices and thus increase production. Input subsidies may influence productivity more than other subsidies, as these specifically lead to investment in productivity-increasing inputs, whereas other subsidies may lead to non-productivity-increasing investments. On the other hand, input subsidies can also lead to overuse of inputs, decreasing productivity. For input taxes the effect is the opposite.

*Exchange rate policy*
In African countries dual or multiple exchange rate systems are often in place, most significantly in the 1970s and 1980s. Through such systems, gaps are created between the exchange rate received by exporters and the exchange rate received by importers. This can alter farmer incentives to produce exportable or import-competing crops (Anderson and Masters, 2009).

*Indirect agricultural distortions*
Indirect distortions affect agricultural incentives indirectly, through distortions in other sectors. This is asserted by the Lerner symmetry theorem, stating that indirect government assistance to non-agricultural production affects farmer incentives. Higher support for non-agricultural-production will lead to more incentives for producers in these sectors to bid up prices of mobile resources that are used in agriculture. This leads to a smaller availability of resources in agriculture than without distortionary policy (Anderson and Masters, 2009).

### 3.1.2 Different commodities

Within African countries, exportables are historically taxed whereas import-competing commodities are subsidized. Exportables, which are mostly perennial crops, are expected to respond less to price distortions. The reason for this is that after incurring start-up costs for a perennial plantation, farmers will be less likely to switch to another crop because of price distortions. This also offers an argument for governments to tax perennials instead of annuals. Because the impact on perennial production is smaller, the cost of raising tax revenue is also smaller (Anderson et al. (2013), De Gorter and Swinnen (2002)). In addition, this study looks at the effect of price distortions for different crops, as it is expected that the extent to which production and productivity are affected by price distortions depends on the commodity, as well as on the supply response to price changes.

### 3.1.3 Intermediating variables

Some variables are expected to affect the strength of the effect of price distortions on productivity:
*Market imperfections*: we can expect that when market imperfections are large, price changes do not always reach farmers. Price transmission in Sub-Saharan African markets has been researched by Minot (2010). Minot (2010) finds very little effect of world price changes on changes in markets in Sub-Saharan Africa. This can for instance be due to poor market access and lack of information. Market imperfections may also directly affect yield through higher input prices.
*Variable policy*: from the data, a large variability in price support can be seen. Farmers may not be induced to change their production decisions in response to fast-changing policy measures. A policy volatility measure can be added to the model. To an extent, this has been researched by Hopenhayn and Muniagurria (1996) who look at the effect of policy variability on economic growth and welfare. They find that high policy variability on subsidization of investment decreases growth. Because those results are focussed on the economy as a whole, it will be interesting to see how African agricultural markets are affected by policy variability.
*Governance*: if a countrys level of governance is low, we can expect that implementation of policy is minimal. Therefore, it can be expected that with a low level of governance, policy changes do not reach farmers. Within the range of distorting measures, border market price support is likely to be least affected by governance, as it is relatively easy to implement. Another important variable that may influence the effect of price

distortions is land tenure security. One can expect that the lower land tenure security, the less incentives farmers will have to increase investment or productivity (Deininger and Jin, 2006). As many measures of governance are based on variables such as tenure security, an overall measure of governance will be sufficient for this analysis.

*Informal sector*: If there is a large informal sector through which agricultural commodities are being traded we can expect that import tariffs or other distortionary measures are not effective. On the other hand, there might be some causal relationship between price distorting measures and the informal sector. When price distortions are high, the informal sector is likely to be large because smuggling may be induced.

## 3.2 Determinants of agricultural price distortions

The second part of this research focuses on the determinants of agricultural price distortions. There are many theoretical and empirical contributions that find indicators for agricultural price distortions in developing and developed countries. One of the main findings, known as the developmental paradox, is that protection to agriculture increases as GDP increases (Anderson et al. (2013), De Gorter and Swinnen (2002)). In addition, empirical studies find a strong correlation between structural changes in the labor market and agricultural protection. The smaller the share of the rural population of the total population, the more agriculture is protected. A given reason for these findings is that the opposition to agricultural protection decreases as the share of labor in agriculture and the share of income spent on food decrease. Thus, in developed countries, where the share of labor in agriculture is lower than in developing countries, the protection to agriculture is generally higher (Anderson et al., 2013). Dennis and İşcan (2011) show that agricultural taxation can also delay structural change. Therefore, some reverse causality may be present in the relationship between agricultural price distortions and structural changes in the labor market.

Governments also have economic arguments to tax agriculture, namely raising tax revenue. For instance, when a countrys public debt is high, there is more reason for governments to raise government revenue (Tanzi, 1992). Especially import tariffs are attractive for governments, as they are less visible to the domestic population than for instance export taxes (Johnson and Antle, 1996). Another often cited reason for governments in developing countries to distort agricultural prices is food self-sufficiency. This also explains the anti-trade bias that African governments generally have (Anderson et al., 2013). An anti-trade bias entails import-tariffs for import-competing products and export taxes for exportables. Such measures form a barrier to trade and stimulate domestic production, thereby increasing food self-sufficiency. A third explanation for price-distorting policies is the protection of the domestic market from international price spikes. This is accomplished by increasing import tariffs or export subsidies when market prices decline and reducing them when market prices rise. This can especially explain year-to-year variation in agricultural price distortions (Anderson et al., 2013). Lastly, if a country has monopoly power in the market for a certain commodity, the government may tax export crops because part of the tax burden is transferred to foreign importers (Johnson and Antle, 1996).

As has been argued by many researchers in this research area, a political-economic framework offers explanations for agricultural price distortions (Anderson et al. (2013), De Gorter and Swinnen (2002), Henning and Struve (2007)). Structural change and the related political cost of agricultural taxation were already touched upon previously. Other political-economic factors include: institutions, rulers preferences and farmer organization. Regarding institutional changes, there are many factors that fall within this domain. The effect of democratization is not straightforward. On the one hand, median-voter theorem predicts that governments will redistribute income from the rich to the poor as they become more democratic (Anderson et al., 2013). However, the less democratic a country is, and thus the less accountable the government is for its actions, the more governments can follow their own preferences (*ibid.*). Thus, agricultural policy may not be easily predicted by looking at democracy alone. In addition, there is evidence that constitutional details are also of influence. Proportional electoral systems and parliamentary regimes are shown to redistribute income more than presidential and majoritarian systems (*ibid.*). The extent to which agricultural interests are able to organize, which in turn is affected by improvements in rural infrastructure, also has an effect on agricultural price distortions (*ibid.*).

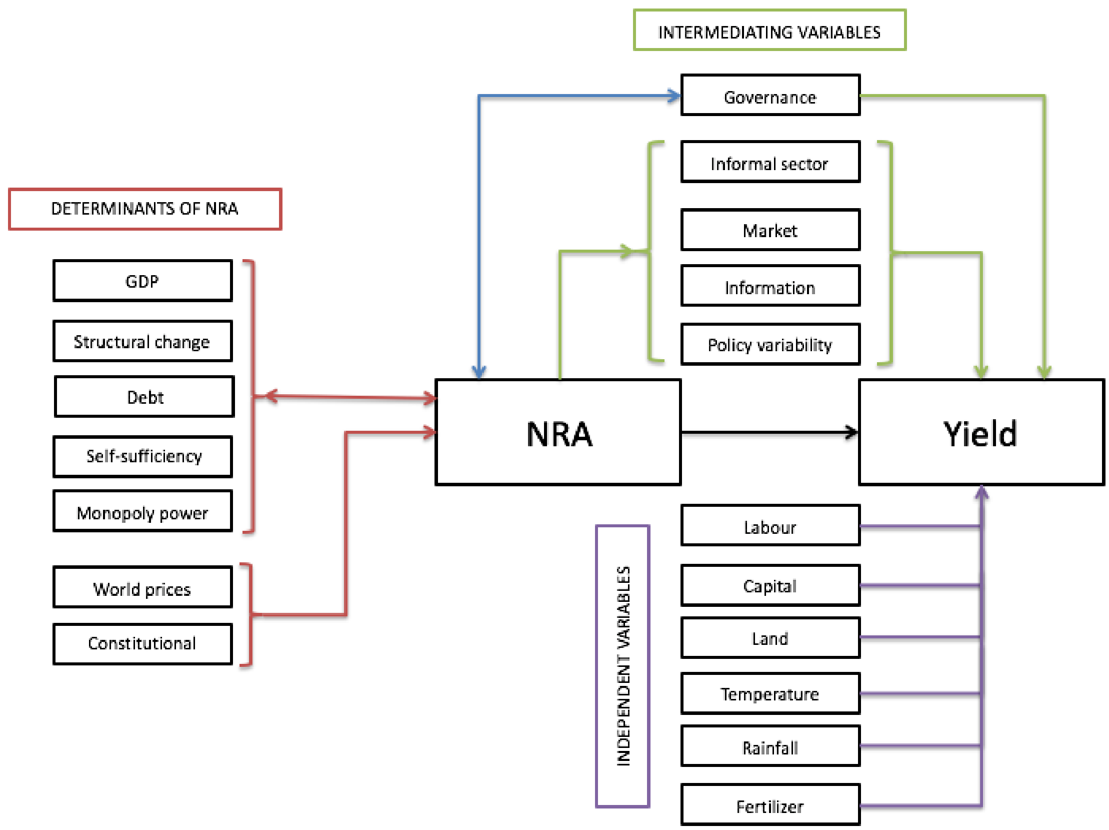


**Figure 2:** Theoretical framework

A full graphical representation of the theoretical framework is given in figure 2. The figure gives an overview of all variables discussed and otherwise relevant for this analysis. The main relationship researched here is the effect of agricultural price distortions (measure by NRA) on productivity (measured by yield). A list of independent variables that affect yield are included in the model as control variables. NRA affects yield directly but the effect may depend on some intermediate variables. One of these, governance is also thought to affect NRA. On the left side of the figure, the hypothesized determinants of NRA are given.

# 4 Methodology

## 4.1 Data and variables

In this section the data collection methods and variable descriptions are given. First, information on the sample of countries and time period is given. This is followed by a section on the selection of commodities. Then, an overview of all variables and their data sources is given. Lastly, the complete estimated model is explained.

### 4.1.1 Country coverage

The following 22 African countries are covered in this research:

| | | | |
|---|---|---|---|
| Benin | Ethiopia | Mozambique | Togo |
| Burkina Faso | Ghana | Nigaria | Uganda |
| Cameroon | Kenya | South-Africa | Zambia |
| Chad | Madagascar | Senegal | Zimbabwe |
| Côte d'Ivoire | Mali | Sudan | |
| Egypt | Morocco | Tanzania | |

These countries are selected because of data availability on agricultural price distortions from the Anderson and Nelgen (2013) dataset. In the Anderson and Masters (2009) report it is stated that the selection of countries covers the largest economies in Sub-Saharan Africa, covering over 90% of agricultural value added, farm households, total population and total GDP. The period covered in this research is from 1961 to 2010, though the dataset is highly unbalanced in terms of data availability in each year.

### 4.1.2 Commodity Coverage

The analysis is done on an aggregate level as well as on commodity level. The aggregate analysis consists of a selection of the most important commodities produced in a country, that makes up 70% of the value of production of the total agricultural sector. This selection is different for each country and follows Anderson and Nelgen (2013) who estimate price distortions for this selection of commodities. For some countries, this selection includes livestock. On commodity-level, a selection of seven crops is made. Because it is expected that farmers respond differently to price policy depending on whether the crop is imported or exported, both types of commodities are covered. As shown in the introduction, price distortions are hardly relevant for non-tradables. Therefore, the analysis is restricted to import-competing and exportable crops. In terms of food crops, the largest sources of calories in Sub-Saharan Africa are: maize, cassava, rice, sorghum, wheat and millet. However, not all of these crops are produced in all countries and some are non-tradable. The import-competing commodities that are most important for the largest sample of countries are: maize (13 countries), rice (10 countries), wheat (9 countries), and sugar (9 countries). The most important exportables are cotton (16 countries), coffee (7 countries), and cocoa (5 countries). Figure 3 gives an overview of the sample and the countries covered for each commodity.

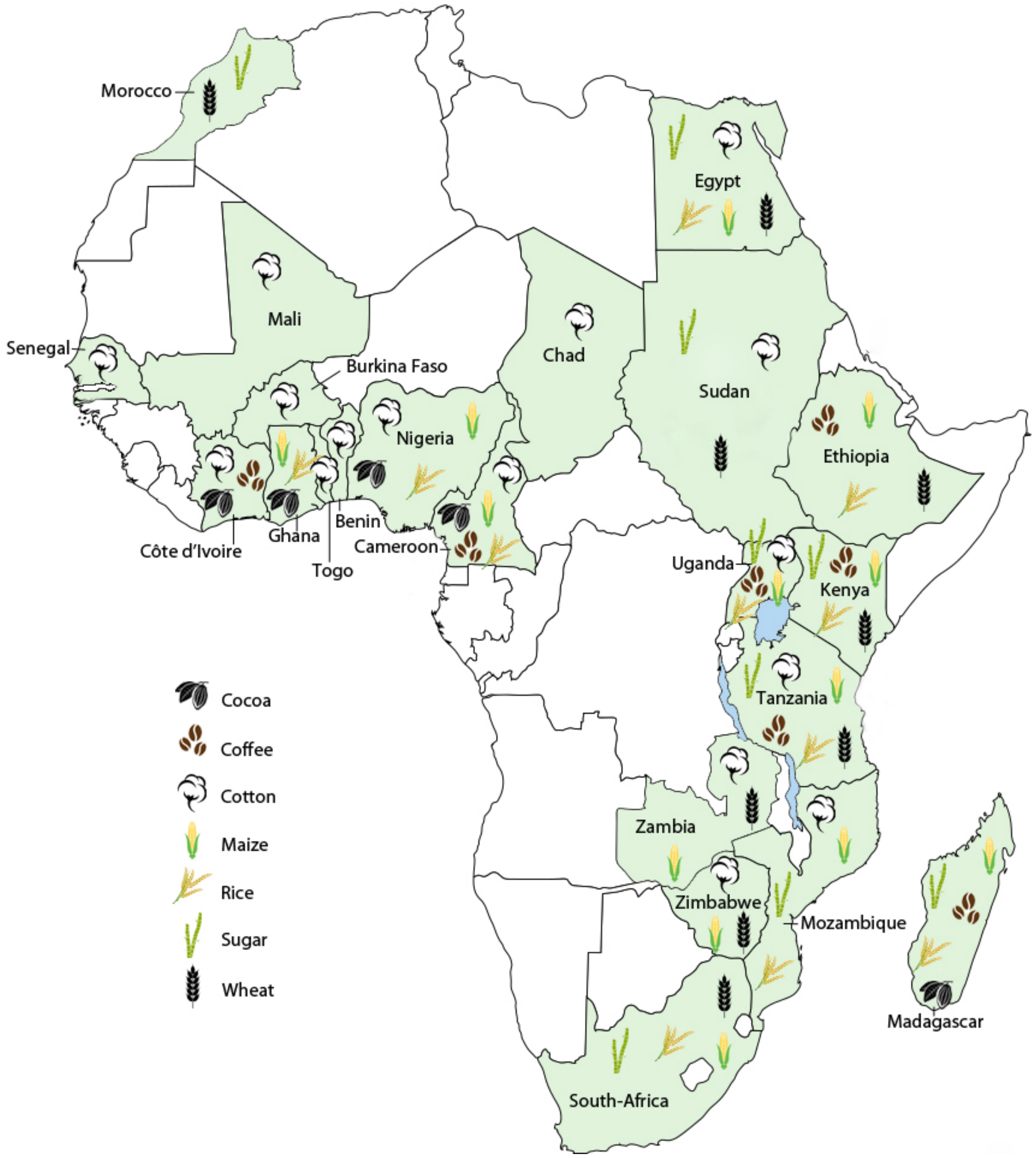


**Figure 3:** Sample and commodity coverage

#### 4.1.3 Variables

This section is divided into two parts: 1) explaining yield, and 2) explaining agricultural price distortions. An overview of the variables from external sources and the methodology of generating some of the other variables is given. All variables are time series ranging somewhere between 1961-2010.

*Effect of agricultural price distortions on yield*

**Yield**: as a measure for productivity or land use efficiency yield is used. For the commodity-level analysis, yield is measured in hectograms per hectare and is available from FAO. For the aggregate analysis a yield-index is constructed for each country. The yield-index is a weighted average of the yield-index for the commodities that are included as covered products in the price distortion calculation by Anderson and Nelgen (2013). The yield indices are created by setting the earliest available yield to 100 and indexing the years following. The weights used for averaging to an aggregate yield-index come from the share of the value of the total production (following Anderson and Nelgen (2013)). When yield data on a certain commodity is not available, the respective calculation of price distortions (NRA) for aggregate commodities is recalculated excluding this commodity.

**NRA**: as a measure of agricultural price distortions the Nominal Rate of Assistance is used. This data is available from Anderson and Nelgen (2013). NRA measures the gap between current domestic prices and free-market prices. It is defined as the percentage change in gross return to farmers compared to the gross return to farmers without intervention (Anderson and Nelgen, 2013). NRA is available for aggregate production, for specific crops and disaggregated into different price distorting policy measures.

**Capital**: gross fixed capital stock in agriculture is available from FAO. The gross fixed capital stock is the value of assets held by a farmer. Assets are valued at 'as new' prices and are the sum of physical assets (land development, livestock, machinery and equipment, and structures for livestock) (FAO). The variable is divided by the rural population to get a per capita measure.

**Land**: arable land plus permanent crops per capita from FAO.

**Rainfall**: yearly rainfall in mm used from Hsiang et al. (2011) for Ethiopia and South-Africa, and Dell et al. (2009) for the other countries.

**Temperature**: average temperature in °C from Dell et al. (2009)

**Fertilizer**: data from FAO is used. Fertilizer use is measured by the quantity of fertilizer in metric tons of plant nutrient consumed in agriculture by a country. Because of unreliability of this data and small year-to-year variation in fertilizer, a dummy variable indicating low and high fertilizer countries will be used, following Schlenker and Lobell (2010). The classification is based on whether countries are below or above the average use of fertilizer per year within the sample.

**Democracy** governance data is not available for African countries, therefore democracy data is used as a proxy. As research shows that a democratic government is likely to score higher on governance (Rivera-Batiz, 2002). Data on democracy is available from PolityIV. The variable from the PolityIV database used is POLITY2, which is a combination of an authoritarian and democratic regime score. Each regime is valued on the basis of a set of indicators measuring the competitiveness of recruitment, openness of executive recruitment, constraint on chief executive and competitiveness of political participation (PolityIV).

**Roads**: data on an overall infrastructure quality is not available, but a proxy on the length of the road network is used instead. Data is available from Calderón and Servén (2004). Calderón and Servén (2004) constructed a measure of logged length of road network in kilometers per square kilometers of land area, for every five years. This value is taken for all five years. Data is available for 16 countries in the sample. The length of the total road network is used as a measure for infrasructure. If the road network is more advanced, farmers will have more access to markets which may have yield-increasing effects.

**Telephone**: data on information access is not available, but a proxy on the amount of telephone lines is used instead. Data is available from Calderón and Servén (2004). Similarly to the roads variable, Calderón and Servén (2004) constructed a variable that measures main telephone lines and mobile phones per 1000 workers. Data is available for 16 countries. This variable serves as a measure for access to information, which, as explained in the theoretical framework, is expected to affect yield.

**Policy variability**: a policy variability variable is constructed by taking the 5-year rolling relative standard

deviation of NRA. This implies that the policy variability for one year is calculated by taking the relative standard deviation (or coefficient of variation) of the previous five years.
**Policy switch**: a policy switch dummy is constructed that indicates whether a country has switched from taxation to subsidization or from subsidization to taxation compared to the previous year.

*Determinants of agricultural price distortions*
**Income**: GDP per capita in 2005 USD from WorldBank (c).
**Economic growth**: annual GDP growth in percentages from WorldBank (c).
**Public debt**: public debt in percentage of GDP from Abbas et al. (2010)
**Inflation**: inflation of average consumer prices in percentage change from IMF. This variable is included because it represents overall macroeconomic stability of a country (together with public debt).
**Population**: total population from Anderson and Nelgen (2013). Population is used as a proxy for food self-sufficiency, because the larger the population, the more food a country needs to produce. Countries with larger populations are less likely to be food self-sufficient.
**Structural change**: share of rural population of total population.
**Food self-sufficiency**: cereal import dependency ratio in percentage (three year average) from FAO . This variable measures how much of the available domestic cereal supply is imported. It indicates how food self-sufficient a country is, where a high cereal import dependency ratio indicates low food self-sufficiency. This variable is available from 1990-2010.
**Elections**: dummy indicating whether there was an election this year. In a presidential system only presidential elections are counted, whereas in parliamentary systems only legislative elections are counted. Data is available from Beck et al. (2001). The variables LEGELEC and EXELEC are combined with information on the political system to construct the election dummy.
**Party orientation**: dummy indicating executive party orientation, where a party can be either right-wing (-1), centrist (0), or left-wing (1). Data comes from Beck et al. (2001). Party orientation is measured with respect to economy policy. If the executive party is defined as conservative, Christian democratic, or right-wing, the party is coded as right-wing. If the executive party is defined as communist, socialist, social democratic, or left-wing, the party is coded as left-wing. If the executive party is defined as centrist or its position can be best described as centrist, the party is coded as centrist.
**System**: dummy indicating whether a country has a presidential or parliamentary executive system. Data comes from Beck et al. (2001). In Beck et al. (2001) a distinction between presidential systems, assembly-elected presidential systems and parliamentary systems is made. Because it is not expected that that an assembly-elected presidential system will have different price policy than presidential systems both are coded as presidential in this research.
**Monopoly power**: the monopoly power of a country in the production of a certain crop, calculated by taking the share of production of world production, using data from FAO. Monopoly power could also be measured with a dummy to indicate monopoly power or not, however, as monopoly power in Africa is relatively irrelevant, the variance of such a dummy will be too low.
**Post-Uruguay**: a year-dummy indicating whether the data is before or after the Uruguay Trade Round in 1986. This dummy is included to take into account the trade liberalization measures that were implemented after the trade round.

### 4.1.4 Model specification

The relationships between the different variables are modelled in a linear panel model for each of the crops and the aggregated model. Democracy, policy variability, roads, and telephone are expected to have an interaction effect with NRA. It is expected that the effect of NRA on yield depends on the extent to which a country has good governance, the pricing policy varies, markets are developed, and rural areas have access to (price) information. Together with NRA, these variables are lagged by one year, as investment decisions of farmers will be based on the previous year. Productivity will thus also depend on the lagged values of these variables. Yield and per capita land are logged because of non-normality. Roads and telephone are already logged. Interaction effects are difficult to interpret, especially when there are multiple interaction effects in one equation. Hence, to get interpretable results, the models are estimated with one interaction effect at the time. Apart from that, different model specifications are tested, ranging from very simple to more complex models. The first equation estimated is:

$$\begin{aligned} \log yield_{i,t} =& \alpha + \beta_1 nra_{i,t-1} + \beta_2 capital_{i,t} + \beta_3 \log land_{i,t} \\ &+ \beta_4 rainfall_{i,t} + \beta_5 temperature_{i,t} \\ &+ \beta_6 democracy_{i,t-1} + \beta_7 (nra_{i,t-1} democracy_{i,t-1}) \\ &+ \beta_8 policyvariability_{i,t-1} + \beta_9 (nra_{i,t-1} policyvariability_{i,t-1}) \\ &+ \beta_{10} roads_{i,t-1} + \beta_{11} (nra_{i,t-1} roads_{i,t-1}) \\ &+ \beta_{12} telephone_{i,t-1} + \beta_{13} (nra_{i,t-l} telephone_{i,t-1}) + \epsilon_{i,t} \\ &\text{for} \quad t = 1, ..., T \quad \text{and} \quad i = 1, ...N \end{aligned} \tag{1}$$

For the second equation, logs are taken for per capita income, population, and debt to create normal distributions and reduce the influence of outliers. Different models are estimated to create a full understanding of the determinants of pricing policy whilst taking data limitations into account. The second equation estimated is:

$$\begin{aligned} nra_{i,t} =& \alpha + \beta_1 \log income_{i,t-1} + \beta_2 growth_{i,t-1} \\ &+ \beta_3 \log population_{i,t-1} + \beta_4 structural_{i,t-1} + \beta_5 \log debt_{i,t-1} \\ &+ \beta_6 fss_{i,t-1} + \beta_7 monopoly_{i,t-1} + \beta_8 democracy_{i,t-1} \\ &+ \beta_9 election_{i,t} + \beta_{10} system_{i,t} + \beta_{11} party_{i,t} \\ &+ \beta_{12} uruguay_{i,t} + \epsilon_{i,t} \\ &\text{for} \quad t = 1, ..., T \quad \text{and} \quad i = 1, ...N \end{aligned} \tag{2}$$

## 4.2 Data analysis

In this section the methods of analysis and diagnostic statistical tests are given. First, some summary statistics for the data are described. Multicollinearity, heteroskedasticity and autocorrelation are tested for both the equations. After this, the different estimation techniques are described. The first technique to estimate both equations separately is a fixed effects model, the second method is a multilevel model. Following this, possible endogeneity issues are described and possible remedies for these issued are discussed.

### 4.2.1 Summary statistics

The summary statistics of the data are given in table 1 and 2. Table 1 gives descriptives on all crop-specific variables: NRA, policy variability, yield and monopoly power. Whereas table 2 gives descriptives on country-specific variables. It is interesting to highlight the NRAs for the the crops studies. Average NRA for aggregate agriculture is negative, as are average NRAs for the exportables. The import-competing crops have positive average NRAs. The policy variability variables are in fact a standard deviation of NRA, but calculated as a five year rolling standard deviation (as explained in the variable descriptions). The average policy variability for all crops is quite high compared to NRA. For monopoly power, this is high for cocoa, and to a lesser extent coffee. The former value can be explained by high production of cocoa in Côte d'Ivoire.

**Table 1:** Summary statistics for crop-specific variables

| **Crop** | **NRA** | **N** | **PV** | **N** | **Yield** | **N** | **Mon** | **N** |
|---|---|---|---|---|---|---|---|---|
| Aggregate | -0.137 | 940 | 1.509 | 856 | 134 | 850 | | |
| | (0.245) | | (5.045) | | (36) | | | |
| Cocoa | -0.373 | 240 | 3.429 | 216 | 4025 | 249 | 12.24 | 250 |
| | (0.267) | | (21.82) | | (1775) | | (10.77) | |
| Coffee | -0.340 | 306 | 0.689 | 278 | 4566 | 349 | 2.271 | 350 |
| | (0.273) | | (1.294) | | (1841) | | (1.476) | |
| Cotton | -0.333 | 667 | 2.160 | 607 | 8845 | 797 | 0.477 | 748 |
| | (0.339) | | (12.28) | | (5305) | | (0.657) | |
| Maize | 0.093 | 578 | 2.702 | 522 | 15988 | 699 | 0.461 | 650 |
| | (0.650) | | (9.731) | | (12551) | | (0.547) | |
| Rice | 0.009 | 442 | 2.321 | 391 | 20235 | 493 | 0.198 | 493 |
| | (0.445) | | (7.567) | | (18124) | | (0.267) | |
| Sugar | 0.142 | 375 | 2.978 | 339 | 609785 | 437 | 0.450 | 537 |
| | (0.546) | | (10.40) | | (320284) | | (0.600) | |
| Wheat | 0.049 | 364 | 3.126 | 328 | 23673 | 450 | 0.228 | 450 |
| | (0.506) | | (9.734) | | (16916) | | (0.287) | |

PV is policy variability, Mon is monopoly power. Standard deviations are in parenthesis

**Table 2:** Summary statistics for country-specific variables

| **Variable** | **Mean** | **Std. Dev.** | **N** |
|---|---|---|---|
| Land | 6835.946 | 6974.054 | 1100 |
| Capital | 14958.74 | 13676.033 | 726 |
| Rainfall | 9.287 | 4.617 | 978 |
| Temperature | 23.912 | 3.432 | 1014 |
| Democracy | -2.913 | 5.347 | 1068 |
| Roads (log) | -2.682 | 0.799 | 624 |
| Telephone (log) | 2.170 | 1.386 | 624 |
| Population | 20115817 | 21534343 | 1092 |
| Rural population share | 0.723 | 0.135 | 1092 |
| GDP per capita | 815.722 | 1031.672 | 998 |
| GDP growth | 3.778 | 5.403 | 995 |
| Inflation | 12.562 | 20.202 | 853 |
| Debt | 64.385 | 45.498 | 853 |
| Food self-sufficiency | 19.878 | 15.466 | 441 |
| Party orientation | 0.533 | 0.827 | 287 |

### 4.2.2 Multicollinearity

When two or more variables in the model are correlated, this leads to larger standard errors and thus can lead to unreliable results. In general, correlation between variables is not a problem but when correlations are too high the problem of multicollinearity can occur. Estimates will still be consistent even when multicollinearity poses a problem (Dougherty (2011), Verbeek (2013)). To test for multicollinearity, pairwise correlation matrices are looked at (see appendix A). Regarding the first equation; for aggregate agriculture, the correlation matrix shows some large correlations; temperature and telephone are strongly negatively correlated (-0.595), as are fertilizer and telephone lines (-0.623). As correlations are not above |0.8|, it is assumed that multicollinearity is not problematic. This is the same for the correlation matrices of the other crops. For the variables of which we would expect high correlations, roads/telephone and democracy, multicollinearity does not seem to be a problem. For the second equation correlations are all below |0.4| except for the share of rural population and GDP per capita (-0.694), which is expected. Multicollinearity is also not a problem for the second equation.

### 4.2.3 Heteroskedasticity

Heteroskedasticity occurs when the disturbances do not have the same variation across all observations (Verbeek, 2013). To test for heteroskedasticity the Modified Wald test for group-wise heteroskedasticity is used (see appendix B). For both equations the null-hypothesis of homoskedasticity is rejected, therefore there is heteroskedasticity in the disturbance terms. This means that when estimating the equations, robust standard errors are needed to control for heteroskedastic errors.

### 4.2.4 Fixed effects model

To analyze panel data a fixed or random effects model can be used. A fixed effects model is a simple linear regression model in which the intercept terms vary over individual units. The fixed effects model is thus focused on differences within individuals. A random effects model is more efficient than a fixed effects model because it also takes into account observed characteristics that remain constant for each country. However to use random effects models the precondition that the observations are randomly drawn from the population needs to be met (Dougherty, 2011). This is not the case for this sample of countries. Formally, a Hausman test shows whether a fixed effects or random effects model is more appropriate (see appendix C). The fixed effects estimator is known to be consistent, whereas the random effects estimator is consistent under the assumption tested. When there is no systemic difference between the two coefficients, a random effects estimator is more appropriate, as it is consistent and efficient. The null-hypothesis of the model is that there are no systematic differences between fixed and random effects estimates. The null-hypothesis is rejected, therefore a fixed effects model is, as expected, more appropriate for this data.

### 4.2.5 Multilevel model

In addition to a fixed effects model for each specific crop, a multilevel or mixed model will also be estimated. In a multilevel model, the data is structured hierarchically. For each country, there is data on each crop, for each year. Thus, instead of having two levels (country and year for each crop), there are now three levels. The advantage of using a multilevel model is more efficiency because of a larger number of observations, as the effect on all crops can be estimated together. However, the disadvantage of this method is that the panel structure of the data is lost. Within a multilevel model, it is possible to fit a random slope and/or random intercepts model. This means that the slope and/or intercept across groups may be varied. A likelihood ratio test can determine which model is more appropriate for the data.

#### 4.2.6 Endogeneity

Within the model, some variables may be endogenous. NRA, as with all policy variables, can be expected to be endogenous. A variable is endogenous when there is reverse causality with the dependent or when it is correlated with the disturbance term (Verbeek, 2013). Institutional factors most likely affect yield, but controlling for institutional factors is difficult as they are barely measured. NRA on the other hand will also be determined by institutional factors. Therefore, econometric techniques to address endogeneity must be used to test the validity of the relationship between productivity and NRA. The methods used in this thesis to account for endogeneity are instrumental variable (IV) regression and a simultaneous equation model. Both will be discussed in the following sections.

#### 4.2.7 Instrumental variable regression

To use IV regression to address endogeneity of NRA, a good instrument for NRA is necessary. A good instrument has two characteristics: it is correlated with the instrumented (endogenous) variable, and it is not correlated with the disturbance terms (Verbeek, 2013). Ideally, all independent variables of the second equation could be instruments for NRA, however, many of these may also be related with yield. Two variables that could theoretically be good instruments are population and rural population as a share of the total population. The first affects NRA in the sense that the larger the population, the lower food self-sufficiency. As explained before, African countries often try to meet food self-sufficiency with the help of pricing policy (Anderson et al., 2013). The share of rural population of the total population is also thought to be related to NRA. The smaller the rural population as a share of the total population, the more agriculture will be subsidized (Anderson et al., 2013).[1] Another possible instrument is the value added of agriculture as a percentage of GDP. This is thought to be related to GDP per capita and economic development, as more developed countries tend to have relatively less important agricultural sectors. Some relation with yield could exist; as yield increases the agricultural value added as a percentage may increase, if value added of other sectors remains constant. To mitigate this possible effect, the agricultural value added will be lagged. For the IV estimator to be consistent, the moment conditions need to hold true. This can be tested by testing for overidentiying restrictions with the Hansen J test. If the moment conditions do not hold true, the instruments are not valid. For this estimation technique, different instruments and combinations of instruments wil be tested.

#### 4.2.8 Simultaneous equation model

The last method used in this research also takes into account the endogeneity problem. Using three-stage-least squares, the two equations are estimated simultaneously.[2] In order for the simultaneous equation model to be valid the order condition for identification must be met. This means that for each equation there must

[1] Whether these possible instruments are also related with yield is questionable. One possible mechanism is that the larger the population, the more land market competition there is. This may affect yield.

[2] As three-stage-least-squares is not possible for panel data, the panel structure of the data is ignored. This has some efficiency loss but does not lead to biased coefficients.

be at least as many exogenous variables which are not included in the particular equation, in order to serve as instruments for each of the endogenous variables (Dougherty, 2011). For the equations estimated the order condition for identification holds.

# 5 Results

The results section is divided into three parts. The first part covers the first equation of the model. Both the results of the fixed effects model and the mixed model are given, and different model specifications are discussed. The second part covers the determinants of agricultural price distortions, again different specifications of the fixed effects and mixed model are discussed. Lastly, the endogeneity issue is addressed by discussing the results of the IV-regression and the simultaneous equation model.

## 5.1 Part I: Effect of agricultural price distortions on land use efficiency

### 5.1.1 Model 1.1: Fixed effects

In this part of the analysis four models are tested. The first is the most simple model, only including inputs and climate data together with NRA. The other three models also include one of the interaction variables with NRA: democracy, policy variability, and infrastructure (roads and telephone). Table 3 gives the results of the most simple model for agriculture as a whole and for the seven separate crops.[3]

**Table 3:** Model 1.1.1 Fixed effects model: effect of NRA on yield

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| | Aggregate | Cocoa | Coffee | Cotton | Maize | Rice | Sugar | Wheat |
| L.nra | 0.00672 | 0.404*** | 0.0680 | 0.297* | 0.0166 | 0.0564 | -0.159 | -0.0420 |
| | (0.14) | (5.85) | (0.28) | (1.99) | (0.76) | (1.68) | (-1.24) | (-0.67) |
| log land | -0.393*** | -0.504 | 0.000511 | -0.338 | -0.977*** | -0.540** | -1.258 | -0.455 |
| per cap | (-3.08) | (-2.00) | (0.00) | (-1.52) | (-3.88) | (-2.36) | (-1.44) | (-1.25) |
| log capital | 0.129 | -1.134*** | -0.00224 | 0.263 | 0.902*** | -0.306 | 0.542* | -0.0595 |
| per cap | (0.71) | (-5.72) | (-0.00) | (0.84) | (3.68) | (-1.37) | (1.91) | (-0.14) |
| rainfall | 0.0104 | 0.00453 | -0.00872 | 0.0139 | 0.0315 | 0.00829 | -0.000412 | 0.0505** |
| | (1.57) | (0.51) | (-0.61) | (0.77) | (1.53) | (0.86) | (-0.05) | (2.34) |
| temperature | 0.00722 | 0.168* | -0.0195 | 0.0181 | -0.164* | 0.0194 | 0.0485 | 0.0813* |
| | (0.22) | (2.58) | (-0.31) | (0.20) | (-1.84) | (0.34) | (0.90) | (2.19) |
| fertilizer | 0.163*** | -0.00456 | 0.0758 | -0.252** | 0.114** | -0.00735 | 0.211 | -0.116** |
| | (4.98) | (-0.62) | (0.53) | (-2.83) | (2.50) | (-0.16) | (1.50) | (-2.31) |
| Constant | 2.342** | -7.767* | 8.930 | 7.721* | 11.35*** | 2.570 | 5.340 | 3.884* |
| | (2.17) | (-2.36) | (1.53) | (2.05) | (6.18) | (1.41) | (0.69) | (1.89) |
| Observations | 574 | 139 | 187 | 444 | 348 | 272 | 223 | 218 |
| $R^2$ | 0.136 | 0.568 | 0.007 | 0.108 | 0.204 | 0.225 | 0.200 | 0.191 |

NRA is crop-specific. Fertilizer is consumption for cocoa, dummy for others.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

As explained in the methodology section, aggregate yield is measured by a weighted yield index, including the most important crops for a specific country. For aggregate agriculture NRA does not have a significant impact on yield. For cocoa and cotton, yield is significantly affected by NRA. For both crops, NRA has

[3] A model with a lagged dependent is also estimated (Appendix D1), but the results are not very different.

a positive effect, meaning that if the nominal rate of assistance to the crop increases in one year (more subsidization or less taxation) yield increases in the following year, keeping the other variables in the model constant. For cocoa, when NRA goes up by one unit, indicating an increase in price of 100%, yield goes up by 40.4%. For cotton, if NRA goes up by one unit, yield increases by 29.7%. To get a more realistic idea about the effect size of NRA on these two crops, the effect of an increase of one standard deviation was calculated and reported in table 4.

**Table 4:** Effect on yield of a standard deviation increase in NRA

| Variable | Std. dev. | Effect size |
|---|---|---|
| Cocoa | 0.267 | 10.8% |
| Cotton | 0.339 | 10.1% |

The other variables in this model are significant depending on the crop. Per capita land is significant for aggregate agriculture, maize, and rice. For all three, land has a negative effect, meaning that if land per capita increases, yield decreases. This implies that the more scarce land is, the more agricultural intensification takes place. The magnitude of the effect is quite small, as a 1% increase in land leads to a 0.040% decrease in yield, keeping other variables constant. Capital is significant for cocoa, maize, and sugar. A 1% increase in capital per capita leads to a decrease of cocoa yield of 0.10%, a 0.09% increase in maize yield, and a 0.05% increase in sugar yield. The climate variables are hardly significant, which could be explained by low variability in climate and the rough nature of annual country climate data. Fertilizer is significant for agriculture as a whole, cotton, maize, and wheat. For aggregate agriculture and maize the effect is positive, a high fertilizer using country has a 16.3% and 11.4% higher yield respectively, *ceteris paribus.* For cotton and wheat, the effect is negative, a high fertilizer using country has a 25,2% and 11,6% lower yield respectively, keeping other variables constant. All inputs: land, capital, and fertilizer are measured on an aggregate level. This means that the inputs are not crop specific. Therefore, an increase in fertilizer use in a country does not necessarily mean that fertilizer is increased for all crops. This could explain why for capital and fertilizer the effects are sometimes not as expected.

The second model (table 5) includes democracy as a direct and interaction effect. The hypothesis being that more democratic countries will have more effective policy and thus the effect of NRA on yield will be stronger in democratic countries than in autocratic countries. From the results it can be seen that democracy has no direct effect on the yield of any of the crops, including aggregate agriculture. For cocoa and wheat, the interaction effect is negative. For cocoa, when the democracy-score is 0, the effect of a one unit increase in NRA is 27.3%. However, when the democracy score increases by 1, the effect of a one unit increase in NRA decreases by 2.95%. A more democratic government decreases the effect of NRA. For wheat the interaction effect is also negative and has a magnitude of -1.93%. A possible explanation for this could be that in democratic countries, property rights are well established. When property rights are well established farmers may focus on other (long-term viable) crops and hence yields of some crops may go down. Another interesting result is that the effect of NRA on wheat becomes significant and negative when controlling for democracy. If assistance to wheat increases by one unit, wheat yield decreases by 12.3%, *ceteris paribus.*

**Table 5:** Model 1.1.2 Fixed effects model: effect of NRA and democracy on yield

| | (1) Aggregate | (2) Cocoa | (3) Coffee | (4) Cotton | (5) Maize | (6) Rice | (7) Sugar | (8) Wheat |
|---|---|---|---|---|---|---|---|---|
| L.nra | 0.0372 | 0.273** | -0.0852 | 0.283* | -0.00968 | 0.0735 | -0.0958 | -0.123** |
| | (0.28) | (4.29) | (-0.37) | (1.83) | (-0.33) | (1.52) | (-1.12) | (-2.82) |
| log land per cap | -0.395*** | -0.402** | 0.0724 | -0.331 | -0.961*** | -0.507** | -1.489 | -0.441 |
| | (-3.05) | (-3.80) | (0.16) | (-1.57) | (-3.97) | (-2.36) | (-1.56) | (-1.16) |
| log capital per cap | 0.133 | -1.179*** | -0.119 | 0.282 | 0.868*** | -0.342 | 0.536 | -0.0695 |
| | (0.70) | (-12.22) | (-0.22) | (0.96) | (3.82) | (-1.43) | (1.14) | (-0.16) |
| rainfall | 0.0104 | 0.00457 | -0.00942 | 0.0105 | 0.0286 | 0.00985 | -0.00402 | 0.0523** |
| | (1.56) | (0.59) | (-0.60) | (0.61) | (1.47) | (0.94) | (-1.34) | (2.31) |
| temperature | 0.00718 | 0.145* | -0.0329 | 0.0198 | -0.182* | 0.0154 | 0.0699 | 0.0842* |
| | (0.23) | (2.16) | (-0.43) | (0.21) | (-2.01) | (0.26) | (1.40) | (2.28) |
| fertilizer | 0.158*** | -0.00389 | 0.0777 | -0.249*** | 0.129* | 0.0139 | 0.269 | -0.144** |
| | (4.19) | (-0.70) | (0.57) | (-3.39) | (2.09) | (0.24) | (1.73) | (-2.70) |
| L.democracy | 0.000562 | -0.00336 | -0.0101 | 0.00256 | 0.00257 | 0.00275 | -0.0256 | -0.00278 |
| | (0.13) | (-0.19) | (-1.19) | (0.41) | (0.31) | (0.51) | (-1.47) | (-0.47) |
| L.nra*democracy | 0.00570 | -0.0295* | -0.0421 | 0.00381 | -0.00749 | 0.00511 | 0.00475 | -0.0193** |
| | (0.27) | (-2.27) | (-1.68) | (0.13) | (-1.74) | (0.72) | (0.53) | (-2.89) |
| Constant | 2.351** | -6.698*** | 8.967 | 7.911* | 11.65*** | 2.670 | 2.846 | 3.879** |
| | (2.46) | (-6.31) | (1.49) | (2.12) | (6.42) | (1.38) | (0.34) | (2.35) |
| Observations | 573 | 139 | 186 | 443 | 347 | 271 | 222 | 218 |
| $R^2$ | 0.136 | 0.588 | 0.050 | 0.104 | 0.212 | 0.232 | 0.263 | 0.216 |

NRA is crop-specific. Fertilizer is consumption for cocoa, dummy for others.

$^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

The third model (table 6) looks at the variability of NRA policy and whether high variability can decrease the effect of NRA[4]. When the interaction effect of NRA and the policy variability of the NRA is added to the model this gives significant effects. It is expected that the effect of NRA on yield depends on how much the NRA varies the previous years. The higher the variability of policy, the less NRA will have impact on farming decisions, farmers cannot count on stable policy and thus changing their farm plan because of a policy change is less likely. For cocoa, cotton, and wheat this is indeed the case. If the policy variability of NRA for cocoa is 0, the effect on yield if NRA increases by one unit is 39,9%, keeping all else constant. However, if the policy variability is one, this effect decreases by 0.74%. Thus, the impact of an increase in support for cocoa decreases as the policy varies more. The partial effect of NRA for cocoa is given by equation 3.

$$\frac{\partial yield}{\partial NRA} = 0.399 - 0.0074 * PV \tag{3}$$

Policy variability also has a direct effect on the yield. A one unit increase in policy variability leads to an increase in yield of 0.156%, if NRA is 0. If NRA increases by one unit, the direct effect of policy variability

[4] A model with the policy switch variable instead of policy variability is also estimated but yields hardly any significant results for policy switch. Refer to appendix D2.

decreases by 0.74%. The partial effect of policy variability is given by equation 4.

$$\frac{\partial yield}{\partial PV} = 0.00156 - 0.0074 * NRA \tag{4}$$

For cotton, the result is much larger than for cocoa. The effect of NRA on yield decreases by 6.87% if policy variability becomes 1, *ceteris paribus*. From theory, we expect that perennial yield will be affected less by policy changes, whereas investment in annuals is more variable. This is not seen in these results, as cotton is planted as an annual crop, we would expect that the effect of NRA on cotton yield would decrease less when policy variability is high. For wheat, the effect of NRA is still insignificant when an interaction with policy variability is added to the model, but the effect of policy variability is significant and negative. The effect of NRA decreases by 0.61% if policy variability is 1.

**Table 6:** Model 1.1.3 Fixed effects model: effect of NRA and policy variability on yield

| | (1) Aggregate | (2) Cocoa | (3) Coffee | (4) Cotton | (5) Maize | (6) Rice | (7) Sugar | (8) Wheat |
|---|---|---|---|---|---|---|---|---|
| L.nra | -0.0109 | 0.399*** | 0.322 | 0.371* | 0.0188 | 0.0610 | -0.172 | -0.0233 |
| | (-0.19) | (5.08) | (1.23) | (2.00) | (0.73) | (1.23) | (-1.17) | (-0.39) |
| log land per cap | -0.402*** | -0.487 | 0.268 | -0.296 | -1.102*** | -0.490* | -1.171 | -0.529 |
| | (-2.96) | (-1.72) | (0.65) | (-1.25) | (-4.11) | (-1.98) | (-1.22) | (-1.31) |
| log capital per cap | 0.126 | -1.174*** | 0.284 | 0.315 | 1.009*** | -0.469 | 0.477 | -0.0205 |
| | (0.66) | (-6.09) | (0.51) | (1.01) | (3.07) | (-1.71) | (1.79) | (-0.05) |
| rainfall | 0.00952 | 0.00540 | -0.0121 | 0.0110 | 0.0340 | 0.0125 | 0.00135 | 0.0414* |
| | (1.48) | (0.61) | (-0.90) | (0.58) | (1.70) | (1.08) | (0.10) | (2.15) |
| temperature | 0.00496 | 0.183* | 0.0226 | 0.0122 | -0.169* | 0.0190 | 0.0449 | 0.0761 |
| | (0.15) | (2.53) | (0.39) | (0.14) | (-1.84) | (0.31) | (0.85) | (1.80) |
| fertilizer | 0.163*** | -0.00468 | -0.0126 | -0.264*** | 0.125** | 0.0134 | 0.194 | -0.113** |
| | (4.89) | (-0.63) | (-0.10) | (-2.96) | (2.57) | (0.28) | (1.22) | (-2.39) |
| L.policyvar | 0.00233 | 0.00156*** | -0.0346* | 0.00439 | -0.000418 | -0.00361 | -0.00279 | 0.00351* |
| | (1.12) | (26.07) | (-2.35) | (0.66) | (-0.42) | (-1.41) | (-0.80) | (2.11) |
| L.nra*policy | 0.00837 | -0.00740*** | -0.177 | -0.0687* | 0.000560 | 0.00595 | 0.00169 | -0.00608*** |
| | (0.33) | (-5.49) | (-1.54) | (-2.07) | (0.26) | (1.26) | (0.77) | (-3.61) |
| Constant | 2.314* | -8.262* | 12.28* | 8.624** | 11.14*** | 1.833 | 5.689 | 3.723* |
| | (2.02) | (-2.70) | (2.29) | (2.26) | (5.68) | (0.88) | (0.70) | (1.96) |
| Observations | 556 | 135 | 179 | 436 | 328 | 253 | 211 | 206 |
| $R^2$ | 0.141 | 0.601 | 0.062 | 0.125 | 0.237 | 0.246 | 0.190 | 0.225 |

NRA is crop-specific. Fertilizer is consumption for cocoa, dummy for others.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

For the fourth model (table 7), two interaction effects are added; NRA with roads and telephone. The aim of this model is to see how infrastructure and information access affect yield, and how these variables affect the impact of NRA on yield. The results show significant effects for some of the crops. For cotton, the direct and interaction effects are all significant and the signs are as expected. The direct effects are positive, indicating that when NRA is 0, yield is positively affected when markets are better developed and there is access to information. The interaction effect for roads is positive, indicating that a more developed infrastructure increases the effect of NRA, telephone however has a negative interaction effect, indicating an decrease in the effect of NRA as information access increases. What also stands out is that the effect of on cotton becomes very large in this model, an increase in NRA of 1% leads to a 273.5% increase in yield, *ceteris paribus*. Moreover, the effects on rice, sugar and wheat become significant in this model. NRA has a positive effect on rice yield (43.8%). For sugar and wheat the effect of NRA is negative, meaning that an increase in subsidies, or decrease of taxation, leads to a decrease in productivity. The magnitudes of the effects are respectively -93.1% and -54.0% due to a 1 unit increase in NRA, keeping other variables constant.

**Table 7:** Model 1.1.4 Fixed effects model: effect of NRA and infrastructure on yield

| | (1)<br>Aggregate | (2)<br>Cocoa | (3)<br>Coffee | (4)<br>Cotton | (5)<br>Maize | (6)<br>Rice | (7)<br>Sugar | (8)<br>Wheat |
|---|---|---|---|---|---|---|---|---|
| L.nra | 0.320 | -0.0289 | 0.788 | 2.735** | -0.0149 | 0.438** | -0.931*** | -0.540*** |
| | (0.84) | (-0.03) | (1.02) | (2.92) | (-0.22) | (2.72) | (-3.98) | (-4.17) |
| log land | -0.0488 | 0.439 | 0.0536 | 0.177 | -0.0611 | -0.0751 | -1.773 | 0.409 |
| per cap | (-0.59) | (0.64) | (0.08) | (0.50) | (-0.33) | (-0.31) | (-1.39) | (1.02) |
| log capital | -0.0743 | -2.568** | -0.278 | -0.0645 | 0.0234 | -0.890** | 1.157 | -0.739** |
| per cap | (-0.55) | (-3.54) | (-0.56) | (-0.17) | (0.17) | (-2.56) | (1.45) | (-2.63) |
| rainfall | 0.00983 | 0.00307 | -0.000656 | 0.0363 | 0.0229 | -0.00102 | 0.0000413 | 0.0284 |
| | (1.53) | (0.39) | (-0.04) | (1.65) | (1.44) | (-0.13) | (0.00) | (1.68) |
| temperature | -0.0390 | 0.0786 | -0.0466 | 0.00763 | -0.201** | -0.0286 | -0.0204 | 0.0562* |
| | (-1.51) | (1.02) | (-0.63) | (0.07) | (-2.40) | (-0.93) | (-0.25) | (2.15) |
| fertilizer | 0.116*** | -0.00769 | 0.160 | -0.171* | 0.108** | 0.0195 | 0.283 | -0.128** |
| | (3.70) | (-1.17) | (1.02) | (-1.87) | (2.47) | (0.38) | (1.42) | (-3.48) |
| L.roads | 0.00123 | 0.170 | -0.383 | 0.898** | 0.211* | -0.0163 | -0.758 | -0.0119 |
| | (0.01) | (0.49) | (-0.82) | (3.10) | (2.18) | (-0.11) | (-1.75) | (-0.04) |
| L.nra*roads | 0.0973 | -0.344 | 0.354 | 0.496** | -0.0805 | 0.176 | -0.174*** | -0.0962* |
| | (0.76) | (-1.92) | (1.14) | (2.35) | (-1.73) | (1.63) | (-4.40) | (-2.02) |
| L.telephone | 0.162*** | 0.174 | 0.129 | -0.269** | 0.205*** | 0.213*** | 0.351* | 0.293* |
| | (3.90) | (0.71) | (0.59) | (-2.48) | (4.84) | (3.64) | (2.25) | (1.97) |
| L.nra*tel | -0.0173 | -0.164 | 0.0581 | -0.454** | -0.0585** | 0.0272 | 0.0747** | 0.0621** |
| | (-0.28) | (-0.55) | (0.17) | (-2.92) | (-2.41) | (0.56) | (2.54) | (3.19) |
| Constant | 4.420*** | -8.098 | 6.710 | 12.71*** | 13.52*** | 3.023 | 3.895 | 6.232*** |
| | (5.73) | (-2.16) | (1.08) | (3.29) | (8.16) | (1.16) | (0.34) | (4.93) |
| Observations | 378 | 103 | 147 | 258 | 272 | 230 | 155 | 174 |
| $R^2$ | 0.213 | 0.650 | 0.152 | 0.221 | 0.256 | 0.333 | 0.333 | 0.389 |

NRA is crop-specific. Fertilizer is consumption for cocoa, dummy for others.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

### 5.1.2 Model 1.2: Mixed model

In addition to the fixed effects model for finding the relationship between NRA and yield a mixed or multilevel model is estimated. As explained in the methodology section, this means the data is structured hierarchically, and instead of two levels (country and year), the data has three levels (country, crop, and year). For this estimation method, a random intercepts and random slope and intercepts model were estimated. For the random slope and intercepts model, the slope and intercept of NRA are allowed to vary across crops. A likelihood ratio test determined that the random intercept and slope model has more explanatory power than the random intercept model. The former is presented in table 8 for four different models, the interaction effect with roads is not shown as it is insignficant.[5] As shown, in none of the different model specification an effect of NRA on yield is found. As for the first, simple model; land, capital, fertilizer and rainfall are significant and all have a sign that is expected. Below in the graph the random-effects parameters are given. As we expect the effect of NRA to vary across crops, we expect different intercepts but also different slopes of the regression line. The coefficients of the parameters give the average slope, the model constant is the average of crop level intercepts. Sd(l.nra) is the standard deviation of the effect of NRA across crops, sd(crop) is the standard deviation of the intercept across crops, and sd(Residual) is the standard deviation at the individual level. As we can see, the standard deviation of the effect of NRA is quite large compared to the effect of NRA. In the first model, the effect of a one unit increase in NRA, leads to an increase in yield of 15%, but the standard deviation of the effect is around 30%. Though NRA is insignificant, the random slope is not, indicating that there is indeed much variance in the effect of NRA across commodities. The random intercept is not significant. As for the intermediating variables, democracy and telephone are significant, but policy variability is not.

[5] A model with a lagged dependent was also estimated, but this leads to invalid results.

**Table 8:** Model 1.2 Mixed models: effect of NRA on yield

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | logyield | logyield | logyield | logyield |
| L.nra | 0.153 | 0.0830 | 0.148 | 0.0888 |
| | (1.27) | (0.71) | (1.21) | (0.60) |
| log land | -0.116*** | -0.109*** | -0.125*** | -0.266*** |
| per cap | (-2.87) | (-2.66) | (-3.06) | (-5.33) |
| log capital | 0.381*** | 0.379*** | 0.359*** | 0.171*** |
| per cap | (14.67) | (14.30) | (13.41) | (5.41) |
| fertilizer | 0.00320*** | 0.00304*** | 0.00311*** | 0.00151*** |
| | (11.52) | (10.48) | (11.03) | (4.09) |
| rainfall | -0.000136 | -0.00253 | -0.000227 | -0.0116*** |
| | (-0.04) | (-0.75) | (-0.07) | (-2.93) |
| temperature | 0.0111** | 0.00691 | 0.0110** | 0.0203*** |
| | (2.08) | (1.29) | (2.03) | (3.20) |
| L.democracy | | -0.00490* | | |
| | | (-1.85) | | |
| L.nra*democracy | | -0.0206*** | | |
| | | (-4.50) | | |
| L.policyvar | | | 0.000492 | |
| | | | (0.44) | |
| L.nra*policyvar | | | -0.00360 | |
| | | | (-1.16) | |
| L.telephone | | | | 0.0543*** |
| | | | | (3.17) |
| L.nra*telephone | | | | -0.00616 |
| | | | | (-0.29) |
| Constant | 11.15*** | 11.31*** | 10.93*** | 8.334*** |
| | (17.73) | (17.97) | (17.17) | (11.30) |
| sd(l.nra) | 0.307*** | 0.294*** | 0.312*** | 0.350*** |
| | (-4.15) | (-4.29) | (-4.09) | (-3.62) |
| sd(crop) | 1.392 | 1.396 | 1.401 | 1.459 |
| | (1.24) | (1.25) | (1.26) | (1.41) |
| sd(residual) | 0.501*** | 0.495*** | 0.501*** | 0.457*** |
| | (-41.72) | (-42.37) | (-40.68) | (-40.27) |
| Observations | 1831 | 1826 | 1748 | 1339 |

*t* statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

## 5.2 Part II: Determinants of agricultural price distortions

### 5.2.1 Model 2.1: Fixed effects

In this section, the results of the fixed effects estimation of the second equation are presented. Two models on the determinants of NRA are reported on, other models that were estimated are reported in the appendix E. The first model is the most simple model, where the independent variables are log GDP per capita, GDP growth, log population, rural population share, log inflation and log debt. All independent effects are lagged by one year, as policy is determined after evaluating certain economic conditions. We expect countries with a high GDP per capita and high economic growth to have a higher NRA, hence a positive effect is expected. Population is used as a proxy for food self sufficiency, we expect the anti-trade bias to be larger for countries with a larger population to feed. The anti-trade bias is large when import competing commodities are subsidized, and export commodities are taxed. From theory on price distortions, we expect the rural share of the population to have a negative effect, as traditionally countries with a smaller rural population will subsidize agriculture more. Inflation is added to the model to account for overall economic stability. Debt can be a reason for countries to levy more taxes; a negative effect is expected.

**Table 9:** Model 2.1.1 Fixed effects model: determinants of NRA

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| | Aggregate | Cocoa | Coffee | Cotton | Maize | Rice | Sugar | Wheat |
| L.log gdp | 0.000329 | 0.465*** | -0.00809 | 0.269* | 0.0279 | -0.479** | 0.344 | 0.210 |
| per cap | (0.00) | (4.96) | (-0.06) | (2.12) | (0.14) | (-2.78) | (1.14) | (0.52) |
| L.gdpgrowth | 0.00239 | 0.00256 | 0.00205 | -0.000594 | -0.00236 | 0.000502 | -0.00580 | -0.00452 |
| | (1.49) | (0.61) | (0.56) | (-0.52) | (-0.72) | (0.13) | (-0.76) | (-0.62) |
| L.log | 0.262** | 0.549** | -0.0504 | 0.162 | 0.161 | 0.771** | -0.0124 | 0.193 |
| population | (2.22) | (2.83) | (-0.67) | (0.92) | (0.67) | (2.93) | (-0.04) | (0.45) |
| L.share | 0.739 | 0.602 | -0.621* | -0.471 | 1.871 | 4.451** | -2.489 | 0.341 |
| rural pop | (1.46) | (0.90) | (-2.36) | (-0.69) | (1.34) | (3.00) | (-1.38) | (0.19) |
| L.log | -0.0261 | -0.000261 | -0.0570** | -0.0334** | -0.0287 | -0.0134 | -0.0885 | -0.0123 |
| inflation | (-1.71) | (-0.01) | (-3.19) | (-2.70) | (-0.54) | (-0.30) | (-1.28) | (-0.18) |
| L.log debt | 0.0327 | 0.0991* | 0.0759 | 0.00913 | 0.0121 | 0.0870* | -0.0331 | 0.0573 |
| | (1.03) | (2.56) | (1.20) | (0.39) | (0.26) | (2.02) | (-0.31) | (1.55) |
| Constant | -5.043** | -13.34** | 0.865 | -4.379 | -3.950 | -13.20** | 0.0791 | -4.989 |
| | (-2.37) | (-3.86) | (0.61) | (-1.66) | (-0.94) | (-2.99) | (0.02) | (-0.81) |
| Observations | 554 | 182 | 202 | 338 | 372 | 273 | 239 | 229 |
| $R^2$ | 0.084 | 0.386 | 0.176 | 0.213 | 0.027 | 0.122 | 0.111 | 0.041 |

NRA is crop-specific.
$^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

The model, presented in table 9, gives mixed results. Many of the independent variables are insignificant, and for the significant variables, effects vary per commodity. For cocoa and coffee, GDP has a significant positive effect. For cocoa, a 1% increase in GDP per capita leads to a 0.000456 unit increase in NRA, for cotton the effect size is 0.000269, keeping other variables constant. For rice however, the effect of GDP per capita leads to a decrease in NRA of 0.000479, *ceteris paribus.* Thus, rice is taxed more as GDP becomes higher. For the other crops, the effect of GDP is insignificant. GDP growth is insignificant for all crops. Population has a significant effect for aggregate agriculture, cocoa, and rice, and is positive for all three. This is as expected for rice, but not for cocoa. As for the magnitude of the population effect, a 1% increase in population leads to a 0.000549 unit increase in NRA of cocoa, a 0.000771 unit increase in NRA of rice, and a 0.000262 unit increase in aggregate NRA, *ceteris paribus*. The rural population variable has a significant negative effect for coffee, which is as expected, and a positive effect for rice. Inflation is negative for coffee and cotton. Debt is positive for cocoa and rice, which is not as expected. Overall there are very little significant effects, and the effects that are present are small and sometimes different from what we expected.

**Table 10:** Model 2.1.2 Fixed effects model: determinants of NRA, full model

| | (1)<br>Aggregate | (2)<br>Cocoa | (3)<br>Coffee | (4)<br>Cotton | (5)<br>Maize | (6)<br>Rice | (7)<br>Sugar | (8)<br>Wheat |
|---|---|---|---|---|---|---|---|---|
| L.log gdp | 0.00103 | 0.934** | 0.598 | 0.800*** | -0.246 | -0.508 | 0.576 | -1.729*** |
| per cap | (0.01) | (3.25) | (1.07) | (3.26) | (-0.60) | (-1.75) | (0.63) | (-4.65) |
| L.gdpgrowth | 0.00243 | 0.00616 | 0.00355 | -0.00343 | 0.00610 | -0.00552 | 0.000490 | 0.0109 |
| | (1.50) | (1.92) | (0.70) | (-0.89) | (1.54) | (-0.48) | (0.04) | (1.45) |
| L.log | 0.0119 | -0.341 | -0.297 | -0.474 | -0.0337 | -0.181 | -1.457* | 1.251** |
| population | (0.10) | (-0.97) | (-0.43) | (-1.35) | (-0.14) | (-0.48) | (-1.98) | (2.61) |
| L.share | 0.0159 | -1.149 | -1.175 | -2.317* | 0.324 | 1.457 | -1.633 | -1.564* |
| rural pop | (0.03) | (-0.52) | (-1.93) | (-2.01) | (0.31) | (1.19) | (-0.92) | (-2.08) |
| L.log | -0.0278*** | 0.0381 | -0.0326** | -0.0277** | -0.0660* | -0.0493 | -0.123 | -0.00539 |
| inflation | (-3.13) | (1.00) | (-2.73) | (-2.48) | (-1.85) | (-1.27) | (-1.24) | (-0.10) |
| L.log debt | 0.00419 | 0.0391 | 0.192** | 0.0485 | 0.0638 | 0.0318 | -0.264 | -0.124 |
| | (0.12) | (1.85) | (2.95) | (1.33) | (0.96) | (0.39) | (-1.20) | (-1.18) |
| L.democracy | 0.000989 | -0.0162 | 0.00925 | -0.0000707 | 0.00164 | 0.0277** | 0.0284** | -0.0185 |
| | (0.20) | (-1.35) | (0.55) | (-0.01) | (0.16) | (2.34) | (2.43) | (-1.27) |
| L.food | 0.00350 | 0.00914 | 0.00672* | -0.00458 | 0.00958* | 0.000525 | 0.0148 | 0.00162 |
| self-sufficiency | (0.99) | (1.37) | (2.41) | (-0.88) | (2.14) | (0.07) | (1.51) | (0.18) |
| election | -0.0409** | -0.0252 | 0.0481 | 0.00477 | 0.0199 | 0.0177 | -0.0380 | -0.0220 |
| | (-2.22) | (-1.71) | (0.74) | (0.14) | (0.36) | (1.00) | (-0.78) | (-0.62) |
| L.monopoly | | 0.00215 | 0.0141 | 0.0655 | 0.0521 | -0.117 | -0.371 | 0.0537 |
| | | (0.32) | (0.30) | (0.77) | (0.31) | (-0.44) | (-1.33) | (0.18) |
| Constant | -0.310 | -0.228 | 1.112 | 4.110 | 1.679 | 5.410 | 24.03* | -8.603 |
| | (-0.16) | (-0.03) | (0.11) | (0.77) | (0.54) | (0.86) | (2.01) | (-1.13) |
| Observations | 337 | 90 | 122 | 211 | 206 | 157 | 148 | 144 |
| $R^2$ | 0.081 | 0.244 | 0.202 | 0.174 | 0.069 | 0.189 | 0.159 | 0.154 |

NRA is crop-specific.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A more complete model is shown in table 10. Unfortunately, not all variables can be included, as some do not have enough variability or data points. This problem is present for the variables party orientation, political system, and the post Uruguay dummy. Because the crop regressions already have a small number of observations, including these variables reduces the number of observations even more. In the next section a mixed model is presented in which all data is used in one regression, increasing the number of observations and making it possible to also include the constitutional and political variables. As for table 10, democracy, food self sufficiency, an election dummy and monopoly power are added to the model. Monopoly power can not be calculated for aggregate agriculture, and is thus left out in the aggregate regression. As shown, hardly any of the effects are significant. Most notably, monopoly power and the election dummy are not significant for any of the crops. Election is significant and negative for aggregate agriculture. meaning that agriculture is taxed more when there are elections in that year. This could mean that the population in agriculture is not the most important voter base. Democracy is significant for rice and sugar only, with a positive effect. Meaning that the higher democracy is, the more rice and sugar are subsidized. This is as expected from theory on price distortions. Overall, the models have little explanatory power.

#### 5.2.2 Model 2.2: Mixed model

In addition to a fixed effects model for each crop, a mixed or multilevel model is estimated for all crops. Specifically, a random intercepts model is estimated for different model specifications. For some variables, a random slope model could theoretically be more appropriate. For instance, we expect the effect of monopoly power to vary across crops. Specifically, we expect a negative effect for exportables. Secondly, a country's aim for food self-sufficiency can have an anti-trade bias as a result, and thus for exportables the effect would be negative, and for import-competing crops it would be positive. However, testing random intercepts and random slope models with a likelihood ratio test does not show random slope models to be more appropriate. The results of the random intercept model are given in table 11. The number of observations greatly increases due to this hierarchical structuring of the data, however, the panel structure is lost when using this method. Five different models were estimated. As compared to the fixed effects models, there are many more significant results. In the most simple model, GDP per capita, population, rural population share, inflation, debt, and democracy are all significant. The signs are as expected for most of the variables. For rural population share we we expect a negative effect, and this is supported by model 1 and 4. For public debt, a significant negative effect is found for models 1 and 4. As for the political variables, election is never significant. Party orientation however, is significant and as expected. The variable is specified as follows: the executive party can either be right-wing (-1), centre (0), or left-wing (1). As the party orientation moves from right to left, NRA becomes lower and thus, agriculture is taxed more. Traditionally, right-wing parties support agriculture more than left-wing parties, which is reflected in the results. Food self-sufficiency does not have a significant effect in this model. This can be explained by the different expectations we have for different crops. The net effect may be smoothed out when compiling all crops. Monopoly power has a significant and negative effect, meaning that when a country has more monopoly power in the production of a commodity, it wil tax the commodity more. This is as expected for export-crops, as taxing these shifts the burden of the tax to foreign importers. In the fixed effects model, monopoly power gave no significant results, which can be due to the low number of observations. The standard deviation of the intercept is significant meaning that for each crop, the regression line has a different intercept. The size of the standard deviation is not very large compared to the intercept, indicating that the variation in intercept across crops is not very large.

**Table 11:** Model 2.2 Mixed models: determinants of NRA

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | nra | nra | nra | nra | nra |
| L.log gdp | -0.0214 | -0.0402 | -0.00121 | -0.0153 | -0.0448 |
| per cap | (-0.98) | (-1.34) | (-0.02) | (-0.69) | (-1.49) |
| L.gdp | 0.00185 | 0.00188 | 0.00744* | 0.00145 | 0.00129 |
| growth | (0.83) | (0.54) | (1.68) | (0.65) | (0.37) |
| L.log | 0.0806*** | -0.0481* | -0.0281 | 0.0848*** | -0.0513* |
| population | (5.57) | (-1.82) | (-0.69) | (5.84) | (-1.94) |
| L.share | -0.419*** | 0.00814 | 0.280 | -0.442*** | -0.118 |
| rural pop | (-2.88) | (0.03) | (0.48) | (-3.04) | (-0.45) |
| L.log | -0.0456*** | -0.105*** | -0.0779*** | -0.0468*** | -0.109*** |
| inflation | (-4.69) | (-5.95) | (-3.32) | (-4.81) | (-6.15) |
| L.log debt | -0.0304* | -0.00436 | 0.0404 | -0.0280* | -0.00483 |
| | (-1.83) | (-0.16) | (0.89) | (-1.68) | (-0.18) |
| L.democracy | 0.00490** | 0.0225*** | 0.0213*** | 0.00439** | 0.0217*** |
| | (2.37) | (6.10) | (3.92) | (2.12) | (5.86) |
| election | | -0.0533 | -0.0397 | | -0.0572 |
| | | (-1.24) | (-0.73) | | (-1.34) |
| party | | -0.132*** | -0.142*** | | -0.135*** |
| orientation | | (-5.46) | (-4.63) | | (-5.62) |
| food | | | 0.00348 | | |
| self-sufficiency | | | (1.43) | | |
| monopoly | | | | -0.00882*** | -0.00828** |
| power | | | | (-2.87) | (-2.05) |
| Constant | -0.765** | 1.276* | 0.169 | -0.847** | 1.475** |
| | (-1.99) | (1.91) | (0.13) | (-2.21) | (2.19) |
| sd(crop) | 0.226*** | 0.272*** | 0.284*** | 0.207*** | 0.254*** |
| | (-5.45) | (-4.71) | (-4.48) | (-5.73) | (-4.90) |
| sd(residual) | 0.461*** | 0.354*** | 0.338*** | 0.461*** | 0.353*** |
| | (-46.75) | (-33.88) | (-27.53) | (-46.65) | (-33.98) |
| Observations | 1829 | 539 | 330 | 1817 | 539 |

$t$ statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

## 5.3 Part III: Addressing endogeneity

### 5.3.1 Model 3.1: IV-regression

In the first IV-model, NRA is instrumented with population, share of rural population and agricultural value added. As discussed in the methodology section, the validity of the instruments is questioned. The value added of agricultural production can theoretically be related to the yield and therefore one of the two conditions of a valid instrument may not be met. Alternatively, a second IV-model is estimated below. NRA is taken in t-1 and the instruments are taken in t-2, because it is expected that pricing policy affects yield with a lag, and on the other hand NRA is affected by the instruments with a lag. The independent variables taken are land, capital, rainfall, temperature, and fertilizer.[6] The first and second stage regressions are reported in table 12. The estimation is done with robust standard errors and with GMM estimation, to increase efficiency. The test for overidentifying restrictions is given by the J-statistic, the p-value is given below. As can be seen, for the aggregate estimation and for coffee, rice, wheat, and sugar the null-hypothesis that all overidentifying restrictions are zero can be rejected. This means that for these models the instruments are invalid. For cocoa, cotton, and maize the instruments are valid. The significant effect of NRA on cocoa and cotton is positive, similarly to the fixed effects models. However, the magnitude of the effect is much larger than in the previous regressions. For cocoa, the increase in yield due to a one unit increase in NRA is 87,1%, ceteris paribus. For cotton, a 1 unit increase in NRA, indicating a 100% increase in prices producers receive for their product, leads to an increase in yield of 104%, keeping the other variables constant. For maize, the sign of NRA is negative and the magnitude is large. An increase of NRA of 1 unit, leads to a decrease in yield of 99.2%. As for the signs of the other variables, these are similar to the fixed effects regressions.

In the second IV-model (table 13) the agricultural value added as a percentage of GDP is left out as an instrument, because as mentioned, there may be some relation with the independent variable yield. Instrumenting with only population and rural population gives a better instrument for NRA. Now, the instruments are valid for cocoa, cotton, maize, sugar, and wheat. For cocoa and cotton, the magnitude of the effect of NRA is significant and even larger, both around 140%. For maize, sugar and wheat the effects are significant and negative. For maize, the decrease in yield due to a one unit NRA increase is 57.2%, keeping all others constant. For sugar, the magnitude of the effect is -122.3%, and for wheat the magnitude is -104.3%. Interestingly, the effects of NRA are positive for the traditional cash crops (exportables) and negative for the import-competing food crops. This indicates that subsidization or taxation does not have a uniform effect on the productivity of crops, but large differences exist between different types of crops.

Other possible instruments that were tried were not valid or dit not have sufficient data. For instance, political factors like elections, party orientation, constitutional system, and the post Uruguay Trade Round Dummy are omitted because of insufficient datapoints and variability. A possible remedy for this is using hierarchical structuring of the data, similar to the data structuring used for the multilevel model. A country-crop grouping variable can be used as panel variable and the IV-regression can be run on the entire dataset (excluding aggregate production). However, different types of instruments were tested and they were never valid (refer to appendix G).

[6] A model with a lagged dependent was also estimated, the only difference is the magnitude of the effects, the significance and sign of the effects remain the same. See appendix F

**Table 12:** Model 3.1.1 IV-model with instruments population, rural population and agricultural value added

| | (1)<br>Aggregate | (2)<br>Cocoa | (3)<br>Coffee | (4)<br>Cotton | (5)<br>Maize | (6)<br>Rice | (7)<br>Sugar | (8)<br>Wheat |
|---|---|---|---|---|---|---|---|---|
| L2.log | -0.0554 | 0.144 | 0.408*** | 0.328*** | -0.882* | -0.823*** | -1.096** | -0.823** |
| population | (-0.61) | (0.59) | (2.80) | (2.95) | (-1.88) | (-3.31) | (-2.46) | (-2.04) |
| L2.share | -0.534 | -1.983 | 1.769** | -0.111 | -1.875* | -4.094*** | -7.579*** | -0.358 |
| rural pop | (-1.39) | (-1.12) | (2.60) | (-0.23) | (-1.75) | (-3.16) | (-4.90) | (-0.30) |
| L2.ag | -0.004** | -0.010* | -0.013*** | -0.009*** | 0.00878 | 0.00192 | -0.00452 | 0.00869 |
| value | (-2.51) | (-1.81) | (-5.32) | (-3.77) | (1.37) | (0.41) | (-0.65) | (0.96) |
| log land | -0.213*** | 0.214 | -0.621** | -0.117 | -0.737 | -0.187 | 0.242 | -1.439*** |
| per cap | (-3.26) | (0.63) | (-2.53) | (-1.23) | (-1.43) | (-0.42) | (0.36) | (-2.99) |
| log capital | -0.114 | -0.864* | 0.211 | 0.251* | -0.277 | -1.287** | -0.580 | 0.0198 |
| per cap | (-1.48) | (-1.89) | (0.90) | (1.96) | (-1.01) | (-2.35) | (-1.14) | (0.07) |
| rainfall | 0.000696 | 0.0128 | 0.000670 | -0.00173 | 0.00141 | 0.00118 | 0.0117 | 0.0577** |
| | (0.13) | (1.38) | (0.07) | (-0.20) | (0.08) | (0.07) | (0.73) | (1.98) |
| temperature | 0.0478** | -0.0679 | -0.0312 | 0.00325 | 0.137* | 0.00456 | 0.278*** | -0.0108 |
| | (2.03) | (-0.87) | (-0.55) | (0.09) | (1.74) | (0.05) | (2.85) | (-0.07) |
| fertilizer | 0.0411 | -0.00450 | 0.0536 | 0.00094 | 0.007*** | 0.003** | 0.006*** | 0.005*** |
| | (0.55) | (-0.48) | (1.10) | (1.60) | (2.79) | (2.43) | (4.55) | (3.15) |
| N | 538 | 130 | 171 | 409 | 318 | 236 | 203 | 202 |

First stage regression. NRA is crop-specific. Fertilizer is consumption for cocoa, dummy for others.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

| | (1)<br>Aggregate | (2)<br>Cocoa | (3)<br>Coffee | (4)<br>Cotton | (5)<br>Maize | (6)<br>Rice | (7)<br>Sugar | (8)<br>Wheat |
|---|---|---|---|---|---|---|---|---|
| L.nra | 1.811*** | 0.871*** | 0.535** | 1.044*** | -0.992*** | -0.581** | -1.458*** | -0.449* |
| | (3.43) | (4.01) | (2.24) | (3.72) | (-3.05) | (-2.49) | (-4.47) | (-1.70) |
| log land | 0.0822 | -0.436** | 0.345 | -0.00265 | -1.210*** | -0.474** | -2.357*** | -0.896** |
| per cap | (0.49) | (-2.03) | (1.22) | (-0.01) | (-5.40) | (-2.17) | (-4.71) | (-2.46) |
| log capital | 0.252* | -1.0920*** | 0.399 | 0.137 | 0.932*** | -0.789** | 1.561** | 0.0493 |
| per cap | (1.73) | (-3.42) | (1.27) | (0.81) | (3.24) | (-2.20) | (2.57) | (0.19) |
| rainfall | 0.0104 | -0.00575 | -0.00321 | 0.0140 | 0.0243 | 0.00236 | 0.0277 | 0.08010** |
| | (1.03) | (-0.50) | (-0.25) | (0.90) | (1.20) | (0.17) | (1.12) | (2.32) |
| temperature | -0.0889* | 0.0865 | 0.0246 | -0.0368 | -0.0445 | -0.0108 | 0.492*** | 0.0908 |
| | (-1.84) | (1.32) | (0.33) | (-0.66) | (-0.45) | (-0.15) | (2.98) | (0.74) |
| fertilizer | 0.0403 | 0.00709 | 0.0181 | -0.355** | 0.225* | 0.0299 | -0.365** | -0.104 |
| | (0.30) | (0.91) | (0.17) | (-2.21) | (1.84) | (0.21) | (-2.10) | (-0.88) |
| N | 538 | 130 | 171 | 409 | 318 | 236 | 203 | 202 |
| j | 14.23 | 2.775 | 10.96 | 0.935 | 2.346 | 8.626 | 7.171 | 7.571 |
| jp | 0.000811 | 0.250 | 0.00418 | 0.627 | 0.309 | 0.0134 | 0.0277 | 0.0227 |

Second stage regression. NRA is crop-specific. Fertilizer is consumption for cocoa, dummy for others.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

**Table 13:** Model 3.1.2 IV-model with instruments population and rural population

| | (1) Aggregate | (2) Cocoa | (3) Coffee | (4) Cotton | (5) Maize | (6) Rice | (7) Sugar | (8) Wheat |
|---|---|---|---|---|---|---|---|---|
| L2.log | -0.0826 | 0.174 | 0.491*** | 0.456*** | -1.111*** | -1.066*** | -0.967** | -1.063** |
| population | (-0.95) | (0.73) | (2.98) | (4.02) | (-3.25) | (-4.93) | (-2.23) | (-2.49) |
| L2.share | -0.835** | -1.893 | 1.501** | 0.511 | -1.539* | -4.828*** | -6.035*** | -1.629 |
| rural pop | (-2.30) | (-1.29) | (2.03) | (1.06) | (-1.72) | (-4.60) | (-4.46) | (-1.31) |
| log land | -0.256*** | 0.00394 | -0.731** | -0.141 | -1.201*** | -0.634* | -0.437 | -2.024*** |
| per cap | (-4.35) | (0.01) | (-2.59) | (-1.58) | (-3.05) | (-1.97) | (-0.65) | (-4.03) |
| log | -0.108 | -0.873** | 0.264 | 0.221* | -0.196 | -1.249*** | -0.336 | 0.148 |
| capital per cap | (-1.42) | (-2.25) | (1.09) | (1.81) | (-0.69) | (-2.65) | (-0.68) | (0.48) |
| rainfall | -0.000795 | 0.0109 | 0.00641 | -0.00340 | -0.00202 | 0.00281 | 0.0179 | 0.0602** |
| | (-0.15) | (1.19) | (0.65) | (-0.38) | (-0.12) | (0.17) | (1.09) | (2.20) |
| temperature | 0.0513** | -0.0945 | -0.00689 | 0.00653 | 0.168** | 0.0188 | 0.330*** | 0.0125 |
| | (2.25) | (-1.22) | (-0.11) | (0.19) | (2.12) | (0.25) | (3.42) | (0.09) |
| fertilizer | 0.0514 | -0.000380 | 0.111** | 0.162 | 0.0119 | -0.0647 | -0.249** | -0.0282 |
| | (0.70) | (-0.04) | (2.15) | (1.34) | (0.13) | (-0.38) | (-2.55) | (-0.18) |
| N | 574 | 139 | 187 | 444 | 348 | 272 | 223 | 218 |

First stage regression. NRA is crop-specific. Fertilizer is consumption for cocoa, dummy for others.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

| | (1) Aggregate | (2) Cocoa | (3) Coffee | (4) Cotton | (5) Maize | (6) Rice | (7) Sugar | (8) Wheat |
|---|---|---|---|---|---|---|---|---|
| L.nra | 1.644*** | 1.434*** | 1.027* | 1.408*** | -0.572*** | -0.319** | -1.223*** | -1.043** |
| | (3.50) | (3.81) | (1.81) | (3.64) | (-2.64) | (-2.25) | (-4.35) | (-2.32) |
| log land | 0.171 | 0.266 | 1.045* | 0.177 | -1.202*** | -0.755*** | -2.449*** | -1.658** |
| per cap | (1.00) | (0.66) | (1.96) | (0.73) | (-7.19) | (-4.97) | (-4.96) | (-2.43) |
| log capital | 0.0918 | -0.919** | 0.460 | -0.0197 | 0.917*** | -0.453* | 1.591*** | 0.311 |
| per cap | (0.80) | (-2.29) | (1.29) | (-0.09) | (3.96) | (-1.91) | (2.69) | (0.81) |
| rainfall | 0.0111 | -0.0107 | -0.00780 | 0.00944 | 0.0238 | -0.000361 | 0.0167 | 0.103** |
| | (1.29) | (-0.74) | (-0.50) | (0.53) | (1.53) | (-0.03) | (0.71) | (2.41) |
| temperature | -0.0901** | 0.138 | 0.0477 | -0.0676 | -0.115 | 0.00262 | 0.397*** | 0.0775 |
| | (-2.08) | (1.43) | (0.54) | (-1.05) | (-1.50) | (0.05) | (2.74) | (0.47) |
| fertilizer | 0.0302 | 0.00105 | -0.150 | -0.393** | 0.169 | 0.0161 | -0.216 | -0.0598 |
| | (0.26) | (0.09) | (-1.22) | (-1.99) | (1.57) | (0.12) | (-1.46) | (-0.34) |
| N | 574 | 139 | 187 | 444 | 348 | 272 | 223 | 218 |
| j | 24.71 | 0.00762 | 11.22 | 0.0534 | 0.0612 | 10.19 | 0.301 | 1.418 |
| jp | 0.000000665 | 0.930 | 0.000810 | 0.817 | 0.805 | 0.00141 | 0.583 | 0.234 |

Second stage regression. NRA is crop-specific. Fertilizer is consumption for cocoa, dummy for others.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

### 5.3.2 Model 3.2: Simultaneous equation model

A last estimation method used to take into account endogeneity is by estimating both equations as a simultaneous equation model. An advantage of this estimation technique is that it does not matter if variables are endogenous. A disadvantage is that is is not possible to estimate with a panel structure, and thus the panel structure is lost. A simple model is shown in tables 15 (aggregate, cocoa, coffee, and cotton) and 16 (maize, rice, sugar, and wheat). The results are interesting and in many ways similar to the fixed effects and IV-regression model. NRA has a significant positive effect for cocoa and coffee, and a significant negative effect for sugar and wheat. The magnitudes of the effects for cocoa and cotton are a 40.4% and 74.2% increase in yield respectively, due to a one unit increase in NRA, ceteris paribus. The effects for sugar and wheat are -15% and -20% respectively. To compare the effect sizes of the IV-regression with the SEM-model, table 14 reports the effect of a standard deviation increase in NRA for both models.

**Table 14:** Effect on yield of a standard deviation increase in the NRA for the IV and SEM model

| Variable | Std. dev. | IV | Effect (%) | SEM | Effect (%) |
|---|---|---|---|---|---|
| Aggregate | 0.245 | invalid | | no effect | |
| Cocoa | 0.267 | + | 38.3 | + | 10.8 |
| Coffee | 0.273 | invalid | | no effect | |
| Cotton | 0.339 | + | 47.7 | + | 25.2 |
| Maize | 0.650 | - | -37.2 | no effect | |
| Rice | 0.445 | invalid | | no effect | |
| Sugar | 0.546 | - | -66.8 | - | -8.2 |
| Wheat | 0.506 | - | -52.8 | - | -10.1 |

As for the determinants of NRA, many significant effects are found. First off, we find a significant negative effect for GDP per capita on coffee, maize, and rice, and a positive effect for cotton. The effect of cotton is as expected, but the other effects are not supported by theory. GDP growth has a significant positive effect for aggregate agriculture and the three export crops, but a negative effect for rice and sugar. Apparently, both taxation and subsidization are decreased as the economy grows. Population is significant and positive for aggregate agriculture, cocoa, coffee, maize and wheat, but negative for cotton. For maize, wheat and cotton this is as expected, as a larger population is hypothesized to lead to an anti-trade bias (taxing exports and imports). The rural population share is significant and negative for most variables, as expected. Inflation is negative for most variable, which coincides with previous findings. The effect of debt is negative for some and positive for others, not giving a clear result.

**Table 15:** Model 3.2.1a Simultaneous equation model: simple

| | (1) Aggregate log yield | nra | (2) Cocoa log yield | nra | (3) Coffee log yield | nra | (4) Cotton log yield | nra |
|---|---|---|---|---|---|---|---|---|
| L.nra | 0.0691 | | 0.404*** | | 0.105 | | 0.742*** | |
| | (1.64) | | (3.67) | | (0.90) | | (7.09) | |
| log land per cap | -0.0888*** | | -0.326*** | | 0.0301 | | -0.155 | |
| | (-3.55) | | (-2.79) | | (0.27) | | (-1.51) | |
| log capital per cap | -0.0327 | | 0.520*** | | -0.497*** | | 0.331*** | |
| | (-1.20) | | (2.71) | | (-5.79) | | (3.81) | |
| rainfall | 0.00769*** | | -0.00959 | | -0.0334*** | | 0.0108 | |
| | (2.77) | | (-0.54) | | (-4.09) | | (0.91) | |
| temperature | -0.00884* | | 0.0182 | | -0.142*** | | 0.0423** | |
| | (-1.81) | | (0.44) | | (-7.13) | | (2.05) | |
| fertilizer | 0.0971** | | 0.0143 | | -0.0818 | | 0.887*** | |
| | (2.38) | | (1.30) | | (-1.04) | | (4.40) | |
| L.log gdp per cap | | 0.00230 | | 0.0643 | | -0.177** | | 0.111* |
| | | (0.08) | | (1.22) | | (-2.31) | | (1.87) |
| L.gdpgrowth | | 0.00426* | | 0.00704* | | 0.00993** | | 0.0118*** |
| | | (1.72) | | (1.84) | | (2.54) | | (3.47) |
| L.log population | | 0.0858*** | | 0.120*** | | 0.203*** | | -0.0964*** |
| | | (6.02) | | (4.51) | | (4.07) | | (-5.18) |
| L.share rural pop | | -0.418** | | -0.478 | | -1.362*** | | 0.217 |
| | | (-2.27) | | (-1.53) | | (-3.32) | | (0.70) |
| L.log inflation | | -0.0678*** | | 0.00243 | | -0.0233 | | -0.0732*** |
| | | (-7.25) | | (0.11) | | (-1.29) | | (-5.74) |
| L.log debt | | -0.0399** | | 0.112*** | | -0.168*** | | 0.0947*** |
| | | (-2.02) | | (3.79) | | (-3.86) | | (2.93) |
| Constant | 4.079*** | -0.957** | 9.090*** | -2.964*** | 8.788*** | -0.866 | 9.167*** | 0.100 |
| | (17.56) | (-2.34) | (6.87) | (-4.09) | (6.19) | (-0.59) | (7.67) | (0.14) |
| Observations | 364 | | 127 | | 136 | | 242 | |
| $R^2$ | 0.169 | 0.232 | 0.363 | 0.352 | 0.641 | 0.322 | 0.386 | 0.240 |

NRA is crop-specific.
$^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

**Table 16:** Model 3.2.1b Simultaneous equation model: simple

| | (1) | | (2) | | (3) | | (4) | |
|---|---|---|---|---|---|---|---|---|
| | Maize | | Rice | | Sugar | | Wheat | |
| | log yield | nra | log yield | nra | log yield | nra | log yield | nra |
| L.nra | -0.0152 | | 0.0620 | | -0.150* | | -0.200*** | |
| | (-0.44) | | (1.28) | | (-1.88) | | (-2.71) | |
| log land per cap | -0.355*** | | -0.966*** | | 0.118* | | -0.738*** | |
| | (-7.41) | | (-14.24) | | (1.76) | | (-5.91) | |
| log capital per cap | -0.000182 | | 0.395*** | | -0.222*** | | 0.462** | |
| | (-0.00) | | (8.80) | | (-2.68) | | (2.54) | |
| rainfall | -0.0161*** | | 0.00224 | | -0.0280*** | | 0.0459*** | |
| | (-2.98) | | (0.31) | | (-3.40) | | (3.32) | |
| temperature | 0.0285 | | 0.0588*** | | 0.0223 | | 0.0472** | |
| | (1.63) | | (5.29) | | (1.30) | | (2.49) | |
| fertilizer | 0.401*** | | -0.138 | | 0.432*** | | -0.0182 | |
| | (5.28) | | (-1.08) | | (4.00) | | (-0.13) | |
| L.log gdp per cap | | -0.148* | | -0.184** | | 0.0502 | | -0.0366 |
| | | (-1.75) | | (-2.10) | | (0.52) | | (-0.30) |
| L.gdpgrowth | | -0.00892 | | -0.0136* | | -0.0187* | | -0.00297 |
| | | (-1.02) | | (-1.94) | | (-1.66) | | (-0.25) |
| L.log population | | 0.481*** | | 0.0408 | | 0.0908 | | 0.188** |
| | | (8.41) | | (1.16) | | (0.85) | | (2.02) |
| L.share rural pop | | -1.207* | | -0.932** | | -0.462 | | -1.000 |
| | | (-1.95) | | (-2.02) | | (-0.63) | | (-1.13) |
| L.log inflation | | 0.0174 | | -0.0828*** | | -0.104** | | -0.0596 |
| | | (0.47) | | (-3.05) | | (-2.54) | | (-1.38) |
| L.log debt | | -0.287*** | | -0.0949** | | 0.0756 | | -0.112 |
| | | (-4.41) | | (-2.24) | | (0.84) | | (-1.09) |
| Constant | 6.160*** | -5.003*** | 3.233*** | 1.830* | 12.45*** | -1.414 | 5.671*** | -1.508 |
| | (13.78) | (-3.32) | (4.71) | (1.71) | (22.15) | (-0.58) | (8.30) | (-0.58) |
| Observations | 239 | | 189 | | 146 | | 140 | |
| $R^2$ | 0.593 | 0.277 | 0.816 | 0.086 | 0.332 | 0.134 | 0.419 | 0.147 |

NRA is crop-specific.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

# 6 Discussion

## 6.1 Discussion of results

The results of the analysis give multiple insights. First off, we find that agricultural price distortions do affect the productivity of some crops. This means that farmers base their production and investment decisions on prices. We find that the productivity of the cash crops cocoa and cotton are affected negatively by taxation, indicating that farmers become less efficient in production of cocoa and cotton when the commodities are taxed. Why farmers are less efficient is not analyzed, but as explained in the theoretical section, there may be several reasons. If production is taxed, farmers may be less stimulated to invest. A reason for this could be that the price is too low to provide for sufficient income or become profitable. The negative effect can also be explained by less investment-induced productivity gains. As Rizov et al. (2013) hypothesizes, subsidies can stimulate credit access. In the case of taxation, credit access could thus be less. The cost of borrowing may also be higher. Additionally, risk aversion may increase also leading to less investment in these crops.

Looked at from a reversed perspective, a reduction of taxation of these crops can thus lead to large increases in productivity. In light of the need for more efficient land usage due to environmental challenges, taxation reduction may be a viable strategy to intensify the production of these crops. Moreover, the idea that perennial crops respond less to price distortions is not supported by the results of this analysis. Cocoa, a perennial crop, is affected strongly by taxation. Cotton, a perennial crop that is often cultivated as an annual, is also affected by taxation. Though results differ per model, the effect of price distortions is not necessarily larger for cotton. Governments taxing perennial crops instead of annuals because the effects are thought to be less strong, due to already incurred start-up costs, may be affecting the productivity of these crops significantly.

On the other hand we have seen the opposite effect for some import-competing crops. This means that these often subsidized crops have not benefitted from subsidies. This can be related to the other side of the theory on the relationship between price distortions and productivity. Subsidization can lead to a lack of effort from farmers leading to a decrease in productivity. A subsidy can be a form of insurance of farmers, leading to moral hazard problems and thus riskier investments which may lead to productivity reductions. In addition, reallocation of resources may be reduced causing allocative inefficiency. These effects can all explain the negative effect of subsidizations of import-competing crops. The results are interesting in the light of the ongoing debate on input subsidies in countries like Malawi, Tanzania, and Zambia. Input subsidies are being implemented in these countries to increase productivity, but as this research shows, subsidization does not necessarily have this effect. From the reverse perspective, the results indicates that a decrease of subsidization (or increase of taxation) may have positive effects on the productivity of import-competing food crops.

Why are the effects of price distortions on import-competing crops like rice, sugar, and wheat diametrically opposed to the effects on export crops like cocoa and cotton? Though this research does not focus on investigating farmer decision-making, a possible explanation is that farmers produce food crops (import-competing crops) up to subsistence level, and generally do not market their food crops. Thus incentives to produce these crops more efficiently because of subsidies are limited. Instead, farmers may switch to other marketable cash crops. Another possible explanation is that the effect of price distortions is not linear. In this research, the effect of NRA on yield is modelled as a linear relationship. However, it could be that the effect of NRA on productivity is different for different levels of NRA. In a linear form, the effect of a subsidy would be the opposite of the effect of taxation (as taxation is identical to a negative subsidy). But, there may be differences in how producers adjust investments depending on whether the crop is being taxed or subsidized. For instance, a quadratic relationship where a negative NRA has a positive yield-effect and a positive NRA has a negative yield-effect could be possible. The closer NRA is to zero, the smaller the yield-effect would be. More research on the possibility of a non-linear relationship between NRA and yield can thus be useful.

This research also focussed on the determinants of NRA. Previous research has shown that there are many factors that can explain taxation and subsidization of agriculture, including economic reasons as well as political reasons. From the results of this research, previous findings and theory are largely supported. These insights into the conditions under which agriculture is taxed can provide information on where policies are likely to be changed in order to increase productivity and land-use efficiency. In countries where macroeconomic conditions are unstable, it is less likely for taxation to be reduced. High public debt and inflation lead to more taxation of agriculture. It should also be taken into consideration that NRA is highly political, and thus unlikely to be changed under certain political conditions. For instance, we have seen that undemocratic governments tax agriculture more, as do governments ruled by leftist parties.

## 6.2 Limitations and further research

There are limitations to this research relating to the quality of the data, measurement issues, sample selection, and the endogeneity problem. One of the major challenges of macro-economic research on African countries is finding reliable and valid data. This has been a recurring issue for many of the control variables used in this analysis. Data on governance, market institutions, informal sector, food self sufficiency, input use are all difficult to find. Often, the quality of the data that can be found is poor, with many missing observations. This greatly reduces the efficiency for some of the models estimated, which in turn reduces the difficulty of finding effects. This could explain why especially in the fixed effects model, where the number of observations is low, the effect of NRA is insignificant for many of the crops.

As Anderson and Masters (2009) explain in their report, there are multiple issues regarding measuring NRA. For instance, distortions to input markets are limited to fertilizer, electrical or diesel power, pesticide, and credit. Water subsidies and distortions to labor and land markets are excluded because of lack of data. Moreover, classifying a commodity into import-competing, exportable or nontradable is not always straightforward. Though Anderson and Masters (2009) calculate NRA thoroughly, the construct may never give a complete picture of agricultural distortions. Fertilizer use data has been taken from FAO and there are also some limitations to this data. FAO presses for caution when comparing fertilizer use across countries due to measurement problems. This research has attempted to overcome this issue by constructing a fertilizer dummy for high and low fertilizer use, which in turn is rather rough. Another problem with fertilizer and inputs in general is that they are measured on an aggregate level whereas estimations are also done on crop level. Thus, some bias can exist when using for instance aggregate land data on a crop which uses relatively low or high amounts of land. Similarly, climate data is averaged out on country level, whereas regional climate data is much more precise. Though not much can be done about these measurement problems because better data is simply not available, it should be taken into account when interpreting the results.

The research is based on data for a sample of 22 African countries. These countries were selected on basis of data availability of agricultural price distortions from Anderson and Nelgen (2013). As stated in their report, the selection covers the largest economies in Sub-Saharan Africa, covering over 90% of agricultural value added, farm households, total population, and total GDP. This may cause a selection bias, by only including the largest economies, the results may not be applicable to smaller African economies. Moreover, a selection of crops was made on the basis relative importance and data availability. For some crops the sample is quite small, cocoa for instance has a sample size of five countries. Reliability of the results could be increased with a larger sample size.

As explained in this research, NRA is a policy variable and can not be considered exogenous. It is likely that there are factors that affect both NRA and yield, leading to a possible bias in the results of estimation methods where endogeneity is not taken into account. Efforts were made to control for endogeneity through an Instrumental Variable-regression but finding suitable instruments is difficult, mostly because many of the factors influencing NRA are endogenous themselves. The instruments chosen (population, rural population and agricultural value added) were not valid for all crop-estimations. Exploring other instruments can be useful in future research to control for endogeneity more effectively.

More opportunities for further research arise from the discussion of the results. First off, this research is limited to looking at the overall NRA for specific crops, and does not distinguish between different types of policy measures. Input subsidies could have a different effect than output subsidies. As also explained in the theoretical framework, this research is focussed on ad valorem taxes. How farmers respond to lump sum taxes could be an interesting question for future research. Moreover, whether taxes and subsidies can be treated as opposites, (having opposites effects) is also a subject for further research. Secondly, the hypothesis that farmers switch from food to cash crops when food crops are subsidized more (or cash crops are taxed less) can be tested. More generally, detailed examination of changes in farmer decision-making due to price distortions can provide more insight into the different effects found for different crops. This may also be related to differences in investment and input use. Cotton is relatively prone to pests and therefore input costs are high, a reduction in taxes may therefore lead to relatively more switching or investment in cotton production. More research into input use of different crops can possibly provide more explanations for the effects found in this research.

## 6.3 Policy implications

What do these results mean in terms of policy implications? From this research we cannot conclude that taxation of cash crops and subsidization of import-competing crops are necessarily bad policies. Governments of course need to raise revenue and it may be that other sources of revenue are not sufficient. A more detailed examination should provide insight into whether price distortions are truly necessary and whether they are harmful or helpful to agricultural productivity. However, when looking at interventions to increase agricultural productivity, it is useful to take into account the policy environment instead of assuming a policy vacuum. Implementing new technologies may not be as effective when the crop is heavily taxed or subsidized. In addition, price distortions, or the reduction thereof, can be a viable strategy for improving agricultural productivity. Increasing productivity is important for various reasons. Food security is expected to become one of the biggest challenges in the coming decades. The environmental impact of agriculture and the need for poverty reduction call for intensification of African agriculture instead of extensification.

# 7 Conclusion

This aim of this research was to analyze the determinants, as well as the effect of agricultural price distortions in African countries. The main research question that has been examined in this thesis is "How do agricultural price distortions affect land use efficiency in Africa, and what are the determinants of price distortions?". The background to this question is the widespread implementation of price distorting policies in agriculture, that has been persistent regardless of international efforts of market liberalization. From the discussed theory and research on price distortions we can derive the expectation that farmers base their investment decisions partly on prices, therefore agricultural productivity may have been affected by these price distortions.

We specifically looked at the difference in the effect of price distortions on import-competing and exportable crops. One reason for this is that import-competing crops are mostly subsidized, whereas exportable crops are heavily taxed. Secondly, there may be a difference in how farmers respond to price changes for the import-competing food crops and the exportable cash crops. This was analyzed with four different estimation techniques: fixed effects estimation, multilevel modelling, IV-estimation and structural equation modelling. The results support that agricultural productivity is indeed affected by price distortions and that this effect differs greatly for different crops. First off, for the export crops cocoa and cotton we find strong positive effects of the nominal rate of assistance, in other words, taxation limits agricultural productivity of these crops. On the other hand, we find a strong negative effect of NRA for the import-competing crops maize, sugar, and wheat, in other words, subsidization of these crops has limited agricultural productivity.

The third aim of this research was to see whether there are contextual variables that influence the effect of NRA on agricultural productivity. Though limited by data availability, we found evidence that for some crops, the more democratic a country is, the smaller the effect of NRA on productivity is. We expected the opposite, as democratic governments have higher governance and thus policy implementation is more effective. It was also tested whether policy variability is an intermediating variable, we find that the more policy varies through the years, the smaller the effect of NRA on productivity is. The third intermediating variable examined is market development, using variables for infrastructure and information access. The results are unclear and often insignificant.

The same econometric techniques were used to examine the determinants of agricultural pricing policy in African countries. From the literature we have seen that there are a range of economic, financial and political variables that may have some influence on these pricing policies. This is supported by the data. Though the fixed effects models were not very conclusive, from the multilevel model we can conclude that macroeconomic stability i.e. low inflation and low public debt leads to less taxation of agriculture. The larger the share of the rural population in agriculture, the more countries tax agriculture. A theoretical explanation for this is that the political resistance to subsidization of agriculture decreases as agriculture becomes less important. Democracy also has an effect on NRA, we find that the more democratic a country is, the more it subsidizes agriculture. This is generally seen, democratic governments tend to redistribute wealth from the rich to the poor. Political preferences are also of importance, left-wing government tax agriculture more than right-wing governments. Whether there are elections does not seem to matter for pricing policy.

The main conclusion of this research is that taxes and subsidies do matter. Cash crop farmers that are taxed have less incentive to make productivity-increasing investments. Furthermore, subsidies of food crops also do not stimulate farmers to invest in productivity. This means that changing pricing policy can be used as an instrument to increase producitivity in African countries.

# 8 Bibliography


Abbas, S. M., Belhocine, N., and ElGanainy, A. A. (2010). A historical public debt database. *IMF Working Papers*, pages 1–26.

Anderson, K. (2009). *Distortions to Agricultural Incentives.* A Global Perspective 1955-2007. World Bank, Washington, DC.

Anderson, K. and Masters, W. A. (2009). *Distortions to Agricultural Incentives in Africa.* World Bank, Washington, DC.

Anderson, K. and Nelgen, S. (2013). Updated national and global estimates of distortions to agricultural incentives, 1955 to 2011. Technical report, World Bank, Washington, DC.

Anderson, K., Rausser, G., and Swinnen, J. (2013). Political Economy of Public Policies: Insights from Distortions to Agricultural and Food Markets. *Journal of Economic Literature*, 51(2):423–477.

Bale, M. D. and Lutz, E. (1981). Price Distortions in Agriculture and Their Effects: An International Comparison. *American Journal of Agricultural Economics*, 63(1):8–22.

Barrett, C. B. (2008). Smallholder market participation: Concepts and evidence from eastern and southern Africa. *Food Policy*, 33(4):299–317.

Bautista, R. M., Robinson, S., Tarp, F., and Wobst, P. (2001). Policy Bias and Agriculture: Partial and General Equilibrium Measures. *Review of Development Economics*, 5(1):89–104.

Beck, T., Clarke, G., Groff, A., and Keefer, P. (2001). New tools in comparative political economy: The Database of Political Institutions. *The World Bank Economic Review*, 15(1):165–176.

Binswanger, H. P., Ruttan, V. W., and Ben-Zion, U. (1978). *Induced innovation: technology, institutions, and development.* John Hopkins University Press, Baltimore.

Brandaõ, A. (1993). Implications of agricultural trade liberalization for the developing countries. *Agricultural Economics*, 8(4):313–343.

Calderón, C. and Servén, L. (2004). *The effects of infrastructure development on growth and income distribution.* Access, Quality, Costs, and Policy Reform. World Bank Publications.

De Gorter, H. and Swinnen, J. (2002). Political economy of agricultural policy. In *Agricultural and Food Policy*, pages 1893–1943. Elsevier.

Deininger, K. and Jin, S. (2006). Tenure security and land-related investment: Evidence from Ethiopia. *European Economic Review*, 50(5):1245–1277.

Dell, M., Jones, B. F., and Olken, B. A. (2009). Temperature and Income: Reconciling New Cross-Sectional and Panel Estimates. *American Economic Journal: Macroeconomics*, pages 66–95.

Denning, G., Kabambe, P., Sanchez, P., Malik, A., and Flor, R. (2009). Input subsidies to improve smallholder maize productivity in Malawi: Toward an African Green Revolution. *PLoS Biol*, 7(1):e1000023.

Dennis, B. N. and İşcan, T. B. (2011). Agricultural Distortions, Structural Change, and Economic Growth: A Cross-Country Analysis. *American Journal of Agricultural Economics*, 93(3):881–902.

Dougherty, C. (2011). *Introduction to Econometrics*. Oxford University Press.

FAO. Fao statistical databse.

FAO (2009). How to Feed the World 2050.

Fulginiti, L. E. and Perrin, R. K. (1993). Prices and Productivity in Agriculture. *The Review of Economics and Statistics*, 75(3):471.

Fulginiti, L. E., Perrin, R. K., and Yu, B. (2004). Institutions and agricultural productivity in Sub-Saharan Africa. *Agricultural Economics*, 31(2-3):169–180.

Henning, C. H. C. A. and Struve, C. (2007). Electoral systems, postelection bargaining and special interest politics in parliamentary systems: The case of agricultural protection. *Topics in Analytical Political Economy*, 17:43–82.

Hopenhayn, H. A. and Muniagurria, M. E. (1996). Policy Variability and Economic Growth. *The Review of Economic Studies*, 63(4):611–625.

Hsiang, S. M., Meng, K. C., and Cane, M. A. (2011). Civil conflicts are associated with the global climate. *Nature*, 476(7361):438–441.

IMF. World economic outlook 2015.

Jayne, T. S., Govereh, J., Wanzala, M., and Demeke, M. (2003). Fertilizer market development: a comparative analysis of Ethiopia, Kenya, and Zambia. *Food Policy*, 28(4):293–316.

Johnson, D. G. and Antle, J. M. (1996). *The economics of agriculture. Vol 2. Papers in honor of D. Gale Johnson*. University of Chicago Press.

Kornai, J. (1986). The soft budget constraint. *Kyklos*.

Krueger, A. O., Schiff, M., and Valdés, A. (1988). Agricultural Incentives in Developing Countries: Measuring the Effect of Sectoral and Economywide Policies. *The World Bank Economic Review*, 2(3):255–271.

Minot, N. (2010). *Transmission of world food price changes to markets in Sub-Saharan Africa*. International Food Policy Research Institute, Washington, DC.

Olper, A. (2001). Determinants of Agricultural Protection: The Role of Democracy and Institutional Setting Alessandro Olper. *Journal of Agricultural Economics*, 52(2):75–92.

Olper, A., Raimondi, V., and Swinnen, J. (2007). Constitutional rules and redistributive policy: evidence from global agricultural protection. *Pue and Piec Working Paper Series*.

Persson, T. and Besley, T. (2009). The origins of state capacity : Property rights, taxation and policy. *The American Economic Review*, 99(4):1218–1244.

PolityIV. Polityiv project, political regime characteristics and transitions, 1800-2014.

Poulton, C., Kydd, J., and Dorward, A. (2006). Overcoming market constraints on pro-poor agricultural growth in Sub-Saharan Africa. *Development policy review*, 24(3):243–277.

Rakotoarisoa, M. A. (2011). The impact of agricultural policy distortions on the productivity gap: Evidence from rice production. *Food Policy*, 36(2):147–157.

Reardon, T., Barrett, C., and Kelly, V. (1999). Policy reforms and sustainable agricultural intensification in Africa. *Development policy review*, 17(4):375–395.

Reardon, T., Delgado, C., and Matlon, P. (1992). Determinants and effects of income diversification amongst farm households in Burkina Faso. *The Journal of Development Studies*, 28(2):264–296.

Rivera-Batiz, F. L. (2002). Democracy, governance, and economic growth: theory and evidence. *Review of Development Economics*, 6(2):225–247.

Rizov, M., Pokrivcak, J., and Ciaian, P. (2013). CAP subsidies and productivity of the EU farms. *Journal of Agricultural Economics*, 64(3):537–557.

Schlenker, W. and Lobell, D. B. (2010). Robust negative impacts of climate change on African agriculture. *Environmental Research Letters*, 5(1):014010.

Swinnen, J. F. M. (1994). A Positive Theory of Agricultural Protection. *American Journal of Agricultural Economics*, 76(1):1–14.

Swinnen, J. F. M. (2010). The Political Economy of Agricultural and Food Policies: Recent Contributions, New Insights, and Areas for Further Research. *Applied Economic Perspectives and Policy*, 32(1):33–58.

Tanzi, V. (1992). *Structural factors and tax revenue in developing countries: a decade of evidence*, chapter 12. Cambridge University Press.

Thiele, R. (2003). Price Incentives, Non-price Factors and Agricultural Production in Sub-Saharan Africa: A Cointegration Analysis. *African Development Review*, 15(2-3):425–438.

Verbeek, M. (2013). *A Guide to Modern Econometrics*. John Wiley & Sons.

WorldBank. Agricultural productivity programs.

WorldBank. World bank database.

WorldBank. World development indicators.

# Appendix

## A. Multicollinearity

**Table A1:** Cross-correlation table equation 1 for aggregate agriculture

| Variables | nra | land | capital | rainfall | temp | fert | dem | pv | roads | tel |
|---|---|---|---|---|---|---|---|---|---|---|
| nra | 1.000 | | | | | | | | | |
| land | -0.086 | 1.000 | | | | | | | | |
| capital | 0.013 | 0.096 | 1.000 | | | | | | | |
| rainfall | -0.047 | 0.408 | -0.360 | 1.000 | | | | | | |
| temperature | 0.041 | 0.280 | -0.333 | 0.171 | 1.000 | | | | | |
| fertilizer | -0.086 | -0.439 | 0.297 | -0.333 | -0.568 | 1.000 | | | | |
| democracy | 0.230 | -0.125 | 0.155 | -0.049 | -0.087 | 0.174 | 1.000 | | | |
| policyvar | 0.144 | -0.045 | 0.044 | 0.008 | 0.062 | 0.033 | 0.138 | 1.000 | | |
| roads | -0.011 | 0.289 | 0.139 | 0.258 | -0.316 | 0.221 | 0.169 | 0.046 | 1.000 | |
| telephone | 0.094 | -0.281 | 0.533 | -0.459 | -0.595 | 0.623 | 0.271 | 0.039 | 0.508 | 1.000 |

Land and capital are in logs.

Correlation matrices for other crops are not reported, no correlations above |0.8| occur.

**Table A2:** Cross-correlation table equation 2

| Variables | GDPpc | growth | pop | ruralpop | inflation | debt | dem | fss | election | party |
|---|---|---|---|---|---|---|---|---|---|---|
| GDP per capita | 1.000 | | | | | | | | | |
| GDP growth | 0.024 | 1.000 | | | | | | | | |
| population | 0.168 | 0.088 | 1.000 | | | | | | | |
| rural population | -0.694 | 0.011 | -0.328 | 1.000 | | | | | | |
| inflation | -0.125 | -0.102 | 0.121 | 0.056 | 1.000 | | | | | |
| debt | -0.090 | -0.064 | 0.128 | -0.265 | 0.099 | 1.000 | | | | |
| democracy | 0.149 | 0.051 | 0.195 | -0.272 | -0.033 | 0.002 | 1.000 | | | |
| food self-sufficiency | 0.242 | -0.131 | -0.020 | -0.489 | -0.101 | 0.028 | -0.072 | 1.000 | | |
| election | -0.016 | 0.004 | -0.020 | -0.001 | 0.018 | 0.025 | 0.099 | -0.021 | 1.000 | |
| party orientation | -0.264 | -0.115 | -0.380 | 0.195 | -0.125 | 0.310 | -0.225 | 0.269 | -0.071 | 1.000 |

GDP, population, inflation, and debt are in logs

## B. Heteroskedasticity

### Equation I

*Modified Wald test for groupwise heteroskedasticity (xttest3):*
H0: homoskedasticity
H1: heteroskedasticity
$\chi^2(22) = 224.63$
Prob $\chi^2 = 0.0000$
The null-hypothesis can be rejected; there is heteroskedasticity in the data. Robust standard errors are needed.

### Equation II

*Modified Wald test for groupwise heteroskedasticity (xttest3):*
H0: homoskedasticity
H1: heteroskedasticity
$\chi^2(22) = 14832.51$
Prob $\chi^2 = 0.0000$
The null-hypothesis can be rejected; there is heteroskedasticity in the data. Robust standard errors are needed.

## C. Fixed or random effects: Hausman test

### Equation I

The hausman test tests whether a random effects model is more appropriate than a fixed effects model.
H0: difference in coefficients of the two models is not systematic.
coefficients FE: consistent under H0 and H1
coefficients RE: inconsistent under H1, efficient under H0
$\chi^2(6) = 19.79$
Prob $\chi^2 = 0.0060$
The H0 is rejected, H1 holds and RE coefficients are inconsistent.

### Equation II

The hausman test tests whether a random effects model is more appropriate than a fixed effects model.
H0: difference in coefficients of the two models is not systematic.
coefficients FE: consistent under H0 and H1
coefficients RE: inconsistent under H1, efficient under H0
$\chi^2(6) = 89.04$
Prob $\chi^2 = 0.0000$
The H0 is rejected, H1 holds and RE coefficients are inconsistent.

## D. Fixed effects equation I: model specifications

**Table D1:** Fixed effects model: effect of NRA on yield with lagged dependent

| | (1) Aggregate | (2) Cocoa | (3) Coffee | (4) Cotton | (5) Maize | (6) Rice | (7) Sugar | (8) Wheat |
|---|---|---|---|---|---|---|---|---|
| L.logyield | 0.385*** | 0.601*** | 0.480*** | 0.654*** | 0.388** | 0.609*** | 0.898*** | 0.358** |
| | (4.13) | (10.46) | (3.83) | (8.25) | (2.99) | (9.09) | (15.14) | (2.48) |
| L.nra | 0.0271 | 0.187** | 0.0863 | 0.178** | 0.0154 | 0.0298 | -0.000960 | -0.0131 |
| | (0.78) | (3.43) | (0.59) | (2.31) | (0.63) | (1.32) | (-0.03) | (-0.30) |
| log land per cap | -0.245*** | -0.239 | 0.0556 | -0.0576 | -0.616** | -0.209** | -0.208 | -0.300 |
| | (-3.20) | (-1.44) | (0.21) | (-0.44) | (-2.82) | (-2.34) | (-1.78) | (-1.28) |
| log capital per cap | 0.0489 | -0.358** | 0.0279 | -0.0332 | 0.438** | -0.203* | 0.0451 | -0.0353 |
| | (0.42) | (-2.89) | (0.09) | (-0.30) | (2.24) | (-1.89) | (0.37) | (-0.13) |
| rainfall | 0.00864 | 0.00150 | -0.00702 | 0.00720 | 0.0340 | 0.00194 | 0.000781 | 0.0447** |
| | (1.29) | (0.27) | (-0.51) | (0.73) | (1.75) | (0.33) | (0.18) | (2.38) |
| temperature | -0.00424 | 0.0954* | 0.00934 | -0.0279 | -0.148 | 0.0120 | -0.0421 | 0.0462 |
| | (-0.18) | (2.21) | (0.18) | (-0.56) | (-1.58) | (0.38) | (-1.02) | (1.21) |
| fertilizer | 0.155*** | 0.000496 | 0.0472 | -0.0865* | 0.144*** | 0.0332 | 0.0524 | -0.0629 |
| | (6.01) | (0.14) | (0.60) | (-2.02) | (3.70) | (1.15) | (1.43) | (-1.30) |
| Constant | 1.385 | -3.420* | 4.893 | 3.136 | 6.877** | 0.340 | 0.851 | 2.508 |
| | (1.50) | (-2.76) | (1.43) | (1.75) | (2.86) | (0.28) | (0.58) | (1.41) |
| Observations | 572 | 139 | 187 | 444 | 348 | 272 | 223 | 218 |
| $R^2$ | 0.284 | 0.738 | 0.226 | 0.499 | 0.317 | 0.514 | 0.837 | 0.294 |

NRA is crop-specific. Fertilizer is consumption for cocoa, dummy for others.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

**Table D2:** Fixed effects model: effect of NRA and policy switch on yield

| | (1) Aggregate | (2) Cocoa | (3) Coffee | (4) Cotton | (5) Maize | (6) Rice | (7) Sugar | (8) Wheat |
|---|---|---|---|---|---|---|---|---|
| L.nra | 0.00143 | 0.450*** | 0.181 | 0.333* | 0.0239 | 0.0694 | -0.154 | -0.0270 |
| | (0.03) | (5.07) | (0.68) | (2.10) | (0.98) | (1.42) | (-1.14) | (-0.40) |
| log land | -0.401*** | -0.469 | 0.109 | -0.329 | -0.983*** | -0.537* | -1.183 | -0.530 |
| per cap | (-3.10) | (-1.77) | (0.24) | (-1.46) | (-3.78) | (-2.22) | (-1.37) | (-1.38) |
| log capital | 0.133 | -1.185*** | 0.0558 | 0.265 | 0.936*** | -0.295 | 0.516 | 0.00276 |
| per cap | (0.73) | (-5.43) | (0.09) | (0.85) | (3.48) | (-1.25) | (1.85) | (0.01) |
| rainfall | 0.00827 | 0.00652 | -0.0100 | 0.0136 | 0.0318 | 0.00773 | 0.000430 | 0.0419* |
| | (1.30) | (0.75) | (-0.68) | (0.74) | (1.54) | (0.80) | (0.05) | (2.25) |
| temperature | 0.00744 | 0.169* | 0.00485 | 0.0206 | -0.161 | 0.0178 | 0.0434 | 0.0767 |
| | (0.23) | (2.51) | (0.08) | (0.23) | (-1.78) | (0.32) | (0.79) | (1.75) |
| fertilizer | 0.163*** | -0.00354 | 0.0587 | -0.249** | 0.114** | -0.0223 | 0.210 | -0.113** |
| | (5.06) | (-0.50) | (0.41) | (-2.87) | (2.47) | (-0.58) | (1.48) | (-2.35) |
| L.policyswi | -0.00832 | -0.120* | -0.139* | -0.0733* | 0.0124 | -0.00421 | 0.00969 | 0.0354 |
| | (-0.58) | (-2.53) | (-1.96) | (-1.94) | (0.40) | (-0.25) | (0.18) | (0.79) |
| L.nra*policy | -0.0124 | -0.116 | -0.729 | -0.131 | -0.145** | -0.111 | -0.203 | -0.117* |
| | (-0.14) | (-0.79) | (-1.15) | (-0.70) | (-2.58) | (-1.06) | (-1.13) | (-1.98) |
| Constant | 2.317** | -7.879* | 9.720 | 7.768* | 11.47*** | 2.728 | 5.891 | 3.855* |
| | (2.14) | (-2.31) | (1.63) | (2.05) | (6.25) | (1.54) | (0.79) | (1.90) |
| Observations | 571 | 139 | 185 | 443 | 346 | 271 | 220 | 215 |
| $R^2$ | 0.138 | 0.580 | 0.047 | 0.113 | 0.217 | 0.231 | 0.202 | 0.203 |

NRA is crop-specific. Fertilizer is consumption for cocoa, dummy for others.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

## E. Fixed effects equation II: model specifications

**Table E1:** Fixed effects model: determinants of NRA, including monopoly power

| | (1) Aggregate | (2) Cocoa | (3) Coffee | (4) Cotton | (5) Maize | (6) Rice | (7) Sugar |
|---|---|---|---|---|---|---|---|
| L.log gdp | 0.474*** | 0.0191 | 0.307** | 0.0479 | -0.318 | 0.431 | 0.257 |
| per cap | (4.76) | (0.13) | (2.34) | (0.23) | (-1.37) | (1.59) | (0.45) |
| L.gdpgrowth | 0.00234 | 0.00208 | -0.00113 | -0.000707 | -0.00250 | -0.00487 | -0.00426 |
| | (0.58) | (0.55) | (-0.96) | (-0.26) | (-0.73) | (-0.67) | (-0.55) |
| L.log | 0.569** | -0.0757 | 0.249 | 0.0818 | 0.891** | -0.134 | 0.199 |
| population | (2.97) | (-0.97) | (1.55) | (0.34) | (3.09) | (-0.43) | (0.47) |
| L.share | 0.633 | -0.608* | -0.0767 | 1.581 | 4.028** | -2.549 | 0.447 |
| rural pop | (0.98) | (-2.10) | (-0.12) | (1.36) | (3.10) | (-1.41) | (0.23) |
| L.log | 0.00234 | -0.0558** | -0.0290** | -0.0274 | -0.0230 | -0.0935 | -0.0142 |
| inflation | (0.05) | (-3.14) | (-2.15) | (-0.56) | (-0.74) | (-1.40) | (-0.22) |
| L.log debt | 0.102* | 0.0811 | 0.0200 | 0.0315 | 0.102* | 0.00571 | 0.0607 |
| | (2.62) | (1.26) | (0.85) | (0.49) | (2.16) | (0.05) | (1.27) |
| L.monopoly | 0.00277 | -0.0207 | 0.0524 | -0.246 | -1.236*** | -0.267 | -0.0939 |
| | (0.71) | (-0.78) | (1.64) | (-0.91) | (-4.87) | (-1.82) | (-0.17) |
| Constant | -13.79** | 1.127 | -6.436** | -2.474 | -15.67** | 1.663 | -5.461 |
| | (-4.13) | (0.84) | (-2.85) | (-0.63) | (-3.21) | (0.36) | (-0.76) |
| Observations | 182 | 202 | 326 | 372 | 273 | 239 | 229 |
| $R^2$ | 0.388 | 0.179 | 0.226 | 0.044 | 0.199 | 0.126 | 0.041 |

NRA and monopoly are crop-specific.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

**Table E2:** Fixed effects model: determinants of NRA, including democracy

| | (1) Aggregate | (2) Cocoa | (3) Coffee | (4) Cotton | (5) Maize | (6) Rice | (7) Sugar | (8) Wheat |
|---|---|---|---|---|---|---|---|---|
| L.log gdp per cap | 0.0484 | 0.445** | 0.196 | 0.301** | 0.130 | -0.450* | 0.512 | 0.420 |
| | (0.43) | (4.16) | (0.77) | (2.52) | (0.65) | (-1.95) | (1.62) | (1.25) |
| L.gdpgrowth | 0.00203 | 0.00280 | 0.00103 | -0.000513 | -0.00223 | 0.00108 | -0.00822 | -0.00593 |
| | (1.38) | (0.62) | (0.26) | (-0.42) | (-0.65) | (0.31) | (-1.18) | (-0.72) |
| L.log population | 0.154 | 0.427* | -0.224 | 0.0751 | 0.0431 | 0.631 | -0.256 | -0.103 |
| | (1.24) | (2.26) | (-0.90) | (0.43) | (0.14) | (1.73) | (-0.64) | (-0.28) |
| L.share rural pop | 0.676 | 0.359 | -0.806* | -0.563 | 1.747 | 4.600** | -2.728 | -0.0913 |
| | (1.25) | (0.67) | (-2.27) | (-0.84) | (1.17) | (3.02) | (-1.48) | (-0.05) |
| L.log inflation | -0.0229 | 0.00276 | -0.0513** | -0.0351*** | -0.0141 | -0.00647 | -0.0750 | 0.00853 |
| | (-1.59) | (0.07) | (-3.16) | (-3.08) | (-0.26) | (-0.14) | (-1.08) | (0.12) |
| L.log debt | 0.0430 | 0.104* | 0.101 | 0.0201 | 0.0300 | 0.103* | -0.0174 | 0.0699 |
| | (1.50) | (2.52) | (1.34) | (0.81) | (0.57) | (2.13) | (-0.16) | (1.62) |
| L.democracy | 0.0103** | 0.00595 | 0.0148 | 0.00735 | 0.0123* | 0.0148 | 0.0190 | 0.0223* |
| | (2.37) | (0.65) | (1.01) | (0.76) | (2.10) | (1.50) | (1.24) | (2.15) |
| Constant | -3.547 | -11.01** | 2.575 | -3.090 | -2.598 | -11.16* | 3.218 | -1.099 |
| | (-1.54) | (-3.20) | (0.89) | (-1.18) | (-0.56) | (-1.97) | (0.56) | (-0.21) |
| Observations | 554 | 182 | 200 | 338 | 370 | 273 | 239 | 227 |
| $R^2$ | 0.111 | 0.394 | 0.225 | 0.221 | 0.031 | 0.139 | 0.119 | 0.057 |

NRA is crop-specific.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

**Table E3:** Fixed effects model: determinants of NRA, including constitutional factors

| | (1) Aggregate | (2) Cocoa | (3) Coffee | (4) Cotton | (5) Maize | (6) Rice | (7) Sugar | (8) Wheat |
|---|---|---|---|---|---|---|---|---|
| L.log gdp | -0.0496 | 0.544** | 0.0158 | 0.222 | -0.0428 | -0.568*** | 0.326 | 0.0322 |
| per cap | (-0.41) | (3.45) | (0.07) | (1.48) | (-0.31) | (-5.01) | (0.97) | (0.07) |
| L.gdpgrowth | 0.000350 | 0.00206 | 0.000805 | -0.00187 | -0.00789* | -0.00835* | -0.00841 | -0.00510 |
| | (0.24) | (0.40) | (0.18) | (-1.32) | (-2.09) | (-1.88) | (-0.93) | (-0.60) |
| L.log | 0.242 | 0.745** | 0.0515 | 0.163 | -0.0102 | 0.529** | -0.0703 | 0.191 |
| population | (1.70) | (4.14) | (0.52) | (0.88) | (-0.04) | (2.85) | (-0.19) | (0.38) |
| L.share | 0.450 | 0.814 | -0.540 | -0.526 | 1.493 | 3.562** | -3.192 | 0.0971 |
| rural pop | (0.74) | (1.46) | (-1.48) | (-0.71) | (1.00) | (2.84) | (-1.60) | (0.05) |
| L.log | -0.0204 | 0.0317 | -0.0447** | -0.0327** | -0.0615 | -0.0474 | -0.0745 | -0.0161 |
| inflation | (-1.58) | (0.78) | (-2.76) | (-2.76) | (-0.82) | (-1.19) | (-0.89) | (-0.23) |
| L.log debt | 0.0235 | 0.104** | 0.0926 | -0.00153 | 0.000604 | 0.0579 | -0.0434 | 0.0123 |
| | (0.79) | (3.83) | (1.22) | (-0.06) | (0.02) | (1.66) | (-0.38) | (0.32) |
| election | -0.0206 | -0.0241 | 0.0247 | -0.00137 | -0.00699 | -0.0695* | -0.0235 | 0.0541 |
| | (-1.13) | (-0.59) | (0.64) | (-0.04) | (-0.11) | (-2.06) | (-0.40) | (0.68) |
| Constant | -4.163* | -17.37*** | -1.166 | -3.982 | -0.157 | -7.695** | 1.736 | -3.386 |
| | (-1.85) | (-5.97) | (-0.39) | (-1.38) | (-0.04) | (-2.29) | (0.28) | (-0.51) |
| Observations | 505 | 163 | 185 | 321 | 325 | 250 | 211 | 201 |
| $R^2$ | 0.076 | 0.428 | 0.200 | 0.180 | 0.036 | 0.144 | 0.115 | 0.013 |

NRA is crop-specific.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

**Table E4:** Fixed effects model: determinants of NRA, including food self-sufficiency

| | (1) Aggregate | (2) Cocoa | (3) Coffee | (4) Cotton | (5) Maize | (6) Rice | (7) Sugar | (8) Wheat |
|---|---|---|---|---|---|---|---|---|
| L.log gdp per cap | -0.00694 | 0.816* | 0.529 | 0.775*** | -0.239 | -0.418* | 0.427 | -1.453*** |
| | (-0.04) | (2.38) | (1.15) | (2.95) | (-0.58) | (-1.89) | (0.48) | (-5.47) |
| L.gdpgrowth | 0.00249 | 0.00611 | 0.00376 | -0.00320 | 0.00621 | -0.00530 | 0.00229 | 0.0118** |
| | (1.62) | (1.84) | (0.86) | (-0.87) | (1.77) | (-0.45) | (0.21) | (2.38) |
| L.log population | 0.0231 | -0.242 | -0.193 | -0.478* | -0.0170 | -0.0949 | -0.820 | 0.799 |
| | (0.17) | (-0.59) | (-0.46) | (-1.77) | (-0.11) | (-0.35) | (-1.09) | (1.30) |
| L.share rural pop | 0.0216 | -0.269 | -1.076 | -2.394* | 0.306 | 0.228 | -1.733 | -1.537* |
| | (0.04) | (-0.11) | (-1.65) | (-2.07) | (0.30) | (0.22) | (-1.05) | (-1.99) |
| L.log inflation | -0.0274*** | 0.0320 | -0.0348** | -0.0327** | -0.0659* | -0.0552 | -0.114 | -0.00916 |
| | (-3.08) | (0.73) | (-2.88) | (-2.82) | (-1.87) | (-1.42) | (-1.23) | (-0.16) |
| L.log debt | 0.00351 | 0.0271 | 0.203** | 0.0460 | 0.0658 | 0.0582 | -0.184 | -0.132 |
| | (0.10) | (0.84) | (2.68) | (1.20) | (1.07) | (0.93) | (-0.78) | (-1.19) |
| L.food self-sufficiency | 0.00375 | 0.00270 | 0.00763 | -0.00481 | 0.00946* | 0.00680 | 0.0149 | 0.000741 |
| | (1.16) | (0.42) | (1.64) | (-1.14) | (1.93) | (0.90) | (1.78) | (0.09) |
| Constant | -0.464 | -1.440 | -0.342 | 4.439 | 1.383 | 3.885 | 13.50 | -2.516 |
| | (-0.25) | (-0.14) | (-0.05) | (1.08) | (0.65) | (0.95) | (1.07) | (-0.26) |
| Observations | 337 | 90 | 122 | 220 | 206 | 157 | 148 | 144 |
| $R^2$ | 0.069 | 0.205 | 0.181 | 0.173 | 0.068 | 0.123 | 0.141 | 0.141 |

NRA is crop-specific.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

# F. IV-model: lagged dependent

**Table F1:** Lagged dependent IV-model with instruments population, rural population and agricultural value added

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| | Aggregate | Cocoa | Coffee | Cotton | Maize | Rice | Sugar | Wheat |
| L.nra | 0.962*** | 0.423* | 0.332 | 0.448** | -0.618*** | -0.192* | -0.377** | -0.301 |
| | (2.93) | (1.85) | (1.42) | (2.50) | (-2.86) | (-1.73) | (-2.46) | (-1.31) |
| L.yield | 0.541*** | 0.429*** | 0.465*** | 0.604*** | 0.417*** | 0.718*** | 0.783*** | 0.307** |
| | (7.06) | (2.98) | (4.72) | (9.30) | (4.07) | (8.56) | (11.03) | (2.42) |
| log land per cap | 0.0644 | -0.332** | 0.201 | 0.0265 | -0.759*** | -0.171 | -0.701*** | -0.648* |
| | (0.56) | (-2.37) | (0.81) | (0.21) | (-3.75) | (-1.30) | (-3.11) | (-1.94) |
| log capital per cap | 0.0767 | -0.591** | 0.191 | -0.0507 | 0.404 | -0.263 | 0.409* | 0.0467 |
| | (0.75) | (-2.13) | (0.69) | (-0.36) | (1.44) | (-1.45) | (1.86) | (0.21) |
| rainfall | 0.00767 | -0.00246 | -0.00321 | 0.0110 | 0.0291* | -0.00156 | 0.00617 | 0.0682** |
| | (1.11) | (-0.30) | (-0.31) | (1.06) | (1.87) | (-0.20) | (0.69) | (2.22) |
| temperature | -0.0621* | 0.0931** | 0.0355 | -0.0481 | -0.0787 | 0.0170 | 0.0831 | 0.0641 |
| | (-1.89) | (2.05) | (0.55) | (-1.19) | (-1.06) | (0.38) | (1.27) | (0.64) |
| fertilizer | 0.0496 | 0.00412 | 0.0375 | -0.149 | 0.239** | 0.0878 | -0.0760 | -0.0577 |
| | (0.47) | (0.79) | (0.46) | (-1.31) | (2.22) | (0.83) | (-1.23) | (-0.62) |
| N | 537 | 130 | 171 | 409 | 318 | 236 | 203 | 202 |
| j | 10.11 | 2.715 | 7.633 | 1.590 | 1.853 | 4.372 | 1.178 | 7.591 |
| jp | 0.00637 | 0.257 | 0.0220 | 0.452 | 0.396 | 0.112 | 0.555 | 0.0225 |

NRA is crop-specific. Fertilizer is consumption for cocoa, dummy for others.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

**Table F2:** Lagged dependent IV-model with instruments population and rural population

| | (1) Aggregate | (2) Cocoa | (3) Coffee | (4) Cotton | (5) Maize | (6) Rice | (7) Sugar | (8) Wheat |
|---|---|---|---|---|---|---|---|---|
| L.nra | 0.874*** | 1.229* | 0.854* | 0.719*** | -0.379** | -0.156* | -0.406*** | -0.911 |
| | (2.64) | (1.78) | (1.78) | (2.98) | (-2.38) | (-1.78) | (-2.74) | (-1.36) |
| L.yield | 0.557*** | 0.145 | 0.523*** | 0.604*** | 0.395*** | 0.647*** | 0.759*** | -0.0111 |
| | (7.72) | (0.46) | (4.64) | (9.72) | (4.31) | (9.31) | (10.38) | (-0.04) |
| log land per cap | 0.128 | 0.218 | 0.804* | 0.173 | -0.770*** | -0.286*** | -0.795*** | -1.498 |
| | (1.00) | (0.54) | (1.76) | (1.19) | (-4.69) | (-2.61) | (-3.06) | (-1.52) |
| log capital per cap | -0.00492 | -0.765* | 0.200 | -0.149 | 0.461** | -0.237 | 0.509** | 0.256 |
| | (-0.05) | (-1.76) | (0.62) | (-0.96) | (2.01) | (-1.56) | (2.08) | (0.64) |
| rainfall | 0.00901 | -0.00907 | -0.00999 | 0.00591 | 0.0288** | -0.000510 | 0.00655 | 0.0978* |
| | (1.39) | (-0.70) | (-0.76) | (0.51) | (2.29) | (-0.07) | (0.71) | (1.90) |
| temperature | -0.0615* | 0.126 | 0.0301 | -0.0638 | -0.115* | 0.0190 | 0.0978 | 0.0858 |
| | (-1.88) | (1.48) | (0.39) | (-1.46) | (-1.83) | (0.49) | (1.47) | (0.54) |
| fertilizer | 0.0299 | 0.00137 | -0.0890 | -0.169 | 0.190* | 0.0758 | -0.0775 | -0.0722 |
| | (0.29) | (0.14) | (-0.91) | (-1.35) | (1.85) | (0.74) | (-1.31) | (-0.43) |
| N | 572 | 139 | 187 | 444 | 348 | 272 | 223 | 218 |
| j | 14.30 | 0.0209 | 6.572 | 0.0498 | 0.274 | 3.758 | 0.0515 | 1.798 |
| jp | 0.000156 | 0.885 | 0.0104 | 0.823 | 0.601 | 0.0526 | 0.821 | 0.180 |

NRA is crop-specific. Fertilizer is consumption for cocoa, dummy for others.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

# G. IV-model: hierarchical

**Table G1:** IV-model with hierarchical data and different instruments

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | logyield | logyield | logyield | logyield |
| L.nra | 0.0774 | 0.0751 | 0.132 | -0.851*** |
| | (0.78) | (0.76) | (1.23) | (-3.25) |
| loglandpc | -0.561*** | -0.560*** | -0.496*** | -1.101*** |
| | (-4.24) | (-4.23) | (-3.61) | (-7.16) |
| logcapita c | 0.250 | 0.248 | 0.256* | 0.476*** |
| | (1.61) | (1.61) | (1.68) | (5.21) |
| rainfall | -0.000818 | -0.000650 | -0.00173 | 0.0195*** |
| | (-0.12) | (-0.09) | (-0.24) | (2.88) |
| temp | -0.0144 | -0.0130 | -0.00779 | 0.0655** |
| | (-0.36) | (-0.33) | (-0.19) | (2.08) |
| fertdummy | -0.121** | -0.119** | -0.129** | -0.00118 |
| | (-2.17) | (-2.16) | (-2.30) | (-0.02) |
| N | 605 | 605 | 570 | 1669 |
| j | 29.35 | 29.41 | 34.44 | 85.26 |
| jp | 0.000000423 | 0.00000183 | 0.000000160 | 3.07e-19 |

*t* statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$